



**HAL**  
open science

# Caractérisation des populations cellulaires normales et tumorales et définition d'épitopes partagés à partir des Rétrovirus Endogènes Humains (HERVs)

Vincent Alcazer

► **To cite this version:**

Vincent Alcazer. Caractérisation des populations cellulaires normales et tumorales et définition d'épitopes partagés à partir des Rétrovirus Endogènes Humains (HERVs). Immunologie. Université de Lyon, 2021. Français. NNT : 2021LYSE1203 . tel-03847755

**HAL Id: tel-03847755**

**<https://theses.hal.science/tel-03847755>**

Submitted on 10 Nov 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



N°d'ordre NNT : 2021LYSE1203

**THESE de DOCTORAT DE L'UNIVERSITE DE LYON**  
opérée au sein de  
**l'Université Claude Bernard Lyon 1**

**Ecole Doctorale N° 340**  
**(Biologie Moléculaire Intégrative et Cellulaire)**

**Spécialité de doctorat** : Aspects moléculaires et cellulaires de la biologie  
**Discipline** : Immunologie

Soutenue publiquement le 15/10/2021, par :  
**Vincent Alcazer**

---

**Caractérisation des populations  
cellulaires normales et tumorales et  
définition d'épitopes partagés à partir  
des Rétrovirus Endogènes Humains  
(HERVs)**

---

Devant le jury composé de :

BACHY Emmanuel  
PORTEU Françoise  
ITZYKSON Raphaël  
LABARRIERE Nathalie  
DEPIL Stéphane

PU-PH, HCL, Lyon  
DR INSERM, IGR, Paris  
PU-PH, AP-HP, Paris  
DR INSERM, CRCINA U892, Nantes  
PU, CRCL U1052, Lyon

Président  
Rapporteuse  
Rapporteur  
Examinatrice  
Directeur de thèse









# UNIVERSITE CLAUDE BERNARD – LYON 1

---

Président de l'Université	M. Frédéric FLEURY
Président du Conseil Académique	M. Hamda BEN HADID
Vice-Président du Conseil d'Administration	M. Didier REVEL
Vice-Président du Conseil des Etudes et de la Vie Universitaire	M. Philippe CHEVALLIER
Vice-Président de la Commission de Recherche	M. Petru MIRONESCU
Directeur Général des Services	M. Pierre ROLLAN

## COMPOSANTES SANTE

Département de Formation et Centre de Recherche en Biologie Humaine	Directrice : Mme Anne-Marie SCHOTT
Faculté d'Odontologie	Doyenne : Mme Dominique SEUX
Faculté de Médecine et Maïeutique Lyon Sud - Charles Mérieux	Doyenne : Mme Carole BURILLON
Faculté de Médecine Lyon-Est	Doyen : M. Gilles RODE
Institut des Sciences et Techniques de la Réadaptation (ISTR)	Directeur : M. Xavier PERROT
Institut des Sciences Pharmaceutiques et Biologiques (ISBP)	Directrice : Mme Christine VINCIGUERRA

## COMPOSANTES & DEPARTEMENTS DE SCIENCES & TECHNOLOGIE

Département Génie Electrique et des Procédés (GEP)	Directrice : Mme Rosaria FERRIGNO
Département Informatique	Directeur : M. Behzad SHARIAT
Département Mécanique	Directeur M. Marc BUFFAT
Ecole Supérieure de Chimie, Physique, Electronique (CPE Lyon)	Directeur : Gérard PIGNAULT
Institut de Science Financière et d'Assurances (ISFA)	Directeur : M. Nicolas LEBOISNE
Institut National du Professorat et de l'Education	Administrateur Provisoire : M. Pierre CHAREYRON
Institut Universitaire de Technologie de Lyon 1	Directeur : M. Christophe VITON
Observatoire de Lyon	Directrice : Mme Isabelle DANIEL
Polytechnique Lyon	Directeur : Emmanuel PERRIN
UFR Biosciences	Administratrice provisoire : Mme Kathrin GIESELER
UFR des Sciences et Techniques des Activités Physiques et Sportives (STAPS)	Directeur : M. Yannick VANPOULLE
UFR Faculté des Sciences	Directeur : M. Bruno ANDRIOLETTI

# REMERCIEMENTS

---

Je tiens tout d'abord à adresser mes plus sincères remerciements aux membres du jury : au **Pr. Emmanuel BACHY** pour me faire l'honneur d'en être le président ; aux **Prs. Françoise PORTEU et Raphaël ITZYKSON** pour avoir accepté d'évaluer cette thèse en tant que rapporteurs et fourni une lecture exhaustive et intéressée du manuscrit. Votre expertise et votre renom dans vos domaines respectifs représentent une contribution capitale garante de la validité globale de mon travail ; A la **Dr. Nathalie LABARRIERE** pour avoir volontiers accepté d'évaluer cette thèse en tant qu'examinatrice.

Je remercie également l'ensemble de nos collaborateurs ayant participé de près ou de loin à ce projet : les **Prs. Tobias HEROLD et Wolfgang HIDDEMANN** ainsi que les **Drs. Klaus METZELER et Aarif BATCH** du groupe Allemand AMLCG pour le partage de leurs données RNAseq et leur réactivité ; La plateforme de bio-informatique Gilles Thomas menée par le **Dr. Alain VIARI** et tout particulièrement à **Laurie TONON**, pour leurs multiples contributions et leur expertise bio-informatique sur l'ensemble du projet ; L'équipe du **Pr. Guy SAUVAGEAU** pour l'accès aux données LEUCEGENE ; Merci également aux **Drs. Nicolas DULPHY et François MALLET** pour leurs précieux conseils, leur aide et leur bienveillance tout au long des comités de suivi de thèse.

Je remercie chaleureusement mon directeur de thèse le **Pr. Stéphane DEPIL** pour avoir tout de suite cru en moi et m'avoir véritablement lancé dans la recherche fondamentale. Merci de m'avoir laissé cette grande liberté m'ayant permis de m'épanouir pleinement sur des domaines très variés au cours de ces 4 années. Ton enthousiasme, tes convictions ainsi que ta faculté à ne jamais baisser les bras sont les principaux moteurs de notre groupe.

Un grand merci à la **Dr. Paola BONAVENTURA** pour sa contribution majeure à ce travail. Le projet n'aurait probablement jamais aussi bien pris sans toi. Ta motivation, ta persévérance ainsi que ta capacité à t'imposer dans un environnement qui était celui de notre groupe au départ force l'admiration. Merci de m'avoir accueilli et formé à la culture cellulaire dans cette période difficile. Bien que la pipette ne soit décidément pas mon instrument de prédilection, je garde de très bons souvenirs de ces journées de manipulations communes à suivre tes protocoles « à l'italienne » !

Je remercie également de façon particulière le **Dr. Christophe CAUX** pour m'avoir accueilli au sein de son laboratoire. Merci de m'avoir fait confiance depuis le M2, de m'avoir laissé le champ libre et de nous avoir toujours soutenu malgré le côté « *outsider* » de notre groupe. Ta grande rigueur scientifique associée à ta vaste culture ont constitué des éléments majeurs autant pour l'évaluation régulière de mon travail que pour ma formation personnelle.

Merci à l'ensemble des membres du groupe de Stéphane, qu'ils soient académiques ou Ervaccine : à **Virginie et Emilie** nos deux ingénieures de choc qui permettent aux belles idées de prendre vie sous forme de solides données ; à **Nicolas, Yann et Rasha**, nouveaux cadres du groupe dont l'implication aux côtés de Paola garantira sans aucun doute la pérennité de la startup ; A **Estelle et Clémentine** qui commencent des travaux prometteurs au sein de cet environnement fertile.

Je présente également mes plus sincères remerciements à l'ensemble des membres et ex-membres de l'équipe CAUX qui m'ont accueilli et soutenu pour toute la durée de ce travail : Aux PIs de l'équipe **Marie-Cécile, Nathalie, Jenny, Christine, Bertrand et Yenkel** pour leur présence, leur disponibilité et leur aide inconditionnelle ; A mes co-thésards, garants de l'esprit d'entraide et de convivialité tout au long de ces années : A **Pacôme**, mon premier partenaire de bio-informatique, de craquage et de point sport/nutrition Yuka sans qui les discussions sur le cosmos auraient sans doute été un brin moins passionnées ; A **Anaïs**, ma référente Spotify et soirée, dont la classe et la culture scientifique permettent aussi bien de publier dans Immunity que dans Santé & Nutrition magazine ; A **Yasmine** pour ces nombreuses discussions passionnantes sur un fond musical d'Idir ; A **Anthony** pour sa disponibilité et ses explications techniques claires et toujours passionnées (quoiqu'un peu trop exhaustives ?!) ; Aux générations précédentes notamment **Marion, Margaux et Coline** qui ont su nous montrer la voie, tout comme à la relève **Valentin, Candice et Manuela** dont l'énergie, la motivation et la curiosité scientifique me rassurent quant à l'avenir de la recherche Française ; Aux petits derniers **Justine et Antoine** qui prendront j'en suis sûr à merveille la suite de tout ce beau monde ; A **Yamila** qui a su illuminer l'open-space chaque jour sans jamais flancher ; A **Maud et Sarah** nos expertes bioinformatiques / analyses d'images qui rendent ce monde accessible aux moldus ; Aux ingénieurs cadres de cette équipe **Anne-Claire, Céline, Lyvia, Justine, Pauline et Laurie** qui tiennent la barre et gardent le cap en préservant l'esprit d'équipe à travers ces générations de doctorants et post-doctorants.

Ces années ne se seraient pas déroulés dans d'aussi bonnes conditions sans mes amis, soupape indispensable de décompression : A **Mathieu et Eloïse**, toujours présents pour moi et sur qui je sais pouvoir compter jusqu'au bout ; Aux **beaubofs**, toujours partant et motivés en toute situation, que ce soit pour un week end ski, randonnée, festival ou simplement pour se retrouver autour d'un thé un dimanche après-midi pluvieux (autour d'un risque Game of Thrones ?). Je me réjouis fortement de l'agrandissement de la troupe à l'approche de la trentaine avec nos petits **Côme et Lison**, les premiers dignes héritiers d'une grande lignée ! Une mention spéciale à **Pierre Bouchard** sans qui ce travail n'aurait jamais été possible ; Aux **Grenoblois** dispersés aux quatre coins de la France mais toujours présents pour les grandes occasions ; Aux **amis d'internat Lyonnais** qui me soutiennent dans mes projets et font que l'internat à 30 ans se passe toujours bien ; aux **Vizillois** qui restent présents malgré nos chemins différents, votre amitié simple et sincère garde une valeur inestimable à mes yeux. Gageons que **Maxime et Livia** sauront nous rendre plus assidus dans nos réunions biennuelles !

Je tiens enfin à remercier **ma famille**, soutien de première ligne et véritable modèle pour moi : A **mes parents**, toujours présents sans jamais faillir même dans les situations difficiles ; A **mon frère**, toujours à l'écoute et prêt à se mettre en quatre pour sa famille ; A mes **oncles, tantes, cousins et cousines** qui me rendent fier de ma famille.

A **Agathe** ma conjointe qui accompagne et illumine ma vie. Mille mercis pour ta présence et ton soutien sans faille, indispensables aussi bien à ma réussite professionnelle qu'à mon équilibre personnel. Un seul de tes sourires suffit à éliminer l'ensemble de mes problèmes.

# TABLE DES MATIERES

---

<b>ABREVIATIONS.....</b>	<b>10</b>
<b>PREFACE .....</b>	<b>13</b>
<b>INTRODUCTION .....</b>	<b>17</b>
I. Généralités sur l'immunité antitumorale.....	19
A. Les concepts d'immunosurveillance et d' <i>immunoediting</i> .....	19
B. Le cycle de reconnaissance immunitaire du cancer .....	20
C. Echappement immunitaire de la tumeur.....	21
II. Les antigènes tumoraux .....	22
A. Bref historique.....	23
B. Les différents types d'antigènes tumoraux.....	23
C. Aperçu des stratégies thérapeutiques.....	28
III. Les Rétrovirus Endogènes Humains (HERVs).....	31
A. Les éléments transposables (ETs) dans le génome humain .....	31
B. Origines et structure des HERVs.....	32
C. Régulation de l'expression des HERVs.....	34
D. Exaptation des ERVs dans le développement des mammifères .....	37
E. Rôle dans l'oncogenèse.....	39
F. Impact immunitaire .....	43
G. Quantification des HERVs : aspects techniques et bioinformatiques.....	51
Conclusions préliminaires.....	61
<b>OBJECTIFS.....</b>	<b>63</b>
<b>REVUES.....</b>	<b>67</b>
REVUE 1 - Cold Tumors: A Therapeutic Challenge for Immunotherapy .....	69
REVUE 2 - Nouvelles approches vaccinales en cancérologie .....	81
REVUE 3 - Neoepitopes-based vaccines: challenges and perspectives.....	91
REVUE 4 - Développement des CAR-T dans les tumeurs solides .....	99
REVUE 5 - Les LT à récepteur antigénique chimérique (CAR-T) : principes physiopathologiques, actualités et perspectives .....	111

REVUE 6 - Toward “off-the-shelf” allogeneic CAR T cells.....	123
REVUE 7 - Human Endogenous Retroviruses (HERVs): Shaping the Innate Immune Response in Cancers .....	135
<b>ARTICLES – RESULTATS.....</b>	<b>149</b>
ARTICLE 1 - Identification of shared tumor epitopes from endogenous retroviruses inducing high avidity cytotoxic T cells for cancer immunotherapy.....	150
ARTICLE 2 - HERVs characterize the different subpopulations of normal and leukemic cells and represent a source of shared epitopes for cancer immunotherapy .....	206
<b>DISCUSSION.....</b>	<b>247</b>
I. Cible thérapeutique.....	249
A. Spécificité et expression par le tissu sain.....	249
B. Sélection des épitopes .....	250
C. Immunogénicité.....	254
D. Applications potentielles .....	258
II. Biomarqueur & quantification .....	260
A. Caractérisation des populations cellulaires.....	260
B. Signature prédictive de réponse thérapeutique.....	261
C. Approches <i>single-cell</i> .....	262
III. Proposition d’un modèle .....	262
<b>CONCLUSION .....</b>	<b>265</b>
<b>REFERENCES .....</b>	<b>269</b>
<b>ANNEXES .....</b>	<b>293</b>
Brevets déposés.....	294
<b>RÉSUMÉ.....</b>	<b>142</b>

# ABREVIATIONS

---

## 0-9

5hmc : 5-hydroxymethylcytosine

5mc : 5-methylcytosine

## A

ADN : Acide Désoxyribonucléique

ADNc : ADN complémentaire

AIM2 : *Absent In Melanoma 2*

AIRE : *Auto Immune Regulator*

ARN : Acide Ribonucléique

ARNm : ARN messenger

ATRA : Acide tout-trans rétinoïque

AZA : 5-azacytidine

## B

BCR : *B Cell Receptor* (Récepteur des Cellules B)

BLAT : *Blast-Like alignment Tool*

## C

CAR-T : *Chimeric Antigen Receptor T-cell* (Lymphocyte T à récepteur antigénique chimérique)

CGAs : *Cancer-Germline Antigens* (Antigènes tumoraux de la lignée germinale)

cGAS : *Cyclic GMP-AMP synthase*

CHIP-seq : *Chromatin-Immunoprecipitation sequencing* (Séquençage après immunoprécipitation de la chromatine)

CMH : Complexe Majeur d'Histocompatibilité

CPM : Comptes par million de *reads* mappés

CSF1R : *Colony-Stimulating Factor 1 Receptor*

CSH : Cellules Souches Hématopoïétiques

cTECs : *Cortical Thymic Epithelial Cells* (cellules corticales épithéliales thymiques)

## D

DAMPs : *Damage Associated Molecular Pattern* (Motifs moléculaires associés aux dégâts)

DNMT : *DNA Methyltransferase* (ADN-méthyltransférase)

DNMTi: Inhibiteur de DNMT

DRIPs : *Defective Ribosomal Products* (Produits Ribosomaux Défectueux)

dsRNA : *Double-stranded RNA* (ARN double brin)

## E

EBV : Virus d'Epstein-Barr

EST : *Expressed Sequence Tag* (Marqueurs de Séquences Exprimées)

ETs : Eléments Transposables

EZH2 : *Enhancer of Zeste Homolog 2*

## F

FPKM : Fragments par kilobase par millions de *reads* mappés

## H

H3Kxme3 : Triméthylation de l'histone H3 sur la lysine x  
HAT : Histone Acétylase  
HDAC : Histone Désacétylase  
HDACi : Inhibiteur de HDAC  
HERVs : *Human Endogenous Retroviruses* (Rétrovirus Endogènes Humains)  
HTLV1 : *Human T-Lymphotropic Virus 1*

## I

ICPI : *Immune Checkpoint Inhibitors* (Inhibiteurs de Points de Contrôle Immunitaires)  
IFN : Interféron  
IRF1 : *Interferon Regulatory Factor 1*  
ISD : *Immunosuppressive Domain* (Domaine immunosuppresseur)  
ISG : *Interferon-Stimulated Genes* (Gènes Stimulés par l'Interféron)

## K

KO : *Knock-Out*  
KoRVs : *Koala Endogenous Retroviruses*  
KRAB-ZFPs : *Krüppel-associated box domain zinc finger proteins*

## L

LAM : Leucémie Aiguë Myéloïde  
LB : Lymphocytes B  
LINE : *Long Interspersed Element*  
LMC : Leucémie Myéloïde Chronique  
LMMC : Leucémie Myéломonocytaire Chronique  
lncRNA : *long non-coding RNA* (ARN long non-codant)  
LPS : Lipopolysaccharide  
LSC : *Leukemic Stem Cell* (Cellule Souche Leucémique)  
LSD1 : *Lysine-specific histone demethylase 1*  
LT : Lymphocytes T  
LTR : *Long Terminal Repeat*

## M

MAIT : *Mucosal-Associated Invariant T cells* (Lymphocytes invariants associés aux muqueuses)  
MAPs : *MHC-I Associated peptides* (peptides associés au CMH de classe I)  
MDM2 : *Murine Double Minute 2*  
miARN : micro-ARN  
MLV : *Murine Leukemia Virus*  
mTECs : *Medullary Thymic Epithelial Cells* (cellules médullaires épithéliales thymiques)

## N

NF- $\kappa$ B : *Nuclear factor-kappa B*  
NGS : *Next Generation Sequencing* (Séquençage Nouvelle Génération)  
NK : *Natural Killer*  
NOD-SCID : *Non-obese diabetic - Severe Combined ImmunoDeficiency*

## O

ORF : *Open Reading Frame* (Cadre Ouvert de Lecture)

## P



pb : Paire de base  
PBMCs : *Peripheral Blood Mononuclear Cells* (Cellules sanguines périphériques mononuclées)  
PBS : *Primer Binding Site*  
PCR : *Polymerase Chain Reaction*  
PHA : *Phytohaemagglutinin*  
pRB : *Retinoblastoma Protein*

## R

RAG2 : *Recombination activating gene 2 protein*  
RPHM : Reads par centaine de millions de reads mappés  
RSEM : *RNA-Seq by Expectation-Maximization*  
RT : *Reverse Transcriptase* (Transcriptase Inverse)

## S

SASP : *Senescence-associated secretory phenotype* (phénotype sécrétoire associé à la senescence)  
SETDB1 : *SET domain bifurcated histone-lysine N-methyltransferase*  
shRNA : *short hairpin RNA* (petit ARN en épingle à cheveu)  
SINE : *Short Interspersed Element*  
SMD : Syndrome Myélodysplasique  
ssRNA : *Single-Stranded RNA* (ARN simple brin)  
STAT1 : *Signal transducer and activator of transcription 1*  
STING : *Stimulator of interferon genes protein*

## T

TAA : *Tumor-Associated Antigens* (Antigènes Associés à la Tumeur)  
TAD : *Topologically Associating Domains*  
TAP : *Transporter associated with Antigen Processing*  
TCGA : *The Cancer Genome Atlas*  
TCR : *T Cell Receptor* (Récepteur de Lymphocyte T)  
TET : *Ten-eleven Translocation*  
TILs : *Tumor Infiltrating Lymphocytes* (Lymphocytes Infiltrant la Tumeur)  
TLR : *Toll-like Receptor* (Récepteur Toll-like)  
TPM : Transcrits par million de reads mappés  
TRECs : *T cell receptor excision circles* (cercles d'excision du TCR)  
TSAs : *Tumor-Specific Antigens* (Antigènes Spécifiques de la Tumeur)

## V

VIH : Virus de l'Immunodéficience Humaine

## **PREFACE**

---



Le cancer est la deuxième cause de mortalité dans le monde, avec un nombre total de décès estimé à 9,6 millions par l'organisation mondiale de la santé (OMS) en 2018 (1). Face à un tel impact, l'oncologie est un domaine en pleine expansion avec de réels défis cliniques motivant la nécessité de nouvelles approches thérapeutiques. La remise du prix Nobel de médecine 2018 aux Dr. James Allison et Tasuku Honjo pour leurs travaux sur les inhibiteurs de *checkpoints* immunitaires (ICPI) vient ainsi marquer l'avènement de l'immunothérapie en tant que « 5<sup>ème</sup> pilier » thérapeutique, aux côtés de la chirurgie, des agents cytotoxiques, de la radiothérapie et des thérapies ciblées (2). Pourtant, l'idée d'utiliser le système immunitaire pour traiter les cancers n'est pas nouvelle, émergeant dès le 19<sup>ème</sup> siècle avec l'approche de William Coley (3), les premiers rapports de cas de patients traités par « infection » de la tumeur pouvant même remonter jusqu'à l'époque de l'Égypte ancienne (4). Malgré ces avancées récentes, de nombreux patients échappent encore à l'immunothérapie, avec par exemple un pourcentage global de patients répondeurs aux ICPI estimé autour de 12% en 2018 (5). Ce long temps de développement et ce pourcentage élevé de résistances illustrent ainsi toute la difficulté d'utiliser le système immunitaire pour traiter le cancer, soulignant le besoin de développer la recherche dans ce domaine pour mieux comprendre les mécanismes de résistance, identifier des facteurs prédictifs de réponse et proposer de nouvelles approches thérapeutiques.

Ce travail de thèse fait le point sur l'analyse des rétrovirus endogènes humains (HERVs) en tant que biomarqueurs originaux et cibles thérapeutiques prometteuses en immunothérapie. Après une introduction générale sur l'immunité antitumorale et les antigènes tumoraux, nous passerons en revue les différents mécanismes permettant à la tumeur d'échapper au système immunitaire. Nous verrons ensuite où en sont les différentes approches thérapeutiques basées sur des cibles issues de mutation des gènes dits « classiques », renforçant le besoin de nouvelles cibles tumorales. Nous pourrons alors introduire l'apport potentiel des HERVs dans ce domaine, en passant en revue les données actuelles concernant le contrôle de leur expression ainsi que leur impact génétique et immunitaire. Nous verrons également pourquoi la quantification de l'expression des HERVs représente un défi aussi bien analytique que bioinformatique.

Les principaux résultats seront ensuite présentés sous la forme de deux articles originaux : dans le premier article, nous avons établi une approche bioinformatique originale pour définir et prioriser des épitopes lymphocytaires T CD8<sup>+</sup> issus de HERVs exprimés dans les tumeurs solides. Dans le second, nous nous sommes concentrés sur l'analyse des HERVs chez les patients présentant une leucémie aiguë myéloïde (LAM), avec une quantification plus fine du rétrotranscriptome que nous retrouvons ainsi comme spécifique de chaque population cellulaire. Ces travaux ouvrent ainsi la voie de l'utilisation des HERVs comme antigènes tumoraux partagés et marqueurs d'identité cellulaire au fort potentiel pour de multiples applications basées sur le séquençage d'acide ribonucléique (ARN) (prédiction de réponse/rechute, signature cellulaire, classification de pathologies...). Nous finirons enfin par une discussion commentant ces résultats à la lumière des connaissances actuelles dans le domaine, en évoquant les défis et perspectives de l'exploitation des HERVs en immuno-oncologie.



# INTRODUCTION

---



## I. Généralités sur l'immunité antitumorale

### A. Les concepts d'immunosurveillance et d'*immunoediting*

Chaque jour, chacune des cellule de notre corps subit environ 20 000 événements à l'origine de dégâts sur son acide désoxyribonucléique (ADN) (6). Si la plupart de ces altérations sont rapidement réparées et n'auront aucun effet sur la cellule, certaines peuvent néanmoins persister et entraîner à terme une transformation cellulaire. Un des rôles du système immunitaire consiste ainsi à reconnaître et détruire ces cellules anormales avant qu'elles n'accumulent trop de dégâts et deviennent potentiellement cancéreuses. Ce concept d'immunosurveillance, proposé initialement par Macfarlane Burnet et Lewis Thomas dans les années 1950, a longtemps été débattu (7,8). De nombreuses études n'allaient en effet pas dans ce sens à l'époque, soulignant plutôt les potentiels effets pro-tumoraux de la réponse inflammatoire (9), le manque de signaux de danger associés à la tumeur pour pouvoir la rendre immunogène (10) ou encore la trop forte similarité avec le tissu normal (11).

Ce n'est qu'à partir des années 1990 avec l'émergence de modèles murins immunodéficients que les premières démonstrations du rôle clé du système immunitaire dans le contrôle tumoral ont pu être apportées. Les travaux initiés par l'équipe de Robert Schreiber, démontrant notamment une croissance tumorale accrue chez des souris présentant différents déficits de l'immunité adaptative auront permis de mettre en lumière le rôle important de cette réponse dans le contrôle antitumoral (12,13), démontrant au passage la présence d'antigènes tumoraux. Le fait qu'une tumeur puisse tout de même se développer malgré cette surveillance immunitaire a par la suite amené à repenser la théorie générale de l'immunosurveillance. Une étude clé (13), montrant que des tumeurs se développant chez des souris immunodéficientes sont plus immunogènes que celles se développant chez des souris immunocompétentes, introduit alors progressivement le concept d'« *immunoediting* » postulant que la pression immunitaire, bien que protégeant en partie de l'émergence de cellules tumorales, peut également façonner ces dernières et mener à la sélection de clones moins immunogènes (14).

Dans sa vision la plus récente, la surveillance immunitaire antitumorale est découpée en 3 phases schématiques (8) :

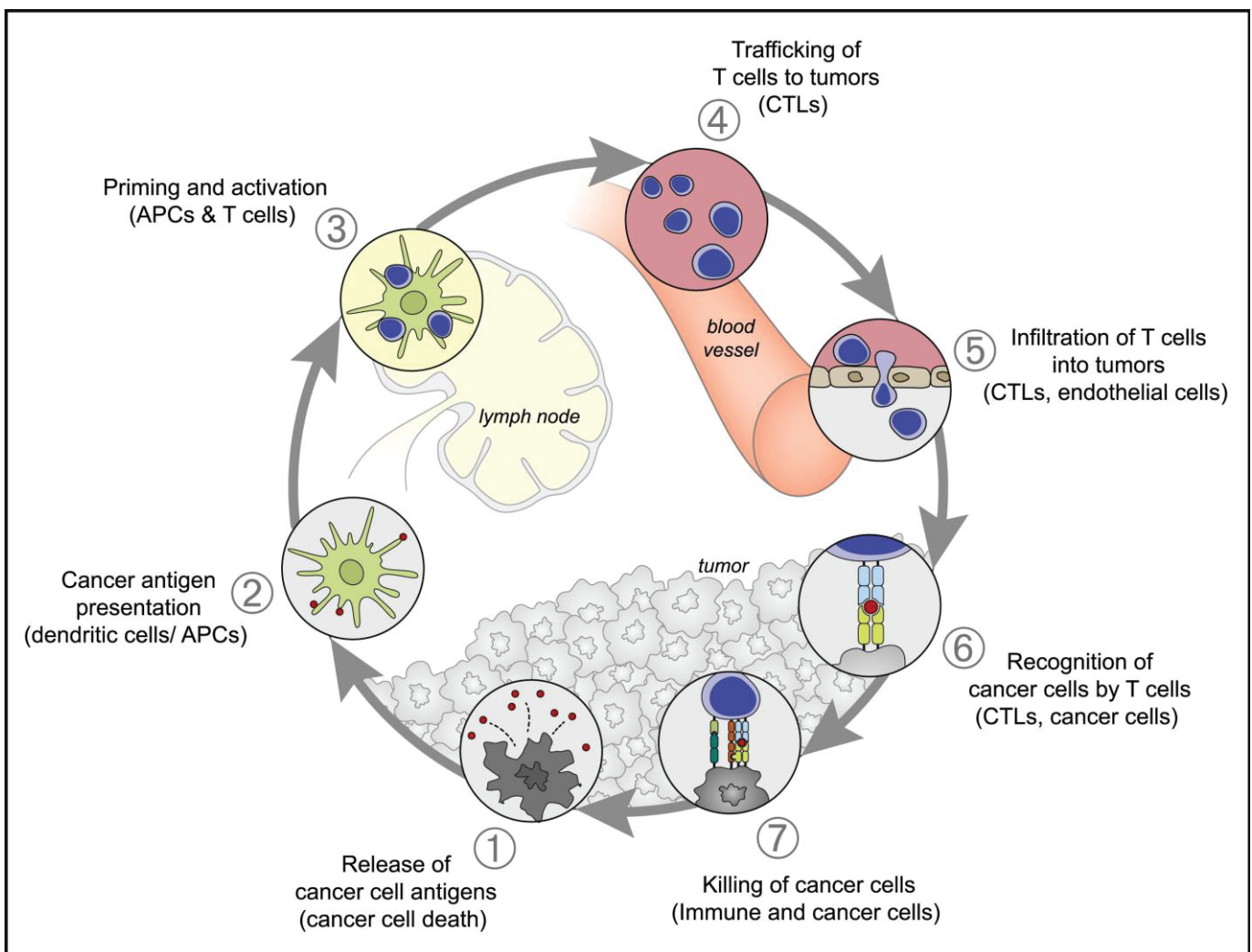
- Pendant la première phase dite « d'élimination », le système immunitaire reconnaît correctement les cellules transformées, aboutissant à leur destruction. Cette reconnaissance est le fruit d'une coopération entre les systèmes immunitaires innés et adaptatifs. Cette phase constitue la phase d'immunosurveillance à proprement parler.
- Certaines cellules transformées peuvent échapper à leur reconnaissance immunitaire, menant alors une deuxième phase dite « d'équilibre ». Durant cette phase, les cellules du système immunitaire adaptatif préviennent la croissance tumorale en éliminant les clones immunogènes émergents, maintenant la tumeur dans un état de dormance (15). C'est au cours de cette phase que se déroule principalement l'*immunoediting*, façonnage immunitaire de la tumeur menant à la sélection de clones non immunogènes.
- Dans la dernière phase dite « d'échappement », les cellules tumorales ayant acquis la capacité d'échapper à la reconnaissance immunitaire deviennent



déTECTABLES cliniquement. Cette phase peut aboutir aussi bien par progression de la tumeur elle-même (avec notamment l'émergence de clones façonnés par l'*immunoediting*) que par l'apparition d'un défaut au niveau du système immunitaire (telle que l'installation d'une immunosuppression par exemple).

### B. Le cycle de reconnaissance immunitaire du cancer

Les travaux sur l'immunosurveillance ont permis de souligner le rôle clé de l'immunité adaptative, et notamment des lymphocytes T (LT) effecteurs CD8<sup>+</sup>, dans la reconnaissance immunitaire antitumorale (8,14). La reconnaissance immunitaire d'une cellule tumorale ou transformée par les LT effecteurs est un processus complexe reposant sur 7 étapes clés résumées par Daniel Chen et Ira Mellman en 2013 (Figure 1) (16). Au cours de la première étape (i), les antigènes tumoraux sont captés par les cellules présentatrices d'antigènes (en premier lieu desquelles les cellules dendritiques) dans le micro-environnement tumoral. Afin d'entraîner une



**Figure 1 : Cycle de reconnaissance immunitaire du cancer**

Décrit en 2013 par Daniel Chen et Ira Mellman (8), ce cycle reprend les 7 étapes clés de la reconnaissance tumorale par les LT. Les antigènes tumoraux sont captés par les cellules présentatrices d'antigènes dans l'environnement tumoral, puis emmenés dans les organes lymphoïdes secondaires pour y être présentés aux LT naifs. Une fois activés, les LT naifs vont se différencier en LT effecteurs et auxiliaires, avant de migrer vers la tumeur par les vaisseaux sanguins. Les LT effecteurs y reconnaissent alors leur cible, aboutissant à la destruction de la cellule tumorale.

D'après Chen D. & Mellman I. Immunity 2013 (16), figure reproduite avec l'autorisation d'Elsevier.

réponse immunitaire, le relargage de ces antigènes tumoraux doit s'accompagner de signaux immunogènes, tels que le relargage de motifs moléculaires associés aux dégâts (DAMPs pour *Damage Associated Molecular Pattern*) ou de cytokines pro-inflammatoires. Une fois l'antigène capté, les cellules présentatrices d'antigènes vont migrer vers les ganglions lymphatiques afférents (ii) pour y présenter les antigènes aux LT naïfs CD4<sup>+</sup> via leur complexe majeur d'histocompatibilité (CMH) de classe II, et aux LT naïfs CD8<sup>+</sup> via leur CMH de classe I (la présentation d'antigènes extra-cellulaires via le CMH de classe I, ou *cross*-présentation, étant uniquement réalisée par les cellules dendritiques). En présence de leurs 3 signaux d'activation (l'antigène présenté par le CMH au récepteur du LT (TCR pour *T Cell Receptor*), le signal de costimulation présent sur la cellule présentatrice d'antigène et le signal cytokinique), les LT naïfs vont s'activer (iii) et se diviser en une population clonale de LT auxiliaires (CD4<sup>+</sup>) ou effecteurs (CD8<sup>+</sup>) spécifiques de la cible tumorale. Une fois activés, les LT effecteurs vont rejoindre le site tumoral via la circulation sanguine (iv), puis pénétrer dans le micro-environnement tumoral (v) afin d'y reconnaître leur cible (vi) et effectuer leur action cytotoxique (vii). A noter que, bien que les LT effecteurs soient classiquement décrits comme étant des LT CD8<sup>+</sup>, de nombreux travaux soulignent désormais le rôle de LT effecteurs CD4<sup>+</sup> dans la réponse antitumorale (17).

### **C. Echappement immunitaire de la tumeur**

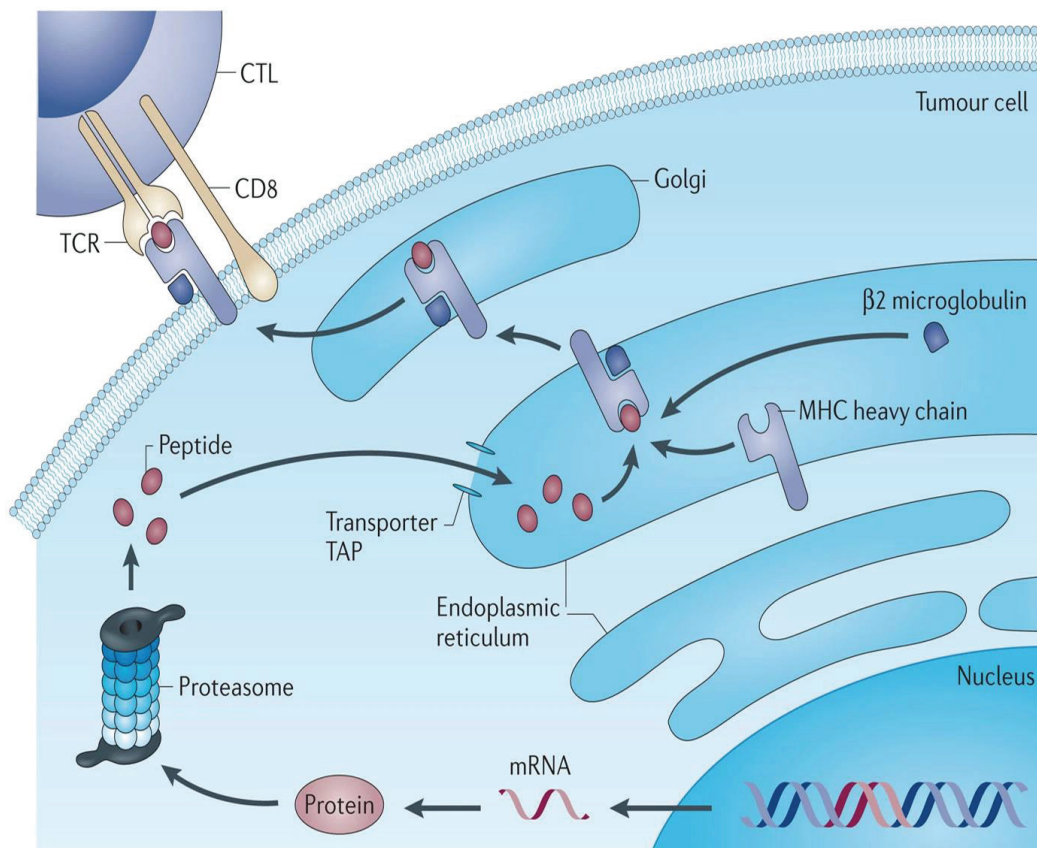
Par définition, la présence d'un cancer implique un échappement de la tumeur à la surveillance immunitaire. Cet échappement tumoral peut être lié à la défaillance d'une ou plusieurs étapes du cycle de la reconnaissance immunitaire. Selon les cas, plusieurs phénotypes tumoraux pourront être distingués (18) : les tumeurs dites « froides », par absence totale de réponse immunitaire (désert immunitaire par défaillance des étapes 1 à 3) ou absence de pénétration des cellules immunitaires dans la tumeur (exclusion immunitaire par défaillance des étapes 4 à 5), et les tumeurs dites « chaudes », infiltrées par les cellules effectrices mais sans élimination efficace des cellules cancéreuses par des phénomènes immunosuppresseurs locaux (défaillance des étapes 6 et 7).

Dans la première revue de cette thèse (**Revue 1 (19)**), nous avons fait le point sur les mécanismes pouvant être à l'origine de tumeurs dites « froides » et proposé différentes approches thérapeutiques selon le mécanisme impliqué. Dans ce contexte, la vaccination thérapeutique associée aux ICPI permettant de faciliter la mise en place de la réponse vaccinale et contrecarrer les phénomènes d'épuisement lymphocytaire apparaît comme une solution prometteuse pour induire ou renforcer la réponse adaptative antitumorale. L'immunothérapie cellulaire apparaît comme une bonne alternative en amenant directement des effecteurs spécifiques antitumoraux.

## II. Les antigènes tumoraux

La présence de cibles immunitaire à la surface de la cellule tumorale est un prérequis indispensable à la reconnaissance d'une cellule maligne par les LT CD8<sup>+</sup>. Les antigènes tumoraux constituant ces cibles sont issus de la dégradation intracellulaire de protéines en peptides de 8 à 10 acides aminés via le protéasome (20). Ces peptides sont ensuite transportés dans le réticulum endoplasmique via la protéine *Transporter associated with Antigen Processing* (TAP) où ils seront chargés sur le CMH de classe I avant d'être présentés à la surface de la cellule aux LT CD8<sup>+</sup> (Figure 2).

Il est estimé qu'une cellule présente environ 2% de son protéome aux LT CD8<sup>+</sup> via le CMH de classe I, ces peptides constituant les MAPs (*MHC-I Associated peptides*) (21). Bien que paraissant faible de prime abord, cet « échantillon » du protéome de la cellule peut cependant s'adapter très rapidement à tout changement de l'homéostasie cellulaire, en cas d'infection, de transformation maligne ou de traitement par exemple (22). Cette adaptation rapide est en partie permise par la production de produits ribosomiaux défectueux (DRIPs pour *Defective Ribosomal Products*), des polypeptides qui n'atteindront jamais leur conformation du fait d'erreurs traductionnelles ou post-traductionnelles (23). Ces DRIPs représentent 30% de la production protéique totale, avec l'avantage de présenter un important *turnover* de production et une



**Figure 2 : Formation et présentation des épitopes T CD8<sup>+</sup>**

Les épitopes T CD8<sup>+</sup> sont issus de la dégradation des protéines intracellulaires par la voie du protéasome. Les peptides de 8 à 10 acides aminés résultant sont alors transportés dans le réticulum endoplasmique via la protéine TAP avant d'être chargés sur le CMH de classe I. Une fois formé, le complexe CMH-peptide peut migrer vers la surface de la cellule via l'appareil de Golgi.

D'après Coulie P. et al. Nat Rev Cancer 2014 (20), figure reproduite avec l'autorisation de Springer Nature.

adaptabilité rapide face à un stress cellulaire (24). Le répertoire de MAPs d'une cellule donnée est ainsi sans arrêt façonné en fonction des conditions qui lui sont imposées. Parallèlement à ces variations physiologiques, la cellule tumorale peut présenter de multiples altérations à chacune de ces étapes, depuis la séquence codante elle-même jusqu'aux modifications post-traductionnelles. Ces différentes modifications participent ainsi à la transformation du répertoire de MAPs de la cellule tumorale, pouvant être plus ou moins spécifique de celle-ci. Dans cette partie, nous passerons en revue les différents types d'antigènes tumoraux, en proposant un aperçu de leurs avantages et inconvénients respectifs résumés dans le Tableau 1. Sans négliger l'importance des épitopes T CD4<sup>+</sup>, nous concentrerons là encore nos exemples sur les épitopes T CD8<sup>+</sup>, au cœur de notre projet de recherche.

### A. Bref historique

L'existence d'antigènes tumoraux a été suggérée dans les années 1960 suite à la démonstration que des souris exposées à une tumeur étaient immunisées contre une réexposition à la même tumeur (25,26). Ce type d'antigène a ainsi d'abord été classé comme antigène de rejet de transplantation (*Transplant Rejection Antigens*). Caractérisé par la formation de tumeurs très immunogènes régressant spontanément, le modèle de tumeurs induites par le virus Moloney sarcoma-leukemia virus (MSV-MoLV) permettra par la suite à plusieurs équipes de mettre en évidence la présence de LT CD8<sup>+</sup> cytotoxiques spécifiques de la tumeur dans les organes lymphoïdes, le sang mais également la tumeur elle-même (27–29). Chez l'homme, ces résultats ont été confirmés plus tard par la lyse de cellules tumorales allogéniques et autologues par des lymphocytes issus du sang de patients, avec la mise au point des premiers protocoles de culture de lymphocyte infiltrant la tumeur (TILs pour *Tumor Infiltrating Lymphocytes*) (30). La nature des antigènes reconnus par les LT CD8<sup>+</sup> cytotoxiques sera identifiée à la fin des années 1980 avec la mise en évidence de peptides de 8 à 10 acides-aminés naturellement présentés par le CMH de classe I dans le cadre de la réponse antivirale (31–33). C'est au cours des années 1990 que les premiers antigènes tumoraux ont ensuite pu être identifiés grâce aux techniques de clonage de l'ADN complémentaire (ADNc), avec notamment l'identification d'antigènes tumoraux issus de la lignée germinale dans le mélanome, dénommés alors *Cancer-Testis Antigens* (34). Les TILs ont ainsi longtemps été utilisés comme « sonde » de détection d'antigènes tumoraux sur des librairies d'ADNc permettant l'identification de différentes cibles (35). L'utilisation d'anticorps à la place des LT pour sonder les librairies de clones d'ADNc via la technologie SEREX (*serological analysis of cDNA expression libraries*) permettra d'accélérer les protocoles, menant à l'identification de nombreux autres antigènes tumoraux (36). Parallèlement, le développement des approches biochimiques visant à purifier puis identifier les peptides présents sur le CMH de classe I via un spectromètre de masse mettra au jour d'autres cibles telles que la protéine de différenciation mélanocytaire GP100 (37). Enfin, la démocratisation des méthodes de séquençage haut débit mènera à l'introduction plus récente des approches dites d'immunologie inversée visant à identifier d'abord des cibles antigéniques potentielles via l'analyse du génome tumoral et ses mutations avant de valider leur immunogénicité *in vitro* voir *in vivo* (20).

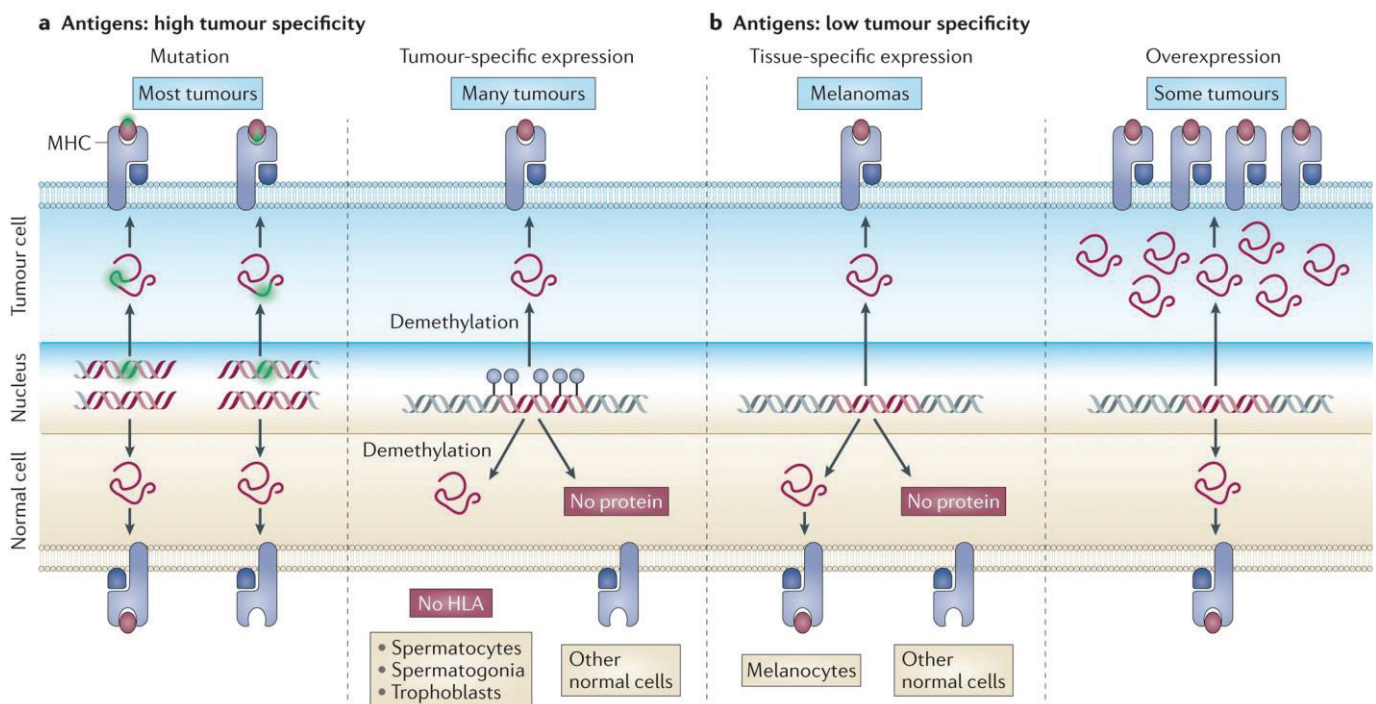
### B. Les différents types d'antigènes tumoraux

Deux grands groupes d'antigènes tumoraux sont classiquement distingués en fonction de leur spécificité d'expression tumorale : les antigènes spécifiques de la tumeur (TSAs pour *Tumor-Specific Antigens*) et les antigènes associés à la tumeur (TAAs pour *Tumor-Associated Antigens*) (Figure 3). Parallèlement à cette distinction TSAs/TAAs, différents sous-types d'antigènes pourront être



distingués selon leur provenance génomique. En effet, si la plupart des recherches d'antigènes tumoraux se sont jusqu'à présent focalisées sur les exons (soit environ 2% du génome), il est désormais bien établi que ces portions dites codantes ne reflètent pas la diversité transcriptionnelle d'une cellule, 75% du génome étant généralement transcrit (38). Ainsi, selon les études, jusqu'à 90% du répertoire de MAPs serait constitué par des peptides dits « cryptiques », c'est-à-dire provenant de localisations génomiques situés en dehors des exons (39). Nous comprenons dès lors le rationnel de s'intéresser aux éléments transposables (ETs) qui, représentant environ 50% du génome, constituent une source formidable de matériel génétique non-exonique pouvant potentiellement être source de MAPs. Bien que leur profil d'expression soit généralement spécifique de la tumeur, certains antigènes peuvent néanmoins être détectés à faible niveau dans le tissu sain, la classification TSA/TAA devant ainsi être discutée pour chaque antigène à l'échelle individuelle.

Sans rentrer dans les détails de l'évolution récente de la classification des antigènes tumoraux pour laquelle aucun consensus réel n'existe, nous reprendrons ici les deux principaux groupes TAA/TSA en développant à part le cas des antigènes cryptiques.



Nature Reviews | Cancer

### Figure 3 : Grandes catégories d'antigènes tumoraux

Les antigènes spécifiques de la tumeur (TSAs) regroupent principalement les antigènes issus de mutations et d'anomalies épigénétiques menant à leur expression tumorale spécifique. Les antigènes associés à la tumeur (TAAs) regroupent des antigènes au profil d'expression restreint à certains tissus ou à l'expression faible dans le tissu normal et très forte dans le tissu tumoral.

D'après Coulie P. et al. Nat Rev Cancer 2014 (20), figure reproduite avec l'autorisation de Springer Nature.

### 1) Les antigènes associés à la tumeur (TAAs)

Cette catégorie regroupe des antigènes de différenciation au profil d'expression très restreint ou dérivés de protéines faiblement exprimées dans le tissu sain et surexprimées dans une tumeur donnée (20,40). De nombreux antigènes ont là encore été d'abord découvert dans le mélanome, avec notamment Melan-A (MART1) et GP100 (41,42), deux protéines exprimées par les mélanocytes normaux. Ces deux antigènes représentent des antigènes de différenciation, un type de TAA particulier relié plus au tissu d'origine de la tumeur qu'à son caractère néoplasique. A l'inverse, les TAA surexprimés reflètent les modifications néoplasiques de la cellule, tels que WT1 dans la LAM (43).

L'expression par le tissu sain expose particulièrement ces antigènes à une tolérance immunitaire aussi bien centrale que périphérique (20). Cependant, l'induction d'une réponse immunitaire spécifique et efficace contre certaines de ces cibles montre que la tolérance centrale face à ces protéines du soi est incomplète et encore mal comprise. D'une façon intrigante, des LT CD8<sup>+</sup> naïfs spécifiques de Melan-A sont retrouvés à des fréquences de  $1.10^{-3}$  des LT CD8<sup>+</sup> totaux circulants chez des donneurs sains HLA-A\*02 :01 (44), soit des fréquences 100 à 1000 fois plus élevées qu'attendues pour des LT CD8<sup>+</sup> naïfs d'autre spécificité et comparables à des fréquences de LT mémoires (45). Bien qu'il ait initialement été suggéré qu'une prolifération périphérique (« post-thymique ») par expansion homéostatique transitoire avec réversion des LT CD8<sup>+</sup> mémoires en LT naïfs soit à l'origine de cette fréquence élevée (46), l'analyse des cercles d'excision du TCR (TREC pour *T cell receptor excision circles*, résidus d'ADN excisés pendant le réarrangement du TCR lors de la maturation thymique des LT) a finalement révélé une sélection positive plus importante dans le thymus, menant à une augmentation du nombre de précurseurs (47). Cette sélection plus importante pourrait être en lien avec une réactivité croisée envers d'autres peptides non identifiés. L'absence d'activation de ces LT CD8<sup>+</sup> naïfs contre Melan-A, exprimé par les mélanocytes, illustre ainsi parfaitement le phénomène d'ignorance immunitaire en lien d'une part avec le profil d'expression particulier de cette cible dans des tissus peu accessibles au système immunitaire, et d'autre part avec la faible expression de la cible dans les tissus en question (47).

La tolérance centrale, caractérisée par l'expression thymique d'un antigène, n'est cependant pas un frein complet à la mise en place d'une réponse immunitaire efficace. Il a ainsi été retrouvé une expression de différents antigènes issus de la lignée germinale (CGAs pour *Cancer-Germine Antigens*) notamment de la famille MAGE par les cellules médullaires thymiques, alors même que des réponses thérapeutiques contre ces cibles avaient pu être décrites (34). Néanmoins, les TCRs spécifiques des protéines du soi présentent des affinités inférieures par rapport aux TCRs d'antigènes étrangers (constante de dissociation Kd de l'ordre de 100µM vs 10µM, respectivement) (48), cette affinité étant directement corrélée à l'activation et l'activité cytotoxique des LT (49,50).

Malgré la possibilité de définir des épitopes partagées, l'expression par le tissu normal fait redouter les atteintes extra-tumorales lors du ciblage des TAAs. Si cette atteinte dite « *on-target / off-tumor* » ne pose que peu de soucis dans certaines situations, comme avec le ciblage du CD19 responsable d'une lymphopénie B collatérale pouvant être gérée cliniquement dans les hémopathies lymphoïdes B ou l'apparition d'un vitiligo avec le ciblage de Melan-A dans le mélanome, d'autres cibles peuvent mener à des toxicités sévères. Ainsi l'utilisation de LT avec

récepteur antigénique chimérique (CAR-T) dirigé contre la protéine ERBB2 dans le cadre d'un cancer du côlon métastatique a conduit au décès d'un patient par atteinte respiratoire sévère en lien avec l'expression initialement non prévue du TAA dans le tissu pulmonaire (51). A noter qu'un cas similaire de toxicité neurologique fatale avait été rapporté suite au ciblage de MAGE-A3 (52), venant remettre en cause la spécificité de l'expression de certains CGAs par la tumeur et la lignée germinale. De même, le ciblage des épitopes issus de mutations n'est pas à l'abri du risque d'auto-immunité, comme cela a été montré chez des patients atteints de sclérodermie pour qui la réponse auto-immune était déclenchée par une mutation du gène *POLR3A* dans une tumeur concomitante (53).

## **2) Les antigènes spécifiques de la tumeur (TSA)**

### *Les antigènes tumoraux issus de la lignée germinale (Cancer-Germline Antigens)*

Les CGAs sont des antigènes exprimés par les tissus embryonnaires et/ou germinaux (gamètes) et non exprimés par les autres tissus normaux (54). Ce type d'antigènes a d'abord été mis en évidence dans le mélanome avec l'identification des antigènes de la famille *Melanoma Antigens* (MAGE) (34), puis *B melanoma antigen* (BAGE) (55) et *G antigen 1* (GAGE) (56). Les CGAs sont également communément désignés sous le terme de *Cancer-Testis Antigens* du fait de leur profil d'expression. Cependant, du fait de l'absence de CMH de classe I dans le testicule (et donc de l'absence de présentation antigénique dans ce tissu) mais également de la présence d'autres tissus germinaux, certains auteurs encouragent à préférer le terme de CGA afin de ne pas limiter la source antigénique au testicule et de ne pas laisser faussement suggérer un potentiel risque d'auto-immunité testiculaire associé (20).

A ce jour, plus de 200 CGAs sont annotés et de nombreux candidats sont en cours d'investigation clinique (57). Pour qu'un antigène puisse être classé comme CGA, il doit être exprimé dans une tumeur et dans la lignée germinale (principalement le testicule ou le placenta) sans être exprimé par plus de deux tissus normaux non germinaux (expression qui, quand elle est détectée, est généralement bien plus faible que dans les tissus germinaux) (54,58). Le fait qu'une cellule tumorale puisse exprimer des antigènes embryonnaires a mené à envisager que cette expression puisse refléter l'activation de tout un programme génique calqué sur celui des gamètes et pouvant participer à la tumorigenèse (54). Ces gènes pourraient effectivement aider la cellule tumorale à contourner certaines barrières pouvant entraîner l'arrêt de prolifération, la sénescence ou son apoptose et de nombreux CGAs ont été rapportés comme associés à un effet pro-tumoral (57). Des modifications épigénétiques, en premier lieu desquelles la déméthylation du promoteur, seraient une des clés menant à la réexpression des CGAs par la tumeur, et plusieurs travaux ont montré une augmentation de l'expression des CGAs après traitement par agents hypométhylants tels que la 5-azacytidine (AZA) (54,59,60).

L'absence d'expression par les tissus sains autres que le testicule, site de privilège immun dépourvu d'expression du CMH de classe I, tend à considérer ces antigènes comme étranger au système immunitaire, avec par conséquent un fort potentiel immunogène (57,61). Dans le cas d'autres tissus embryonnaires tels que le placenta, l'expression du CMH de classe I serait contrebalancée par un fort environnement immunosuppresseur, comme nous le verrons plus tard. Leur relative spécificité d'expression, leur implication dans le développement tumoral ainsi que leur forte immunogénicité potentielle font des CGAs des cibles de choix pour l'immunothérapie. Le principal inconvénient des CGAs vient finalement du nombre limité de

tumeurs dans lesquelles ils sont détectés, limitant l'applicabilité finale des approches ciblant ce type d'antigènes (49).

### *Les antigènes issus de mutations (« néoépitopes »)*

C'est en cherchant à identifier les principales cibles antigéniques des LT activés par les ICPI que l'équipe de Schreiber a pu montrer en 2014 le rôle important joué par les protéines mutées spécifiques de la tumeur dans la reconnaissance immunitaire de cette dernière (62). En utilisant un modèle murin de sarcome, l'équipe a pu ensuite démontrer un taux de rejet tumoral immunitaire similaire entre l'utilisation d'ICPI et d'une approche vaccinale incorporant ces cibles antigéniques dans des peptides longs (autour de 30-40 acides aminés, avec la capacité d'induire une réponse aussi bien LT CD4<sup>+</sup> que CD8<sup>+</sup>), ouvrant la voie de la vaccination thérapeutique (62). D'autres équipes ont confirmé en parallèle la reconnaissance de ces épitopes issus de protéines mutées par les LT CD8<sup>+</sup> au cours de la réponse antitumorale et/ou de l'*immunoediting*, renforçant le rationnel et l'intérêt pour ces antigènes spécifiques de la tumeur (63,64).

Si les efforts ont jusque-là été concentrés sur les néoépitopes issus de mutations ponctuelles, d'autres types de mutations (insertion et/ou délétion) pouvant décaler le cadre ouvert de lecture (ORF pour *Open Reading Frame*) peuvent amener à des protéines très différentes du soi avec un important potentiel immunogène (65,66). Il était initialement considéré que la contribution de ces épitopes à la réponse antitumorale était directement corrélée au taux de mutation d'une tumeur. Ainsi, une importante réponse T CD8<sup>+</sup> dirigée contre des antigènes issus de mutations était attendue dans des tumeurs fortement mutées telles que le cancer du poumon et le mélanome, ou encore dans les 15% de cancers colorectaux hyper-mutés en lien avec une déficience dans la machinerie de réparation des *mismatch* de l'ADN (67). Le rationnel de cette association s'est vu renforcé par la suite avec la corrélation retrouvée entre la réponse clinique aux ICPI et la charge mutationnelle des tumeurs (68). Le nombre de cibles nécessaires pour aboutir à une réponse clinique n'est cependant pas bien défini, et il a par ailleurs été montré récemment que le ciblage par des approches vaccinales peut entraîner des réponses cliniques même pour des tumeurs faiblement mutées telles que les glioblastomes (69). Le nombre important d'autres facteurs prédictifs de réponse ou de toxicité aux ICPI, tels que le ratio neutrophiles/lymphocytes, la protéine C-réactive ou encore la composition du microbiote amènent à nuancer l'importance relative de la charge mutationnelle (70). Plus que la quantité elle-même, il semblerait donc que la qualité des épitopes issus de mutations ainsi que l'environnement dans lequel ils sont exprimés soient primordiaux.

La très grande spécificité et l'immunogénicité potentielle de ces cibles absentes des tissus normaux (et donc théoriquement non sujettes à la tolérance centrale) constituent les deux grands avantages des antigènes issus de mutations. En pratique, l'identification de ces épitopes est cependant une tâche laborieuse, nécessitant de coupler des techniques de séquençage complet de l'exome des cellules somatiques du patient ainsi que de la tumeur afin d'identifier les mutations. Il est ensuite estimé que seules 10% des mutations conduiront à un peptide de forte affinité pour le CMH de classe I, parmi lesquelles seule une fraction sera effectivement immunogène (64). La spécificité de ce type de cible constitue également sa principale faiblesse, les approches d'immunothérapies devant être adaptées à chaque patient (approche « personnalisée »), avec un temps d'identification des cibles et de mise au point des thérapies pouvant être relativement long.



### 3) Les antigènes cryptiques

Les antigènes dits cryptiques sont des peptides provenant de localisation génomique habituellement non transcrites ou non traduites, c'est-à-dire principalement situées en dehors des exons (40). Comme évoqué précédemment, certaines études montrent que jusqu'à 90% du répertoire de MAPs peut être constitué par ce type de peptides (39). De nombreux événements peuvent mener à la production d'un antigène cryptique notamment en contexte tumoral :

- Des phénomènes modifiant l'ARNm ont largement été décrits (66): anomalies de l'épissage de l'ARNm (souvent par mutation de parties non codantes intervenant dans l'épissage) (71,72), rétention d'intron (73), transcription antisens via un promoteur cryptique (74) ou encore dérégulation de la machinerie du spliceosome (75).
- Des anomalies ribosomiques peuvent également être source d'anomalies au niveau traductionnel, avec décalage du cadre de lecture (76), initiation traductionnelle sur un site non-canonique (77) ou non-respect des codons start et stop canoniques (78) et modifications post-traductionnelles (79).
- Enfin, la transcription aberrante de zones habituellement silencieuses peut être une source d'antigène cryptique, comme nous le verrons avec les ETs et les HERVs.

Ces anomalies semblent représenter une portion non négligeable de l'immunopeptidome : une étude publiée dans Science en 2016 montre que jusqu'à 1/4 de l'immunopeptidome du CMH de classe I d'une lignée de lymphoblastes est représenté par des peptides issus d'anomalies post-traductionnelles générées au niveau du protéasome (telle que la découpe et le collage de peptides issus de protéines différentes) (80). Ces peptides pourraient avoir une importance particulière notamment dans les tumeurs faiblement mutées telles que les hémopathies malignes, certaines pathologies comme les syndromes myélodysplasiques étant de plus connues pour être la cible de nombreuses mutations affectant le spliceosome (75,81). L'équipe de Claude Perreault a ainsi très récemment montré que, parmi 58 TSAs identifiés par spectrométrie de masse dans 19 échantillons primaires de LAM, 48% résultent de rétention ou traduction d'introns (82). Ces antigènes sont par ailleurs partagés par de nombreux patients, et surexprimés en cas de mutations de *DNMT3A*, *NPM1* et *FLT3-ITD*.

Proposant la spécificité et l'immunogénicité d'un TSA avec l'avantage de pouvoir être partagé et donc connu à l'avance comme un TAA, ce type d'antigène possède des caractéristiques très intéressantes. Le profil de chaque épitope doit cependant être contrôlé à l'échelle individuelle, certains antigènes pouvant être exprimés par le tissu sain. Non détectable par les approches classiques de séquençage, l'identification de tels peptides nécessite le développement d'outils spécifiques tels que le séquençage de l'ARN total avec des algorithmes d'alignement particuliers, le séquençage de l'ARN en cours de traduction (Riboseq) ou encore la spectrométrie de masse (83). Bien que très performantes, ces techniques peuvent néanmoins manquer de sensibilité pour certaines cibles comme nous le discuterons plus tard.

### C. Aperçu des stratégies thérapeutiques

Par opposition aux immunothérapies ne ciblant pas d'antigène en particulier (*antigen-unaware therapies*, telles que l'activité immunogène des chimio/radiothérapies, les ICPI ou encore la

greffe allogénique de cellules souches hématopoïétiques), la connaissance des antigènes tumoraux permet de proposer des immunothérapies spécifiques d'un antigène donné (*antigen-aware therapies*) (40). L'efficacité de ces dernières repose ainsi en partie sur les propriétés de l'épitope ciblé (Tableau 1) et sur les caractéristiques de la thérapie en elle-même.

Dans les **Revues 2 et 3** (84,85), nous avons fait le point sur les principales approches actuelles de vaccination thérapeutique. Après les résultats décevants obtenus avec le ciblage de TAAs (86), celles-ci sont actuellement dominées par les stratégies visant à cibler les antigènes issus de mutations spécifiques de la tumeur (néoépitopes) en association avec des ICPI. Les résultats obtenus sont encourageants bien que la lourdeur de la mise en place d'une vaccination personnalisée rende complexe sa généralisation à grande échelle.

Dans les **Revues 4 et 5** (87,88), nous avons passé en revue les principes de fonctionnement des LT à récepteur antigénique chimérique (CAR-T), ainsi que les résultats des principaux essais récents et en cours notamment dans les tumeurs solides. Nous avons également abordé le développement récent des lymphocytes CAR-T allogéniques (**Revue 6** (89)).

L'ensemble de ces travaux bibliographique nous aura permis de faire le point sur les différentes approches thérapeutiques actuelles visant à cibler des antigènes tumoraux. Plus que l'approche proposée en elle-même, le point majeur réside ainsi dans le choix de la cible thérapeutique. Cela s'illustre parfaitement avec les lymphocytes CAR-T, montrant des résultats spectaculaires avec le ciblage de CD19 dans les hémopathies lymphoïdes B, et pourtant décevant sur d'autres cibles tumorales. Il existe donc un véritable besoin d'identifier et mettre au jour de nouvelles cibles thérapeutiques afin de bénéficier au mieux des thérapies proposées. Dans ce contexte, nous nous sommes intéressés aux antigènes dérivés de rétrovirus endogènes humains (HERVs) (sur)exprimés spécifiquement par les cellules tumorales.

Catégorie	Avantages globaux	Type	Avantages	Inconvénients
TAAs	<ul style="list-style-type: none"> <li>- Partage entre plusieurs patients</li> <li>- Possibilité de développer des thérapies prêtes à l'emploi</li> </ul>	Différenciation tissulaire	<ul style="list-style-type: none"> <li>- Profil de toxicité potentiellement maîtrisé par le profil d'expression restreint</li> <li>- Expression très forte par la tumeur</li> </ul>	<ul style="list-style-type: none"> <li>- Elimination des T effecteurs les plus réactifs</li> <li>- Risque important d'auto-immunité « <i>on-target/off-tumor</i> »</li> </ul>
		Surexpression	<ul style="list-style-type: none"> <li>- Pas d'expression théorique dans les tissus sains non germinaux</li> <li>- Partage entre plusieurs patients</li> <li>- Possibilité de développer des thérapies prêtes à l'emploi</li> </ul>	<ul style="list-style-type: none"> <li>- Faible nombre de cancer éligibles</li> <li>- Spécificité non assurée en pratique</li> </ul>
TSAs	<ul style="list-style-type: none"> <li>- Absence complète des tissus sains avec risque théorique minimal de toxicité « <i>on-target/off-tumor</i> »</li> <li>- Mutations drivers partagées pouvant potentiellement être ciblées (rares)</li> </ul>	Lignée germinale (CGAs)	<ul style="list-style-type: none"> <li>- Nombreuses études</li> <li>- Prédiction simple</li> <li>- Fréquence élevée</li> </ul>	<ul style="list-style-type: none"> <li>- Traitement personnalisé avec un temps de détection potentiellement important</li> <li>- Faible partage des cibles</li> <li>- Faible dissimilarité avec le soi</li> </ul>
		Mutation (SNV)	<ul style="list-style-type: none"> <li>- Nombreux peptides par mutation</li> <li>- Importante dissimilarité avec le soi</li> </ul>	<ul style="list-style-type: none"> <li>- Traitement personnalisé avec un temps de détection potentiellement important</li> <li>- Fréquence faible</li> </ul>
		Mutation ( <i>frameshift</i> )	<ul style="list-style-type: none"> <li>- Fréquence élevée</li> </ul>	<ul style="list-style-type: none"> <li>- Peu d'outils disponible pour la détection</li> <li>- Peu d'études</li> </ul>
Cryptique	<ul style="list-style-type: none"> <li>- Possibilité de trouver à la fois des cibles spécifiques et partagées entre plusieurs tumeurs</li> <li>- Importante dissimilarité avec le soi</li> <li>- Immunogénicité</li> </ul>	Anomalie d'épissage	<ul style="list-style-type: none"> <li>- Partage entre plusieurs patients</li> <li>- Nombreux peptides</li> </ul>	<ul style="list-style-type: none"> <li>- Fréquence faible</li> </ul>
		Protéine de fusion	<ul style="list-style-type: none"> <li>- Fréquence élevée</li> <li>- Origine virale avec forte immunogénicité potentielle</li> <li>- Partage entre plusieurs patients</li> </ul>	<ul style="list-style-type: none"> <li>- Peu d'études</li> <li>- Profil d'expression par le tissu sain à contrôler de façon stricte</li> <li>- Traduction protéique à confirmer</li> </ul>
		Rétroélément		

**Tableau 1 : Avantages et inconvénients des principaux types d'antigènes tumoraux.**

*Frameshift* : Décalage du cadre de lecture, *SNV* : *Single Nucleotide Variant* (variation d'un seul nucléotide), *TAA* : *Tumor-associated antigen* (antigène associé à la tumeur), *TSA* : *Tumor-specific Antigen* (antigène spécifique de la tumeur)

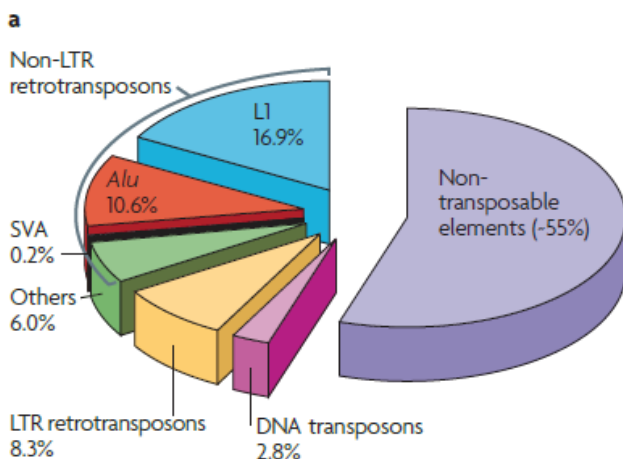
### III. Les Rétrovirus Endogènes Humains (HERVs)

#### A. Les éléments transposables (ETs) dans le génome humain

Environ 45% du génome humain est constituée d'ETs (Figure 3) (90). Considérées initialement comme inactives (« *Junk-DNA* »), nous savons désormais que ces séquences d'ADN mobiles jouent un rôle majeur dans la variation structurale du génome, source importante d'innovation évolutive (91). Deux grandes classes d'ETs sont distinguées selon leur mécanisme de transposition : les éléments de classe I (rétrotransposons) et les éléments de classe II (transposons à ADN). Les transposons à ADN représentent environ 3% du génome humain et se mobilisent par des mécanismes d'excision directe de l'ADN (mécanisme dit « couper-coller ») (92,93). Désormais inactifs chez l'Homme, leur période majeure d'activité et d'intégration remonte à la période d'évolution précoce des primates et jusqu'à il y a environ 37 millions d'années (94). Les rétrotransposons utilisent quant à eux une transcriptase inverse (RT pour *Reverse Transcriptase*) pour intégrer leur intermédiaire ARN dans une autre localisation génomique, se propageant ainsi par rétrotranscription (mécanisme de transposition dit « copier-coller ») (92).

Deux grandes familles sont distinguées parmi les rétrotransposons : caractérisés par la présence de séquences promotrices *Long Terminal Repeat* (LTR), les LTR-rétrotransposons représentent environ 8% du génome humain. Cette catégorie regroupe majoritairement les rétrovirus endogènes, mais également les *Mammalian Apparent LTR Transposons* (MaLR). La majorité des ETs est finalement constituée de rétrotransposons non-LTR, représentant environ 1/3 du génome humain et regroupant deux principaux groupes: les *Long Interspersed Elements* (LINEs) et les *Short Interspersed Elements* (SINEs) comprenant notamment les éléments Alu, et SVA (90). Cette famille est actuellement la plus active dans le génome humain, avec de nombreuses insertion *de novo* rapportées à l'origine de différentes pathologies (95,96) comme cela

a été montré par exemple avec l'insertion d'éléments LINE1 dans l'exon 14 du gène codant pour le facteur VIII, responsable d'une forme d'hémophilie A (97). Apparus il y a plus de 150 millions d'années, les LINE1 sont les ETs les plus nombreux en termes de quantité d'ADN (>500 000 copies de LINE1 dans le génome humain pour une moyenne de 6 000 paires de bases (pb)) (92). Les SINEs ont la particularité de ne pas être autonomes, utilisant ainsi la RT des LINEs pour se rétrotransposer. Parmi les SINEs, les éléments Alu disposent de plus d'1 millions de copies dans le génome humain, avec une moyenne de 300pb, faisant de ces éléments les plus nombreux en termes de nombre de copies (92). Les éléments SVA sont des rétrotransposons composites regroupant une séquence SINE, un



**Figure 3 : Classes d'éléments transposables dans le génome humain**

Environ 45% du génome humain est composé d'ETs. La famille des rétrotransposons (éléments de classe I), regroupant les LTR-rétrotransposons, les LINEs et les SINEs, se mobilise par rétrotransposition via une transcriptase inverse. La famille des transposons à ADN (classe II) se mobilise par excision de l'ADN. D'après Cordaux R. & Batzer M. *Nat Rev Genet* 2009 (91), figure reproduite avec l'autorisation de Springer Nature.

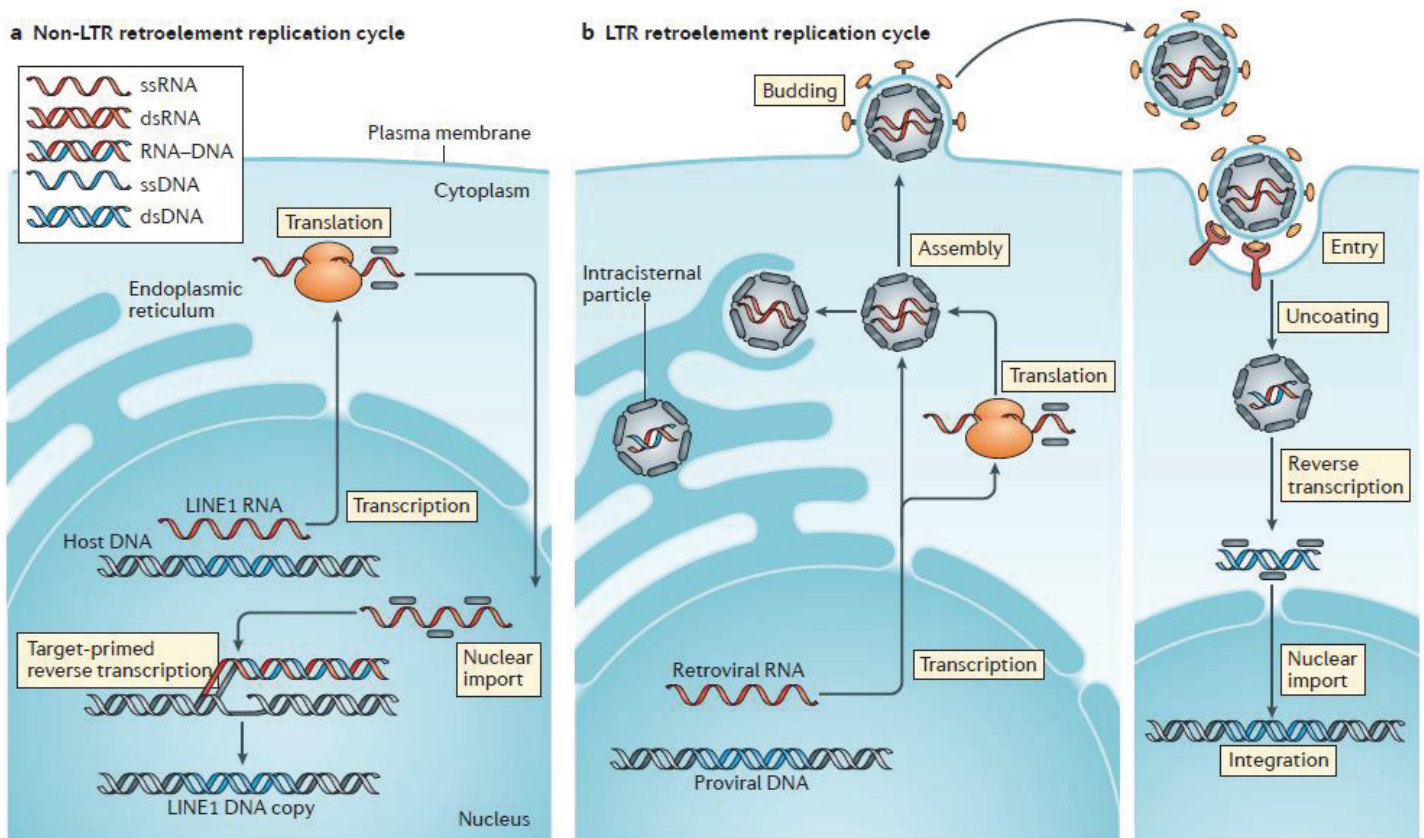
nombre variable de répétitions en tandem (VNTR pour *Variable Number of Tandem Repeats*) et deux fragments Alu inversés (**S**INE-**V**NTR-**A**lu) (98). Du fait de leur intégration récente, ces séquences spécifiques des hominidés ne représentent qu'environ 2 700 copies dans le génome humain. Leur activité importante de transposition en fait cependant une grande source de variation génétique à l'origine là encore de différentes pathologies (98).

Bien que les ETs non-LTR présentent un intérêt non négligeable, nous avons fait le choix de nous concentrer sur les rétrotransposons LTR et plus particulièrement les HERVs. Ces derniers présentent en effet de multiples atouts, tels que leur structure virale au fort potentiel immunogène, en faisant de prometteurs antigènes tumoraux cryptiques partagés comme nous le détaillerons plus loin. Dans cette troisième partie, nous verrons d'où viennent ces HERVs et quelles sont les données actuelles sur les mécanismes contrôlant leur expression dans le génome humain. Nous nous attarderons sur les quelques données disponibles concernant le rôle des HERVs dans le développement des mammifères. Nous ferons ensuite un point sur le rôle des HERVs dans l'oncogenèse puis leur impact immunitaire, avant de finir sur les considérations techniques et bioinformatiques de leur quantification. Sur le plan sémantique, nous parlerons d'ETs lorsque les études portent sur l'ensemble des ETs, de rétroéléments ou rétrotransposons pour regrouper ERVs, LINEs et SINEs, de ERVs pour les études portant sur les rétrovirus en général (notamment chez la souris ou d'autres modèles murins) et de HERVs pour les études et données obtenues plus particulièrement chez l'Homme.

## **B. Origines et structure des HERVs**

Des études phylogénétiques basées sur la divergence des séquences de RT montrent que l'ensemble des rétrotransposons (LTR et non-LTR) dérivent d'un ancêtre commun virus à ARN (99). La séquence classique d'un ERV est similaire à celle d'un rétrovirus exogène : deux séquences LTR entourant les ORFs *gag*, *pro*, *pol* et *env* (100). La séquence *gag* code pour les différents composants de la capsid du virion, les séquences *pro* et *pol* pour les enzymes nécessaires à la synthèse protéique et nucléotidique (notamment la RT) tandis que la séquence *env* code pour la protéine d'enveloppe permettant l'entrée du rétrovirus dans la cellule cible. Un rétrovirus complet est constitué de deux copies d'ARN simple brin (ssRNA pour *single-stranded RNA*) contenues dans une capsid avec la RT, le tout entouré d'une membrane exprimant la protéine d'enveloppe (101). Les ERVs sont ainsi les seuls ETs pouvant infecter une autre cellule que celle d'origine, particularité permise notamment par la protéine d'enveloppe (Figure 4). Cependant, si il est possible de trouver des ERVs pleinement fonctionnels conservant une certaine infectiosité chez la souris, aucun HERV complet n'a été rapporté comme infectieux chez l'Homme (102). Ceci est lié aux nombreux dégâts successifs, aux mutations et recombinaisons acquis sur les ORFs des HERVs, les rendant défectueux pour la plupart. Il est d'ailleurs estimé que la grande majorité des HERVs correspondent à des séquences LTR seules ayant subi une excision complète de leurs ORFs (102). Malgré tout, certains HERVs conservent la capacité de produire des protéines virales, voir même des particules virales intactes pour les familles les plus récemment intégrées telles que la famille HML-2 / HERV-K (103). L'excision de certains ORFs peut même procurer un avantage dans certains cas : il a par exemple été estimé que les ERVs dépourvus de protéine d'enveloppe ont une dissémination génomique accrue d'un facteur 30 par rapport aux ERVs fonctionnels possédant une protéine Env (104). En empêchant la dissémination intercellulaire, l'absence de protéine d'enveloppe pourrait donc forcer les rétrovirus à se propager par rétrotransposition intracellulaire sur un modèle similaire aux autres ETs.





**Figure 4 : Mécanismes de réplication des rétrotransposons.**

- A. Rétrotransposons non-LTR (LINE et SINE) : Une fois transcrit, l'ARN issu de la séquence génomique du rétrotransposon va être traduit en protéines dans le cytoplasme. Ces protéines vont permettre le transport de ce même ARN dans le noyau d'une part (protéine de liaison à l'ARN codée par l'ORF1p) puis son intégration par transcription inverse (endonucléase et RT codées par l'ORF2p). A noter que les SINEs utilisent la RT des LINEs pour se rétrotranscrire.
- B. Rétrotransposons LTR : l'ARN transcrit à partir de la séquence du ERV va mener à la formation d'un virion via la traduction des protéines Gag (capside virale), Pro, Pol (RT) et Env (enveloppe). Une fois complet, le virion va pouvoir encapsuler deux copies d'ARN simple brin, sortir de la cellule via la protéine d'enveloppe, puis infecter une nouvelle cellule où il pourra rétrotranscrire ses ARN en ADNc via la RT puis l'intégrer au génome de sa cellule hôte.

D'après Kassiotis G. & Stoye J. Nat Rev Immunol. 2016 (102), figure reproduite avec la permission de Springer Nature.

### *Mobilisation et polymorphisme*

Contrairement aux LINEs possédant encore une importante activité de mobilisation, l'absence de HERVs capables de rétrotransposition chez l'Homme minimise le polymorphisme interhumain. La grande majorité des HERVs ont en effet été intégrés il y a plusieurs millions d'années, bien avant l'émergence d'*Homo sapiens*, et sont désormais fixes dans le génome humain (105). Cette fixité relative n'est pas la règle chez tous les mammifères et doit faire interpréter les modèles animaux avec précaution, comme chez la souris par exemple où le niveau d'activité transpositionnelle des ERVs peut être très important (106) ou bien encore chez le Koala comme nous le verrons plus loin.

Un certain polymorphisme existe malgré tout pour la famille HERV-K, la plus récemment intégrée au génome humain. L'analyse de données issues du séquençage complet de plus de 2 500 génomes humains de 26 pays différents a permis de mettre en évidence la présence de 36 sites d'intégration polymorphique de HERV-K, avec des fréquences allant de <0,05% à >75% selon les populations étudiées (107). Parmi ces sites d'intégration, plusieurs loci présentaient des séquences provirales complètes, et certaines de ces insertions survenaient autour ou même à l'intérieur d'un gène, tels que *GLUT9* par exemple, cible de TP53. Les auteurs de

cette étude estiment que les dernières insertions de HERV-K remontent à environ 0,67 millions d'années. D'autres méthodes d'analyse des sites d'intégration polymorphiques de HERV-K présentent des résultats légèrement différents mais convergeant vers les mêmes conclusions (108).

### *Classification et nomenclature*

La nomenclature et classification des ERVs a pendant longtemps été hasardeuse et très hétérogène. Initialement, les ERVs furent classés selon les rétrovirus exogènes les plus proches, en ajoutant une lettre pour préciser l'espèce source de l'identification (MERV pour mouse, HERV pour humain, etc.) (109). Rapidement par la suite, les ERVs ont été classés selon le type d'ARN de transfert (ARNt) se liant à la séquence PBS (*Primer Binding Site*) pour débiter la transcription inverse (110). La famille HERV-K regroupe ainsi tous les HERVs nécessitant une lysine ARNt. Certains HERVs dépourvus de séquence PBS ont été nommés en fonction de gènes proches (ex : HERV-ADP (111)), du numéro de clone ayant mené à leur identification (ex : HERV-S71 (112)) ou encore de motif d'acides aminés (ex : HERV-FRD (113)). Parallèlement à cette classification interne des ERVs, une classification « externe » en tant que virus a également été proposée afin de les regrouper selon la taxonomie des rétrovirus acceptée par les nomenclatures internationales. Les ERVs sont ainsi regroupés en 3 grands groupes : *Gammaretrovirus*, *Betaretrovirus* et *Spumaretrovirus*, selon leurs homologies de séquences (et notamment de la séquence *pol* par rapport aux rétrovirus exogènes) (110,114).

Ces différentes nomenclatures ont conduit à de nombreux problèmes de classification. L'utilisation d'une lettre en fonction de l'espèce d'origine de découverte du ERV pose soucis dans le cas de ERVs intégrés avant la divergence de son hôte, et donc présents dans plusieurs espèces. L'utilisation du type d'ARNt a également montré ses limites, certains ERVs de la même lignée phylogénétique pouvant utiliser des ARNt différents, d'autres ERVs étant dépourvus de PBS. De ces différentes problématiques ont émergé plusieurs tentatives de systématisation de la classification des ERVs. En 2011, Mayer J. *et al.* proposent une classification des HERVs référencés dans les bases de données génomiques majeures Américaines (Genbank), Européenne (European Molecular Biology Laboratory (EMBL)) et Japonaise (DNA Databank of Japan (DDBJ)) (115). Consistant plus en une harmonisation des annotations de loci déjà décrits, cette nomenclature se concentre sur le côté pratique au détriment de son exhaustivité. Gifford R. *et al.* ont développé quant à eux une nomenclature plus systématisée en 2018, proposant d'identifier chaque ERV par sa catégorie, son groupe taxonomique, un numéro d'identification puis l'espèce (ex : ERV – K(HML.2).113 – Hsa) (110). Là encore, les auteurs proposent de se concentrer sur les provirus au détriment des LTR seuls. A ce jour, il n'existe pas de classification ni de nomenclature universelle, et l'identification d'un ERV donné par sa famille suivie de sa localisation génomique à partir d'une base de référence (telle que Repbase) semble être la solution la plus adaptée en vue d'une annotation reproductible des ERVs dans un génome donné.

### **C. Régulation de l'expression des HERVs**

L'expression des HERVs est sujette à d'importants mécanismes de régulation épigénétique (116). Il est considéré de façon classique que la méthylation des histones joue un rôle prépondérant dans la répression des HERVs dans les cellules souches et embryonnaires, tandis que la méthylation de l'ADN serait le mécanisme prépondérant dans les cellules

différenciées. Cette dichotomie est cependant remise en cause à la lumière de données récentes comme nous le verrons.

### *Méthylation des cytosines*

La méthylation du 5<sup>ème</sup> carbone des cytosines (5-méthylcytosine ou 5mC) est un des processus importants de répression de la transcription de l'ADN (117). Chez les mammifères, la méthylation de l'ADN est établie *de novo* dès le développement embryonnaire par des enzymes de la famille des ADN-méthyltransférases (DNMT pour *DNA-Methyltransferase*), DNMT3A et DNMT3B. Ces empreintes de méthylation sont ensuite maintenues à travers les divisions cellulaires par l'enzyme DNMT1, permettant ainsi la transmission germinale et donc l'héritabilité des empreintes de méthylation (118). La déméthylation active de l'ADN est finalement rendue possible par les enzymes de la famille *Ten-eleven Translocation (TET) methylcytosine dioxygenases*, permettant l'oxydation de la 5mC en 5-hydroxyméthylcytosine (5hmC). La méthylation de l'ADN est impliquée dans les phénomènes d'inactivation génique tels que l'empreinte parentale génomique (inactivation d'une copie génique transmise par l'un des deux parents) ou l'inactivation de l'X, et joue un rôle majeur dans la répression des ETs (117). Il a été démontré chez la souris un biais de méthylation par DNMT1 en faveur des régions contenant des rétroéléments (119). Chez l'homme, les HERVs sont particulièrement méthylés dans les tissus sains, et il semblerait que les phénomènes de déméthylation au sein de la tumeur affectent de façon prioritaire les rétroéléments (120). Les HERVs récemment intégrés seraient plutôt contrôlés par méthylation de l'ADN tandis que les plus anciens seraient plutôt sujet à la méthylation des histones (121). Ces « jeunes » HERVs seraient en théorie particulièrement sensibles à la réactivation par les agents hypométhylants tels que l'AZA.

### *Méthylation des histones*

La méthylation des histones est un autre processus important de contrôle de l'expression génique. Les deux principales marques habituellement distinguées sont la triméthylation de l'histone H3 sur la lysine 9 (H3K9me3), marqueur de répression transcriptionnelle, et la triméthylation de l'histone H3 sur la lysine 4 (H3K4me3), marqueur d'activation (122). La répression des ERVs dans les cellules embryonnaires est en grande partie assurée par la protéine *SET domain bifurcated histone-lysine N-methyltransferase* (SETDB1), assurant la formation d'hétérochromatine (marque H3K9me3) (123). Chez la souris, la démonstration initiale que les rétroéléments réprimés par SETDB1 dans les cellules souches embryonnaires ne se réactivent pas dans les cellules différenciées SETDB1<sup>-/-</sup> suggérait initialement un rôle négligeable de la méthylation des histones dans le contrôle des ERVs dans les cellules matures (124). Les travaux qui ont suivi ont cependant montré que SETDB1 joue également un rôle important dans le contrôle de l'expression de ERVs dans certaines cellules matures telles que les lymphocytes B (LB) (125), les LT (126) et les fibroblastes (127) de souris. Les protéines clés à l'origine de la méthylation des histones sont de la famille des *Krüppel-associated box domain zinc finger proteins* (KRAB-ZFPs), une des plus grandes familles de facteurs de transcription chez les mammifères (128). Avec leur cofacteur TRIM28 (également dénommé KAP1 pour *Krab-Associated Protein 1*), les KRAB-ZFP participent à la formation de l'hétérochromatine via le recrutement SETDB1 et d'ADN méthyltransférases telle que DNMT1 (129). Chez l'homme, les trois-quarts des sites de liaison de TRIM28 résident dans des séquences HERVs (129,130).



D'autres mécanismes de répression des HERVs ont été décrits, tels que la triméthylation de l'histone H3 sur la lysine 27 (H3K27me3) assurée par la protéine pRB (*retinoblastoma protein*) via le recrutement de EZH2 (*enhancer of zeste homolog 2*) (131). *Lysine-specific histone demethylase 1* (LSD1 ou KDM1A), une protéine de la famille des Lysines Déméthylases (KDMs), participerait également à la répression des ERVs comme cela a été démontré plus récemment chez l'Homme et la souris (132).

### Acétylation des histones

L'acétylation des histones, catalysée par les Histone AcetylTransférases (HAT) entraîne une déstabilisation de la chromatine favorisant l'activité transcriptionnelle. À l'inverse, la désacétylation, catalysée par les Histone désacétylases (HDAC) stabilise la chromatine aboutissant à une répression transcriptionnelle (122). La démonstration de la réactivation du rétrovirus de l'immunodéficiência humaine 1 (VIH-1) après utilisation de traitement inhibiteur des HDACs a mené plusieurs équipes à étudier le rôle de l'acétylation des histones dans le contrôle des HERVs (133). Hurst T. *et al.* n'ont pas retrouvé de transcription de HERVs après traitement par différents inhibiteurs de HDAC (vorinostat, panobinostat et romidepsine) de monocytes et LT primaires infectés par le VIH-1 à l'état latent (133). Cette étude n'analysait cependant que très partiellement l'expression des HERVs, se concentrant uniquement sur certaines cibles (séquences *pol* et *env* de HML-2 et *env* de HERV-W et HERV-FRD) par une technique de *Reverse-Transcriptase Polymerase Chain Reaction* (RT-PCR). Une analyse plus systématique réalisée dans notre laboratoire avec l'approche Telescope (cf partie H) sur un jeu de données publiques de séquençage ARN d'une lignée de LAM (HL60) exposée à différents traitements épigénétiques (134) montre une forte stimulation des HERVs dans l'ensemble des conditions, la plus forte étant obtenue sous combinaison décitabine + romidepsine (données non présentées).

### Régulation post-transcriptionnelle

Outre la régulation transcriptionnelle, plusieurs mécanismes d'inactivation des transcrits issus de ERVs ont été décrits, tels que la régulation des transcrits ERVs par des protéines de liaison à l'ARN (RBPs pour *RNA Binding Proteins*) (135,136), ou encore la méthylation des ARNm ERVs (137).

Les mécanismes épigénétiques à l'origine du contrôle de l'expression des HERVs sont donc complexes et vont bien au-delà de la dichotomie méthylation de l'ADN dans les cellules matures / des histones dans les cellules embryonnaires. Ainsi, le modèle initial considérant les HERVs comme exprimés uniquement dans les cellules embryonnaires et de la lignée germinale et réprimés dans les cellules somatiques normales est actuellement remis en cause, comme le souligne le travail récent de l'équipe de Claude Perreault détectant l'expression de rétroéléments dont des HERVs dans 32 tissus sains différents, bien que celle-ci soit très diverse et hétérogène en termes de nombre et de niveau d'expression (138). La répression des HERVs par régulation épigénétique est un des nombreux mécanismes de défense développés pour lutter contre la rétrotransposition et maintenir une certaine stabilité génomique, certains auteurs avançant même que la méthylation de l'ADN a évolué dans le but principal de réprimer les rétroéléments avant d'être cooptée pour d'autres fonctions de contrôle génique (123,139). La nécessité de contrôler l'invasion génomique des ERVs pourrait donc avoir été le moteur évolutif de multiples mécanismes physiologiques, comme cela est désormais connu pour la famille des KRAB-ZFPs dont les rétroéléments sont le principal moteur d'expansion dans le génome des mammifères

(128). Dans la prochaine partie, nous allons voir comment certaines fonctions des ERVs ont pu être utilisées pour le développement des mammifères (phénomène d'exaptation), puis comment ces mêmes fonctions ont pu être détournées durant l'oncogenèse.

#### **D. Exaptation des ERVs dans le développement des mammifères**

L'expansion des rétroéléments dans le génome constitue une formidable source de matériel génétique pour l'évolution, favorisant aussi bien la variabilité génétique et la plasticité génomique que la régulation génique elle-même. La dissémination des rétroéléments et des ERVs en particulier aura permis *in fine* l'accélération de l'évolution des génomes de leurs hôtes (140). Ainsi, tout comme les virus ont pu acquérir des gènes de leurs hôtes à travers leur évolution (141), les mammifères se sont largement servis des ERVs pour leur propre développement, avec de nombreux exemples à ce jour concernant le développement cérébral (142), la maintenance de la pluripotente des cellules souches (143) ou encore le contrôle de certaines voies de l'immunité adaptative avec notamment l'équilibre  $T_{H1}/T_{H2}$  dans les LT  $CD4^+$  (144). En plus des fonctions de régulation génique à proprement parler, les ETs pourraient fournir des moyens de réorganiser la chromatine, comme cela a été démontré pour un rétrovirus de la famille HERV-H impliqué dans des domaines d'association topologique (TAD pour *Topologically Associating Domain*) (145). Certaines protéines résiduelles produites par les HERVs pourraient également être utiles : outre l'exemple des Syncitines que nous verrons plus loin, nous savons que plusieurs HERVs de la famille HML-2 sont exprimés durant l'embryogenèse, avec la détection de protéines Gag jusqu'au stade de blastocyte, sans pour autant connaître leur rôle exact (146). Présents uniquement chez l'Homme, ces HERVs pourraient contribuer au développement précoce spécifique d'*Homo sapiens* ou encore protéger l'embryon de l'infection par d'autres rétrovirus, la protéine Rec de HERV-K étant suffisante pour entraîner la production d'*Interferon Induced Transmembrane Protein 1* (IFITM1) à la surface de la cellule, un facteur de restriction viral induit par l'interféron (IFN) (146). Sans être exhaustif sur l'ensemble de ces mécanismes, nous reprendrons ici les principaux rôles attribués aux ERVs dans leur contribution au développement des mammifères en nous focalisant sur leur rôle de régulateur génique et le potentiel impact des protéines résiduelles issues d'ORFs conservés.

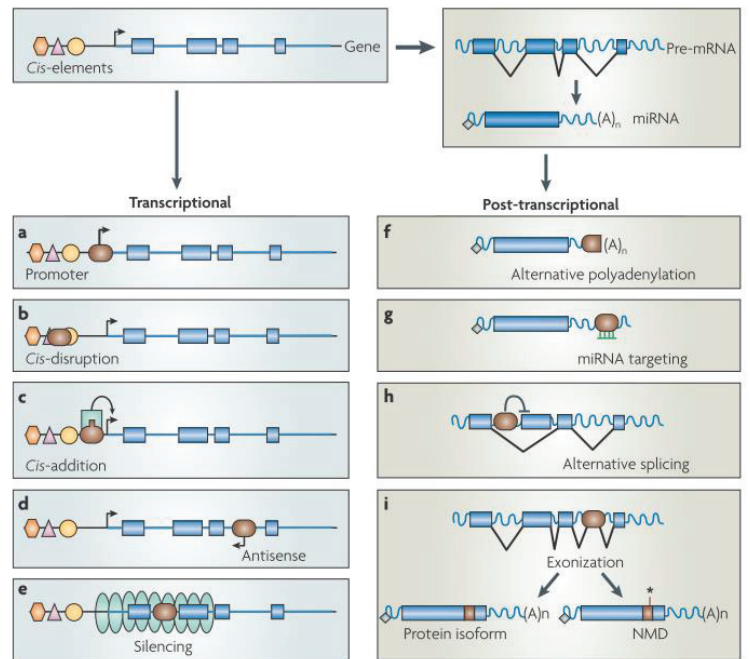
#### *Régulation génique*

Les ETs peuvent impacter l'expression d'autres gènes par de nombreux moyens (Figure 5): au niveau génique, l'insertion d'un ET peut (a) fournir un promoteur alternatif, (b) interrompre un promoteur existant, (c) introduire un motif permettant la liaison d'un nouveau facteur de transcription au sein d'un promoteur, (d) favoriser une transcription antisens interférant avec la transcription canonique, ou bien (e) entraîner la répression de gènes voisins en induisant la formation d'hétérochromatine (147). Au niveau post-transcriptionnel, l'intégration d'un ET peut (f) introduire un signal alternatif de polyadénylation, (g) fournir un site de liaison pour un micro-ARN (miARN) ou pour une protéine, (h) entraîner un épissage alternatif ou encore (i) agir en tant qu'exon alternatif. Ainsi, chez l'Homme, de nombreux promoteurs et signaux de polyadénylation sont dérivés d'ETs (148). Outre ces actions intervenant sur le gène en lui-même, la possibilité pour les ETs d'agir à distance en tant qu'*enhancer* a également été démontrée (149,150).

En fournissant des motifs de régulation génique « prêts à l'emploi », les ETs et en particulier les ERVs avec leurs LTRs ont ainsi pu fournir un moyen de choix pour la formation

de nouveaux réseaux de régulation génique. En effet, si l'on considère un modèle évolutif classique par accumulation de mutations, la régulation de plusieurs gènes par un seul et même facteur de transcription nécessiterait le développement de plusieurs mutations selon le même schéma à plusieurs localisations génomiques différentes (151). Ce modèle, en plus d'être très coûteux, serait également très long à développer. L'intégration de ERVs et notamment de LTRs possédant des motifs de liaison pour des facteurs de transcription à plusieurs endroits dans le génome fournit ainsi des motifs de régulation prêts à l'emploi pour de nombreux gènes, permettant une évolution des gènes en réseau de façon bien plus rapide et moins coûteuse.

Ce modèle a été élégamment démontré dans le cas du réseau de l'IFN- $\gamma$  (152): en analysant des données de séquençage après immunoprécipitation de la chromatine (CHIP-seq) au sein de cellules immunitaires primaires humaines, Chung E. *et al.* ont retrouvé un enrichissement en ETs dans les zones ouvertes se liant à *Signal transducer and activator of transcription 1* (STAT1) et *Interferon Regulatory Factor 1* (IRF1) en réponse à l'IFN- $\gamma$ . Parmi 27 familles d'ETs, 20 correspondaient à des LTRs issus de HERVs. Ces HERVs présentent en fait des motifs de liaison pour STAT1 et IRF1 et sont très fortement enrichis dans les zones proches de gènes stimulés par l'IFN (ISG pour *Interferon-Stimulated Genes*). Les auteurs se sont ensuite intéressés à *Absent In Melanoma 2* (AIM2), protéine clé de la voie de l'inflammasome à l'origine de l'activation de la caspase-1 et de la pyroptose secondaire à la reconnaissance d'ADN double brin d'origine microbienne. Alors que le gène *AIM2* est constitutivement exprimé chez la souris, celui-ci est activé par l'IFN- $\gamma$  chez l'Homme, le seul motif de liaison STAT1 dans les 50kb entourant le gène étant un HERV de la famille MER41 (MER41.AIM2, localisé 220 pb en amont du gène AIM2). La liaison de STAT1 à *MER41* entraîne une acétylation des histones (H3K27) entraînant la transcription d'*AIM2*. Inversement, la délétion de *MER41* par CRISPR-Cas9 entraîne une absence totale d'expression d'*AIM2* avec un déficit majeur de la réponse inflammatoire après traitement par IFN- $\gamma$ . Ainsi, l'intégration de *MER41* en tant qu'élément régulateur d'*AIM2* est un mécanisme de contrôle de la voie de l'inflammasome acquis par les primates et participant à la divergence de lignées avec les autres mammifères tels que la souris.



**Figure 5 : Régulation génique par les éléments transposables.**

Au niveau génique, l'insertion d'un ET peut (a) fournir un promoteur alternatif, (b) interrompre un promoteur existant, (c) introduire un motif permettant la liaison d'un nouveau facteur de transcription au sein d'un promoteur, (d) favoriser une transcription antisens interférant avec la transcription canonique, ou bien (e) entraîner la répression de gènes voisins en induisant la formation d'hétérochromatine. Au niveau post-transcriptionnel, l'intégration d'un ET peut (f) introduire un signal alternatif de polyadénylation, (g) fournir un site de liaison pour un micro-ARN (miARN) ou pour une protéine, (h) entraîner un épissage alternatif ou encore (i) agir en tant qu'exon alternatif.

D'après Feschotte C. Nat Rev Genet. 2008 (136), figure reproduite avec l'autorisation de Springer Nature.

Ces interactions entre LTRs et gènes peuvent également être bénéfiques pour le ERV en lui-même. Chez l'Homme, plus d'un tiers des domaines de liaison pour la protéine p53 sont localisées au sein de LTRs de HERVs (153). En plus de l'apport des LTRs sur le contrôle de la voie p53, cette intégration pourrait également fournir aux HERVs un signal de danger permettant la transcription et l'échappement du rétrovirus d'une cellule sous stress activant cette voie de protection du génome.

#### *Fonctions liées aux protéines des ERVs*

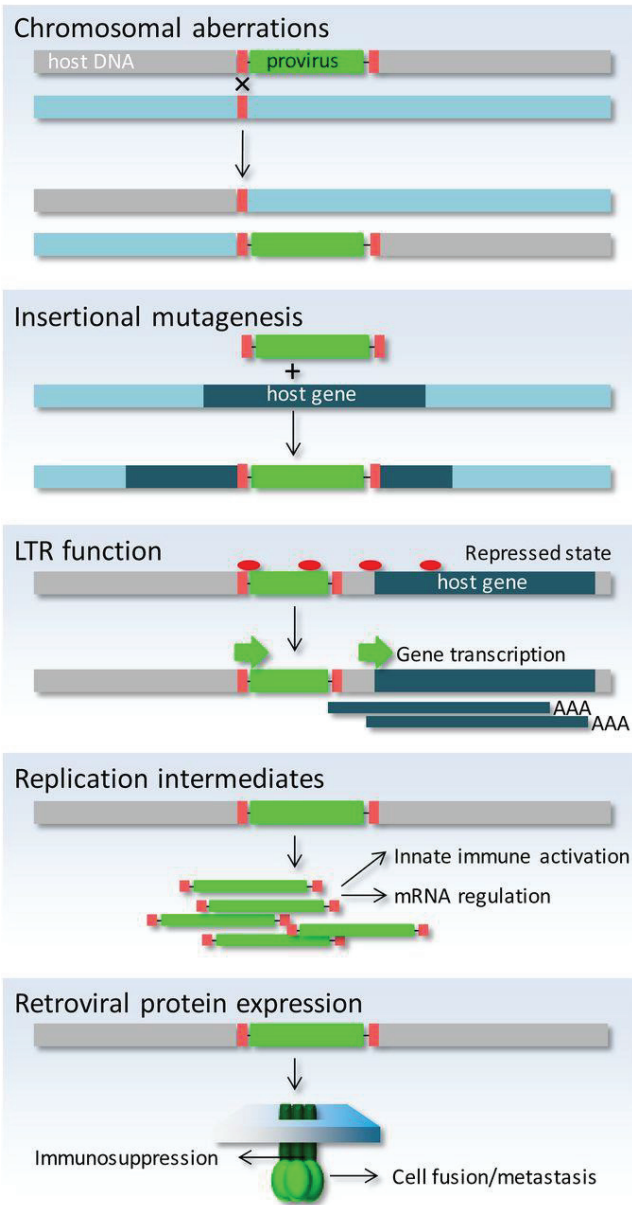
Outre leur activité sur le génome en lui-même, les ERVs ont également pu impacter le développement de leur hôte via les protéines qu'ils produisent. Le meilleur exemple connu à ce jour chez l'Homme est celui de la Syncytine, une protéine permettant la fusion des cellules à l'origine du syncytiotrophoblaste dans le placenta. Cette protéine est en fait dérivée de la protéine d'enveloppe d'un HERV (de la famille HERV-W pour la Syncytine 1, et HERV-FRD pour la Syncytine 2) s'étant intégré dans notre génome (154). La fonction première de la protéine d'enveloppe permettant l'entrée et la sortie cellulaire a ainsi été détournée pour permettre la fusion cellulaire à l'origine de la formation du syncytiotrophoblaste (154,155). En parallèle de cet effet fusogénique, les Syncytines permettraient également de promouvoir la tolérance materno-fœtale. Bien que les mécanismes exacts ne soient pas encore bien établis, cet effet serait principalement médié par le domaine immunosuppresseur propre à la protéine d'enveloppe des rétrovirus, par action locale directe ou à distance via la formation d'exosomes placentaires (156).

Corollaire de cet effet, une baisse de l'expression des Syncytines placentaires est associée à un surrisque de survenue de prééclampsie, une pathologie de la grossesse caractérisée par une réduction de la taille du syncytiotrophoblaste associée à une réponse inflammatoire placentaire exacerbée (157). Les Syncytines pourraient ainsi constituer à la fois un biomarqueur prédictif et une voie thérapeutique dans le dépistage et la prise en charge de la prééclampsie. De nombreux autres gènes dérivés de protéines d'enveloppe de ERVs et impliqués dans la formation du placenta ont été décrits chez une grande variété de mammifères (156). Cette convergence évolutive souligne l'importance potentielle des ERVs dans le développement des mammifères.

### **E. Rôle dans l'oncogenèse**

Dès le début des années 1900, les travaux pionniers de Francis Peyton Rous démontrent la possibilité de transmission tumorale entre des poulets de la même couvée via des extraits acellulaires issus de sarcomes (158). Le rôle d'agents transmissibles de petites tailles, en premier lieu de virus, était dès lors suspecté. Les liens entre virus, en particuliers rétrovirus, et cancérogénèse ont par la suite fait l'objet de multiples travaux ayant joué un rôle pivot dans la compréhension de la biologie du cancer, menant notamment à la découverte de nombreux oncogènes ainsi qu'au mécanisme de transcription inverse et des ERVs (158,159). Les modèles murins ont grandement contribué à ces découvertes, révélant de nombreux mécanismes par lesquels les ETs et en particulier les ERVs peuvent participer à l'oncogenèse. Il faut cependant garder à l'esprit que ces modèles murins sont très limités dans ce contexte et leur extrapolation à l'Homme doit être prudente. En effet, comme nous l'avons évoqué, de nombreux ERVs sont encore complets et très actifs chez la souris, conservant un potentiel infectieux proche de rétrovirus exogènes murins tandis qu'aucun HERV à ce jour n'a été décrit comme actif (c'est-à-dire capable de se rétrotransposer seul) et conservant un potentiel infectieux chez l'Homme (102,159). Il est ainsi estimé qu'environ 0.1% des mutations survenant *de novo* sont liées à des ETs





**Figure 6 : Impact des ERVs sur l'oncogénèse.**

Les ERVs peuvent favoriser l'oncogénèse de multiple façon, en promouvant l'instabilité génomique (A et B), en induisant la transcription de gènes habituellement silencieux via leur fonction LTR (C), en régulant différents processus comme la formation d'ARNm ou la réponse immunitaire via la formation d'ARN, ou encore en activant directement des processus oncogéniques via leurs dérivés protéiques.

D'après Kassiotis G. *J Immunol.* 2014 (159), figure reproduite avec la permission de l'American Association of Immunologists.

les rétrovirus du Koala (KoRVs) sont encore pleinement fonctionnels, produisant des particules virales et se transmettant de façon active. Les Koalas sont par ailleurs connus pour développer un taux importants de tumeurs bénignes et malignes, notamment de lymphomes et de leucémies, l'association entre la charge virale de KoRV et le développement de néoplasies lymphoïdes ayant été démontrée dans certaines populations de Koala (165). Une étude récente s'est penchée sur les

chez l'Homme, contre 10% chez la souris, en grande partie liée aux ERVs (160). Par ailleurs, l'élevage intensif de souris à visée de recherche associé à la consanguinité fréquente nécessaire au maintien de lignées spécifiques a contribué au développement d'un certain polymorphisme insertionnel des ERVs chez la souris, contribuant à l'importante activité des ERVs chez les souris de laboratoire (159,161). Ces limites en tête, nous verront ici les principaux mécanismes décrits à ce jour concernant l'implication des ERVs dans les mécanismes d'oncogénèse (Figure 6).

### *Anomalies chromosomiques et mutagenèse insertionnelle*

La promotion d'une instabilité génomique en lien avec les capacités de mobilisation des ETs est le premier mécanisme généralement évoqué concernant l'oncogénèse liée aux rétrotransposons (Figure 6A et 6B). La transformation d'une cellule par mutagenèse insertionnelle entraînant l'activation d'oncogènes tels que *wnt1* ou *notch1*, ou provoquant l'interruption de gènes suppresseurs de tumeurs, a été largement décrite dans le cadre des ERVs chez la souris (162). Cette instabilité chromosomique peut également être favorisée par des événements environnementaux externes. En étudiant des cellules souches hématopoïétiques (CSH) murines après irradiation corporelle totale, Barbieri *et al.* ont ainsi démontré non seulement qu'une partie des dégâts génomiques induits par l'irradiation est liée à la mobilisation de LINE1, mais également que ces dégâts peuvent être limités par l'action de la thrombopoïétine via l'activation d'une réponse IFN de type 1 (163).

Traversant actuellement une période active d'invasion de sa lignée germinale par des rétrovirus, le Koala est un modèle souvent utilisé pour analyser l'impact de l'intégration de nouveaux ERVs (164). Ayant débuté leur invasion il y a environ 50 000 ans,

liens entre insertion rétrovirale et oncogénèse dans différentes populations de Koala (166). En analysant le génome d'échantillons de tumeurs et de tissus sains en provenance d'une dizaine de Koalas, les auteurs ont retrouvé 1 002 sites d'intégration de KoRV, enrichis au niveau de gènes associés au cancer. Bien que la raison de l'intégration préférentielle au niveau de ces sites en particulier ne soit pas bien connue, le fait que ces gènes soient pour la plupart des gènes actifs jouant des rôles généraux dans l'homéostasie et la prolifération expliquent en tout cas que ces zones de chromatine soient ouvertes et particulièrement accessibles aux phénomènes d'intégration. Cette intégration permanente de KoRVs entraîne ainsi une charge mutationnelle liée à ces ERVs particulièrement importante chez le Koala. D'une façon très intéressante, certains sites d'intégration de KoRVs sont partagés entre plusieurs Koala, le nombre de sites partagés étant d'autant plus important que les Koalas sont proches géographiquement. Cette instabilité génomique liée à l'invasion de ERV devrait en théorie faire l'objet d'une sélection négative avec le temps, comme cela l'a été dans le cas des autres mammifères présentant des ERVs anciens et intégrés, fixés et dégénérés. La survie du Koala à cette « invasion » génomique passera inévitablement par l'adaptation de son génome et l'évolution de ses mécanismes de suppression virale.

L'existence de ce mécanisme d'oncogénèse nécessitant l'insertion d'un rétrovirus complet a fait l'objet de nombreux débats chez l'Homme (167). Bien que cette possibilité existe avec des rétrovirus exogènes comme cela est désormais bien démontré pour certains virus tels que HTLV1 (*Human T-Lymphotropic Virus-1*), aucun provirus autonome issu de HERVs n'a été découvert à ce jour. Dans une analyse d'activité des ETs chez l'Homme sur des données de séquençage complet de génome issues de 43 tumeurs avec leurs cellules sanguines appariées, 194 nouveaux sites d'insertion ont été retrouvés, dont 183 issus de LINE1, 10 de séquence Alu et 1 seul de HERV, ce dernier résultant par ailleurs probablement d'une homologie médiée par anomalie de réparation de l'ADN plus que par une réelle rétrotransposition (168). Bien que ce mécanisme d'oncogénèse paraisse moins plausible chez l'Homme, il n'en reste pas moins théoriquement possible, médié non pas par un seul HERV complet mais par plusieurs HERVs conservés. Ainsi, la famille HML-2 récemment intégrée conserve encore de nombreux ORFs intacts qui, de façon collective, contiennent l'information nécessaire à la formation d'un provirus apte à la réplication et la transposition (169). Les recombinaisons entre éléments de la famille HML-2 surviennent naturellement et pourraient être facilitées par des mécanismes exogènes, comme lors d'une infection par un rétrovirus tel que le VIH (170). Bien que la probabilité que de telles recombinaisons aboutissent à la formation d'une copie complète et infectieuse de HERV soit très faible, celle-ci pourrait être augmentée de façon significative en contexte d'immunosuppression comme cela a été montré chez la souris. En utilisant une souche de souris B6 dépourvue de copie entière fonctionnelle de MLVs (*Murine Leukemia Viruses*), Young G. *et al.* ont démontré l'émergence spontanée de MLVs fonctionnels issus de recombinaisons dans des conditions de déficit immunitaire affectant la production d'anticorps (171). Ces MLVs présentaient alors leur plein potentiel infectieux menant à la formation de lymphomes chez ces souris immunodéprimées. D'une façon intéressante, l'émergence de ces MLVs était abolie chez les mêmes souris élevées dans des conditions menant à une réduction ou une absence complète de microbiote intestinal, suggérant un contrôle de l'expression des ERVs par des facteurs environnementaux supplémentaires.

### Modulation génique

Les différentes possibilités de contrôle génique offertes par les ERVs que nous avons déjà passées en revue peuvent également être exaptées par la cellule tumorale elle-même. CSF1R (*Colony-Stimulating Factor 1 Receptor*), un proto-oncogène myéloïde fortement exprimé par les cellules de Reed-Sternberg dans le lymphome de Hodgkin, constitue un exemple caractéristique. Ces cellules ayant perdu l'expression de gènes spécifiques de la lignée B dont celui codant pour le BCR (*B Cell Receptor*) sont caractéristique de ce type de lymphome, dépendant d'autres voies telles que CSF1R pour leur survie. Il a ainsi pu être montré que l'expression de CSF1R dans le lymphome de Hodgkin était liée à l'activation d'un LTR de la famille MaLR (*THE1B*) localisé ~6,2kB en amont du gène (172). L'activation de *CSF1R* médiée par *THE1B* a été détectée dans la quasi-totalité des prélèvements tumoraux de lymphome de Hodgkin testés et également dans certains échantillons tumoraux de lymphomes anaplasiques, mais pas dans d'autres types de lymphomes non-hodgkiniens (172). Les auteurs ont par ailleurs montré que cette activation est secondaire à la perte d'expression de *CBFA2T3*, un corépresseur permettant de maintenir un état d'hétérochromatine au niveau de *THE1B*, cette perte d'expression étant systématique dans les cellules de Hodgkin. En lien avec ces observations, l'acétylation des histones combinée à la déméthylation globale subie par les lymphocytes B induit leur transformation progressive vers un phénotype de cellule de Hodgkin (173). Cet exemple constitue un des premiers cas bien documenté de changement de lignée et d'induction tumorale secondaire à l'activation d'un proto-oncogène médiée par un LTR. Dans une étude plus générale portant sur 7 769 tumeurs de 15 types de cancer différents, Hyo Sik Jang *et al.* ont identifié 129 activations cryptiques d'ETs menant à l'expression de 106 oncogènes différents dans quasiment la moitié des tumeurs analysées, ces événements allant de 12 à 87% des tumeurs selon les types de cancers (174). Parmi les familles d'ETs, la grande majorité étaient des LTRs.

Des nombreuses questions que l'on peut se poser sur ce dialogue entre HERVs et expression d'oncogènes, celle de la causalité est probablement la plus récurrente : comment savoir si le LTR entraîne la transformation initiale (*i.e.* l'expression de l'oncogène), ou si ce sont simplement les phénomènes de déméthylation globaux propres à la tumeur qui entraînent une transcription large de HERVs accompagnant celle des oncogènes (simple effet « *bystander* ») ? Dans une étude récente, Özgen *et al.* ont identifié plusieurs ERVs activant l'expression de gènes participant au blocage de maturation et au développement tumoral dans un modèle murin de LAM (175). Plusieurs démonstrations fournies dans cet article, dont le fait que la répression de ces ERVs inhibe par ailleurs la croissance des cellules leucémiques *in vitro* en induisant leur apoptose, permettent de confirmer le rôle régulateur de ces ERVs plutôt que simple épiphénomène accompagnant la transcription de gènes impliqués dans le développement tumoral. L'étude sur CSF1R vient par ailleurs apporter toute la mécanistique nécessaire à la démonstration de la causalité en faveur du LTR entraînant la transformation tumorale. Il est ainsi très probable que les deux phénomènes coexistent en pratique, certains HERVs participant véritablement à l'oncogénèse là où d'autres ne sont que marqueurs d'une certaine activité transcriptionnelle et/ou ouverture chromatinienne.

### Synthèse de protéines oncogènes

Bien que déficients pour la plupart, de nombreux HERVs conservent un ou plusieurs ORFs intacts pouvant produire des protéines d'origine rétrovirales (159,176). Ces protéines peuvent être reconnues par le système immunitaire, comme nous le verrons plus en détail dans la



section F. En plus de cet impact immunitaire, certaines protéines issues de HERVs peuvent avoir des effets oncogéniques directs. La protéine d'enveloppe de HERV-K peut par exemple activer la voie ERK et promouvoir ainsi la transformation cellulaire (177). L'inactivation de la transcription du gène de la protéine d'enveloppe de HERV-K par interférence ARN dans des lignées de cancer du sein (178) ou de mélanome (179) diminue par ailleurs leur prolifération, leur migration et leur caractère invasif, suggérant un rôle direct des protéines rétrovirales endogènes dans le phénotype tumoral. Np9, une protéine issue d'un épissage alternative de la séquence *env* de HERV-K, agit en tant qu'oncogène leucémique en activant les voies ERK, AKT et NOTCH1 ainsi qu'en favorisant la production de  $\beta$ -Caténine, essentielle pour la survie des cellules souches leucémiques (180). L'expression de Np9 est par ailleurs détectée dans 78% des lignées cellulaires testées et 56% d'échantillons primaires de patients présentant une LAM ou une leucémie myéloïde chronique (LMC), contre seulement 4,5% des échantillons de donneurs sains, suggérant un profil d'expression spécifique du cancer. Là encore, la répression de *np9* inhibe la croissance tumorale de lignées leucémiques myéloïdes et lymphoïdes ainsi que la croissance tumorale dans des modèles murins immunodéficients NOD-SCID (*Non-obese diabetic - Severe Combined ImmunoDeficiency*) (180).

L'onco-exaptation des HERVs est donc un mécanisme fréquent utilisé par la tumeur pour activer des voies de signalisation oncogénique qui lui sont indispensables. Toutefois, bien que la dérégulation de certains HERVs puisse constituer une force pour la cellule tumorale en lui fournissant de nombreux moyens de moduler l'expression génique et favoriser l'instabilité génomique, la réactivation de ces mêmes HERVs fournit également plusieurs voies thérapeutiques prometteuses comme nous le verrons dans la prochaine partie.

## F. Impact immunitaire

Les relations entre les différents produits des ERVs (autant nucléotidiques que protéiques) et le système immunitaire sont complexes et encore mal comprises. Plus qu'une relation unilatérale, c'est un véritable dialogue bilatéral qui est en place, avec d'une part une régulation de l'expression des ERVs par le système immunitaire, et d'autre part une régulation du système immunitaire par les ERVs.

### Réponse innée

Dans la **Revue 7** (181), nous avons passé en revue l'état des connaissances sur l'impact de la réponse immunitaire innée sur l'expression des HERVs, ainsi que sur la modulation de la réponse immunitaire innée par les HERVs eux-mêmes. Nous avons ensuite discuté l'utilisation des HERVs en tant qu'« adjuvant » potentiel dans la réponse immunitaire antitumorale, en évoquant notamment les différentes stratégies visant à réactiver les HERVs pour augmenter l'immunogénicité de la cellule tumorale.

Plus récemment, il a été montré que les HERVs participent à la régulation de la balance entre réponse homéostatique et inflammatoire de l'hôte envers son propre microbiote (182). En utilisant un modèle murin de colonisation cutanée à *Staphylococcus epidermidis*, les auteurs ont montré que les kératinocytes murins expriment différents ERVs par reconnaissance de motifs moléculaires associés au pathogène (notamment l'acide téichoïque et la lipoprotéine, via les Récepteur Toll-like (TLR) 2 et 4). Ces ERVs possèdent tous un ORF intact *pol* permettant de coder une RT et de générer de l'ADNc stimulant la voie *cyclic GMP-AMP synthase/stimulator of interferon genes protein* (cGAS/STING). Il en résulte une accumulation tonique de différentes

cellules immunitaires au niveau cutané (LT CD8<sup>+</sup>, CD4<sup>+</sup> T<sub>H</sub>1 et T<sub>H</sub>17, lymphocytes invariants associés aux muqueuses (MAIT pour *Mucosal-Associated Invariant T cells*)) intervenant à la fois dans le contrôle des bactéries commensales mais également dans différents processus physiologiques tels que la réparation tissulaire. D'une façon intéressante, les auteurs montrent qu'un régime riche en lipides entraîne une augmentation de l'expression des ERVs en réponse à la colonisation par *Staphylococcus epidermidis*, faisant cette fois-ci pencher la balance vers une réponse inflammatoire. Dans les deux situations, les différentes réponses observées étaient abolies par l'inactivation de la voie cGAS/STING ou l'utilisation d'un traitement inhibiteur de la RT réduisant la production d'ADNc sans impacter l'expression des ERVs. Les souris exposées à un régime riche en lipide et traitées par antirétroviraux ne présentaient ainsi pas les signes d'inflammation cutanée précoce observées chez les souris non traitées. Ce dialogue entre ERVs et microbiote pourrait ainsi constituer un des mécanismes clés de régulation de la réponse homéostatique, contrôlant notamment la balance immunitaire.

### Réponse adaptative

La proportion exacte de HERVs contenant un ORF intact et pouvant coder pour une protéine rétrovirale est difficile à estimer. Il existe à ce jour 71 références validées de protéines issues de HERVs dans la base de données Uniprot (183), mais ce chiffre est largement sous-estimé. Dans une analyse utilisant l'outil Retrotector qui permet de détecter et localiser les séquences provirales dans un génome donné, Vargiu L. *et al.* ont identifié 3 178 loci HERVs contenant au moins un ORF (100). Combinant plusieurs outils, la base gEVE rapporte elle plus de 5 000 ORFs dérivés de HERVs dans le génome humains (184). Du fait de leur origine virale, les protéines issues de ces ORFs pourraient constituer une source importante de stimulation antigénique pour le système immunitaire.

Les superantigènes sont des antigènes capables d'activer une large portion de LT (10-40% versus 0,01% pour les antigènes conventionnels) via une liaison directe au CMH de classe II ainsi qu'à certaines chaînes bêta du TCR (185). Sutkowski N. *et al.* ont montré vers la fin des années 1990 que les LB infectés par le virus d'Epstein-Barr (EBV) expriment un superantigène entraînant une expansion en 4-6h de LT- $\alpha\beta$  possédant la chaîne bêta TCRBV13, suivi d'une expansion polyclonale de LT après 48h d'une façon dépendante du CMH de classe II (186). Cette activation participerait à la persistance virale de l'EBV dans les lymphocytes B mémoires en leur prodiguant notamment les signaux de costimulation nécessaires à leur formation (CD40-CD40L) via les LT activés. Tandis qu'aucun superantigène n'a pu être retrouvé dans le génome de l'EBV, les auteurs ont finalement démontré que cette activité superantigène était liée à une protéine d'enveloppe issue de l'activation transcriptionnelle de HERV-K18, un HERV récemment intégré au niveau du premier intron du gène *CD48* transactivé par l'EBV (186). D'une façon remarquable, les primates dépourvus de HERV-K18 dans leur génome ne peuvent pas établir d'infection persistante à l'EBV (186,187).

En parallèle de cette mise en lumière des HERVs comme potentiels superantigènes, de nombreux travaux portant sur HERV-K10 ont permis la découverte de la production de protéines virales issues de HERVs par la cellule tumorale. Il a ainsi été montré dans les années 1990 la présence de la protéine Gag de HERV-K10 par microscopie électronique dans des lignées puis échantillons primaires de tératocarcinome (188,189). La fonctionnalité de la protéase produite par HERV-K10 a ensuite été démontrée (190). La mise en évidence directe de protéines

rétrovirales issues principalement de Gag et Env de la famille HERV-K suivront notamment dans le sein (113) et le mélanome (191). Les approches d'immunologie inversée ont par la suite permis d'identifier la production d'anticorps anti-Gag en Env survenant naturellement chez les patients atteints de séminomes (189,192). La mise en évidence de réponses T spécifiques se développera quant à elle à partir du début des années 2000 avec l'identification du premier peptide HERV antigénique présenté par le HLA-A\*02 dans le mélanome (HERV-K-MEL - MLAVISCAV) (193). Malgré tout l'intérêt porté au sujet, un nombre relativement faible d'études suivra (sélection des principales études dans le tableau 2), principalement en lien avec les difficultés d'identification de nouvelles cibles (toutes les études reprenant pour la plupart la même protéine Gag ou Env issue de HERV-K10) ainsi que la lourdeur des méthodes de validation des réponses spécifiques, passant notamment par l'établissement de clones via des bibliothèques d'ADNc.

La découverte de réponses spontanées chez les patients a conduit à la recherche d'induction de réponse T et/ou B spécifiques en vue d'approches immunothérapeutiques. Malgré quelques études concluantes notamment chez la souris (194) ou le macaque rhésus (195), peu d'exemples ont suivi chez l'Homme, hormis l'exemple notable de l'équipe de Pierre Coulie ayant établi des clones de LT dirigés contre l'antigène HERV-K-MEL et prouvé leur cytotoxicité face à des cellules primaires autologues issues de mélanomes de patients (193). Des lymphocytes CAR-T dirigés contre la protéine d'enveloppe de HERV-K ont même été développés par l'équipe de Feng Wang-Johanning dans le mélanome et le cancer du sein, avec des résultats encourageants *in vitro* et *in vivo* chez la souris (196,197). De façon notable, la régulation négative de l'expression de la protéine d'enveloppe de HERV-K chez les souris traitées par lymphocytes CAR-T était associée à une surexpression de la protéine p53 et une diminution des voies oncogéniques ERK et RAS (197).

Plus récemment, l'avènement des méthodes d'analyse de biologie moléculaire a signé le début d'une nouvelle période d'identification plus large d'épitopes, se basant sur un grand nombre de candidats identifiés dans des bases de séquençage ARN suivi d'une prédiction des épitopes ayant le plus d'affinité pour un HLA donné via des outils bioinformatiques. La technologie des multimères identifiés par code-barres ADN vient illustrer ce changement : les multimères sont des polymères de dextrose couplés à un fluorochrome et comprenant plusieurs molécules de CMH liées à un peptide d'intérêt. Leur utilisation permet la détection de populations lymphocytaires T spécifiques d'un épitope donné en cytométrie de flux. En couplant un code-barre ADN unique à chaque multimère, il devient désormais possible de rechercher la présence de LT spécifiques de plus de 1 000 couples HLA-peptide dans un seul échantillon (198). L'équipe de Sine Hadrup au Danemark a utilisé cette technique pour rechercher des réponses LT CD8<sup>+</sup> spécifiques pour 1 169 épitopes HERV potentiels dans différentes hémopathies myéloïdes (principalement syndrome myélodysplasiques (SMD) et LAM), permettant d'identifier des réponses spécifiques pour 29 de ces peptides (199). Bien que séduisante, cette approche peut manquer de sensibilité et de spécificité comme nous le verrons dans la discussion.

Type de réponse	Famille	Protéine	Épitope	Maladie	Référence
Anticorps IgG (Immunoblot & IHC)	HERV-K10	Gag	Partie N-ter de Gag	Séminome (45%), tératome (11%), hémopathies (LAM, LAL, LMC, LLC, LH) (~2-5%), Cancer du sein (2,9%)	Sauter et al. 1995
Anticorps IgG (IF)	HERV-K10	Env	OM (56-465 aa) et TM (463-699 aa)	Séminome (85%)	Sauter et al. 1996
CTL (Clones par librairie ADNc et tri Tétramères, Cr51)	HERV-K	Env	MLAVISCAV (HERV-K-MEL) (HLA-A*02)	Mélanome	Schiavetti et al. 2002
LT (ELISPOT & Tetramères)	HERV-K	Gag	FLQFKTWWI (HLA-A*02)	Séminome	Rakoff-Mahoum et al. 2006
Anticorps IgG (ELISA) LT (ELISPOT) CTL (PBMCs, Cr51)	HERV-K	Env, Gag	Env: 30-kDa SU Gag: NP	Cancer du sein (88%)	Wang-Johanning et al. 2008
CTL (Clones par librairie ADNc, Cr51)	HERV-E	Env	ATFLGSLTWK (CT-RCC-1) HLA-A*11	Cancer rénal à cellules claires	Takahashi et al. 2008
LT (ELISPOT) CTL (PBMCs, Cr51)	HERV-H	Env	CLYPFSAFL (RVP03) & SLNFNSFHFL (RVP09) (HLA-A*02)	Cancer colorectal	Mullins et al. 2012
LT (ELISPOT)	HERV-E	Env	FLHKTSVREV (CR1), SLNITSCYV (SU1) & LLLQIMRSFV (TM1) (HLA-A*02)	Cancer rénal à cellules claires	Cherkasova et al. 2016
LT (Tétramères)	HERV-E	Gag, Pol	Pool de tétramères NSWQEMVPV & MVGPWPRPV HLA-A*02	Cancer rénal à cellules claires	Smith et al. 2018
LT (Dextramères code-barres)	HERV-K, HERV-FRD1 & 2, HERV-E, HERV-S	NP	29 épitopes pour différents HLA	SMD, LAM	Saini et al. 2020

**Tableau 2 : Sélection d'études mettant en évidence une réponse adaptative contre des épitopes issus de HERVs.**

Seules les études réalisées chez l'Homme sont reprises ici. Concernant les réponses T, LT réfère à la description d'une réponse spécifique avec éventuellement des tests de fonctionnalité, CTL à la mise en évidence d'une réponse cytotoxique via l'utilisation de lymphocytes polyclonaux stimulés (PBMCs) ou l'établissement d'un clone.

CTL : Lymphocytes T Cytotoxiques, IF : Immunofluorescence, IHC : Immuno-Histo-Chimie, LAL : Leucémie Aiguë Lymphoblastique, LAM : Leucémie Aiguë Myéloïde, LH : Lymphome de Hodgkin, LMC : Leucémie Myéloïde Chronique, NP : Non précisé, SMD : Syndrome Myélodysplasique, TM : Transmembrane Unit

### Tolérance immunitaire

La tolérance centrale se déroule dans le thymus et est un des principaux mécanismes de tolérance immunitaire (200). La première étape de sélection dite « positive » se déroule dans le cortex thymique et permet de procurer un signal de survie aux LT- $\alpha\beta$  capables de reconnaître un complexe CMH-peptide à la surface des cellules corticales thymiques (cTECs pour *Cortical Thymic Epithelial Cells*). La deuxième étape de sélection dite « négative », garante de la production de LT tolérants pour les antigènes du soi, se déroule dans la médullaire thymique. Les cellules médullaires épithéliales thymiques (mTECs pour *Medullary Thymic Epithelial Cells*) y jouent un rôle majeur en présentant des antigènes du soi exprimés hors de leurs tissus d'origine grâce à la protéine AIRE (*Auto-Immune REgulator*), permettant d'éliminer les LT- $\alpha\beta$  présentant une forte affinité pour les antigènes du soi (201). Afin d'éviter une sélection négative trop stringente qui pourrait entraîner des trous dans le répertoire TCR- $\alpha\beta$ , ce processus de sélection négative ne conduit pas à l'élimination complète des LT auto-réactifs mais plutôt à un élagage de ceux-ci (202). Afin de compenser ce déficit, un des rôles de la médullaire thymique est également de convertir une partie des thymocytes CD4<sup>+</sup> en LT régulateurs permettant de contrôler en périphérie la réactivité des LT ayant échappé à la tolérance centrale (203).

Dans l'étude de l'équipe de Claude Perreault précédemment citée, la plus forte expression de HERVs parmi les tissus sains a été retrouvée dans deux tissus n'exprimant pas le CMH de classe I (le testicule et les cellules souches embryonnaires), ainsi que dans les mTECs où leur expression semble indépendante de la protéine AIRE (138). Certains HERVs pourraient ainsi être directement impliqués dans les phénomènes de tolérance centrale thymique. Dans une étude plus ancienne de 2009, Ebert P. *et al.* ont identifié un polypeptide Gag-Pol-Env permettant la sélection thymique positive de LT CD4<sup>+</sup> chez la souris (204). La présence de tels peptides au niveau thymique pourrait ainsi participer au maintien de l'homéostasie des LT, sur le modèle des peptides dits « privés » ou « publics » (201). Contrairement aux peptides « publics » présentés en périphérie, les peptides « privés », générés par une voie alternative du protéasome dans les cTECs (voie dite du thymoprotéasome via la sous-unité catalytique  $\beta 5t$ ), sont des peptides à expression thymique exclusive et absents des tissus périphériques. Leur expression par les cTECs permet de promouvoir la sélection positive d'un grand nombre de thymocytes via des interactions plus faibles avec le TCR, permettant une perte de la liaison TCR-peptide plus rapide. A l'issue de la sélection thymique, deux populations distinctes de LT pourraient ainsi être distinguées : les LT sélectionnés par des peptides « publics », présentant un TCR de plus forte affinité et exprimant fortement le CD5, maintenus dans un état basal d'activation via une stimulation périphérique chronique de leur TCR (205), et les LT sélectionnés par des peptides « privés », présentant un TCR de faible affinité et exprimant faiblement le CD5, présentant un plus faible potentiel d'activation et d'expansion et ne bénéficiant pas de stimulation basale chronique du fait de l'absence de cibles en périphérie. L'existence d'une telle population de LT CD5<sup>faible</sup> moins réactive pourrait présenter plusieurs avantages : en effet, un répertoire ne présentant que des clones fortement réactifs pourrait favoriser l'apparition de clones auto-réactifs. Par ailleurs, la population de LT CD5<sup>faible</sup> pourrait entraîner une réponse immunitaire retardée mais plus persistante dans le temps, là où la population hautement réactive de LT CD5<sup>forte</sup> pourrait être plus sujette à la mort induite par l'activation suite à une activation et expansion très rapide (205). Le rôle des ERVs comme source d'épitopes privés ou publics dans le thymus reste à préciser chez l'homme.



La présence de peptides HERV dans le thymus pourrait également avoir un rôle dans la sélection négative des LT. L'équipe de Georges Kassiotis a ainsi montré qu'un peptide issu de ERV participe à la sélection négative d'une fraction polyclonale de LT CD4<sup>+</sup> dirigés contre le rétrovirus exogènes de Friend chez la souris (206). En analysant de plus près le répertoire et l'avidité des LT sélectionnés, les auteurs se sont rendu compte que cette sélection négative permettait d'éliminer de façon préférentielle les clones de faible affinité pour le virus de Friend, présentant une plus forte cross-réactivité pour les peptides du soi. Cet élagage des clones de faible affinité permet ainsi la sélection de clones de plus forte avidité pour le rétrovirus exogène. Chez l'Homme, il a été démontré que l'expression thymique du superantigène de HERVK-18 est nécessaire et suffisante pour la sélection négative des thymocytes CD4<sup>+</sup> Vβ7 (207). La sélection plus ou moins complète au niveau thymique pourrait expliquer les niveaux variables de LT CD4<sup>+</sup> Vβ7 retrouvés chez les sujets sains, et les différences de réactivité périphérique face au superantigène qui en découlent.

Bien que les études soient peu abondantes, ces quelques exemples illustrent comment l'expression endogène d'un ERV peut façonner le répertoire TCR et l'activation des LT. L'identification des HERVs exprimés par les différentes zones thymiques, notamment par les cTECs et les mTECs, est ainsi un élément majeur à prendre en compte lors de la recherche de cible immunothérapeutique. Le contrôle épigénétique de l'expression des HERVs, associé à la possibilité d'immunisation contre des peptides HERVs issus de la protéine Gag ou Env prouve toutefois que cette tolérance n'est pas complète, amenant certains auteurs à parler d'ignorance immunitaire plus que de tolérance (102). Cela explique également le grand intérêt porté aux HERVs dans les maladies auto-immunes, différents auto-antigènes issus de HERVs ayant déjà été identifiés comme potentiellement responsable de réactions auto-immunes (102).

#### *Activité immunosuppressive des HERVs*

Comme nous l'avons déjà évoqué avec l'exemple des Syncitines, la protéine d'enveloppe de nombreux rétrovirus aussi bien endogènes qu'exogènes contient un domaine immunosuppresseur (ISD pour *ImmunoSupressive Domain*) (208). Bien que cet effet soit désormais communément admis, l'effet immunosuppresseur de la protéine Env est encore mal compris. Morozov *et al.* ont montré qu'un peptide synthétique correspondant à la séquence ISD de la protéine d'enveloppe de HERV-K suffit à reproduire les effets de la protéine complète, entraînant une inhibition de l'activation et de la prolifération de cellules sanguines périphériques mononuclées (PBMCs pour *Peripheral Blood Mononuclear Cells*) face à un mitogène ainsi qu'une augmentation de la synthèse de nombreuses cytokines, notamment de l'IL-10, IL-6 et IL-8 (209). Cette séquence ISD est par ailleurs très conservée avec d'autres rétrovirus exogènes notamment le VIH-1. L'exposition de monocytes purifiés au domaine ISD du VIH (protéine gp41) reproduit l'augmentation d'IL-10 associée à une baisse d'IL-2, d'IL-4 et une synthèse d'IFN-γ, qui ne se retrouve pas en cas d'exposition des lymphocytes T, B ou *Natural Killer* (NK) seuls (210). Ces effets sont par ailleurs totalement abolis en présence d'anticorps anti-IL-10, suggérant une médiation de l'effet immunomodulateur de la protéine d'enveloppe passant par la synthèse d'IL-10 par les monocytes/macrophages.

Les études sur l'activité immunosuppressive de la protéine d'enveloppe peuvent parfois être discordantes : ainsi la Syncitine-1 a été mise en avant comme non immunosuppressive comparée à la Syncitine-2, suite à des résultats obtenus dans un modèle murin de rejet de tumeurs

(211). En revanche, l'exposition de PBMCs humains à des exosomes contenant de la Syncytine-1 inhibe la synthèse de TNF- $\alpha$ , IFN- $\gamma$  et CXCL10 en réponse à une stimulation par lipopolysaccharide et phytohaemagglutinin (LPS/PHA) (212). Plus étonnant encore, une protéine recombinante dérivée de la Syncytine 1, ne comprenant pas le fameux domaine ISD, inhibe la réponse au LPS et la synthèse d'IL-1- $\beta$  dans des PBMCs de donneurs sains (213). Hummel *et al.* ont par ailleurs montré que des lignées de choriocarcinome (Bewo ou JEG) exprimant trois protéines d'enveloppe distinctes (Syncytine-1, -2 et la protéine d'enveloppe de HML-2) inhibent l'expansion de LT lors d'une coculture, contrairement à des cellules CHO (lignée de cellules épithéliales ovariennes de hamster) transfectées pour exprimer l'une de ces trois protéines ou bien les trois en même temps (214). En réalisant une coculture avec des cellules dendritiques induites à partir de monocytes via l'utilisation de LPS, les auteurs ont observé une stimulation de la synthèse de TNF- $\alpha$  et d'IL-12 dans les conditions CHO, alors qu'une induction de la synthèse d'IL-10 était observé avec les lignées de choriocarcinome. Les deux types cellulaires entraînaient cependant une réduction de l'expansion de LT lors d'une coculture incluant les cellules dendritiques, en lien avec une diminution des contacts entre les cellules dendritiques et les LT ainsi qu'une diminution des flux calciques. Ces résultats soulignent le rôle potentiel de facteurs complémentaires dans l'activité immunosuppressive des protéines d'enveloppe, agissant en partie de façon indirecte via une modulation de l'activité des cellules dendritique.

L'action immunorégulatrice des HERVs peut également jouer de façon indirecte. Ainsi, l'intégration d'un HERV en amont du gène codant pour CD5 entraîne la production d'un transcrite alternatif (E1B) responsable d'une baisse du transcrite classique (E1A) et donc du niveau de CD5 membranaire dans des lymphocytes B normaux (215). Cette diminution du niveau de CD5 membranaire empêche l'action régulatrice de CD5 via son domaine SHP-1 (*SH2-containing Phosphatase 1*), favorisant *in fine* la production d'anticorps. La balance entre les deux transcrits E1A et E1B jouerait ainsi un rôle central dans la régulation de l'activation des lymphocytes B.

#### *Régulation génique de la réponse immunitaire*

Outre l'impact immunitaire de leurs produits protéiques et nucléotidiques, les HERVs participent au contrôle de la réponse immunitaire à un niveau supérieur via le contrôle génique qu'ils exercent. Chez la souris, l'inactivation de *SETDB1* dans les LT CD4<sup>+</sup> entraîne un biais de différenciation vers un profil T<sub>H</sub>1 par la dérégulation de ERVs contrôlant des gènes associés à ce phénotype (agissant soit en tant qu'*enhancers*, soit régulateurs de gènes du réseau T<sub>H</sub>1, soit simplement en permettant l'ouverture de la chromatine à proximité de ces gènes) (144). En réprimant certains ERVs spécifiques contrôlant la voie T<sub>H</sub>1, *SETDB1* permet donc de maintenir la stabilité des lymphocytes CD4<sup>+</sup> T<sub>H</sub>2. Toujours chez la souris, une équipe japonaise a récemment montré que les cellules dendritiques murines perdent progressivement l'expression de TRIM28 avec l'âge, ce qui induit une perte des marques répressives H3K9me3 au niveau de différents ERVs (216). Cette dérégulation des ERVs entraîne l'expression de plusieurs gènes à proximité de leurs LTRs, entraînant notamment une réponse inflammatoire et une expansion des LT CD4<sup>+</sup> T<sub>H</sub>17 se traduisant par une augmentation des manifestations auto-immunes et inflammatoires. La perte progressive du contrôle des ERVs pourrait ainsi être un des mécanismes participant à l'immunosénescence et la réponse inflammatoire liée à l'âge (217). Enfin, il a également été montré que l'intégration d'ETs et en particuliers de HERVs au niveau des portions introniques des gènes codant pour le CMH de classe II (notamment DRB) a permis de



promouvoir la diversité des molécules du CMH (218). La mobilisation d'ETs serait également source de diversité pour le CMH de classe I, à l'origine notamment de la duplication génique ayant conduit à la divergence des HLA-B et -C (219).

L'activité *enhancer* des HERVs pourrait également jouer un rôle important dans la mémoire immunitaire innée. Une équipe chinoise a ainsi montré que l'exposition de cellules rénales humaines embryonnaires HEK293F au TNF- $\alpha$  entraîne une déméthylation médiée par des enzymes de la famille TET de régions contenant des motifs de liaison pour *Nuclear factor-kappa B* (NF- $\kappa$ B), potentialisant ainsi la réponse aux expositions futures (220). Parmi les zones déméthylées, deux ERVs agissant en tant qu'*enhancers* ont été identifiés comme permettant la plus forte mémoire et potentialisation de la réponse au TNF- $\alpha$ .

### *Mimétisme viral (Viral mimicry)*

Bien que déjà discutée dans notre revue, le lien entre HERVs et réponse aux agents hypométhylants mérite d'être repris ici. Depuis les travaux de Chiappinelli L. *et al.* et Roulois D. *et al.* en 2015 (221,222), il est bien établi que l'un des mécanisme d'action des agents hypométhylants passe par la réactivation de HERVs par la cellule tumorale, imitant un état d'infection virale avec production d'ARN double brin (dsRNA) stimulant la réponse IFN de type I et III (181). Il semblerait que les HERVs plus récemment intégrés soient préférentiellement réactivés par les phénomènes de déméthylation induits par l'AZA (223).

En parallèle de cette stimulation immunitaire « innée », il a été montré que les inhibiteurs de DNA-méthyltransférase (DNMTi) et de Histone désacétylase (HDACi) entraînent la transcription *de novo* de LTRs cryptiques menant à la production d'isoformes de protéines générant des néoépitopes (224). Cette production accrue de protéines « tronquées », associée à l'augmentation de la quantité de molécules du CMH de classe I liée à la réponse IFN de type I, pourrait permettre de favoriser la présentation d'épitopes T CD8<sup>+</sup>, posant le rationnel d'une combinaison thérapie épigénétique – ICPI en particulier dans les tumeurs faiblement mutées telles que les hémopathies (225). Par ailleurs, l'effet des DNMTi ne se limiterait pas aux cellules tumorales et pourrait également permettre d'agir sur les cellules effectrices elles-mêmes : ainsi, le blocage de la méthylation de l'ADN des LT CD8<sup>+</sup> activés permet de prévenir leur épuisement, permettant d'intensifier et de prolonger la réponse immunitaire sous ICPI (226).

Désormais documentés pour les agents dits hypométhylants, d'autres molécules pourraient agir en partie par cet effet de mimétisme viral : citons par exemple l'acide tout-trans rétinoïque (ATRA), molécule connue principalement pour son effet différenciant dans le traitement des leucémies aiguës promyélocyaires. Des études pré-cliniques suggèrent un intérêt de l'ATRA dans le traitement des cancers du sein, cet effet passant en partie par l'activation de HERVs entraînant une activation de la réponse IFN de type I (227). La réponse inflammatoire associée à la radiothérapie pourrait également agir en partie par cet effet comme cela a été montré avec l'expression de différents HERVs entraînant une réponse IFN de type I secondairement à l'irradiation de lignées de monocytes humaines (228). Enfin, un article récent publié dans *Cancer Discovery* montre que l'effet des inhibiteurs de MDM2, molécules activatrices de la voie p53, passe par l'inhibition de LSD1 et DNMT1 entraînant la dérégulation de ERVs induisant une réponse IFN de type I et une réponse antitumorale par mimétisme viral (229).

Les relations entre les HERVs et le système immunitaire sont donc complexes et se jouent à plusieurs niveaux, ouvrant de nombreuses voies à la fois de compréhension physiopathologique et d'approche thérapeutique. Le décryptage de ces différentes interactions passe inévitablement par une appréciation et une quantification juste de l'expression des HERVs, ce que nous discuterons dans la dernière partie.

## **G. Quantification des HERVs : aspects techniques et bioinformatiques**

Du fait de leur nature même de séquence répétée et de leur abondance dans le génome, les ETs représentent un véritable défi en termes d'identification et de quantification (230). De plus, la transcription d'ETs peut refléter non seulement la transcription d'un ET en lui-même, mais également celle d'un gène contenant une portion d'ET par un phénomène de co-transcription passive ou de transcription du brin situé en sens inverse (230). Si les anciens ETs tendent à devenir uniques par accumulation de mutations, les ETs plus récemment acquis présentent très peu de divergences de séquences, majorant la problématique d'alignements multiples (*multimap*) (230). L'activité de certaines familles d'ETs (notamment les LINEs) entraîne par ailleurs un polymorphisme non négligeable. Chez l'Homme, environ 20% du polymorphisme total du génome est lié à de nouvelles insertions d'ETs (231). Ces différentes difficultés expliquent pourquoi les ETs ne sont habituellement pas analysés dans les analyses de séquençage ARN classique, ceux-ci étant généralement masqués.

Dans cette partie, nous passerons en revue les principaux aspects théoriques et techniques de l'analyse et la quantification des ETs, en nous concentrant sur les problématiques plus spécifiques liées à l'analyse des HERVs. Nous verrons d'abord quelles sont les différentes bases de données référençant les ETs et les HERVs, avant de nous concentrer sur les techniques de quantification elles-mêmes et leurs limites. Nous finirons par comparer 5 pipelines d'analyse des HERVs que nous avons sélectionné et retenu comme particulièrement intéressants.

### **1) Base de données / référence**

La première question à se poser avant toute analyse de HERVs est de savoir sur quelle base de données / référence nous nous reposons. Il existe trois principaux type de base de donnée d'ETs: les bases centrées sur les ETs, les bases centrées sur le génome et celles centrées sur les polymorphismes (232). Les plus souvent, ces bases de données sont associées à un outil permettant de les alimenter.

#### *Bases centrées sur les éléments transposables*

Ces bases se concentrent sur les séquences d'ETs en elles-mêmes, en référençant les séquences consensus regroupées en sous-familles et familles. Ce type de base vise généralement à regrouper de façon exhaustive l'ensemble des séquences possibles appartenant à la famille des ETs, et sont utilisées comme base de travail pour les programmes de localisation génomique. Parmi les différentes bases existantes, Repbase constitue incontestablement une des plus grandes références en termes de référencement des ETs (233). Publiée dans sa première version en 1992 et mise à jour régulièrement depuis, Repbase référencie toutes les séquences existantes d'ETs sous forme de séquences consensus ou de séquence représentative.

Le caractère participatif et contributif de Repbase (les séquences publiées étant soumise par différents auteurs et contributeurs suite à l'identification de nouvelles séquences) constitue autant la force qu'une potentielle faiblesse de cette base, ne proposant pas de *screening*

systématique des génomes. Dfam propose une approche complémentaire, se basant sur les séquences connues et publiées d'ETs (notamment celles de Repbase) pour aller identifier dans les génomes les séquences potentielles d'ETs via un modèle de Markov caché (234). L'approche de Dfam permet une annotation plus systématique des différentes familles d'ETs et facilite également l'annotation de séquences anciennes ayant fortement divergé de leur séquence d'origine par accumulation de mutations. Le caractère gratuit et *open-source* de Dfam le rend désormais très populaire depuis le passage au format payant de Repbase.

Développé initialement pour masquer les séquences répétées, Repeatmasker est devenu peu à peu un outil de référence pour la détection et le référencement des ETs (235). Utilisant une base de données de référence (telle que Repbase ou Dfam), le programme permet de détecter les séquences répétées dans une séquence ADN via différents algorithmes tels que nhmmer (modèle de Markov caché) ou WUBlast (alignement). Cet outil permet ainsi de compléter les bases de référence en cherchant de façon plus exhaustive de nouveaux ETs dans un génome donné en fonction des familles d'ETs déjà décrites.

Ces trois outils ont joué un rôle majeur dans l'identification et l'annotation des ETs dans le génome humain (et également d'autres organismes eucaryotes). Actuellement, plus de 50% du génome humain peut être annoté et masqué (ou identifié) par Repeatmasker.

### *Bases centrées sur le génome*

Au lieu de proposer des séquences consensus regroupées au sein de familles et sous-familles, ce type de base vise plutôt à identifier les séquences exactes des différents ETs ainsi que leurs localisations génomiques. Ce type de référence utilise généralement une base de référence centrée sur les ETs associée à une méthodologie statistique pour aller identifier et localiser les séquences dans le génome. De très nombreux programmes existent, rendant difficile la comparaison exhaustive des différentes méthodes lors d'une simple analyse (236,237).

Une première approche vise à identifier les séquences correspondant à des ETs dans le génome en fonction de leur similarité et homologies par rapport aux références d'ETs existantes décrites ci-dessus. Repeatmasker est l'outil le plus populaire, proposant une annotation des séquences retrouvées par les différents modes de recherche et permettant ainsi d'annoter plus de 50% du génome humain en tant qu'ET en couplant les bases Repbase et Dfam (233–235). Basée sur Repbase et Repeatmasker, HERVd est une base publiée par l'académie des sciences de République Tchèque proposant de définir des loci regroupant des fragments provenant d'une même famille de HERVs (238).

Une autre approche vise à fournir une annotation *de novo* en se basant sur des homologies de séquences. P-clouds détecte les ETs potentiels d'un génome en détectant des clusters de nucléotides proches et répétés (239). C'est un des outils les plus sensibles à ce jour, annotant 66 à 69% du génome humain comme ET. Bien que permettant l'annotation de plusieurs séquences non identifiés par les approches « classiques » par homologie de séquence, ce type d'approche semble cependant trop sensible et proposer un nombre important de faux positifs (240). Concernant les HERVs, l'outil Retrotector vise à identifier et localiser des séquences provirales complètes dans le génome (241). L'outil utilise une base de motifs rétroviraux (séquences *LTR*, *gag*, *pol*, *env*...) pour localiser les HERVs potentiels dans le génome humain. Chaque résultat est ensuite vérifié par un algorithme dit de « *fragment threading* », prenant en compte la distance entre

chaque motif comparativement à la distance retrouvée dans d'autres rétrovirus exogènes. Les séquences possédant une structure provirale complète avec deux LTR entourant au moins une séquence *gag*, *pol* ou *env* répondant aux critères de distance sont alors identifiés comme HERV. Retrotector favorise ainsi l'identification de séquences « longues » (de plus de 1000 pb). Cette approche a permis l'identification de 3 173 provirus dont 1 214 sont considérés canoniques (l'ensemble des motifs appartenant à la même famille de HERV) (100). Bien que très intéressante pour la détection de provirus, cette approche sous-estime grandement le nombre de HERVs, certaines familles (notamment les plus dégénérées) étant sous-représentées.

Une dernière approche vise à identifier directement la présence d'ETs par l'assemblage de séquences brutes. Bien que moins sensible que les deux approches précédentes (annotation de 20 à 30% du génome humain en tant qu'ET pour l'outil Replong par exemple (242)), ce type d'approche prend tout son intérêt pour l'identification de séquences au sein de nouveaux génomes encore mal annotés. L'utilisation de technologies menant à la production de séquences plus longues (*long reads*) est particulièrement importante pour gagner en sensibilité avec ce type d'approche, permettant de surmonter plusieurs problématiques liées aux séquences dites courtes (*short reads*, 50-250 pb).

Pour combiner les forces et palier aux faiblesses individuelles de ces approches, certains auteurs proposent de combiner plusieurs outils. Ainsi la base gEVE propose un référencement des HERVs possédant des ORFs de plus de 80 acides aminés correspondant à des motifs rétroviraux pour plus de 19 espèces de mammifères en se basant sur les annotations de Retrotector et Repeatmasker combinées à un alignement avec recherche d'homologies via *Blast-Like alignment Tool* (BLAT) (184).

### *Bases centrées sur les polymorphismes*

Ce troisième type de base de données vise à répertorier les différents polymorphismes existants et non présents dans le génome de référence. Là encore de nombreux outils ont été développés pour détecter le polymorphisme des ETs, se basant sur plusieurs méthodes telles que le fractionnement des séquences alignées, l'alignements discordants de paires de séquences ou la recherche de motifs correspondant à des ETs (232). En plus du référencement, ces bases essaient de proposer un aspect exploratoire en répertoriant les associations entre polymorphismes et phénotypes. Deux bases principales existent pour l'Homme : la base dbRip (243), répertoriant tous les polymorphismes de rétrotransposons par loci (mais plus mise à jour depuis 2012), et la base euL1db (244), développée spécifiquement pour répertorier les insertions de LINE1 dans le génome humain.

## **2) Techniques de quantification pré-NGS**

### *RT-PCR*

Pendant plusieurs années les techniques de RT-PCR ont constitué la seule méthode d'analyse moléculaire de l'expression des ETs. Une grande partie des études quantifiant l'expression des HERVs repose ainsi sur ce type de technique, permettant généralement d'apprécier l'expression d'un HERV dans un locus bien particulier (245–247) ou bien à l'échelle d'une famille (248).

Ces techniques présentent plusieurs problématiques : le manque de spécificité tout d'abord, le design d'un *primer* vraiment spécifique d'un ET donné étant difficile d'autant plus que

l'élément en question a été récemment intégré dans le génome. Ainsi le signal obtenu pour l'expression d'un HERV d'une famille donnée peut provenir de plusieurs localisations génomiques, regroupant à la fois des provirus et des transcrits chimériques partagés avec des gènes. Par ailleurs, l'utilisation d'ARN total ne permet pas de distinguer les ARN messagers des ARN issus de transcription passive (séquences introniques, co-transcription...). Ces problématiques rencontrées avec les HERVs sont exacerbées avec les autres ETs tels que les LINE1, certaines études montrant que jusqu'à 99% du signal ARN LINE1 détecté ne provient pas d'un promoteur de LINE1 (249). La RT-PCR expose également à des risques de recombinaisons *ex-vivo* et de génération de transcrits chimériques bien documentés dans le cadre de l'analyse des HERVs (250). En outre, aucune de ces techniques n'est exhaustive et applicable à l'échelle du transcriptome HERV entier (251). L'utilisation de la RT-PCR pour quantifier les HERVs est donc désormais plutôt utilisée dans certaines analyses fonctionnelles ou mécanistiques, à visée de confirmation pour montrer l'absence d'expression de quelques loci ou d'une famille suite à la répression d'un gène régulant l'épigénétique par exemple.

### *Puces microarrays et ESTs (Expressed Segment Tag)*

L'ère des puces à ADN/ARN s'est ensuite accompagnée du développement de plusieurs puces visant à quantifier l'expression d'ETs. Concernant les HERVs, l'une des plateformes les plus abouties (HERV-V3) a été développée par l'équipe de François Mallet (252). Après plusieurs versions plus restreintes, HERV-V3 permet finalement de quantifier plus de 350 000 loci différents (dont des HERVs, MaLR et LINE mais également 1 559 gènes relatifs à l'immunité) dont les localisations génomiques ont été manuellement annotées en se basant sur les séquences consensus de Repbase ainsi que sur des prototypes de HERVs, formant la base HERVgDB4. En parallèle, la quantité de données générées avec le séquençage d'EST (Marqueurs de Séquences Exprimées ou *Expressed Sequence Tag*) s'est accompagnée de la mise au point de stratégie de quantification des HERVs dans les bases d'EST, en utilisant des modèles statistiques tels que des modèles mixtes génératifs pour la problématique des séquences s'alignant à plusieurs endroits (253).

Malgré un intérêt certain, la difficulté de mettre au point des sondes à la fois courtes et spécifiques, la faible adaptabilité des puces une fois mises en place (ne permettant notamment pas de détecter les éléments non ciblés par l'analyse initiale) et surtout le développement majeur des techniques de séquençage dit « nouvelle génération » (NGS pour *Next Generation Sequencing*) a relayé ces techniques au second plan.

### **3) NGS**

L'arrivée des techniques NGS a permis un véritable bond dans l'analyse et la quantification des ETs. Cette avancée majeure s'est accompagnée de son lot de limites techniques, en lien avec la technologie dominante et la nature même des ETs. Brièvement, lorsque l'on veut séquencer l'ARN ou l'ADN d'un échantillon, les techniques désormais classiquement utilisées passent par une fragmentation du matériel génétique avant le séquençage de séquences courtes (*short reads*, entre 50 et 250 bp). Des techniques de séquençage de longues séquences existent mais sont pour l'instant moins standardisées et proposent surtout un débit plus faible. Une fois les séquences obtenues, l'approche bioinformatique classique cherche à aligner ces *reads* à un génome de référence pour savoir de quelle portion du génome elles proviennent. Si cette approche peut être relativement bien maîtrisée dans le cas de l'étude des



gènes canoniques, l'analyse des ETs introduit de nombreuses problématiques, la principale étant la correspondance multiple des *reads* de ces séquences répétées dans le génome (ou *multimap*). Les différents outils d'analyse des HERVs sur donnée NGS proposent ainsi de multiples solutions de gestion des *multimaps* et *mismatches* ainsi que différents types de quantification (par famille ou par locus spécifique) qui seront plus ou moins adaptés selon les objectifs de l'étude.

### *Alignements multiples (multi-mapping)*

La gestion des *reads* présentant de multiples alignements dans le génome (*multimap*) est un point crucial dans la quantification des ETs. En effet, du fait des fortes homologies de séquences, il est fréquent de voir un *read* provenant d'un ET s'aligner à de multiples localisations dans le génome, et ce d'autant plus que (i) l'ET est récemment intégré dans le génome (ayant donc accumulé peu de mutations et présentant de nombreuses copies similaires), (ii) les *reads* sont courts et ne sont pas séquencés à partir de leurs deux extrémités (*paired-end*) et (iii) l'ET présente de multiples copies dans le génome. Certaines approches telles que ERVmap proposent de s'affranchir de cette problématique en ne conservant que les *reads* présentant un alignement unique, éliminant les *multimaps* (254). L'utilisation de *reads* séquencés à partir de leurs deux extrémités de 50pb, 75pb et 100pb permet respectivement l'alignement unique de 68, 85 et 88% des ETs de la base Repeatmasker, en autorisant 3 erreurs (*mismatches*) par *reads* (255). Bien que permettant d'obtenir un signal très spécifique, ce type d'approche tend à largement sous-estimer le signal en provenance des familles les plus récentes d'ETs, revenant à quantifier la « capacité » d'alignement (*mappability*) plus que le niveau du transcrit d'un ET donné (256). Ainsi dans l'étude précédemment citée, si l'utilisation de *reads* de 75bp permettait l'alignement unique de 84% des HERVs en moyenne, le taux allait de 28% pour la famille HERV-K à 94% pour les HERVs des familles les anciennes (255).

Une technique alternative pour prendre en compte les *multimaps* consiste à agréger les *reads* par familles de HERVs plutôt que par loci en les alignant sur une séquence consensus. Si cette méthode permet de gagner fortement en sensibilité pour la détection des *reads* provenant de familles plus récentes, les familles plus anciennes pâtissent cette fois-ci de la quantification du fait de leur divergence plus importante par rapport à la séquence consensus, nécessitant de pousser la tolérance des *mismatches* (et donc de récupérer potentiellement du signal non spécifique).

Plusieurs autres stratégies ont été développées pour permettre la prise en compte des *multimaps* dans l'analyse des HERVs. Une approche courante consiste à diviser le nombre de *reads* entre chaque localisation à l'origine du *multimap*, soit de façon égale soit en favorisant la localisation avec le meilleur score d'alignement. Le nombre limite de *multimap* à considérer doit être fixé à l'avance avec ce type d'approche. La méthode *hervQuant* par exemple a défini un seuil optimal de *multimap* combiné à l'alignement via l'aligneur STAR (257). Combiné à la base *Retrotector*, les auteurs montrent qu'un seuil de *multimap* fixé à 10 permet de minimiser la perte de signal des classes les plus récentes. Les *multimaps* inférieurs à 10 sont ainsi distribués de façon équitable entre chaque localisation, tandis que les *multimaps* supérieurs à 10 sont éliminés. D'autres algorithmes permettent de réattribuer les *multimaps* via des modèles statistiques : *Telescope* par exemple propose de conserver le maximum de *multimaps* établis par l'aligneur *Bowtie2* (jusqu'à 100) puis de les réassigner via un modèle mixte Bayésien dont les paramètres sont calculés par un algorithme d'espérance-maximisation (que nous développerons un peu plus loin) (258).

### *Co-transcription et transcription passive*

Selon l'objectif de l'analyse, les phénomènes de co-transcription (transcription d'un HERV liée à la transcription du gène le contenant) peuvent vouloir être évités. En se basant sur une référence affinée ne regroupant que des provirus et en filtrant les *reads* de faible complexité, *hervQuant* et *ERVmap* diminuent ainsi les chances de trouver des *reads* issus d'une co-transcription (254,257). Certains outils permettent à l'inverse de précisément capturer le signal provenant d'une transcription passive en comparant les ratio sens/antisens sur des données de séquençage ARN orientés (*stranded*), ou bien en séquençant uniquement l'ARN double brin via une digestion des ARN simple brins associée à une immunoprécipitation des ARN double brins (230,259).

Ce signal peut également être plus ou moins bien capté selon le protocole de séquençage ARN utilisé. Ainsi le type de sélection d'ARN (capture polyA *vs.* ARN total) impacte grandement la présence de transcrits issus de la transcription passive d'ETs et donc souvent dépourvus de signal de polyadénylation (260). Si l'utilisation d'un protocole polyA permet tout de même de capter la majorité des transcrits HERVs et une bonne partie des LINE1, la quantification de autres ETs et des ARN non codants (type *long non-coding RNA (lncRNA)*) est très fortement perturbée. Enfin, l'approche polyA permet de s'assurer de l'absence toute de contamination par de l'ADN génomique, l'utilisation de DNase dans les protocoles ARN total n'étant pas efficace à 100%. Bien qu'une telle contamination sur un protocole ARN total puisse être corrigée par la suite lors de l'analyse des gènes canoniques (en retirant par exemple les *reads* d'origine intronique / non-exonique), celle-ci est plus difficile à corriger lors de l'analyse des HERVs. De nombreux auteurs utilisent toutefois une approche ARN total sans se soucier de cette problématique, considérant le biais de la contamination ADN comme négligeable. Aucune étude ne vient malheureusement faire le point sur cette problématique précise de la contamination ADN dans l'étude des HERVs dans les protocoles réalisés sur ARN total.

#### **4) Comparaison d'une sélection d'approche**

Trouver la bonne méthode d'analyse parmi ces différentes approches n'est pas une tâche aisée. En effet il n'existe pas une méthode universelle qui surpasserait toutes les autres, mais plutôt des méthodes plus ou moins adaptées en fonction des objectifs visés. Nous passerons ici en revue 5 outils populaires et couramment utilisés que nous avons jugés comme représentatif des différentes approches.

#### *Approche « Rooney » (base Mayer)*

La base de Mayer J. *et al.* regroupe 66 références de HERVs issues des principales bases de données génomiques américaines (Genbank), européennes (EMBL) et Japonaise (DDBJ) (115). Cette base constitue en fait une compilation de séquences HERVs publiées par des équipes différentes. Afin d'être retenus dans cette base, les HERVs devaient répondre à 3 critères : (i) avoir un ARNm détecté au moins une fois dans une base de données publique, (ii) avoir un alignement unique non ambigu dans le génome humain et (iii) représenter un provirus plutôt qu'un LTR seul. Cette base présente donc l'avantage de ne regrouper que des HERVs déjà étudiés, chaque séquence correspondant la plupart du temps à une protéine existante référencée dans Uniprot. Cependant, elle ne fournit qu'une vision très restreinte du rétrotranscriptome HERV, et souffre de l'hétérogénéité de ses sources, regroupant ainsi parfois plusieurs séquences dupliquées pour la famille HERV-K.



L'étude de Rooney M. *et al.* (261), un des papiers les plus cités à ce jour sur l'impact immunitaire adaptatif des HERVs en oncologie, utilise la base de Mayer comme référence pour quantifier les HERVs dans différents types de tumeurs. Dans cet article, les données de séquençage ARN pancancer de l'atlas du génome du cancer (TCGA pour *The Cancer Genome Atlas*) ont été alignées à la référence de Mayer avec Bowtie2 avec des paramètres de *multimap* stringents (1 max). Les HERVs ainsi quantifiés ont été corrélés au score d'activité cytolitique représenté par la moyenne géométrique de l'expression des gènes *GZMA* et *PRF1* (Granzyme A et Perforine). Les auteurs ont ainsi identifié une corrélation positive entre plusieurs HERVs de la famille HERV-K et l'activité cytolitique dans les cancers du sein et du rein. Bien que ces résultats isolés présentent de nombreuses limites (simples corrélations, coefficients R faibles autour de 0,2-0,4, p-values ajustées proches de 0,05, résultats contradictoires entre plusieurs HERV-K présentant des séquences quasiment identiques avec des protéines/peptides partagés), leur publication dans une grande revue associée à d'autres résultats de très grande qualité a permis une large couverture de l'association peptides HERV – réponse cytotoxique dans le cancer, ouvrant la voie de plusieurs projets de recherche plus spécifiques.

La base de Mayer associée à une approche telle que celle de Rooney a été la première méthode de quantification des HERVs que nous avons utilisée. Cette méthode permet un aperçu rapide de l'expression de différents représentants prototypiques issus des différentes familles de HERVs référencées. Un des principaux avantages est qu'en cas de résultat positif, une protéine identifiée et déjà décrite est attachée à chaque séquence de HERV. Si nous avons pu obtenir des résultats encourageants avec cette approche notamment dans le cancer du sein, ses nombreuses limitations (au premier rang desquelles son manque d'exhaustivité et l'hétérogénéité de ses sources) nous ont finalement poussés à abandonner cette méthode. La gestion stricte des *multimaps* nous a par ailleurs amené à documenter la perte de sensibilité dans la quantification de certains HERV-K, certaines zones très conservées ne présentant parfois aucun *reads* du fait de leur élimination.

### *Approche hervQuant (base Retrotector)*

Afin d'affiner notre recherche de peptides issus de HERVs, nous nous sommes ensuite tournés vers la base de Retrotector précédemment présentée et ses 3 173 références (100,241). La méthodologie d'identification systématique à partir de motifs rétroviraux ainsi que l'enrichissement en séquences provirales (contenant au moins un ORF *gag*, *pol* ou *env*) s'intégrait ainsi parfaitement dans une stratégie d'identification de cibles thérapeutiques. Utilisant cette base comme référence, la méthode hervQuant (257) précédemment exposée propose une analyse fiable et robuste des HERVs. Un des points forts de hervQuant est la proposition d'une vraie méthodologie de quantification des HERVs basée sur une analyse bioinformatique robuste ayant comparé de multiples approches afin de sélectionner des outils et paramètres optimaux.

Dans le premier article de résultats de cette thèse (**Article 1**), nous avons utilisé l'approche hervQuant pour quantifier l'expression des HERVs dans le jeu de donnée pancancer du TCGA. Comme nous le verrons, cette approche nous a permis d'identifier de nombreux épitopes T CD8<sup>+</sup> avec de bons résultats, confirmant la validité de cette méthode pour l'identification de peptides issus d'ORFs de HERVs. Le seul point négatif réside finalement dans le manque d'exhaustivité de cette référence ne captant que les provirus et pouvant manquer certaines familles plus anciennes. Si cette approche peut donc largement convenir à certaines analyses

(identifications d'épitopes issus d'ORFs de HERVs, aperçu de l'expression de certaines familles basée sur un échantillon de HERVs...) elle reste limitée pour les approches visant à tirer le maximum du rétrotranscriptome HERV, telles que les approches cherchant à identifier des biomarqueurs.

#### *Approche par famille (base RepeatMasker)*

Plutôt que de chercher à quantifier précisément certains loci, de nombreuses équipes choisissent une approche plus globale par famille. En partant d'une base de référence complète type RepeatMasker, cette approche vise à quantifier le nombre de *reads* s'alignant aux éléments d'une même famille de façon globale. Bien que cette technique permette en théorie de s'affranchir du problème de *multimap* (les *reads* s'alignant à plusieurs endroits appartenant *a priori* à la même famille), un filtre supplémentaire ne conservant que les *reads* ne présentant pas de *multimap* est souvent appliqué (262). Le nombre important de références, avec notamment la possibilité d'étudier les autres ETs tels que les LINEs, est ici contrebalancé par le manque de spécificité du signal. Ce type d'approche est donc surtout utilisée afin d'obtenir un aperçu global de la transcription des HERVs par famille.

#### *Approche Telescope (base RepeatMasker modifiée)*

Pour pallier au manque de spécificité de l'approche précédente, Bendall M. *et al.* proposent Telescope (258), un outil complet permettant de quantifier le rétrotranscriptome HERV à partir de donnée de séquençage ARN. Le premier point fort de Telescope est sa base de référence : les auteurs sont en effet partis d'une annotation exhaustive des HERVs du génome à partir de RepeatMasker. Ils ont ensuite procédé à un regroupement des fragments de HERVs sous forme d'unité transcriptionnelle, en identifiant d'abord les LTRs puis en recherchant les séquences proches d'une même famille. Les loci discordants ont été manuellement revus, et l'annotation a été validée par comparaison des HERVs de la famille HML-2 avec l'annotation de référence de Subramanian R. *et al.* (263). Cette approche a abouti à l'annotation de 14 968 loci HERVs, regroupant aussi bien des provirus complets que des séquences LTRs seules.

Telescope fournit ensuite un algorithme de quantification permettant de conserver le maximum de signal des familles récentes sujettes au *multimap*. Dans un premier temps, chaque *read* est aligné de façon indépendante au génome de référence, en cherchant à chaque fois les différents alignements possibles sous réserve d'un score minimum de qualité d'alignement. Ensuite, les *reads* sont assignés à un locus, chaque assignement étant pondéré par le score d'alignement. Cette première étape montre que moins de 50% des *reads* s'alignant aux HERVs peut être assigné à une localisation unique, de nombreux *reads* pouvant avoir jusqu'à 20 localisations avec de bon scores d'alignement. La proportion exacte de transcrits à un endroit donné est ensuite représentée par un modèle mixte Bayésien prenant en compte l'incertitude de l'alignement et la quantité relative de transcrit à cet endroit. Les paramètres de ce modèle sont calculés par un algorithme d'espérance-maximisation qui converge vers un modèle final maximisant la vraisemblance de la distribution.

Ce type de méthode s'approche d'autres méthodes connues et largement utilisées en analyse de séquençage ARN telle que la méthode *RNA-Seq by Expectation-Maximization* (RSEM) permettant notamment de quantifier les transcrits issus d'isoformes (264). Ainsi, Telescope fournit non seulement une annotation relativement exhaustive du rétrotranscriptome HERV, mais sa méthode de calcul permet d'avoir une quantification sensible de l'ensemble des familles

HERVs, particulièrement des familles récemment intégrées. Cette méthode se présenterait donc comme plus sensible que les autres et permettrait une meilleure appréciation du rétrotranscriptome HERV. Cette référence comprend par ailleurs tous les provirus de la référence de Retrotector, permettant également de faire une analyse centrée sur les ORFs. En outre, Telescope est un logiciel *open-source* avec un code très bien documenté et accessible, en faisant un outil modifiable et reproductible.

Ces différents avantages nous ont poussé à considérer Telescope comme une méthode de référence. De plus, une comparaison directe des approches pour les besoins de l'article 2 que nous développerons plus tard a révélé que Telescope surpassait les approches hervQuant et « Kassiotis » pour la classification non supervisée des différentes sous-populations cellulaires de moelles osseuses normales et leucémiques.

#### *Approche « Kassiotis » (base RepeatMasker)*

Plus récemment, l'équipe de George Kassiotis a développé une approche innovante pour l'analyse du rétrotranscriptome HERV (265). Les auteurs se sont servis des données de séquençage ARN de 768 patients de la base TCGA représentant 31 types de cancers différents pour réaliser un assemblage de transcriptome *de novo*. Ils ont ensuite procédé à une annotation des transcrits correspondant à des ETs via la base RepeatMasker. Sur 753 166 transcrits exprimés à au moins 1 Transcrits par million de *reads* mappés (TPM) dans au moins un échantillon, presque la moitié n'était pas annotée, correspondant pour 37% à des SINEs, 30% à des LINEs, 21% des LTRs et 12% des éléments non ETs. Pour comparaison, la portion annotée comprenait 31% de SINEs, 19% de LINEs, 14% de LTRs et 36% d'éléments non ETs (correspondant principalement à des isoformes de gènes issus d'épissages alternatifs). Concernant les HERVs (LTRs), 58% ne présentaient pas de chevauchement avec un autre gène, 32% présentaient un chevauchement avec un gène codant et 10% avec un autre ARN (type *lncRNA*).

Sur les 630 356 LTRs annotés, 108 946 (~17%) étaient exprimés à au moins 1 TPM dans au moins 1 échantillon, et 20 164 (~3%) étaient considérés spécifique du cancer (non retrouvés dans une sélection de tissus sains de la base GTEx). Parmi ces HERVs exprimés spécifiquement dans le cancer, 17% correspondaient à des HERVs complets et 39% à des transcrits chimériques débutant par le LTR (et donc potentiellement exprimés via celui-ci). Les autres cas correspondaient à une séquence HERV complètement contenue dans un exon ou bien précédant celui-ci en fournissant une zone d'épissage alternatif.

Ces données nous montrent que l'expression des HERVs dans le cancer est en partie due à une dérégulation de séquences LTR mais également à des transcrits alternatifs d'autres gènes menant à la transcription passive de HERVs localisés dans des portions exoniques ou introniques. Ces transcrits alternatifs pourraient avoir une importance particulière, comme l'a montré une étude réalisée dans les SMD publiée par la même équipe (266): en analysant le transcriptome (assemblé *de novo* selon cette méthode) de cellules souches hématopoïétiques (CSH) de patients présentant un SMD ou une Leucémie Myélomonocytaire Chronique (LMMC) et traités par AZA, les auteurs montrent que, bien que l'AZA entraîne la dérégulation d'ETs, ceux-ci ne sont pas prédictif de la réponse au traitement contrairement à différents transcrits issus de phénomènes d'épissage alternatif induits par le traitement. A noter que le séquençage ARN réalisé dans cette étude était réalisé sur populations CD34<sup>+</sup> triées, ne captant ainsi pas tout le signal de l'environnement pouvant être informatif dans un but de mise au point de signature prédictive.

	Rooney	hervQuant	Telescope	Kassiotis	RepeatMasker
<b>Référence</b>	Mayer J. <b>66 loci</b>	Retrotector <b>3 173 loci</b>	RepeatMasker (regroupés en loci) <b>~15 000 loci</b>	RepeatMasker (assemblage de transcriptome <i>de novo</i> ) <b>~630 000 transcrits</b>	RepeatMasker, basé sur Repbase et Dfam
<b>Algorithme d'alignement</b>	Bowtie 2 -q--end-to- end --no-unal	STAR (default)	Bowtie2 --very sensitive local	Bwa mem ou aln (default)	Variable
<b>Alignement multiple (<i>multimap</i>)</b>	1 (distribution aléatoire des <i>multimaps</i> )	Jusqu'à 10	Jusqu'à 100	Distribution aléatoire des <i>multimaps</i> guidée par le score d'alignement	Généralement limité à 1
<b>Mismatch</b>	2	<=7	5% max	2	Variable
<b>Quantification</b>	CPM ou RPM	Salmon sur BAM filtré avec uniquement les références HERVs. CPM	Modèle mixte bayésien avec algorithme EM pour l'estimation des paramètres. Proche du RSEM	Salmon, TPM	HTseq ou Featurecount CPM, RPKM ou TPM
<b>Avantages</b>	- Regroupe uniquement des HERVs validés avec une protéine connue	- Identification systématique des provirus - Quantification de provirus et séquences complètes	- Base exhaustive regroupée par loci - Méthode de quantification sensible	-Analyse plus large des événements (épissage alternatif, transcrits chimériques LTR- gènes)	- Base de référence très vaste - Aperçu global du rétrotranscriptome
<b>Inconvénients</b>	- Annotation hétérogène - Séquences duplicées - Manque de sensibilité pour les familles récentes du fait de l'élimination des <i>multimaps</i>	- Aperçu partiel du rétrotranscriptome minimisant l'utilisation des HERVs type biomarqueur - Sous- représentation de certaines classes	- Pas d'annotation des transcrits chimériques chevauchant des gènes	- Analyse spécifique nécessitant une reconstitution du transcriptome à chaque cancer (voir chaque patient)	- Compte par famille ne donnant qu'un aperçu global du rétrotranscriptome

**Tableau 3 : Comparaison d'une sélection d'approches bioinformatiques pour la quantification des HERVs dans les bases de données de séquençage ARN.**

BAM : *Binary Alignment Map* (fichier d'alignement de séquences), CPM : Comptes par million de *reads* mappés, RPM : *Reads* par million de *reads* mappés, RPKM : *Reads* par kilobases par million de *reads* mappés, TPM : Transcrits par million de *reads* mappés

Cette approche propose une analyse exhaustive du rétrotranscriptome dans le cancer et présente donc un intérêt majeur pour la compréhension des mécanismes d'oncogenèse. Elle a l'avantage de pouvoir se pencher sur les transcrits chimériques associés à d'autres gènes. Pour apprécier ce type de transcrit avec l'approche de Telescope, il faudrait réaliser une analyse spécifique pour chaque HERV exprimé : regarder les gènes situés autour et leur expression, et vérifier si un gène proche exprimé correspond bien à un transcrit unique avec le LTR, ce qui est long et fastidieux et expose également à un manque de sensibilité.

## **Conclusions préliminaires**

Cette introduction nous aura permis de resituer les défis actuels dans le domaine de l'immunité antitumorale et notamment de la découverte d'antigènes tumoraux. Nous avons pu souligner les limites actuelles des cibles « conventionnelles » que représentent les antigènes mutés, introduisant de fait les avantages potentiels des épitopes issus de HERVs. Nous avons également abordé les difficultés bioinformatiques liées à la quantification des HERVs, et proposé une brève comparaison des approches existantes. L'ensemble de ces données viennent poser les bases théoriques et le rationnel de l'utilisation des HERVs en immuno-oncologie. Nos travaux, présentés en résultats de cette thèse, viennent en présenter les aspects pratiques.



## OBJECTIFS

---





De par leur profil d'expression particulier et leur origine virale, les protéines issues de HERVs pourraient donc constituer de très bons antigènes tumoraux. Le nombre important de séquences présentes dans le génome, représentées en grande majorité par des séquences LTR isolées, vient néanmoins perturber le processus d'identification de cibles thérapeutiques, menant à la sélection de nombreux faux positifs. Ainsi, au début de ce projet, aucun outil bio-informatique ne permettait d'analyser le profil d'expression des HERVs dans un cancer donné afin d'en déterminer les cibles épitopiques d'intérêt.

L'objectif principal de cette thèse était de mettre au point un pipeline d'analyse bio-informatique permettant, pour une tumeur donnée, de définir les épitopes LT CD8<sup>+</sup> issus de HERVs les plus pertinents. Cette mise au point impliquait (1) d'identifier les HERVs exprimés spécifiquement par une tumeur donnée (c'est-à-dire fortement exprimés dans une tumeur et non exprimés dans le tissu sain), (2) de détecter les épitopes LT CD8<sup>+</sup> potentiels à partir de ces HERVs et (3) de définir une stratégie pour filtrer ces épitopes afin de diminuer la liste de peptides à valider et de ne retenir que les plus prometteurs. Toute une série de validation biologique *in vitro* était prévue dans un second temps afin de confirmer l'immunogénicité des épitopes identifiés, depuis la validation d'affinité des peptides pour le HLA et l'induction de réponses LT CD8<sup>+</sup> spécifiques chez le donneur sain jusqu'à la confirmation de la cytotoxicité de clones spécifiques en contact de cellules tumorales. Le cancer du sein étant la pathologie de prédilection de l'équipe, c'est naturellement vers cette pathologie que se sont tournées nos premiers travaux avant d'élargir le champ progressivement aux LAM.

L'acquisition d'une certaine expertise dans le domaine d'analyse des HERVs ainsi que l'adoption progressive d'une méthode permettant une capture plus complète et fine du rétrotranscriptome nous ont par la suite poussés à élargir ces objectifs, avec un nouvel axe concernant la caractérisation des populations cellulaires via le rétrotranscriptome HERV. Ce deuxième axe, visant à souligner l'importance des HERVs dans la constitution de l'identité d'une population cellulaire et questionnant sur leur rôle causal via notamment leur activité de régulateur génique, a ainsi pleinement émergé au cours de cette thèse.



## REVUES

---



## REVUE 1 - Cold Tumors: A Therapeutic Challenge for Immunotherapy

*Paola Bonaventura<sup>1,2</sup>, Tala Shekarian<sup>1,2</sup>, **Vincent Alcazer<sup>1,2</sup>**, Jenny Valladeau-Guilemond<sup>2</sup>, Sandrine Valsesia-Wittmann<sup>1,2</sup>, Sebastian Amigorena<sup>3</sup>, Christophe Caux<sup>1,2</sup> and Stéphane Depil<sup>1,2,4\*</sup>*

1 Centre Léon Bérard, Lyon, France

2 INSERM U1052, Centre de Recherche en Cancérologie de Lyon, Lyon, France,

3 Institut Curie, PSL Research University, INSERM, U932, Paris, France

4 Université Claude Bernard Lyon 1, Lyon, France

\* Corresponding author: Stéphane Depil, Centre Léon Bérard, 28 Prom. Léa et Napoléon Bullukian, 69008 Lyon, France;

Tel: +33 (0) 4 78 78 28 28; Mail: [stephane.depil@lyon.unicancer.fr](mailto:stephane.depil@lyon.unicancer.fr)

Front Immunol. 2019 Feb 8;10:168. doi: 10.3389/fimmu.2019.00168. PMID: 30800125; PMCID: PMC6376112.





# Cold Tumors: A Therapeutic Challenge for Immunotherapy

Paola Bonaventura<sup>1,2</sup>, Tala Shekarian<sup>1,2</sup>, Vincent Alcazer<sup>1,2</sup>, Jenny Valladeau-Guilemond<sup>2</sup>, Sandrine Valsesia-Wittmann<sup>1,2</sup>, Sebastian Amigorena<sup>3</sup>, Christophe Caux<sup>1,2</sup> and Stéphane Depil<sup>1,2,4\*</sup>

<sup>1</sup> Centre Léon Bérard, Lyon, France, <sup>2</sup> INSERM U1052, Centre de Recherche en Cancérologie de Lyon, Lyon, France,

<sup>3</sup> Institut Curie, PSL Research University, INSERM, U932, Paris, France, <sup>4</sup> Université Claude Bernard Lyon 1, Lyon, France

Therapeutic monoclonal antibodies targeting immune checkpoints (ICPs) have changed the treatment landscape of many tumors. However, response rate remains relatively low in most cases. A major factor involved in initial resistance to ICP inhibitors is the lack or paucity of tumor T cell infiltration, characterizing the so-called “cold tumors.” In this review, we describe the main mechanisms involved in the absence of T cell infiltration, including lack of tumor antigens, defect in antigen presentation, absence of T cell activation and deficit of homing into the tumor bed. We discuss then the different therapeutic approaches that could turn cold into hot tumors. In this way, specific therapies are proposed according to their mechanism of action. In addition, “supra-physiological” therapies, such as T cell recruiting bispecific antibodies and Chimeric Antigen Receptor (CAR) T cells, may be active regardless of the mechanism involved, especially in MHC class I negative tumors. The determination of the main factors implicated in the lack of preexisting tumor T cell infiltration is crucial for the development of adapted algorithms of treatments for cold tumors.

**Keywords:** cold tumors, T cells, tumor antigen, presentation, priming, trafficking, immunotherapy

Immune checkpoint inhibitors (ICIs) have changed the treatment landscape of many tumors, inducing durable responses in some cases, Tumor mutational load, CD8<sup>+</sup> T cell density and Programmed cell Death Ligand–1 (PD-L1) expression have each been proposed as distinct biomarkers of response to PD-1/-L1 antagonists. The lymphocyte infiltration and IFN- $\gamma$  status may be key factors for effective anti-PD-1/-L1 therapy by defining a “T cell inflamed” phenotype (“hot tumors”). In contrast, lack of T cells infiltrating the tumor characterizes “non-inflamed” or “cold tumors” (in which other immune populations or myeloid cells can however be observed). Immunological treatment of cold tumors is a great challenge as no adaptive immune response has been set up or maintained. In this review, we discuss the possible issues that the immune system could encounter at different steps of the anti-tumor immune cycle (1), leading to the absence of T cell infiltration: lack of tumor antigens, defect in Antigen Presenting Cells (APCs), absence of T cell activation and deficit of homing into the tumor bed (**Figure 1**). The potential therapeutic strategies to overcome these problems will be described in the second part of this review. We will not discuss here the mechanisms of immune escape developed by inflamed tumors, reviewed elsewhere (2).

## LACK OF TUMOR ANTIGENS

Tumor antigens can be divided into three main classes: tumor specific antigens (TSA), cancer-germline antigens (CGA), and tumor associated antigens (TAA) (3). TSA are expressed only by

## OPEN ACCESS

### Edited by:

Salem Chouaib,  
Institut Gustave Roussy, France

### Reviewed by:

Viktor Umansky,  
German Cancer Research Center  
(DKFZ), Germany  
Daniel Olive,  
Aix Marseille Université, France

### \*Correspondence:

Stéphane Depil  
stephane.depil@lyon.unicancer.fr

### Specialty section:

This article was submitted to  
Cancer Immunity and Immunotherapy,  
a section of the journal  
Frontiers in Immunology

**Received:** 04 December 2018

**Accepted:** 21 January 2019

**Published:** 08 February 2019

### Citation:

Bonaventura P, Shekarian T, Alcazer V,  
Valladeau-Guilemond J,  
Valsesia-Wittmann S, Amigorena S,  
Caux C and Depil S (2019) Cold  
Tumors: A Therapeutic Challenge for  
Immunotherapy.  
Front. Immunol. 10:168.  
doi: 10.3389/fimmu.2019.00168

cancer cells and not in healthy tissues (4). TSA include mutation-associated neoantigens (MANA) and viral antigens. MANAs arise from DNA mutation/rearrangement in a gene coding sequence and play a crucial role in the recognition of tumor cells by CD8<sup>+</sup> T cells after immune checkpoint treatment (5). Viral antigens may also represent the target for immune recognition of virus-associated tumors (6). CGA are expressed in tumor cells of different histological origins, but they are silent in normal adult tissues, except in the male germ line and trophoblastic cells (7). Their expression is associated to the demethylation of their promoter. TAA correspond to antigens with low expression in normal tissues and overexpressed in tumor cells, like HER2 (8), or melanocyte differentiation proteins (9).

Tumor mutation burden (TMB) is a quantitative measure of the total number of mutations per coding area of a tumor genome that has been shown to predict responses to ICIs in a range of advanced cancers (10). Tumors with a high TMB are believed to express more MANA. Interestingly, a correlation between MANA and CGA has been observed (11). However, even if some tumors are characterized by low expression of both MANA and CGA, the quantity of tumor neoantigens does not seem to be the main limiting factor for the induction of a T cell response. A recent study by Spranger et al. analyzed the impact of the presence of differentiation antigens, CGA and MANA on T cell infiltration in malignant melanoma. They reported that non-T-cell-inflamed melanomas do not lack antigens for T-cell recognition, arguing for other mechanisms causing the lack of T cell priming and recruitment. Moreover, the number of neoantigens and the mutational load was still comparable between non-T cell-inflamed and T cell-inflamed subtypes in other solid tumors (12). Finally, the kinetics of tumor antigens

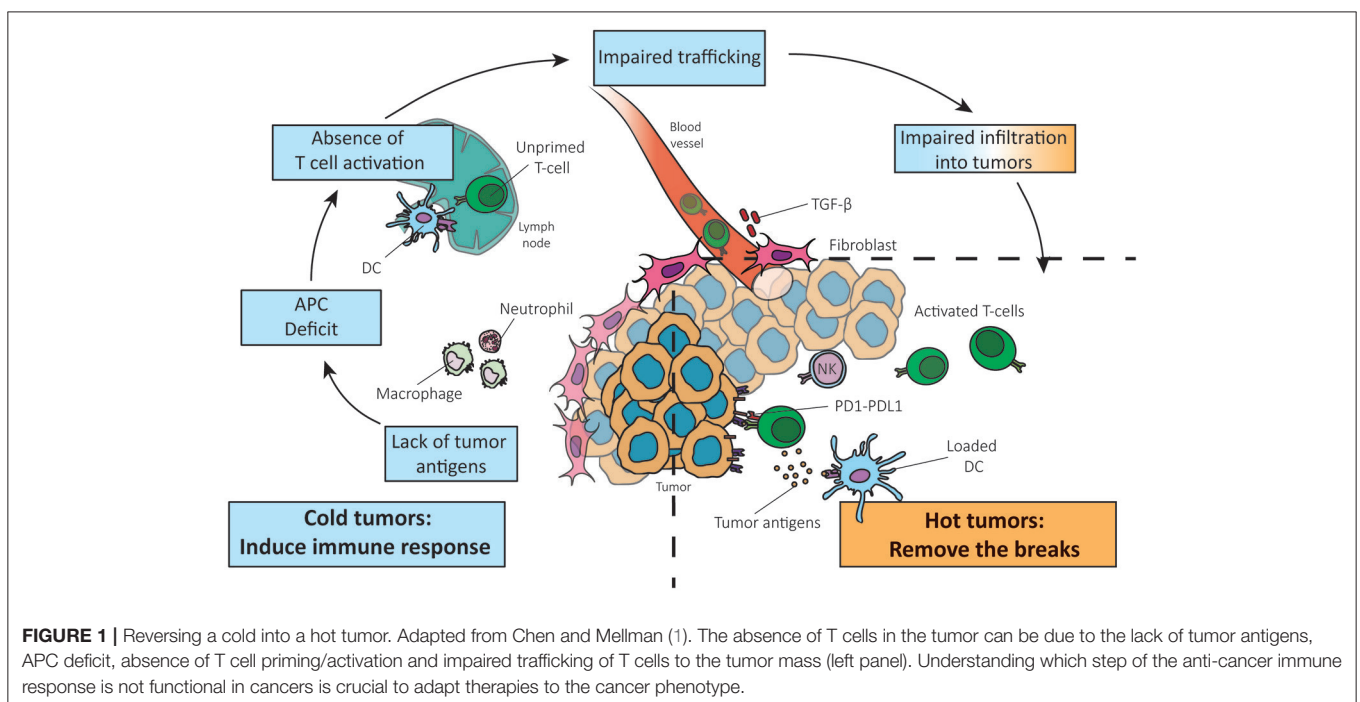
release, associated with different modalities of spontaneous or induced cancer cell death, may also influence the quality of the T cell response (13).

## ABSENCE OF T CELL PRIMING/ACTIVATION

### Defective Recruitment of APCs

The second step of the anti-tumor immune response consists in the presentation of tumor antigens by dendritic cells (DCs), resulting in the priming and activation of specific effector T cells. Several DCs subsets such as classical DCs (type 1 cDC1s and type 2 cDC2s), Langerhans cells, inflammatory DCs and plasmacytoid DCs (pDCs) exist and are specialized in different functions to shape the immune response and cope with the threat of diversity (14). In particular some evidences exist showing that a higher ratio of cDC1s over monocytes/macrophages in the tumor bed favors protective anti-tumoral adaptive immune responses (15). Spranger et al. have also shown in melanoma that MHC expression and DCs infiltration is associated with T cell infiltration (12).

Among DCs, cDC1s excel at inducing anti-tumoral CD8 T cell responses through cross-presentation of exogenous antigens on MHC-I (16–18). A strong correlation between CD8 gene transcript and cDC1s markers was observed, suggesting that lack of T cell activation and infiltration in the non-T-cell-inflamed tumor microenvironment is mainly associated with a defective recruitment and activation of cDC1 (19–21). Moreover, recent papers published by M. Krummel and C. Reis e Sousa teams recently demonstrated a critical role of the cross-talk between cDC1s and NK cells for the CTL infiltration in melanoma



(22, 23). In mice, several reports have shown that cDC1s are necessary for the natural rejection of transplanted tumors and for the efficiency of anti-tumoral immunotherapies including ICIs or adoptive transfer of anti-tumoral CD8 T cells (24).

## Lack of T Cell Co-stimulation and Activation After Antigen Presentation

The maturation and activation of antigen-presenting DCs is a critical step for activating an efficient T cell-response. In this context, the DC activation marker DC-LAMP is a good prognostic marker in solid tumors (25). Naive T cells require contact with activated APCs to be primed in an appropriate context of “danger signal” (26). APCs expressing Pattern Recognition Receptors (PRRs) can be directly activated by Pathogen-Associated Molecular Patterns (PAMPs) or Danger Associated Molecular Patterns (DAMPs) to become competent to prime T cell responses (27). Engagement of PRRs on DCs induces NF- $\kappa$ B activation, up-regulation of co-stimulatory molecules, production of cytokines and promotion of cross-priming (28, 29). Various DAMPs are produced by tumor cells undergoing immunological cell death [e.g., calreticulin, HighMobility Group Box 1 protein (HMGB1) or Sin3A Associated Protein 130 (SAP130)] (30). The absence or low production of DAMPs could induce a lack of DCs maturation as well as production of immunosuppressive factors such as transforming growth factor beta (TGF- $\beta$ ) leading to the absence of CD4<sup>+</sup> T cell help (30, 31). Recent works demonstrate the importance of the protein Formyl Peptide Receptor 1 (FPR1) expressed by tumoral DCs in the anthracycline-induced immunogenic cell death. DCs lacking or presenting a variant of FRP1, failed in antigen presentation and activation of T cells, resulting in poor anticancer immune responses and reduced overall survival in breast and colon cancer (32).

Stimulation of CD40 on APCs through CD40L expressed on helper CD4<sup>+</sup> T cells is another crucial step for the activation of APCs to prime CD8<sup>+</sup> T cells. Moreover, the stimulation of CD40 on DCs regulates the expression of the co-stimulatory molecules CD80 and CD86, enhances the production of cytokines (most notably IL-12 and IFN-I) and promotes the cross-priming to exogenous antigens (33). As a consequence, reduced CD8<sup>+</sup> T cell responses are largely due to impaired activation of APCs or to the absence of co-stimulation.

## DEFICIT OF HOMING TO THE TUMOR BED

### CD8<sup>+</sup> T Cell Exclusion by the Immunosuppressive Peritumoral Stroma and Tumor Cell Alterations

When DCs are mature and T-cells correctly primed and activated, the access of T cell to the tumor bed could be compromised by the stromal compartment (34). The exclusion of CD8<sup>+</sup> T cells from the vicinity of cancer cells was shown to correlate with a poor long-term clinical outcome in colorectal cancer, ovarian cancer and pancreatic ductal adenocarcinoma (35, 36). Interestingly, Spranger et al. reported an inverse relationship between intrinsic  $\beta$ -catenin signaling of tumor cells and intra-tumoral T cells

in melanoma. Using a genetically engineered mouse model they showed that melanomas arising from mice with active  $\beta$ -catenin were characterized by an almost complete absence of both CD8<sup>+</sup> T cells and cDC1 subsets (37). A second pathway identified to play a role in T cell exclusion is PI3K pathway activation/PTEN loss. Loss of PTEN in tumor cells in preclinical models of melanoma was shown to increase the expression of immunosuppressive cytokines, inhibit T cell-mediated tumor killing and decrease T cell trafficking into tumors. Furthermore, in patients PTEN loss correlated with decreased T cell infiltration at tumor sites and inferior outcome after PD-1 inhibitor therapy (13). PTEN-deficient prostate tumors similarly induce an immunosuppressive tumor microenvironment by upregulating PTPN11/SHP2 and inducing activity of the Jak2-Stat3 pathway (38). Loss of PTEN was recently associated with resistance to anti-PD1 therapy in metastatic uterine leiomyosarcoma (39) and the blockade of this pathway *in vivo* contributed to an improved tumor control (13).

Tauriello et al. investigated how genetic alterations and the tumor microenvironment (TME) interact in a metastatic colorectal carcinoma (CRC) model. A Tumor Growth Factor (TGF)- $\beta$  activity correlating with T cell exclusion and a low TMB was described (40). Recently, a study associated a TGF- $\beta$  signature of stromal cells with lack of response to anti PD-L1 in the excluded tumor-immune phenotype (41). Blockade of TGF- $\beta$  in a pancreatic ductal adenocarcinoma model improved the cure rate of mice by decreasing the presence of immune suppressive cells in the TME and enhancing CD8<sup>+</sup> T cell infiltration within the tumor (42).

### Modified Production of Chemokines and Cytokines Affecting Cell Trafficking and Activation

Cytokines and chemokines may influence cell trafficking to the tumor bed. Besides the steady-state influx of immature dendritic cells (iDCs) within tissues, chemokines, abundantly secreted under inflammatory conditions, can provoke influx of iDCs in the tumor bed (43). Lack of those chemokines and the consequent reduced influx of iDCs in the tumor bed can be the cause of the reduced activation and migration of T cells at the tumor site. Chemokines acting on iDCs are the Monocyte Chemoattractant Proteins (CCL2, CCL7, CCL8) as well as CCL3/MIP-1 $\alpha$ , CCL5/RANTES, and CCL4/MIP-1 $\beta$  (44). Cytokines are also necessary to generate active DCs: as an example type I interferon (IFN-I) produced by DCs can act in an autocrine manner to generate fully active DC1s (45). Moreover, DC1s are a source of CXCL-9/10 and their absence lead to a reduced production of these chemokines (20). The chemokine CXCL16, produced by DCs, and its receptor CXCR6 for example have been associated with an increased CD4<sup>+</sup> and CD8<sup>+</sup> T cell recruitment and a good prognosis in CRC (46). The disruption of the CXCL16/CXCR6 pathway could lead to a reduced tumor T cell infiltration.

The deregulation of trafficking can directly involve T cells: DCs-activated T cells against tumor antigens have to reach the tumor bed to perform their anti-cancer activity. Tumors can disrupt chemokine expression to deregulate the immune

response and chemokines involved in effector T-cell recruitment is significantly reduced in tumors lacking a CD8<sup>+</sup> T-cell infiltrate. CXCL9 and CXCL10 (CXCL11 in humans) are key chemokines in the recruitment of CD8<sup>+</sup> T cells engaging the CXCR3 on their surface and their production is generally deregulated in “non-inflamed” tumors (47). CXCL9/10 can be produced by the tumor cell itself where a methylation of chemokine genetic loci results in a reduced CD8<sup>+</sup> T cell infiltration. The use of demethylating agents restores chemokine production and T-cell recruitment, showing that epigenetic modification is a mechanism of tumor escape which could lead to the lack of immune cells infiltration (48). Tumors can also alter the chemistry of certain chemokines to preferentially recruit myeloid cells: as an example the nitrosylated CCL2 eliminates the ability to recruit CTLs and Th1 effector cells (49), while selectively recruiting myeloid dendritic stem cells (MDSCs) to tumor sites.

## THERAPEUTIC APPROACHES

Different therapeutic approaches can theoretically be used to overcome the absence of T cell infiltration in tumors. These strategies are summarized in **Figure 2**. The demonstration that these therapies can effectively transform a cold into hot tumor remains to be done in the clinic in most instances.

### Specific Therapies for Tumors Expressing Few Antigens

#### Demethylating Agents

It has been shown that DNA methyltransferase inhibitors (DNMTi) and histone deacetylase inhibitors can enhance the expression of tumor antigens and components of antigen processing and presenting machinery pathways, as well as other immune related genes (50, 51). These agents can also induce the expression of retroelements such as endogenous retroviruses (ERVs), usually silent and able to induce a type I IFN response (52). Epigenetic drugs have been reported to induce transcription from normally repressed ERV LTR, that may cause ectopic expression of transcripts with canonical or novel open reading frames, leading to the production of immunogenic peptides (53, 54). DNMTi and Histone-lysine N-methyltransferase EZH2 inhibitors have also been shown to reverse epigenetic silencing

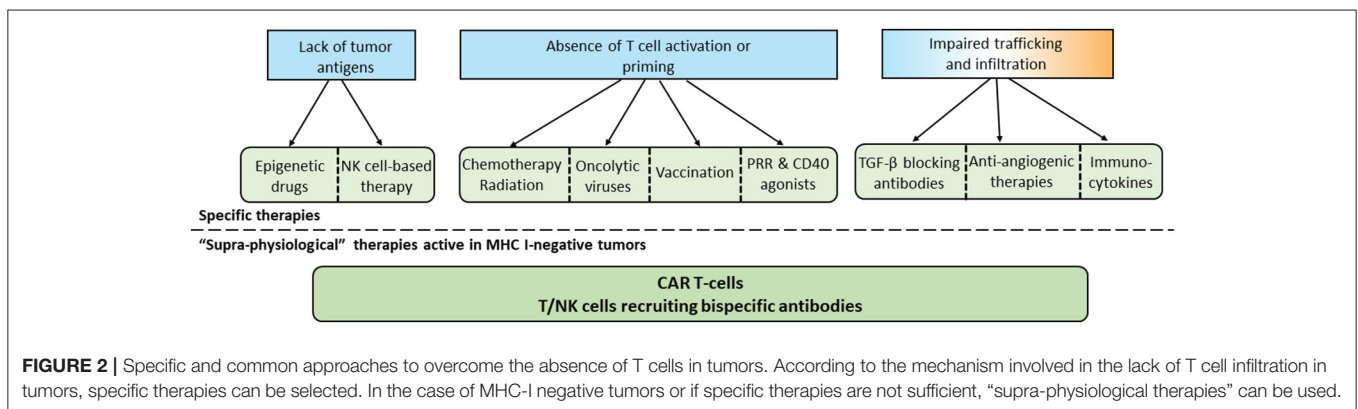
of Th1-type chemokines in tumor cells, which is negatively associated with CD8<sup>+</sup> T cells in tumors and patient outcome (48). There is thus a strong rationale to combine epigenetic therapy and immunotherapy and many clinical trials are currently ongoing (55).

### NK Cell-Based Approaches

Natural killer (NK) cells are lymphocytes of the innate immune system able to recognize and kill tumors lacking self-MHC class I molecules, by recognizing stressed cells. For this reason NK approaches could be suitable in the absence of tumor antigens or in case of deficient antigen presentation machinery (e.g., lack of MHC class I). While a large portion of cancer immunotherapies focus on targeting T cells, NK cell system for therapeutic intervention stays relatively underexplored. Nevertheless, different NK cell-based approaches have been described, such as *ex vivo* activated NK cells or NK cells transduced with a chimeric antigen receptor (CAR) to target specific cancer cell surface antigen (56, 57). Antibodies-mediated targeting of NK activating receptors such as NKG2D and NKP46, or inhibitory receptors such as KIR and NKG2a, is under deep investigation (58). Lirilumab is a fully human antibody directed against KIR2DL-1,-2,-3 inhibitory receptors expressed predominantly on NK cells and is being tested in combination with ipilimumab or nivolumab for the treatment of patients bearing advanced solid malignancies (59).

### Specific Therapies for Tumors With Defective Priming or T Cell Activation Chemo/Radiotherapy Inducing Immunogenic Cell Death (ICD)

Several chemotherapies were found to work mainly in immunocompetent subjects, accumulating evidences that tumor inhibition partially relies on the immune system competence and not only on the direct anti-tumor toxicity of chemotherapy (60). In this regard ICD inducing chemotherapeutic agents can be classified within cancer immunotherapy strategies. Radiotherapy was initially designed to selectively kill tumor cells within the irradiated field. However, emerging evidence indicates that radiotherapy, by inducing ICD, harnesses the host's immune system to attack the tumor cells outside the



**FIGURE 2 |** Specific and common approaches to overcome the absence of T cells in tumors. According to the mechanism involved in the lack of T cell infiltration in tumors, specific therapies can be selected. In the case of MHC-I negative tumors or if specific therapies are not sufficient, “supra-physiological therapies” can be used.



irradiation field, explaining the “abscopal” effect (regression of tumor lesions outside the irradiation field) (61). Based on this rationale, many trials are ongoing to combine chemotherapy and radiotherapy with PD-1/PD-L1 antibodies. Of note, the same rationale also applies to antibody-drug conjugates (ADCs) with cytotoxic payloads capable of inducing ICD, justifying the initiation of clinical trials combining ADCs and immunotherapy (62).

### Oncolytic Viruses

Oncolytic viruses are common viruses that can selectively target, replicate in and destroy cancer cells (63). Most oncolytic viruses can induce cancer cell death and directly eliminate tumor cells, but they also initiate systemic immune responses through different mechanisms such as induction of ICD and release of danger signals (DAMPs) and tumor antigens from virus-infected cells. They also release viral PAMPs contributing to APCs maturation that conduct to activation of antigen-specific CD4<sup>+</sup> and CD8<sup>+</sup> T cell responses (64). Moreover, the infected cells are directly recognized by the innate immune system such as NK cells or macrophages (65). It has been recently shown in a melanoma phase I trial that use of oncolytic viro-therapy was able to convert a cold into hot tumor, as patients with a low level of immune cell infiltrate and a negative IFN $\gamma$  signature before treatment responded well to the combination of talimogene laherparepvec with the PD-1 antagonist pembrolizumab (66).

### PRR Agonists

PRRs consist of five families including Toll-like receptors (TLRs), RIG-I-like receptors (RLRs), nucleotide-binding oligomerization domain (NOD)-like receptors (NLRs), C-type lectin receptors (CLRs), and cytoplasmic DNA sensors. PRRs agonists have notably the ability to activate PRR pathways inducing antigen presentation by myeloid cells residing in the tumor micro-environment (67–69).

TLR agonists showed controversial results in pre-clinical studies by either promoting or inhibiting tumor progression depending on the TLR and the tumor type (70). Multiple TLR agonists are currently in clinical development. As an example, the intra-tumoral injections of a TLR-9 agonist, the CpG-rich oligonucleotide PF-3512676, showed significant activity in both injected and non-injected lesions of B- and T-cell lymphomas when used concomitantly with low-dose (2 × 2 Gy) local irradiation (71).

STING has been shown to play an important role in the innate immune response against cancer. In the TME, tumor cell DNA detected by APCs is correlated with activation of STING pathway that leads to IFN- $\beta$  production (25), enhancing CD8<sup>+</sup> T cell priming and trafficking of effector T cells. One STING agonist, ADU-S100 (Aduro Biotech/Novartis), is currently evaluated in phase I clinical trial by intra-tumoral administration in cutaneously accessible tumors (72).

### CD 40 Agonistic Antibodies

CD40 is broadly expressed on immune cells, predominantly on DCs, B cells and macrophages. A major role of CD40 signaling is

to activate and “license” DCs to prime effective cytotoxic CD8<sup>+</sup> T cell responses. CD40 signaling can also be effectively triggered using agonistic antibodies or CD40L, thus bypassing the need for CD4<sup>+</sup> helper T cells (73). In preclinical studies, agonistic CD40 antibodies have demonstrated T cell-dependent anti-tumor activity, in particular in combination with conventional chemotherapy and immune checkpoint inhibitors (ICIs). Many CD40 antibodies are under clinical development (74). Toxicity profile is acceptable in monotherapy and combination trials are ongoing (75).

### Tumor Vaccines

The therapeutic breakthrough provided by ICIs and the demonstration of the role of MANA in T-cell mediated antitumor response have paved the way for a next generation of personalized cancer vaccines based on the use of MANA specific of the tumor. Preclinical results showed induction of efficient antitumor response and clinical trials providing a clinical proof of concept in melanoma have been published (76–78). A specific CD8<sup>+</sup> and CD4<sup>+</sup> T cell immune response characterized by the induction or the amplification of a preexisting response against MANA has been shown. Encouraging clinical activity seems to be associated but larger trials are awaited to firmly demonstrate the clinical activity of this therapeutic approach. Theoretically, cancer vaccines may either reinforce the activity and therapeutic margin of ICIs by increasing the number of specific effector T cells, or convert cold into inflamed tumors. The potential limitations are the availability of T cell repertoire in cancer patients and the risk of specific loss of heterozygosity (LOH) of the HLA presenting MANA in advanced metastatic tumors (79).

### Specific Therapies for Tumors With Impaired T Cell Trafficking to the Tumor TGF- $\beta$ Blocking Antibodies and TGF- $\beta$ -Receptor Antagonists

TGF- $\beta$  has been involved in cell proliferation, angiogenesis, epithelial-to-mesenchymal transition, immune infiltration, metastases dissemination, and drug resistance (80). Of note TGF- $\beta$  produced by the tumor cells mediates alterations in tumor-associated pDCs functions, e.g., impaired capacity to produce IFN type I, leading to a lacking/unbalanced T cell recruitment (81, 82). Recent reports have shown that TGF- $\beta$  in the peritumoral area was a major factor involved in T cell exclusion from the tumor. Mariathasan et al. used a preclinical model recapitulating T cell exclusion and showed that combination of a TGF $\beta$  blocking antibody with a PD-L1 antibody induced T cell penetration into the center of tumors, allowing anti-tumor immunity and tumor regression (41). Several TGF- $\beta$  antibodies or small molecules TGF- $\beta$ -receptor antagonists are in clinical development and could be tested in this setting.

### Anti-angiogenic Therapies

The clinical activity of anti-angiogenic drugs is modest when used as single agent. However, it has been shown that anti-angiogenic drugs normalize the tumor vasculature and

induce the upregulation of the leukocyte adhesion molecules ICAM-1 and VCAM-1 on tumor endothelial cells (83), leading to increased T cell infiltration (84). These therapies may thus represent a treatment of choice for tumors characterized by T cells blocked in the periphery of the tumor in order to enhance intratumoral penetration of T cells. It has been shown in the clinic that anti-angiogenic therapies could synergize with ICI in metastatic melanoma (85).

### Immunocytokines

Cytokines such as IL-2, TNF, IL-12 mediate the influx and expansion of leukocytes at the tumor site. However, these cytokines, and especially IL-12, induce significant toxicity when administered systemically in clinical trials (86). Next generation immunocytokines combining a fragment from a specific tumor antigen antibody with a modified cytokine are being developed with the aim of activating specifically the immune system inside the tumor to reduce systemic side effects. For example, cergutuzumab amunaleukin (CEA-IL2v), is a novel monomeric CEA-targeted immunocytokine that comprises a single IL-2 variant (IL2v) moiety with abolished CD25 binding (to avoid activation of regulatory T cells) fused to the C-terminus of a high affinity bivalent carcinoembryonic antigen (CEA)-specific antibody devoid of Fc-mediated effector functions. A superior efficacy over the respecting monotherapies was observed with CEA-IL-2v in combination with PD-L1 antibody and ADCC competent antibodies in CEA-positive solid tumor models (87).

## Therapeutic Approaches Active in Different Immune Contexts

### Adoptive T Cell Therapy and Chimeric Antigen Receptor (CAR) T Cells

Historically, cell-based therapies have focused on cytotoxic T cells targeting MHC-restricted antigens. This approach remains promising, in particular with the development of T cell receptor (TCR)-engineered T cells (88) and improvement of tumor infiltrating lymphocytes (TILs) infusion (89). However, their efficacy may be limited by the tendency of tumors to downregulate MHC molecules. CARs are engineered receptors made of the combination of an antigen binding domain of a monoclonal antibody specific for a cancer antigen (not MHC restricted) together with an intracellular domain of the CD3-zeta chain (90). CAR T cells responses can be further enhanced by addition of costimulatory domains, such as CD28 and CD137 (4-1BB) to support the expansion and persistence of genetically engineered cells *in vivo* (91). The reinfusion of CAR T-cells is preceded by a “lymphodepleting” chemotherapy regimen used to physically create enough space for the expansion and persistence of CAR T cell clones. Over the past decade multiple tumor antigens have been targeted by CARs. Outstanding activity of CAR T cells targeting CD19 has been observed in hematological malignancies, in particular in acute lymphoblastic leukemia and Diffuse Large B cell lymphoma with two CAR T-cell therapies approved by the Food and Drug Administration (FDA) in 2017 (92). The clinical activity of CAR in solid tumors is still to

demonstrate and will probably require optimization of CAR functions [for review see (93)].

### T-Cell Recruiting Bi-specific Antibodies

Bi-specific antibodies (bsAbs) are engineered antibodies that can bind two different antigens. One of the main strategies in the development of bsAbs is the recruitment and activation of immune effector T cells by targeting CD3 domain of the TCR complex (T-cell recruiting bsAbs) together with another antigen abnormally expressed on the tumor cell surface. The approval of catumaxomab (anti-epithelial cell adhesion molecule EpCAM and anti-CD3) and blinatumomab (anti-CD19 and anti-CD3) has become a major milestone in the development of bsAbs. BsAbs can be divided into two categories: immunoglobulin G (IgG)-like molecules and non-IgG-like molecules. Non-IgG-like bsAbs are smaller in size, leading to enhanced tissue penetration, but shorter half-life (94). Currently, more than 60 different bsAbs formats exist, some of them making their way into clinical trials. As for CAR T-cells, activity of T cell-recruiting bsAbs is not dependent on MHC class I expression on tumor cells and both approaches represent thus very promising treatments for MHC I-negative cold tumors. It is however not known whether a minimum threshold of T cells inside the tumor is required for the activity of bsAb, alone or in combination with other therapies. A comparison between CAR T cells and T-cell recruiting bsAbs is proposed in **Table 1**.

## PERSPECTIVES

Changing the natural history of a tumor characterized by the absence of T cells remains a great therapeutic challenge. However, as discussed above, many therapeutic approaches can be evaluated in this context. A major issue will be to determine the origin of this lack of T cell response to adapt the therapy to the physiology of the tumor. **Figure 2** summarizes the possibilities of treatment according to this tumor context.

Conventional anticancer approaches like chemotherapy and radiotherapy have still a room in the therapeutic armamentarium, not only to potentially induce ICD but also to reduce the tumor burden and thus potentially decrease the selection of immune-resistant clones. In other cases, “*in situ*” vaccination using PRR or CD40 agonists may be used to induce a specific immune antitumor response against naturally presented tumor antigens. Personalized cancer vaccines also represent a very promising strategy especially in highly mutated tumors. Finally, “supra-physiological” approaches like CAR T cells or T-cell recruiting bsAbs could be efficient in tumors characterized by the absence of MHC expression or even LOH of the HLA alleles presenting tumor antigens. It is likely that in the near future, algorithms of treatment will be developed to adapt the therapeutic strategy to the immune context of the tumor, considering also space and time evolution for adequate sequential strategies. At the end, while the efficacy of the therapy in inflamed tumors depends principally in the remove of the breaks induced by the immune activation itself, the conversion of a cold into an inflamed tumor will require a prior combination of therapies to induce immune infiltration



**TABLE 1** | Comparison between CAR T-cells and T-cell recruiting bi-specific antibodies.

	CAR T-Cells	Bi-specific antibodies
Mechanism of action	<ul style="list-style-type: none"> <li>- Direct cancer antigen recognition</li> <li>- Non MHC-restricted</li> <li>- Does not require pre-existing T cell infiltration, independent of receiver T cell characteristics</li> </ul>	<ul style="list-style-type: none"> <li>- Recruitment of immune effector T cells by their CD3 with another antigen expressed on the tumor cell</li> <li>- Non MHC-restricted</li> <li>- May be more dependent on quantity/quality of patients' T cells</li> </ul>
Administration	<ul style="list-style-type: none"> <li>- Single administration</li> <li>- Long half-life (months/years)</li> </ul>	<ul style="list-style-type: none"> <li>- Repeated administration (continuous infusion for non-IgG-like)</li> <li>- Short/intermediate half-life (hours/days)</li> </ul>
Tissue penetration	<ul style="list-style-type: none"> <li>- Homing of the T cells for blood, lymph nodes, and bone marrow</li> </ul>	<ul style="list-style-type: none"> <li>- Non IgG like: enhanced tissue penetration</li> </ul>
Toxicity	<ul style="list-style-type: none"> <li>- Acute reversible neurotoxicity (CD19)</li> <li>- Cytokine release syndrome (CRS)</li> </ul>	<ul style="list-style-type: none"> <li>- Lower toxicity expected</li> <li>- Acute reversible neurotoxicity (CD19)</li> <li>- Cytokine release syndrome (CRS)</li> </ul>
Main diseases	<ul style="list-style-type: none"> <li>- Outstanding activity in some hematological malignancies: B cell acute lymphoblastic leukemia, Diffuse Large B cell lymphoma (CD19)</li> <li>- Clinical trials in solid tumors</li> </ul>	<ul style="list-style-type: none"> <li>- B cell acute lymphoblastic leukemia (CD19)</li> <li>- Clinical trials for many solid tumors including colorectal, ovarian, breast and prostate cancer</li> </ul>
Other limitations	<ul style="list-style-type: none"> <li>- Clinical activity in solid tumors is still to demonstrate</li> <li>- Immunosuppressive microenvironment (rationale for combination with ICIs or use of optimized CAR T-cells)</li> <li>- Target specificity: risk of escape by loss of the target</li> </ul>	<ul style="list-style-type: none"> <li>- Clinical activity in solid tumors is still to demonstrate</li> <li>- Immunosuppressive microenvironment (rationale for combination with ICIs)</li> <li>- Target specificity: risk of escape by loss of the target</li> </ul>
Cost and availability	<ul style="list-style-type: none"> <li>- Long process, manufacturing issues</li> <li>- Cost<sup>+++</sup></li> </ul>	<ul style="list-style-type: none"> <li>- Immediate availability, less manufacturing and regulatory issues</li> <li>- Cost<sup>+</sup></li> </ul>

and then different immune checkpoint modulators to remove the breaks.

## AUTHOR CONTRIBUTIONS

PB, TS, VA, JV-G, and SD contributed to the writing. SD and CC contributed to the conception. SV-W, SA, CC, and SD participated in the lecture and corrections.

## REFERENCES

- Chen DS, Mellman I. Oncology meets immunology: the cancer-immunity cycle. *Immunity* (2013) 39:1–10. doi: 10.1016/j.immuni.2013.07.012
- Gajewski TF, Schreiber H, Fu YX. Innate and adaptive immune cells in the tumor microenvironment. *Nat Immunol.* (2013) 14:1014–22. doi: 10.1038/ni.2703
- Yarchoan M, Hopkins A, Jaffee EM. Tumor mutational burden and response rate to PD-1 Inhibition. *N Engl J Med.* (2017) 377:2500–1. doi: 10.1056/NEJMc1713444
- Kelderman S, Kvistborg P. Tumor antigens in human cancer control. *Biochim Biophys Acta BBA Rev Cancer* (2016) 1865:83–9. doi: 10.1016/j.bbcan.2015.10.004
- Coulie PG, Van den Eynde BJ, van der Bruggen P, Boon T. Tumour antigens recognized by T lymphocytes: at the core of cancer immunotherapy. *Nat Rev Cancer* (2014) 14:135–46. doi: 10.1038/nrc3670
- Mitch L. Mutation Burden Predicts Anti-PD-1 Response. *Cancer Discov.* (2018) 8:258. doi: 10.1158/2159-8290.CD-NB2018-005
- Chomez P, De Backer O, Bertrand M, De Plaen E, Boon T, Lucas S. An overview of the MAGE gene family with the identification of all human members of the family. *Cancer Res.* (2001) 61:5544–51.
- Fisk B, Blevins TL, Wharton JT, Ioannides CG. Identification of an immunodominant peptide of HER-2/neu protooncogene recognized by ovarian tumor-specific cytotoxic T lymphocyte lines. *J Exp Med.* (1995) 181:2109–17.
- Coulie PG, Brichard V, Van Pel A, Wölfel T, Schneider J, Traversari C, et al. A new gene coding for a differentiation antigen recognized by autologous cytolytic T lymphocytes on HLA-A2 melanomas. *J Exp Med.* (1994) 180:35–42.
- Yarchoan M, Johnson BA, Lutz ER, Laheru DA, Jaffee EM. Targeting neoantigens to augment antitumor immunity. *Nat Rev Cancer* (2017) 17:209–222. doi: 10.1038/nrc.2016.154
- Ilyas S, Yang JC. Landscape of tumor antigens in T-cell immunotherapy. *J Immunol.* (2015) 195:5117–22. doi: 10.4049/jimmunol.1501657
- Spranger S, Luke JJ, Bao R, Zha Y, Hernandez KM, Li Y, et al. Density of immunogenic antigens does not explain the presence or absence of the T-cell-inflamed tumor microenvironment in melanoma. *Proc Natl Acad Sci USA.* (2016) 113:E7759–68. doi: 10.1073/pnas.1609376113
- Peng W, Chen JQ, Liu C, Malu S, Creasy C, Tetzlaff MT, et al. Loss of PTEN promotes resistance to T cell-mediated immunotherapy. *Cancer Discov.* (2016) 6:202–216. doi: 10.1158/2159-8290.CD-15-0283
- Merad M, Sathe P, Helft J, Miller J, Mortha A. The dendritic cell lineage: ontogeny and function of dendritic cells and their subsets in the steady state and the inflamed setting. *Annu Rev Immunol.* (2013) 31:563–604. doi: 10.1146/annurev-immunol-020711-074950
- Guilliams M, Ginhoux F, Jakubczik C, Naik SH, Onai N, Schraml BU, et al. Dendritic cells, monocytes and macrophages: a unified nomenclature

- based on ontogeny. *Nat Rev Immunol.* (2014) 14:571–8. doi: 10.1038/nri3712
16. Jongbloed SL, Kassianos AJ, McDonald KJ, Clark GJ, Ju X, Angel CE, et al. Human CD141+ (BDCA-3)+ dendritic cells (DCs) represent a unique myeloid DC subset that cross-presents necrotic cell antigens. *J Exp Med.* (2010) 207:1247–60. doi: 10.1084/jem.20092140
  17. Bachem A, Güttler S, Hartung E, Ebstein F, Schaefer M, Tannert A, et al. Superior antigen cross-presentation and XCR1 expression define human CD11c+CD141+ cells as homologues of mouse CD8+ dendritic cells. *J Exp Med.* (2010) 207:1273–81. doi: 10.1084/jem.20100348
  18. Fuertes MB, Kacha AK, Kline J, Woo SR, Kranz DM, Murphy KM, et al. Host type I IFN signals are required for antitumor CD8+ T cell responses through CD8 $\alpha$ + dendritic cells. *J Exp Med.* (2011) 208:2005–16. doi: 10.1084/jem.20101159
  19. Hildner K, Edelson BT, Purtha WE, Diamond M, Matsushita H, Kohyama M, et al. Batf3 deficiency reveals a critical role for CD8 $\alpha$ + dendritic cells in cytotoxic T cell immunity. *Science* (2008) 322:1097–100. doi: 10.1126/science.1164206
  20. Spranger S, Dai D, Horton B, Gajewski TF. Tumor-residing Batf3 dendritic cells are required for effector t cell trafficking and adoptive t cell therapy. *Cancer Cell* (2017) 31:711–23.e4. doi: 10.1016/j.ccell.2017.04.003
  21. Broz ML, Binnewies M, Boldajipour B, Nelson AE, Pollack JL, Erle DJ, et al. Dissecting the tumor myeloid compartment reveals rare activating antigen-presenting cells critical for T cell immunity. *Cancer Cell* (2014) 26:638–52. doi: 10.1016/j.ccell.2014.09.007
  22. Böttcher JP, Bonavita E, Chakravarty P, Blees H, Cabeza-Cabrero M, Sammicheli S, et al. NK cells stimulate recruitment of cDC1 into the tumor microenvironment promoting cancer immune control. *Cell* (2018) 172:1022–37.e14. doi: 10.1016/j.cell.2018.01.004
  23. Barry KC, Hsu J, Broz ML, Cueto FJ, Binnewies M, Combes AJ, et al. A natural killer–dendritic cell axis defines checkpoint therapy–responsive tumor microenvironments. *Nat Med.* (2018) 24:1178–91. doi: 10.1038/s41591-018-0085-8
  24. Sánchez-Paulete AR, Cueto FJ, Martínez-López M, Labiano S, Morales-Kastresana A, Rodríguez-Ruiz ME, et al. Cancer immunotherapy with immunomodulatory anti-CD137 and anti-PD-1 monoclonal antibodies requires BATF3-dependent dendritic cells. *Cancer Discov.* (2016) 6:71–9. doi: 10.1158/2159-8290.CD-15-0510
  25. Woo SR, Fuertes MB, Corrales L, Spranger S, Furdyna MJ, Leung MYK, et al. STING-dependent cytosolic DNA sensing mediates innate immune recognition of immunogenic tumors. *Immunity* (2014) 41:830–42. doi: 10.1016/j.immuni.2014.10.017
  26. Dieu-Nosjean MC, Antoine M, Danel C, Heudes D, Wislez M, Poulot V, et al. Long-term survival for patients with non-small-cell lung cancer with intratumoral lymphoid structures. *J Clin Oncol.* (2008) 26:4410–7. doi: 10.1200/JCO.2007.15.0284
  27. Akira S, Uematsu STO. Pathogen recognition and innate immunity. *Cell* (2006) 124:783–801. doi: 10.1016/j.cell.2006.02.015
  28. Janeway CA, Medzhitov R. Innate immune recognition. *Annu Rev Immunol.* (2002) 20:197–216. doi: 10.1146/annurev.immunol.20.083001.084359
  29. Reis e Sousa C. Dendritic cells in a mature age. *Nat Rev Immunol.* (2006) 6:476–83. doi: 10.1038/nri1845
  30. Kroemer G, Galluzzi L, Kepp O, Zitvogel L. Immunogenic cell death in cancer therapy. *Annu Rev Immunol.* (2013) 31:51–72. doi: 10.1146/annurev-immunol-032712-100008
  31. Green DR, Ferguson T, Zitvogel L, Kroemer G. Immunogenic and tolerogenic cell death. *Nat Rev Immunol.* (2009) 9:353–63. doi: 10.1038/nri2545
  32. Vacchelli E, Ma Y, Baracco EE, Sistigu A, Enot DP, Pietrocola F, et al. Chemotherapy-induced antitumor immunity requires formyl peptide receptor 1. *Science* (2015) 350:972–8. doi: 10.1126/science.aad0779
  33. Ridge JP, Di Rosa FMP. A conditioned dendritic cell can be a temporal bridge between a CD4+ T-helper and a T-killer cell. *Nature* (1998) 393:474–8. doi: 10.1038/30989
  34. Rosenberg SA, Sherry RM, Morton KE, Scharfman WJ, Yang JC, Topalian SL, et al. Tumor progression can occur despite the induction of very high levels of self/tumor antigen-specific CD8+ T cells in patients with melanoma. *J Immunol.* (2005) 175:6169–76. doi: 10.4049/jimmunol.175.9.6169
  35. Lohneis P, Sinn M, Bischoff S, Jühling A, Pelzer U, Wislocka L, et al. Cytotoxic tumour-infiltrating T lymphocytes influence outcome in resected pancreatic ductal adenocarcinoma. *Eur J Cancer* (2017) 83:290–301. doi: 10.1016/j.ejca.2017.06.016
  36. Peranzoni E, Lemoine J, Vimeux L, Feuillet V, Barrin S, Kantari-Mimoun C, et al. Macrophages impede CD8 T cells from reaching tumor cells and limit the efficacy of anti-PD-1 treatment. *Proc Natl Acad Sci USA.* (2018) 115:E4041–50. doi: 10.1073/pnas.1720948115
  37. Spranger S, Bao R, Gajewski TF. Melanoma-intrinsic  $\beta$ -catenin signalling prevents anti-tumour immunity. *Nature* (2015) 523:231–5. doi: 10.1038/nature14404
  38. Toso A, Revandkar A, Di Mitri D, Guccini I, Proietti M, Sarti M, et al. Enhancing chemotherapy efficacy in Pten-deficient prostate tumors by activating the senescence-associated antitumor immunity. *Cell Rep.* (2014) 9:75–89. doi: 10.1016/j.celrep.2014.08.044
  39. George S, Miao D, Demetri GD, Adegbe D, Rodig SJ, Shukla S, et al. Loss of PTEN Is Associated with resistance to Anti-PD-1 checkpoint blockade therapy in metastatic uterine leiomyosarcoma. *Immunity* (2017) 46:197–204. doi: 10.1016/j.immuni.2017.02.001
  40. Tauriello DVF, Palomo-Ponce S, Stork D, Berenguer-Llergo A, Badiá-Ramentol J, Iglesias M, et al. TGF $\beta$  drives immune evasion in genetically reconstituted colon cancer metastasis. *Nature* (2018) 554:538–43. doi: 10.1038/nature25492
  41. Mariathasan S, Turley SJ, Nickles D, Castiglioni A, Yuen K, Wang Y, et al. TGF $\beta$  attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. *Nature* (2018) 554:544–8. doi: 10.1038/nature25501
  42. Soares KC, Rucki AA, Kim V, Foley K, Solt S, Wolfgang CL, et al. TGF- $\beta$  blockade depletes T regulatory cells from metastatic pancreatic tumors in a vaccine dependent manner. *Oncotarget* (2015) 6:43005–15. doi: 10.18632/oncotarget.5656
  43. Randolph GJ, Ochando J, Partida-Sánchez S. Migration of dendritic cell subsets and their precursors. *Annu Rev Immunol.* (2008) 26:293–316. doi: 10.1146/annurev.immunol.26.021607.090254
  44. Luther SA, Cyster JG. Chemokines as regulators of T cell differentiation. *Nat Immunol.* (2001) 2:102–7. doi: 10.1038/84205
  45. Montoya M, Schiavoni G, Mattei F, Gresser I, Belardelli F, Borrow P, et al. Type I interferons produced by dendritic cells promote their phenotypic and functional activation. *Blood* (2002) 99:3263–3271. doi: 10.1182/blood.V99.9.3263
  46. Hojo S, Koizumi K, Tsuneyama K, Arita Y, Cui Z, Shinohara K, et al. High-level expression of chemokine CXCL16 by tumor cells correlates with a good prognosis and increased tumor-infiltrating lymphocytes in colorectal cancer. *Cancer Res.* (2007) 67:4725–31. doi: 10.1158/0008-5472.CAN-06-3424
  47. Harlin H, Meng Y, Peterson AC, Zha Y, Tretiakova M, Slingluff C, et al. Chemokine expression in melanoma metastases associated with CD8+ T-cell recruitment. *Cancer Res.* (2009) 69:3077–85. doi: 10.1158/0008-5472.CAN-08-2281
  48. Peng D, Kryczek I, Nagarsheth N, Zhao L, Wei S, Wang W, et al. Epigenetic silencing of TH1-type chemokines shapes tumour immunity and immunotherapy. *Nature* (2015) 527:249–53. doi: 10.1038/nature15520
  49. Molon B, Ugel S, Del Pozzo F, Soldani C, Zilio S, Avella D, et al. Chemokine nitration prevents intratumoral infiltration of antigen-specific T cells. *J Exp Med.* (2011) 208:1949–62. doi: 10.1084/jem.20101956
  50. Srivastava P, Paluch BE, Matsuzaki J, James SR, Collamat-Lai G, Karbach J, et al. Immunomodulatory action of SGI-110, a hypomethylating agent, in acute myeloid leukemia cells. *Leuk Res.* (2014) 38:1332. doi: 10.1016/j.leukres.2014.09.001
  51. Dunn J, Rao S. Epigenetics and immunotherapy: the current state of play. *Mol Immunol.* (2017) 87:227–39. doi: 10.1016/j.molimm.2017.04.012
  52. Chiappinelli KB, Strissel PL, Desrichard A, Li H, Henke C, Akman B, et al. Inhibiting DNA Methylation Causes an Interferon Response in Cancer via dsRNA including endogenous retroviruses. *Cell* (2015) 162:974–86. doi: 10.1016/j.cell.2015.07.011

53. Brocks D, Schmidt CR, Daskalakis M, Jang HS, Shah NM, Li D, et al. DNMT and HDAC inhibitors induce cryptic transcription start sites encoded in long terminal repeats. *Nat Genet.* (2017) 49:1052–60. doi: 10.1038/ng.3889
54. Licht JD. DNA Methylation inhibitors in cancer therapy: the immunity dimension. *Cell* (2015) 162:938–9. doi: 10.1016/j.cell.2015.08.005
55. Chiappinelli KB, Zahnow CA, Ahuja N, Baylin SB. Combining epigenetic and immune therapy to combat cancer. *Cancer Res.* (2016) 76:1683. doi: 10.1158/0008-5472.CAN-15-2125
56. Zhang C, Oberoi P, Oelsner S, Waldmann A, Lindner A, Tonn T, et al. Chimeric antigen receptor-engineered NK-92 cells: an off-the-shelf cellular therapeutic for targeted elimination of cancer cells and induction of protective antitumor immunity. *Front Immunol.* (2017) 8:533. doi: 10.3389/fimmu.2017.00533
57. Chu J, Deng Y, Benson DM, He S, Hughes T, Zhang J, et al. CS1-specific chimeric antigen receptor (CAR)-engineered natural killer cells enhance *in vitro* and *in vivo* anti-tumor activity against human multiple myeloma. *Leukemia* (2014) 28:917–27. doi: 10.1038/leu.2013.279
58. Mahoney KM, Rennert PD, Freeman GJ. Combination cancer immunotherapy and new immunomodulatory targets. *Nat Rev Drug Discov.* (2015) 14:561–84. doi: 10.1038/nrd4591
59. Aranda F, Vacchelli E, Eggermont A, Galon J, Fridman WH, Zitvogel L, et al. Trial watch: immunostimulatory monoclonal antibodies in cancer therapy. *Oncoimmunology* (2014) 3:e27297. doi: 10.4161/onci.27297
60. DeNardo DG, Brennan DJ, Rexhepaj E, Ruffell B, Shiao SL, Madden SF, et al. Leukocyte complexity predicts breast cancer survival and functionally regulates response to chemotherapy. *Cancer Discov.* (2011) doi: 10.1158/2159-8274.CD-10-0028
61. Golden EB, Apetoh L. Radiotherapy and immunogenic cell death. *Semin Radiat Oncol.* (2015) 25:11–7. doi: 10.1016/j.semradonc.2014.07.005
62. Parslow AC, Parakh S, Lee FT, Gan HK, Scott AM. Antibody-drug conjugates for cancer therapy. *Biomedicines* (2016) 4:14 doi: 10.3390/biomedicines4030014
63. Russell SJ, Peng KW, Bell JC. Oncolytic virotherapy. *Nat Biotechnol.* (2012) 30:658–70. doi: 10.1038/nbt.2287
64. Kaufman HL, Kohlhapp FJ, Zloza A. Oncolytic viruses: a new class of immunotherapy drugs. *Nat Rev Drug Discov.* (2015) 14:642–62. doi: 10.1038/nrd4663
65. Seth RB, Sun L, Chen ZJ. Antiviral innate immunity pathways. *Cell Res.* (2006) 16:141–7. doi: 10.1038/sj.cr.7310019
66. Ribas A, Dummer R, Puzanov I, VanderWalde A, Andtbacka RHI, Michielin O, et al. Oncolytic virotherapy promotes intratumoral T cell infiltration and improves anti-PD-1 immunotherapy. *Cell* (2017) 170:1109–19.e10. doi: 10.1016/j.cell.2017.08.027
67. Bevers RFM, Kurth K-H, Schamhart DHJ. Role of urothelial cells in BCG immunotherapy for superficial bladder cancer. *Br J Cancer* (2004) 91:607–12. doi: 10.1038/sj.bjc.6602026
68. Jahrsdörfer B, Hartmann G, Racila E, Jackson W, Mühlhoff L, Meinhardt G, et al. CpG DNA increases primary malignant B cell expression of costimulatory molecules and target antigens. *J Leukoc Biol.* (2001) 69:81–8. doi: 10.1189/jlb.69.1.81
69. Smits ELJM, Cools N, Lion E, Van Camp K, Ponsaerts P, Berneman ZN, et al. The Toll-like receptor 7/8 agonist resiquimod greatly increases the immunostimulatory capacity of human acute myeloid leukemia cells. *Cancer Immunol Immunother.* (2010) 59:35–46. doi: 10.1007/s00262-009-0721-8
70. Pradere JP, Dapito DH, Schwabe RF. The Yin and Yang of toll-like receptors in cancer. *Oncogene* (2014) 33:3485–95. doi: 10.1038/ncr.2013.302
71. Brody JD, Ai WZ, Czerwinski DK, Torchia JA, Levy M, Advani RH, et al. *In situ* vaccination with a TLR9 agonist induces systemic lymphoma regression: a phase I/II study. *J Clin Oncol.* (2010) 28:4324–32. doi: 10.1200/JCO.2010.28.9793
72. *Safety and Efficacy of MIW815 (ADU-S100) +/- Ipilimumab in Patients With Advanced/Metastatic Solid Tumors or Lymphomas - Full Text View - ClinicalTrials.gov.* Available online at: <https://clinicaltrials.gov/ct2/show/NCT02675439> (Accessed May 25, 2018).
73. Khong A, Nelson DJ, Nowak AK, Lake RA, Robinson BWS. The use of agonistic anti-CD40 therapy in treatments for cancer. *Int Rev Immunol.* (2012) 31:246–66. doi: 10.3109/08830185.2012.698338
74. Vonderheide RH, Burg JM, Mick R, Trosko JA, Li D, Shaik MN, et al. Phase I study of the CD40 agonist antibody CP-870,893 combined with carboplatin and paclitaxel in patients with advanced solid tumors. *Oncoimmunology* (2013) 2:e23033. doi: 10.4161/onci.23033
75. Vonderheide RH. The immune revolution: a case for priming, not checkpoint. *Cancer Cell* (2018) 33:563–9. doi: 10.1016/j.ccell.2018.03.008
76. Ott PA, Hu Z, Keskin DB, Shukla SA, Sun J, Bozym DJ, et al. An immunogenic personal neoantigen vaccine for patients with melanoma. *Nature* (2017) 547:217–221. doi: 10.1038/nature22991
77. Sahin U, Derhovannessian E, Miller M, Kloke B-P, Simon P, Löwer M, et al. Personalized RNA mutanome vaccines mobilize poly-specific therapeutic immunity against cancer. *Nature* (2017) 547:222–6. doi: 10.1038/nature23003
78. Carreno BM, Magrini V, Becker-Hapak M, Kaabinejadian S, Hundal J, Petti AA, et al. Cancer immunotherapy. A dendritic cell vaccine increases the breadth and diversity of melanoma neoantigen-specific T cells. *Science* (2015) 348:803–8. doi: 10.1126/science.aaa3828
79. McGranahan N, Rosenthal R, Hiley CT, Rowan AJ, Watkins TBK, Wilson GA, et al. Allele-specific HLA loss and immune escape in lung cancer evolution. *Cell* (2017) 171:1259–71.e11. doi: 10.1016/j.cell.2017.10.001
80. Gramont A de, Faivre S, Raymond E. Novel TGF- $\beta$  inhibitors ready for prime time in onco-immunology. *Oncoimmunology* (2017) 6:e1257453. doi: 10.1080/2162402X.2016.1257453
81. Labidi-Galy SI, Sisirak V, Meeus P, Gobert M, Treilleux I, Bajard A, et al. Quantitative and functional alterations of plasmacytoid dendritic cells contribute to immune tolerance in ovarian cancer. *Cancer Res.* (2011) 71:5423–34. doi: 10.1158/0008-5472.CAN-11-0367
82. Sisirak V, Faget J, Vey N, Blay JY, Ménétrier-Caux C, Caux C, et al. Plasmacytoid dendritic cells deficient in IFN $\alpha$  production promote the amplification of FOXP3+ regulatory T cells and are associated with poor prognosis in breast cancer patients. *Oncoimmunology* (2013) 2:e22338. doi: 10.4161/onci.22338
83. Kandalaf LE, Motz GT, Busch J, Coukos G. Angiogenesis and the tumor vasculature as antitumor immune modulators: the role of vascular endothelial growth factor and endothelin. *Curr Top Microbiol Immunol.* (2011) 344:129–48. doi: 10.1007/82\_2010\_95
84. Shrimali RK, Yu Z, Theoret MR, Chinnasamy D, Restifo NP, Rosenberg SA. Antiangiogenic agents can increase lymphocyte infiltration into tumor and enhance the effectiveness of adoptive immunotherapy of cancer. *Cancer Res.* (2010) 70:6171–80. doi: 10.1158/0008-5472.CAN-10-0153
85. Ott PA, Hodi FS, Buchbinder EI. Inhibition of immune checkpoints and vascular endothelial growth factor as combination therapy for metastatic melanoma: an overview of rationale, preclinical evidence, and initial clinical data. *Front Oncol* (2015) 5:202. doi: 10.3389/fonc.2015.00202
86. Lasek W, Zagożdżon R, Jakobisiak M. Interleukin 12: still a promising candidate for tumor immunotherapy? *Cancer Immunol Immunother.* (2014) 63:419–35. doi: 10.1007/s00262-014-1523-1
87. Klein C, Waldhauer I, Nicolini VG, Freimoser-Grundschober A, Nayak T, Vugts DJ, et al. Cergutuzumab amunaleukin (CEA-IL2v), a CEA-targeted IL-2 variant-based immunocytokine for combination cancer immunotherapy: overcoming limitations of aldesleukin and conventional IL-2-based immunocytokines. *Oncoimmunology* (2017) 6:e1277306. doi: 10.1080/2162402X.2016.1277306
88. Matsuda T, Leisegang M, Park JH, Ren L, Kato T, Ikeda Y, et al. Induction of neoantigen-specific cytotoxic t cells and construction of T-cell receptor-engineered T cells for ovarian cancer. *Clin Cancer Res.* (2018) 24:5357–67. doi: 10.1158/1078-0432.CCR-18-0142
89. Doherty M, Leigh NB, Feld R, Bradbury PA, Wang L, Nie J, et al. Phase I/II study of tumor-infiltrating lymphocyte (TIL) infusion and low-dose interleukin-2 (IL-2) in patients with advanced malignant pleural mesothelioma (MPM). *J Clin Oncol.* (2015) 33:TPS7586. doi: 10.1200/jco.2015.33.15\_suppl.tps7586

90. Irving BA, Weiss A. The cytoplasmic domain of the T cell receptor zeta chain is sufficient to couple to receptor-associated signal transduction pathways. *Cell* (1991) 64:891–901.
91. Carpenito C, Milone MC, Hassan R, Simonet JC, Lakhal M, Suhoski MM, et al. Control of large, established tumor xenografts with genetically retargeted human T cells containing CD28 and CD137 domains. *Proc Natl Acad Sci USA*. (2009) 106:3360–5. doi: 10.1073/pnas.0813101106
92. Research AA for C. CAR T-cell therapies produce durable remissions. *Cancer Discov*. (2018) 8:379. doi: 10.1158/2159-8290.CD-NB2018-017
93. Fesnak AD, June CH, Levine BL. Engineered T cells: the promise and challenges of cancer immunotherapy. *Nat Rev Cancer* (2016) 16:566–81. doi: 10.1038/nrc.2016.97
94. Kontermann RE, Brinkmann U. Bispecific antibodies. *Drug Discov Today* (2015) 20:838–47. doi: 10.1016/j.drudis.2015.02.008

**Conflict of Interest Statement:** SD is also an employee for Cellectis and reports personal fees from AstraZeneca, Elsalys, Erytech Pharma, and Netris Pharma.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Bonaventura, Shekarian, Alcazer, Valladeau-Guilemond, Valsesia-Wittmann, Amigorena, Caux and Depil. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



## REVUE 2 - Nouvelles approches vaccinales en cancérologie

*Stéphane Depil<sup>1,2</sup>, Paola Bonaventura<sup>1</sup>, **Vincent Alcazer<sup>1</sup>**, Laurie Tonon<sup>3</sup>*

1 Centre Léon-Bérard, 28, Prom. Léa-et-Napoléon-Bullukian, 69008 Lyon, France

2 Université Claude-Bernard-Lyon-1, 43, boulevard du 11-Novembre-1918, 69100 Villeurbanne, France

3 Fondation Synergie-Lyon-Cancer, 28, rue Laennec, 69008 Lyon, France

\* Corresponding author: Stéphane Depil, Centre Léon Bérard, 28 Prom. Léa et Napoléon Bullukian, 69008 Lyon, France;

Tel: +33 (0) 4 78 78 28 28; Mail: [stephane.depil@lyon.unicancer.fr](mailto:stephane.depil@lyon.unicancer.fr)

Bull Cancer. 2018 Dec;105 Suppl 1:S113-S120. French. doi: 10.1016/S0007-4551(18)30396-5. PMID: 30595192.



# Nouvelles approches vaccinales en cancérologie

Stéphane Depil<sup>1,2</sup>, Paola Bonaventura<sup>1</sup>, Vincent Alcazer<sup>1</sup>, Laurie Tonon<sup>3</sup>

1. Centre Léon-Bérard, 28, Prom. Léa-et-Napoléon-Bullukian, 69008 Lyon, France
2. Université Claude-Bernard-Lyon-1, 43, boulevard du 11-Novembre-1918, 69100 Villeurbanne, France
3. Fondation Synergie-Lyon-Cancer, 28, rue Laennec, 69008 Lyon, France

## Correspondance :

Stéphane Depil  
stephane.depil@lyon.unicancer.fr

## Mots clés

Vaccin  
Néoantigènes  
Checkpoints  
Adjuvant

## ■ Résumé

Le concept de vaccination antitumorale repose sur la mise en place d'une réponse immunitaire spécifique, dirigée contre des antigènes présentés par la tumeur, et capable de contrôler et détruire les cellules tumorales exprimant ces antigènes. Les premières générations de vaccins antitumoraux ont en général été associées à des résultats cliniques assez décevants. Le changement de paradigme amené par les inhibiteurs de checkpoints de type anti-*Programmed Cell-Death 1* (PD-1/*Programmed Death-Ligand 1* (PD-L1) et la démonstration du rôle clé joué dans la réponse T antitumorale par les néoantigènes associés aux mutations somatiques spécifiques de la tumeur ont permis de remettre en avant le concept de vaccination antitumorale. Cette nouvelle génération de vaccins repose donc sur l'utilisation de néoantigènes mutés propres à chaque tumeur (approche personnalisée). Les modèles précliniques ont montré que ces néoantigènes permettaient de générer des réponses T antitumorales efficaces, en lien avec l'absence de sélection thymique des clones T réactifs. Plusieurs études cliniques sont actuellement en cours, avec des premiers résultats prometteurs. La démonstration clinique d'un effet antitumoral reste néanmoins à réaliser. D'autre part, l'optimisation du processus de sélection de ces néoantigènes apparaît comme un enjeu majeur. À terme, la vaccination pourrait soit renforcer la réponse aux anti-PD-1/PD-L1, soit permettre d'induire une réponse T antitumorale dans les tumeurs immunologiquement « froides ». Enfin, la vaccination pourrait avoir un positionnement privilégié en adjuvant pour permettre le contrôle au long cours de la maladie résiduelle.

## Keywords

Vaccine  
Neoantigens  
Checkpoints  
Adjuvant

## ■ Summary

### The new vaccine approaches in oncology

*The concept of cancer vaccine relies on the induction of an adaptive immune response against antigens presented by tumor cells, in order to allow a long-term control of the tumor. First generations of cancer vaccines have been associated with relatively disappointing clinical results. The therapeutic breakthrough provided by the immune checkpoint inhibitors targeting PD-1/PD-L1, and the demonstration of the key role played by mutation-associated neoantigens in anti-tumor T cell response, have led to a renaissance of cancer vaccination. This next generation of cancer*

*vaccines is thus based on the use of mutated neoantigens specific to each tumor (personalized approach). Preclinical results have shown that these neoantigens are able to induce potent anti-tumor immune responses, explained by the absence of thymic selection against the reactive T cells. Several clinical trials are ongoing, with promising preliminary results. Nevertheless, the demonstration of an anti-tumor effect remains to be done in the clinic. The optimization of neoantigen selection process is still a major challenge. Finally, cancer vaccines may either reinforce the activity of anti-PD-1/PD-L1, by increasing the number of effector T cells in the tumor, or induce an adaptive T cell response in non-T cell infiltrated "cold tumors". Vaccination may also have a strong interest in an adjuvant setting to allow a long-term control of the residual disease.*

Cet article fait partie du numéro supplément *Nouvelles perspectives dans l'immunothérapie des cancers* réalisé avec le soutien institutionnel des partenaires AstraZeneca, Bristol-Myers Squibb, Roche.

## Historique rapide de la vaccination en cancérologie

Le concept de vaccination antitumorale repose sur la mise en place d'une réponse immunitaire spécifique, dirigée contre des antigènes présentés par la tumeur, et capable de contrôler et détruire les cellules tumorales exprimant ces antigènes. L'identification d'antigènes tumoraux de type *Melanoma-Associated Antigen-A1* (MAGE-A1) dans le mélanome, suivie par la découverte d'autres antigènes dans d'autres types tumoraux, a conduit à la réalisation de très nombreux essais cliniques. Malgré l'inclusion de milliers de patients, les résultats cliniques sont restés assez limités avec des taux de réponses objectives inférieurs à 7 % [1] et un bénéfice clinique estimé aux alentours de 20 % [2]. Un gain de survie de 4 mois a néanmoins pu être démontré avec le vaccin Provenge (cellules dendritiques autologues chargées avec des antigènes tumoraux prostatiques) en phase III [3]. La vaccination avec un peptide dérivé de l'antigène de mélanome gp100 a pu aboutir à une discrète amélioration de la survie sans progression de 0,6 mois, sans gain de survie significatif [4]. En revanche, les essais de phase III récents de vaccination avec l'antigène *Melanoma-Associated Antigen-A3* (MAGE-A3) dans le mélanome et le cancer du poumon se sont avérés négatifs [5].

Différentes explications peuvent être avancées pour expliquer ces résultats cliniques plutôt décevants, et souvent corrélés à des réponses immunitaires assez faibles : (i) utilisation d'antigènes du soi et stimulation de lymphocytes T possédant un récepteur de faible affinité pour l'antigène, incapables de générer une réponse antitumorale efficace ; (ii) utilisation d'antigène monovalent favorisant la sélection de clones tumoraux résistants n'exprimant pas (ou plus) cet antigène ; (iii) support vaccinal suboptimal induisant des réponses lymphocytaires T de faible intensité et de faible durée ; (iv) microenvironnement immunosuppresseur bloquant la phase effectrice de la réponse immune induite ; (v) utilisation dans un contexte clinique non adapté de maladie avancée métastatique rapidement progressive [6].

Le changement de paradigme associé aux résultats cliniques des inhibiteurs de checkpoints de type anti-PD-1/PD-L1 et la démonstration du rôle joué dans les réponses T antitumorales par les néoantigènes associés aux mutations présentes dans la tumeur (*Mutation-Associated Neoantigens* [MANA], que nous appellerons aussi ici néoantigènes mutés) ont permis de remettre en avant le concept de vaccination antitumorale basée sur l'utilisation de néoantigènes mutés spécifiques de la tumeur, avec donc le développement d'une vaccination personnalisée.

## Néoantigènes mutés et réponse immunitaire

### Définition

La présence de mutations non synonymes dans les cellules tumorales peut conduire à l'expression de protéines mutées spécifiques de la tumeur, non présentes dans les cellules normales (néoantigènes). Les processus de dégradation protéasomale vont, pour certaines de ces protéines, générer de nouveaux peptides, dont certains présenteront une affinité accrue pour le complexe majeur d'histocompatibilité (CMH) ou pour le récepteur T (*T Cell Receptor* [TCR]) par rapport au peptide natif non muté (néoépitopes). Ces peptides néoépitopiques étant spécifiques de la tumeur, ils peuvent en théorie être reconnus par des lymphocytes T de forte affinité (absence de tolérance centrale par délétion thymique de ces lymphocytes) (*figure 1*).

## Réponse aux inhibiteurs de checkpoints et néoantigènes mutés

Plusieurs études ont permis de mettre en évidence un lien entre la réponse aux inhibiteurs de checkpoints et la quantité de mutations présentes dans la tumeur (charge mutationnelle), dans le mélanome [7], le cancer du poumon [8] ou le cancer colorectal [9]. Une relation linéaire a ainsi pu être mise en évidence,

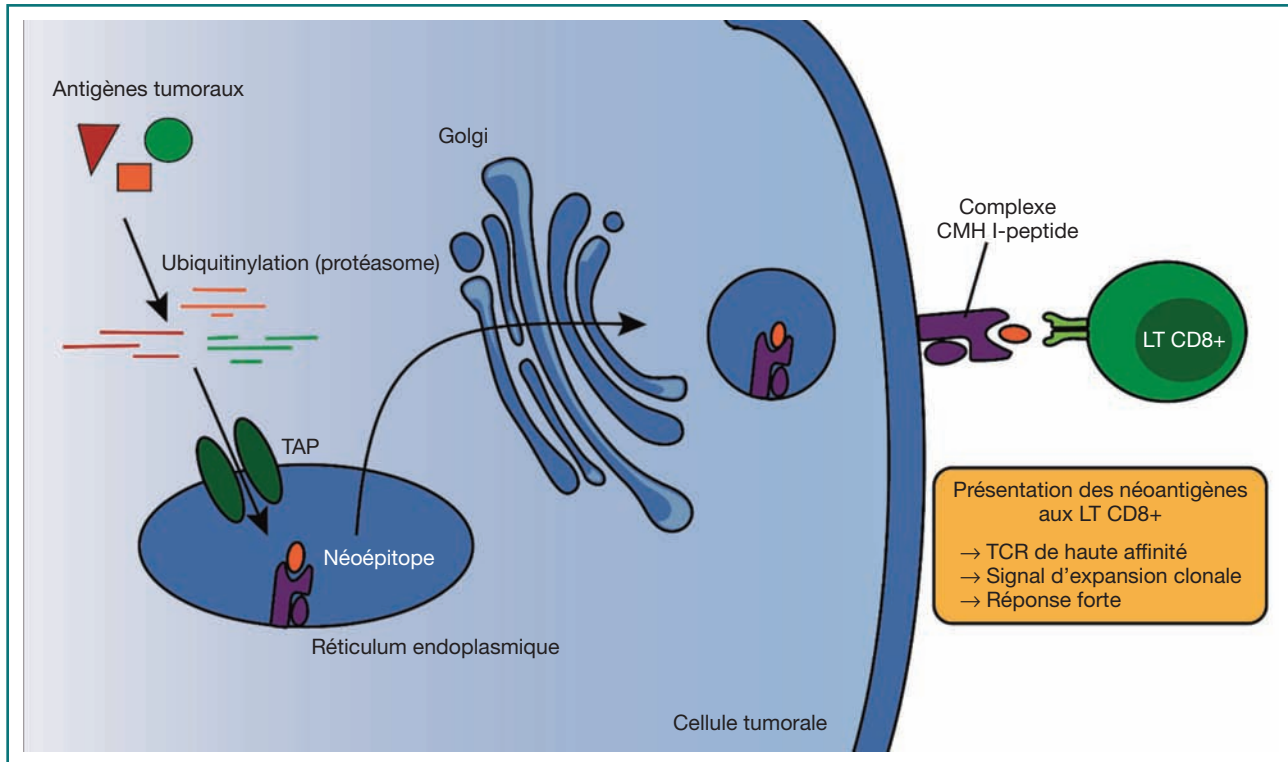


FIGURE 1

**Processing et présentation des antigènes tumoraux au niveau du CMH de classe I.** Les antigènes tumoraux sont dégradés par le protéasome, générant ainsi de courts peptides entrant dans le réticulum endoplasmique *via* la protéine TAP. Les peptides clivés y sont chargés sur le CMH, puis le complexe CMH-peptide migre à la surface de la cellule après un passage par l'appareil de Golgi. Dans le cadre d'un néoantigène muté spécifique de la tumeur, les néoépitoques présentés par le complexe CMH I sont reconnus par des lymphocytes T *via* un TCR de forte affinité, pouvant entraîner une prolifération clonale accompagnée d'une forte réponse cytotoxique. TAP : *Transporter Associated with Antigen Processing*, LT : lymphocyte T, TCR : T Cell Receptor. CMH : complexe majeur d'histocompatibilité.

dans une analyse rétrospective et multi-pathologies, entre le taux de réponse objective aux anti-PD-1/PD-L1 et le nombre médian de mutations somatiques codantes par mégabase [10]. La charge mutationnelle étant un reflet de la quantité possible de néoantigènes mutés, ces éléments suggèrent que les néoantigènes mutés représentent des cibles majeures des lymphocytes T activés après traitement par inhibiteurs de checkpoints [11]. Il s'agit néanmoins d'études corrélatives n'éliminant pas le rôle d'autres antigènes tumoraux.

**Preuve de concept préclinique**

Une première approche a consisté à utiliser des modèles murins syngéniques. Il a ainsi pu être montré dans le modèle de mélanome B16F10 que la vaccination avec de longs peptides contenant des néoépitoques mutés induisait une réponse immunitaire spécifique (16/50 peptides immunogènes en test *Enzyme-Linked Immunospot* [ELISPOT] interféron [IFN]- $\gamma$ ) et avait un effet antitumoral sur un mode aussi bien préventif que thérapeutique [12]. De manière assez inattendue – mais à corrélérer aux résultats cliniques aujourd'hui connus (voir ci-dessous) –, il

a pu être montré que la plupart des néoantigènes mutés dans les modèles B16F10 et CT26 correspondaient à des épitoques T CD4 et que l'utilisation d'acide ribonucléique messager (ARNm) vaccinal contenant des épitoques T CD4+ était capable d'induire une réponse immunitaire antitumorale efficace [13]. Une étude a été réalisée dans le modèle de sarcome murin T3 et a permis de montrer que l'utilisation de longs peptides contenant deux néoantigènes connus pour être impliqués dans la réponse aux inhibiteurs de checkpoints dans ce modèle était capable d'induire un effet antitumoral comparable à un inhibiteur de checkpoints [14]. Une autre étude, utilisant la spectrométrie de masse pour confirmer la liaison au CMH I des néoépitoques prédits, a également pu montrer l'efficacité antitumorale de peptides contenant ces néoépitoques dans les modèles MC38 et *Transgenic Adenocarcinoma Mouse Prostate* (TRAMP)-1 [15]. La deuxième approche préclinique repose sur des modèles de souris immunodéprimées de type *NOD Scid Gamma* (NSG) avec implantation de xélogreffe de tumeur humaine (*Patient Derived Xenograft* [PDX]). Trois modèles de PDX de sein ont ainsi été mis en place et des néoantigènes mutés de ces tumeurs prédits. Le

transfert adoptif de cellules mononucléées sanguines stimulées *in vitro* avec ces néoantigènes est associé à un effet antitumoral à l'encontre de la tumeur correspondante [16].

## Méthodes d'identification des néoantigènes mutés

La prédiction de néoépitopes est un point critique qui soulève encore de nombreuses questions quant à son optimisation. En ce qui concerne les épitopes T CD8, la mutation doit être exprimée pour aboutir à une protéine mutée qui sera clivée par le protéasome pour générer de courts peptides qui seront transportés dans le réticulum endoplasmique par les protéines *Transporter Associated with Antigen Processing* (TAP). Le peptide muté doit se lier à la molécule de CMH et le complexe CMH-peptide reconnu par un lymphocyte T *via* son récepteur (*figure 1*). Les différentes étapes de prédiction (*figure 2*) reposent donc sur :

- (i) la détection précise des mutations somatiques (c'est-à-dire propres à la tumeur) ainsi que leur clonalité. Des outils de plus en plus performants existent pour détecter ces mutations dans des échantillons souvent contaminés par de l'acide désoxyribonucléique (ADN) non tumoral. Toutefois l'attribution d'un statut clonal (partagé par tous les clones de la tumeur) ou sous-clonal (propre à un sous-clone) à cette mutation reste délicate, et nécessite des données d'excellente qualité ;
- (ii) l'utilisation de logiciels de prédiction (type netMHC) destinés à sélectionner des peptides possédant une bonne affinité pour le CMH. Les prédictions sont actuellement plus performantes pour les allèles *Human Leukocyte Antigen* (HLA) de classe I (davantage de données disponibles pour les épitopes T CD8, complexité des épitopes T CD4 liés aux variations de taille possible et la plus grande diversité des molécules HLA de classe II). Un *cut-off* de 500 nM est souvent utilisé pour une sélection basée sur l'affinité prédite. Certains auteurs sélectionnent des

peptides mutés avec un gain d'affinité par rapport au peptide natif correspondant. Le *cut-off* de 500 nM reste arbitraire et la sélection basée sur la forte affinité repose sur des données obtenues à partir des antigènes viraux [17]. Une autre méthode de sélection moins arbitraire consiste à classer l'affinité du peptide par rapport à celle d'une base de peptides naturels aléatoires [18]. Cette approche ne permet cependant pas d'identifier des néoépitopes présentant non pas un gain d'affinité pour le CMH mais pour le TCR par rapport au peptide natif. Il est également possible que la tolérance périphérique élimine certains lymphocytes T reconnaissant des néoépitopes de forte affinité [19] ;

- (iii) ajout possible de la prédiction du clivage par le protéasome, du transport par les TAP, ou de la vérification d'absence de réactivité croisée avec des protéines du soi [20,21] ;
- (iv) vérification de l'expression du transcrite correspondant par *Ribonucleic Acid Sequencing* (RNA-seq) ou autre technique (type *Complementary Deoxyribonucleic Acid* [cDNA] capture). Récemment une étape d'estimation de la capacité du néoépitope à être reconnu par des lymphocytes T a été proposée, grâce à un modèle basé sur la similarité de sa séquence avec celle d'antigènes connus [22]. Certains auteurs ont aussi ajouté une étape de validation par spectrométrie de masse des peptides réellement présentés par les cellules tumorales au niveau des molécules de CMH [15]. Il est également possible de vérifier la liaison des épitopes prédits aux molécules de CMH (test de *binding*) ou de tester leur immunogénicité par des tests de stimulation *in vitro*. Néanmoins, compte tenu du nombre d'épitopes sélectionnés et des délais nécessaires entre séquençage et production de vaccin, il reste difficile en pratique d'ajouter de lourdes étapes de validation (elles-mêmes non standardisées). Il persiste donc une problématique associée au fait qu'un nombre plus ou moins important d'épitopes prédits risque de ne pas être présenté de manière effective par les molécules de CMH. Un enjeu essentiel

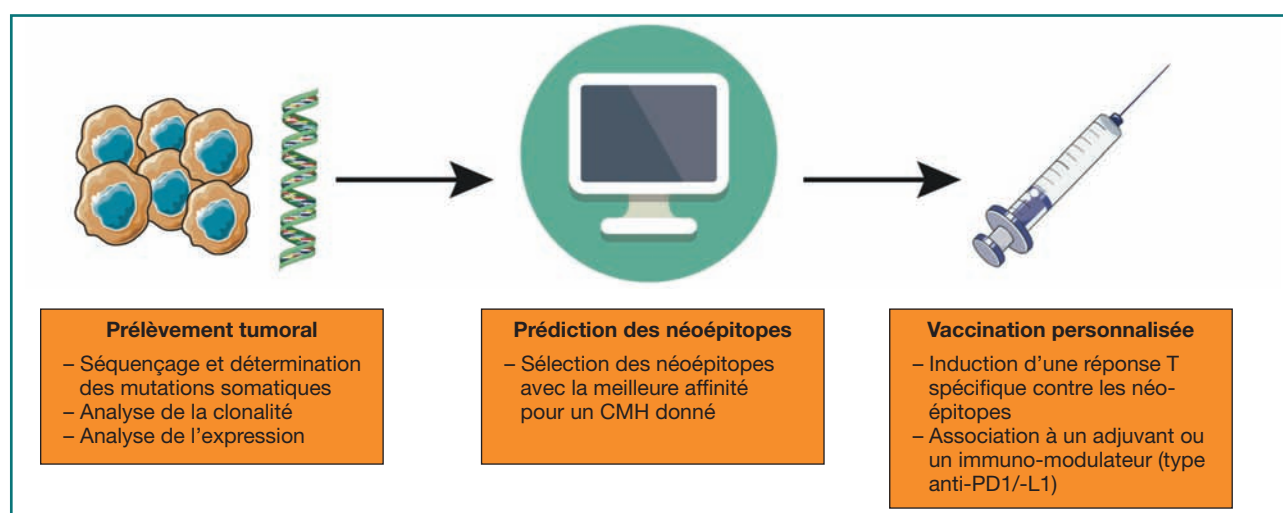


FIGURE 2  
Processus global de vaccination néoépitopique personnalisée

aujourd'hui est donc l'optimisation des outils de prédiction de néoépitopes, notamment pour les molécules HLA de classe II.

## Résultats cliniques obtenus avec les approches vaccinales néoépitopiques

Trois essais de vaccination utilisant des néoépitopes mutés ont été publiés. Une première preuve de concept a été réalisée chez 3 patients HLA-A\*02:01-positifs atteints de mélanome, avec utilisation de cellules dendritiques autologues pulsées avec des peptides [23]. La sélection des épitopes reposait sur la prédiction *in silico* avec validation *in vitro* de la liaison sur les molécules HLA-A2. Trois sites tumoraux ont été séquencés pour 2 patients et 1 seul pour le dernier, pour aboutir au final à 7 épitopes sélectionnés pour chaque patient (à partir de 219 à 598 mutations non synonymes identifiées). À noter 3 épitopes communs (épitopes clonaux) entre les 3 lésions pour le premier patient, et 2 pour le deuxième. Après 3 injections intraveineuses (IV), une réponse T CD8 a pu être mise en évidence après stimulation *ex vivo* pour 3 peptides sur 7 pour chaque patient, avec soit augmentation d'une réponse préexistante (pour 1 peptide), soit induction d'une réponse immune (pour 2 peptides). L'analyse de clonalité a montré une diversification du répertoire T contre les néoantigènes ciblés. Il n'y a pas eu de démonstration d'activité clinique associée dans cet essai.

Deux autres études ont été publiées récemment, toujours dans le mélanome avancé. P. A. Ott et al. ont vacciné 6 patients avec 7 administrations sous-cutanées de longs peptides synthétiques de 15 à 30 *mers* avec *Polyinosinic-Polycytidylic acid with Polylysine and Carboxymethylcellulose* (poly-ICLC) (Hiltonol) comme adjuvant. Pour chaque patient, entre 13 et 20 néoépitopes ont été sélectionnés à partir de la prédiction de liaison aux molécules HLA de classe I exprimées. La vaccination a induit une réponse T CD4 contre 60 % des néoantigènes, et une réponse T CD8 contre 16 % des néoantigènes cibles. Quatre patients sur 6 n'ont pas présenté de signes de récurrence avec un suivi allant jusqu'à 32 mois après vaccination. Deux patients présentant toujours des lésions tumorales après vaccination ont répondu à un traitement par anti-PD-1 [24].

Dans la troisième étude, 13 patients avec un mélanome avancé, dont 5 avec une maladie métastatique évaluable, ont été immunisés avec un vaccin ARN. Pour chaque patient, 10 épitopes mutés ont été sélectionnés à partir de la prédiction d'affinité pour les molécules HLA de classe I et II. Les ARN synthétiques vaccinaux codent chacun pour 5 peptides de 27 *mers* liés par un *linker*. Les patients ont reçu entre 8 et 20 injections de vaccin administré au niveau des ganglions inguinaux. La vaccination a induit une réponse T CD4 contre 66 % des néoantigènes, et une réponse T CD8 contre 25 % des néoantigènes cibles (analyses par ELISPOT, cytokines intracellulaires, multimères). En termes d'activité clinique, 2 patients parmi les 5 avec lésions métastatiques ont présenté une réponse objective. Une réponse complète a pu être obtenue chez un patient évolutif traité par la suite par anti-PD-1.

À noter une rechute avec mutation de la  $\beta$ 2-microglobuline au niveau de la tumeur (chez un patient ayant pourtant développé une réponse polyépitopique à la fois CD4 et CD8), suggérant que la perte de CMH de classe I pourrait représenter un mécanisme de résistance [25].

Les deux dernières études confirment la capacité des néoépitopes mutés à induire une réponse immune spécifique. Le profil de tolérance est très satisfaisant (absence d'auto-immunité) et il existe une impression d'activité clinique très encourageante, qui reste à confirmer compte tenu du faible nombre de patients et de l'absence de bras contrôle pour les critères de progression. On retrouve dans les deux cas une majorité de réponse T CD4. Le rôle des CD4 avec le vaccin ARN avait déjà été observé en modèle murin [13]. Il est probable que les régions sélectionnées contiennent à la fois des épitopes T CD4 et CD8 en raison de la présence de clusters épitopiques au sein d'une protéine. Le rôle des CD4 reste à préciser : rôle helper, activité antitumorale liée à la sécrétion d'IFN- $\gamma$  ? Il sera également utile de vérifier l'absence de réponse T CD4 régulatrice associée.

## Mécanismes de résistance et limitations possibles

Compte tenu de l'hétérogénéité tumorale, notamment à un stade avancé, il existe un risque élevé de sélectionner des épitopes présents au niveau de certains clones uniquement (épitopes dits « sous-clonaux »). La publication de B. M. Carreno va dans ce sens avec une majorité de néoépitopes non retrouvés au niveau des différents sites lésionnels étudiés [23]. Compte tenu de l'impact associé à la présence de néoépitopes sous-clonaux par rapport à des néoépitopes clonaux dans la réponse immune [26], il conviendra en théorie de favoriser les néoépitopes clonaux ou de privilégier un mélange de néoépitopes représentatifs des principaux sous-clones, ce qui ajoute une complexité à la méthode de sélection des néoépitopes candidats.

La plupart des mutations intervenant de manière aléatoire au niveau de gènes de type *passenger*, la fréquence de néoantigènes mutés partagés entre différentes tumeurs est très faible. Il a ainsi été retrouvé que, sur un total de 911 548 néoantigènes mutés, seulement 24 étaient partagés chez au moins 5 % des patients [11]. Une étude récente ne retrouve, parmi 3 760 néoantigènes prédits, que 0,42 % présents dans plus d'une tumeur [21]. Il convient donc de définir pour chaque patient un cocktail de néoépitopes adapté dans une démarche purement personnalisée. Il existe donc un délai lié à la prédiction des néoépitopes et la production du vaccin, ce qui limite son utilisation dans des formes avancées, métastatiques progressives. Une possibilité est de débiter avec un vaccin contenant des antigènes tumoraux partagés avant d'y associer le vaccin personnalisé [25].

Une limite possible de l'approche est la disponibilité du répertoire T disponible chez les patients. Il a ainsi pu être observé que le répertoire T de donneurs sains contenait des lymphocytes T



reconnaissant des néoantigènes mutés pour lesquels il n'était pas retrouvé de lymphocytes T intratumoraux spécifiques dans les tumeurs correspondantes [27]. Une approche complémentaire réside donc dans le transfert adoptif de cellules T génétiquement modifiées et transduites avec un TCR spécifique de néoantigène [28].

Un mécanisme de résistance déjà identifié dans l'étude de U. Sahin et al. est la perte de CMH par exemple *via* la perte de  $\beta$ 2-microglobuline (ce qui induit l'absence d'expression de molécules HLA de classe I à la surface de la cellule tumorale) [25]. Ce mécanisme a également été identifié dans la résistance acquise aux anti-PD-1/PD-L1 [29]. De manière plus subtile, la perte d'hétérozygotie des allèles HLA correspondant aux CMH présentant les néoépitopes mutés pourrait représenter un mécanisme de résistance initiale ou acquise à la vaccination basée sur des néoépitopes mutés. Il a ainsi été montré une perte d'hétérozygotie des allèles HLA dans 40 % des cas de cancer du poumon non à petites cellules, associée à une forte quantité de néoantigènes sous-clonaux [30]. Il conviendra donc de vérifier la qualité de l'expression du CMH, et même idéalement d'adapter le choix des néoépitopes sélectionnés en fonction de l'expression du CMH correspondant, dans ce type d'approche thérapeutique. Enfin, l'immunosuppression associée à la tumeur avancée représente un mécanisme de résistance général aux approches d'immunothérapie. Il existe dans ce contexte un rationnel solide à combiner d'emblée les approches vaccinales à un anti-PD-1/PD-L1, de manière à inhiber la résistance induite par la réponse IFN- $\gamma$  [31], voire favoriser une réponse immune *de novo* et l'activation T initiale [32].

## Perspectives

### Positionnement clinique

La plupart des essais sont réalisés dans un contexte avancé, et doivent faire face aux problématiques évoquées ci-dessus : forte hétérogénéité tumorale, immunosuppression, perte de HLA, etc. Le délai avant administration, lié à la détermination des néoépitopes et la phase de production, nécessite soit d'utiliser un antigène partagé au début [25], soit de débiter un traitement préalable, par exemple par anti-PD-1/PD-L1, et de rajouter la vaccination ultérieurement (exemple : étude NCT02897765, qui ajoute un vaccin néoépitopique personnalisé après 3 mois de traitement par nivolumab dans les cancers du poumon, vessie et mélanome). À noter que l'avantage dans ce contexte de tumeur avancée est la possibilité d'utiliser des biopsies tumorales séquentielles pour évaluer les changements en termes d'infiltrat immunitaire suite à la vaccination, en plus de la mise en évidence de l'induction d'une réponse immunitaire spécifique détectée de manière classique au niveau sanguin.

Certains essais se positionnent en situation néoadjuvante ou adjuvante. Le cancer du sein de type triple négatif, en non-réponse pathologique complète après chimiothérapie néoadjuvante, se prête bien à ce type d'approche par exemple. Le risque de rechute est en effet important et le traitement vaccinal est mis en place dans un contexte de maladie résiduelle après traitement conventionnel combinant chimiothérapie néoadjuvante, chirurgie et radiothérapie. Deux essais de phase I sont en cours dans ce contexte :

TABLEAU I  
Principaux essais cliniques en cours

N° d'étude	Promoteur	Cancer	Type de vaccin	Positionnement clinique
NCT02897765	Neon Therapeutics	Poumon, vessie et mélanome	Longs peptides	Avancé/métastatique
NCT02035956	BioNTech RNA Pharmaceuticals	Mélanome	ARN	Avancé/métastatique
NCT02316457	Biontech AG	Sein triple négatif	ARN	Adjuvant
NCT01970358	Patrick Ott, MD	Mélanome	Longs peptides	Adjuvant
NCT02427581	Washington University School of Medicine	Sein triple négatif	Longs peptides	Adjuvant
NCT02348320		Sein triple négatif	ADN	Adjuvant
NCT03122106		Pancréas	ADN	Après chimiothérapie adjuvante
NCT02600949	MD Anderson Cancer Center	Pancréas, colorectal	Longs peptides	Avancé/métastatique
NCT02721043	Nina Bhardwaj	Divers	Longs peptides	Rémission complète à risque de rechute
NCT02992977	Agenus Inc.	Divers	Longs peptides	Avancé/métastatique,
NCT02933073	UConn Health	Ovaire	Longs peptides	Adjuvant
NCT03068832	Washington University School of Medicine	Cérébral (pédiatrie)	Longs peptides	Rechute/réfractaire
NCT02950766	Patrick Ott, MD	Rénal	Longs peptides	1 <sup>re</sup> ligne – Vaccination adjuvante avec ipilimumab concomitant



NCT02427581 utilisant de longs peptides synthétiques et NCT02348320 utilisant un vaccin ADN. Un essai de vaccination incluant néoantigènes et épitopes dérivés de la mésothéline vient également de débiter dans le cancer du pancréas après chirurgie et chimiothérapie adjuvante (NCT03122106). L'étude NCT02933073 prévoit quant à elle l'ajout d'un schéma de vaccination personnalisée chez les patientes présentant un cancer de l'ovaire de stade III/IV en rémission après traitement standard par chirurgie et chimiothérapie. Il conviendra cependant de mettre en place dans le contexte adjuvant des groupes contrôles pour démontrer un gain de survie associé. L'absence de lésion tumorale biopsiable limite les analyses immunologiques au seul niveau sanguin. Néanmoins, il est probable que le contexte adjuvant dans les pathologies à fort risque de rechute représente le positionnement de choix pour les approches vaccinales, utilisées seules ou en combinaison avec un anti-PD-1/PD-L1. Les principaux essais en cours sont présentés dans le *tableau 1*.

### Autres antigènes tumoraux

Pour le moment les études se sont focalisées sur les épitopes mutés issus de variants nucléotidiques (*Single Nucleotide Variants* [SNV]) mais d'autres néoantigènes pourraient s'avérer tout à fait intéressants, tels ceux liés à des anomalies d'épissage, propres à la tumeur, ou associés aux insertions-délétions (*indels*). Les *indels* pourraient être une source importante de néoantigènes et une étude récente de l'équipe de C. Swanton suggère une fréquence plus élevée de peptides possédant une forte affinité pour le CMH pour les *indels* comparativement aux SNV. De plus, leur probabilité de présence au niveau de gènes suppresseurs de tumeurs pourrait également être associée à une plus forte probabilité d'obtention d'antigènes partagés [33]. Enfin, la combinaison avec les anti-PD-1/PD-L1 remet également en avant la possibilité d'utiliser aussi certains antigènes tumoraux de type *cancer germline* [34].

### Conclusion

La vaccination à base de néoantigènes spécifiques dans une approche de type personnalisée ouvre de nouvelles perspectives thérapeutiques en cancérologie. Dans les formes avancées métastatiques, la vaccination pourrait soit renforcer la réponse aux anti-PD-1/PD-L1, en augmentant la quantité d'effecteurs au sein de la tumeur, soit permettre d'induire une réponse T antitumorale dans les tumeurs immunologiquement « froides » – caractérisées par l'absence d'infiltrat T<sup>+</sup>, condition préalable à l'activité d'un anti-PD-1/PD-L1. L'une des limites reste dans ce contexte l'absence d'expression de certains HLA par les cellules tumorales (avec dans ce cas un intérêt plus marqué par exemple pour des approches « supra-physiologiques » de type anticorps bispécifiques T cell recruiting ou l'utilisation de *Chimeric Antigen Receptor T-cells* (CAR T-cell), qui permettent une activation T non CMH-dépendante). Néanmoins, même en l'absence de molécules HLA de classe I sur la cellule tumorale, il est possible que les lymphocytes T CD4 induits par la vaccination puissent favoriser une réponse antitumorale efficace de par la sécrétion d'IFN- $\gamma$  et la stimulation d'autres cellules de l'environnement immun. Cet aspect reste assez mal connu et mérite d'être précisé compte tenu de l'importance de la réponse T CD4 générée lors des premiers essais de vaccination néoépitopique. Enfin, le traitement adjuvant représente un positionnement privilégié pour la vaccination, de manière à permettre un contrôle au long cours de la maladie résiduelle suite à la génération d'une réponse immune mémoire.

**Déclaration de liens d'intérêts** : Stéphane Depil est employé de Collectis et a réalisé des activités de consultation ou des prestations pour : AstraZeneca, BMS, Servier, Eisai, Erytech, PDC<sup>®</sup>line Pharma, EpiVax, Netris Pharma, Celyad. Paola Bonaventura, Laurie Tonon : aucun.

Vincent Alcazer : participation congrès Amgen, Novartis ; actions de formation : Gilead, Novartis, Amgen.

Cet article fait partie du numéro supplément *Nouvelles perspectives dans l'immunothérapie des cancers* réalisé avec le soutien institutionnel des partenaires AstraZeneca, Bristol-Myers Squibb, Roche.

### Références

- [1] Rosenberg SA, Yang JC, Restifo NP. Cancer immunotherapy: moving beyond current vaccines. *Nat Med* 2004;10:909-15.
- [2] Romero P, Banchereau J, Bhardwaj N, Cockett M, Disis ML, Dranoff G, et al. The Human Vaccines Project: A roadmap for cancer vaccine development. *Sci Transl Med* 2016;8:334ps9.
- [3] Graff JN, Chamberlain ED. Sipuleucel-T in the treatment of prostate cancer: an evidence-based review of its place in therapy. *Core Evid* 2014;10:1-10.
- [4] Schwartzenuber DJ, Lawson DH, Richards JM, Conry RM, Miller DM, Treisman J, et al. gp100 Peptide Vaccine and Interleukin-2 in Patients with Advanced Melanoma. *N Engl J Med* 2011;364:2119-27.
- [5] Vansteenkiste JF, Cho BC, Vanakesa T, De Pas T, Zielinski M, Kim MS, et al. Efficacy of the MAGE-A3 cancer immunotherapeutic as adjuvant therapy in patients with resected MAGE-A3-positive non-small-cell lung cancer (MAGRIT): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet Oncol* 2016;17:822-35.
- [6] Melief CJM, van Hall T, Arens R, Ossendorp F, van der Burg SH. Therapeutic cancer vaccines. *J Clin Invest* 2015;125:3401-12.
- [7] Chan TA, Wolchok JD, Snyder A. Genetic Basis for Clinical Response to CTLA-4 Blockade in Melanoma. *N Engl J Med* 2015;373:1984.
- [8] Rizvi NA, Mazières J, Planchard D, Stinchcombe TE, Dy GK, Antonia SJ, et al. Activity and safety of nivolumab, an anti-PD-1 immune checkpoint inhibitor, for patients with advanced, refractory squamous non-small-cell lung cancer (CheckMate 063): a phase 2, single-arm trial. *Lancet Oncol* 2015;16:257-65.
- [9] Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, Eyring AD, et al. PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *N Engl J Med* 2015;372:2509-20.
- [10] Yarchoan M, Hopkins A, Jaffee EM. Tumor Mutational Burden and Response Rate to PD-1 Inhibition. *N Engl J Med* 2017;377:2500-1.
- [11] Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, et al. Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. *Cell Rep* 2017;18:248-62.

- [12] Castle JC, Kreiter S, Diekmann J, Löwer M, van de Roemer N, de Graaf J, et al. Exploiting the mutanome for tumor vaccination. *Cancer Res* 2012;72:1081-91.
- [13] Kreiter S, Vormehr M, van de Roemer N, Diken M, Löwer M, Diekmann J, et al. Erratum: Mutant MHC class II epitopes drive therapeutic immune responses to cancer. *Nature* 2015;523:370.
- [14] Gubin MM, Zhang X, Schuster H, Caron E, Ward JP, Noguichi T, et al. Checkpoint blockade cancer immunotherapy targets tumour-specific mutant antigens. *Nature* 2014;515:577-81.
- [15] Yadav M, Jhunjhunwala S, Phung QT, Lupardus P, Tanguay J, Bumbaca S, et al. Predicting immunogenic tumour mutations by combining mass spectrometry and exome sequencing. *Nature* 2014;515:572-6.
- [16] Zhang X, Kim S, Hundal J, Herndon JM, Li S, Petti AA, et al. Breast Cancer Neoantigens Can Induce CD8+T-Cell Responses and Antitumor Immunity. *Cancer Immunol Res* 2017;5:516-23.
- [17] Srivastava PK. Neoepitopes of Cancers: Looking Back, Looking Ahead. *Cancer Immunol Res* 2015;3:969-77.
- [18] Andreatta M, Nielsen M. Gapped sequence alignment using artificial neural networks: application to the MHC class I system. *Bioinforma Oxf Engl* 2016;32:511-7.
- [19] Sette A, Vitiello A, Reheman B, Fowler P, Nayarsina R, Kast WM, et al. The relationship between class I binding affinity and immunogenicity of potential cytotoxic T cell epitopes. *J Immunol* 1994;153:5586-92.
- [20] Calis JJA, Maybeno M, Greenbaum JA, Weiskopf D, De Silva AD, Sette A, et al. Properties of MHC class I presented peptides that enhance immunogenicity. *PLoS Comput Biol* 2013;9:e1003266.
- [21] Kim S, Kim HS, Kim E, Lee MG, Shin E, Paik S, et al. Neopepsee: accurate genome-level prediction of neoantigens by harnessing sequence and amino acid immunogenicity information. *Ann Oncol* 2008;29:1030-6.
- [22] Łuksza M, Riaz N, Makarov V, Balachandran VP, Hellmann MD, Solovyyov A, et al. A neoantigen fitness model predicts tumour response to checkpoint blockade immunotherapy. *Nature* 2017;551:517-20.
- [23] Carrero BM, Magrini V, Becker-Hapak M, Kaabinejadian S, Hundal J, Petti AA, et al. Cancer immunotherapy. A dendritic cell vaccine increases the breadth and diversity of melanoma neoantigen-specific T cells. *Science* 2015;348:803-8.
- [24] Ott PA, Hu Z, Keskin DB, Shukla SA, Sun J, Bozym DJ, et al. An immunogenic personal neoantigen vaccine for patients with melanoma. *Nature* 2017;547:217-21.
- [25] Sahin U, Derhovanessian E, Miller M, Kloke BP, Simon P, Löwer M, et al. Personalized RNA mutanome vaccines mobilize poly-specific therapeutic immunity against cancer. *Nature* 2017;547:222-6.
- [26] McGranahan N, Furness AJS, Rosenthal R, Ramskov S, Lyngaa R, Saini SK, et al. Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. *Science* 2016;351:1463-9.
- [27] Strønen E, Toebe M, Kelderman S, van Buuren MM, Yang W, van Rooij N, et al. Targeting of cancer neoantigens with donor-derived T cell receptor repertoires. *Science* 2016;352:1337-41.
- [28] Leisegang M, Engels B, Schreiber K, Yew PY, Kiyotani K, Idel C, et al. Eradication of Large Solid Tumors by Gene Therapy with a T-Cell Receptor Targeting a Single Cancer-Specific Point Mutation. *Clin Cancer Res* 2016;22:2734-43.
- [29] Zaretsky JM, Garcia-Diaz A, Shin DS, Escuin-Ordinas H, Hugo W, Hu-Lieskovan S, et al. Mutations Associated with Acquired Resistance to PD-1 Blockade in Melanoma. *N Engl J Med* 2016;375:819-29.
- [30] McGranahan N, Rosenthal R, Hiley CT, Rowan AJ, Watkins TBK, Wilson GA, et al. Allele-Specific HLA Loss and Immune Escape in Lung Cancer Evolution. *Cell* 2017;171:1259-71.e11.
- [31] Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. *Nat Rev Cancer* 2012;12:252-64.
- [32] Hui E, Cheung J, Zhu J, Su X, Taylor MJ, Wallweber HA, et al. T cell costimulatory receptor CD28 is a primary target for PD-1-mediated inhibition. *Science* 2017;355:1428-33.
- [33] Turajlic S, Litchfield K, Xu H, Rosenthal R, McGranahan N, Reading JL, et al. Insertion-and-deletion-derived tumour-specific neoantigens and the immunogenic phenotype: a pan-cancer analysis. *Lancet Oncol* 2017;18:1009-21.
- [34] Finn OJ, Rammensee HG. Is It Possible to Develop Cancer Vaccines to Neoantigens, What Are the Major Challenges, and How Can These Be Overcome? Neoantigens: Nothing New in Spite of the Name. *Cold Spring Harb Perspect Biol* 2017.



### REVUE 3 - Neoepitopes-based vaccines: challenges and perspectives

***Vincent Alcazer**<sup>1,2,3</sup>, Paola Bonaventura<sup>3,4</sup>, Laurie Tonon<sup>5</sup>, Sandrine Wittmann<sup>3,4</sup>, Christophe Caux<sup>3,4</sup>, Stéphane Depil<sup>2,3,4\*</sup>*

1 Hospices Civils de Lyon; Service d'hématologie Clinique, Centre Hospitalier Lyon Sud, Pierre-Bénite, France

2 Université Claude Bernard Lyon 1, Lyon, France

3 Inserm U1052/CNRS 5286, Centre de Recherche en Cancérologie de Lyon, Lyon, France

4 Centre Léon Bérard, Lyon, France

5 Synergie Lyon Cancer, Centre Léon Bérard, Lyon, France

\* Corresponding author: Stéphane Depil, Centre Léon Bérard, 28 Prom. Léa et Napoléon Bullukian, 69008 Lyon, France;

Tel: +33 (0) 4 78 78 28 28; Mail: [stephane.depil@lyon.unicancer.fr](mailto:stephane.depil@lyon.unicancer.fr)

Eur J Cancer. 2019 Feb;108:55-60. doi: 10.1016/j.ejca.2018.12.011. Epub 2019 Jan 12. PMID: 30648630.



Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

ScienceDirect

journal homepage: [www.ejcancer.com](http://www.ejcancer.com)



Current Perspective

## Neoepitopes-based vaccines: challenges and perspectives



Vincent Alcazer<sup>a,b,c</sup>, Paola Bonaventura<sup>c,d</sup>, Laurie Tonon<sup>e</sup>,  
Sandrine Wittmann<sup>c,d</sup>, Christophe Caux<sup>c,d</sup>, Stéphane Depil<sup>b,c,d,\*</sup>

<sup>a</sup> Hospices Civils de Lyon, Service D'hématologie Clinique, Centre Hospitalier Lyon Sud, Pierre-Bénite, France

<sup>b</sup> Université Claude Bernard Lyon 1, Lyon, France

<sup>c</sup> Inserm U1052/CNRS 5286, Centre de Recherche en Cancérologie de Lyon, Lyon, France

<sup>d</sup> Centre Léon Bérard, Lyon, France

<sup>e</sup> Synergie Lyon Cancer, Centre Léon Bérard, Lyon, France

Received 20 September 2018; received in revised form 27 November 2018; accepted 5 December 2018

### KEYWORDS

Personalised vaccine;  
Cancer;  
Neoantigens;  
Neoepitopes

**Abstract** First generations of cancer vaccines using shared tumour antigens have been associated with disappointing clinical results. However, the paradigm shift introduced by immune checkpoint inhibitors has led to a renewed interest on anti-tumoural vaccination based on mutation-associated neoantigens. First clinical results are encouraging with some signs of clinical activity associated with induction of a specific immune response. In advanced or metastatic diseases, vaccination may either enhance the response to Programmed cell death 1 (PD-1/L1) antagonists by increasing the number of effectors within the tumour or induce an anti-tumoural T-cell response in immunologically 'cold' tumours. There is also a strong rationale to use cancer vaccines in an adjuvant setting to induce a long-term control of the residual disease. Prediction of neoepitopes efficiently presented by Human Leukocyte Antigen (HLA) molecules remains a challenge, as well as identification of clonal neoantigens. Some mechanisms of resistance are already identified, such as tumour loss of neoepitopes-presenting HLA class I molecules. In this context, the role of CD4<sup>+</sup> T cells induced by different cancer vaccines should be clarified. Finally, although studies have focused on mutated epitopes corresponding to single nucleotide variants, other neoantigens could be of strong interest such as those linked to tumour specific RNA-splicing abnormalities or associated with insertions-deletions.

© 2019 Elsevier Ltd. All rights reserved.

\* Corresponding author: Centre Léon Bérard, 28 Prom. Léa et Napoléon Bullukian, 69008 Lyon, France.  
E-mail address: [stephane.depil@lyon.unicancer.fr](mailto:stephane.depil@lyon.unicancer.fr) (S. Depil).

## 1. Towards personalised vaccines

First generation cancer vaccines have shown relatively disappointing clinical results, with less than 7% objective clinical responses and an overall clinical benefit rate estimated around 20% [1]. The use of non-tumour-specific antigens, which elicit low affinity T-cell response due to elimination of the most reactive T cells by central tolerance, is one of the multiple factors that explain these results. Tumour-specific antigens (neoantigens), derived from tumour gene mutations or rearrangements, have been reported to induce stronger immune responses in the absence of central tolerance [2]. The abundant literature around mutation-associated neoantigens (MANA) together with the paradigm shift introduced by immune checkpoint inhibitors (ICIs) led to a renewed interest in the field of cancer vaccines. Indeed, multiple studies showed a link between clinical responses to ICIs and tumour mutational burden in melanoma [3], lung cancer [4] or colorectal cancer [5]. Furthermore, a linear relation was observed between anti-PD1/-L1 objective response rate and the median number of coding somatic mutations per megabase of DNA [6]. Because mutational burden is a reflection of the quantity of MANA [7], these latter may represent the main targets of activated T cells after ICI therapy.

## 2. Neoepitope-based vaccines: current challenges

### 2.1. Neoepitope selection

The identification of neoepitopes relies on different steps. First, a precise detection of tumour somatic mutations is performed by combined whole exome sequencing of tumour and normal cells [8]. A matched RNA expression analysis (such as RNA sequencing) of tumour cells is added to select expressed mutations. Prediction softwares are then used to identify potential neoepitopes with a high affinity for the individual's major histocompatibility complex molecules (MHCs) [9]. However, classical algorithms have a limited predictive value as they do not take into account the different steps of epitope processing. Moreover, the affinity threshold needed to elicit a cytotoxic T-cell response has been mostly validated on viral epitopes and may not be adapted for mutated self-peptides [10]. In a set of 448 potential CD8-T-cell epitopes identified in a patient with melanoma, van Rooij *et al.* [11] found that less than 1% were effectively recognised by the patient's T cells. The positive predictive value of such tools is, therefore, pretty low and different additional steps (such as proteasome cleavage/TAP transport predictions or self-proteins cross-reactivity verification) can be added to the prediction [12,13]. Recently, a quality-control step considering the probability of recognition of a neoepitope by a T-cell receptor has been proposed, using a

model based on sequence similarity with that of known antigens [14]. It is also possible to verify predicted epitopes' MHC affinity by binding assay or to test their immunogenicity by *in vitro* stimulation assays.

The fact that a majority of predicted class I epitopes may not be effectively presented by MHC class I molecules on tumour cells remains a major concern. A validation step has been added by some authors, using mass spectrometry for identification of peptides eluted from tumour MHC class I molecules [15]. The sensitivity of such approach is, however, limited to neoepitopes with relatively high expression, preventing the detection of less abundant but still immunogenic peptides [16]. The optimisation of MHC class II neoepitopes prediction is another challenge (MHC class II epitopes harbour a high diversity with variables sizes rendering prediction more complicated than for MHC class I epitopes), considering the potential role of CD4+ T-cell response after neoepitope-based vaccine [17,18].

### 2.2. Formulation

Different formulations or antigen sources currently exist. DNA vaccines consist in transfecting a DNA sequence encoding for the neoepitopes of interest. Although unmethylated CG-rich DNA was supposed to provide an immune adjuvant by Toll-like receptor (TLR) stimulation, poor immunogenicity and clinical activity were observed [19]. In the opposite, RNA vaccines harbour a convincing profile with a potential adjuvant effect (by TLR 7/8 stimulation) and multiple epitope encoding possibility [20]. However, this TLR stimulation precludes the combination with other adjuvants. Conversely, peptide-based vaccine can be combined with any adjuvant. Of note, the diversity of possible adjuvants has already been reviewed elsewhere and is beyond the scope of this article [2,21]. As single peptide formulation is associated with a higher risk of immune escape, multi-peptides approaches are usually developed [22]. With their potential to stimulate both CD4+ and CD8+ T-cell response, long peptides (20–30 mer) are indeed being more and more considered [23]. Other formulations such as viral vectors, *ex vivo*-generated dendritic cells or whole cell have also been developed. The impressive results obtained in infectious diseases with viral vectors, eliciting a strong cytotoxic T-cell response in terms of quality and number, merit a particular attention [24]. Considering the importance of T-cell numbers in clinical responses after adoptive T-cell therapies, there is no doubt that vectors inducing a higher number of effector T cells compared to peptides or mRNA approaches would be of a great interest.

### 2.3. Tumour heterogeneity

Given tumour heterogeneity, especially at an advanced stage, there is a high risk of selecting epitopes present in



certain tumour clones only ('subclonal' epitopes). In Carreno's study, a majority of the identified neoepitopes were not found at the different tumour sites [25]. Assigning a clonal (shared by all the tumour clones) or subclonal (specific to a subclone) status to a mutation remains difficult and requires data of excellent quality. Moreover, as they do not confer any survival advantages, passenger mutations could be lost without any counterpart for tumour cells.

As most mutations occur randomly in passenger-type genes, the frequency of mutated neoantigens shared between different tumours is very low. It was found that out of a total of 911,548 mutated neoantigens, only 24 were shared in at least 5% of patients [7]. A recent study found that, among 3760 predicted neoantigens, only 0.42% were found in more than one tumour [13]. It is, therefore, necessary to define for each patient a panel of neoepitopes in a purely personalised approach. Considering the impact of the presence of clonal versus subclonal neoepitopes in the quality of the anti-tumour immune response [26], it will theoretically be appropriate to favour clonal neoepitopes or a mixture of neoepitopes representative of the main subclones, prioritising when possible neoepitopes derived from driver mutations, which adds complexity to the neoepitopes selection method. The delay related to this complex process of production may be an issue for the use in advanced progressive diseases. One possibility is to start with a vaccine containing shared tumour antigens before combining the personalised vaccine [27].

#### 2.4. Mechanisms of resistance

A possible limitation of this approach is the availability of the patient's T repertoire. It was observed that the T-cell repertoire of healthy donors contained T cells recognising MANA for which specific intra-tumour T lymphocytes were not found in the corresponding tumours [28].

A resistance mechanism already identified in Sahin's study is the  $\beta$ 2-microglobulin loss leading to the absence of MHC class I molecules on tumour cells [27]. This mechanism has also been identified in acquired resistance to PD1/L1 antagonists [29]. More subtly, the loss of heterozygosity (LOH) of HLA alleles corresponding to MHCs presenting mutated neoepitopes could represent an initial or acquired resistance mechanism. HLA LOH has been shown in 40% of non-small cell lung cancers and is associated with a large amount of subclonal neoantigens [30]. It will, therefore, be necessary to verify the quality of MHC expression and ideally to select neoepitopes in accordance with the corresponding MHC expression. Optimisation of high affinity MHC class II neoepitopes prediction will also help mounting a diversified response by adding CD4+ T-cell anti-tumoural effects.

Finally, advanced tumour-associated immunosuppression represents a general resistance mechanism to

immunotherapy. In this context, there is a strong rationale for combining vaccine approaches with anti-PD-1/L1 from the outset, so as to inhibit the resistance induced by IFN- $\gamma$  response [31] or even favour a *de novo* immune response and initial T-cell activation [32].

### 3. Perspectives

#### 3.1. Addition of other tumour antigens

Although ongoing studies have focused on mutated epitopes corresponding to single nucleotide variants (SNVs), other neoantigens could be of great interest such as those derived from tumour specific RNA splicing abnormalities [33] or from insertions-deletions (indels). Indels could be a major source of neoantigens, and a recent study suggests a higher frequency of epitopes with high MHC affinity for epitopes derived from indels compared with SNVs. In addition, the localisation of indels in tumour suppressor genes could also be associated with a higher probability of obtaining shared antigens [34]. The combination with PD-1/L1 antagonists also highlights the possibility of using some cancer germline antigens [35].

#### 3.2. Improving T-cell response

Targeting multiple neoepitopes, together with stimulating a CD4+ T-cell response, is a promising way to induce an efficient immune response against the tumour. Sahin *et al* [27] reported the feasibility and immunogenicity of an mRNA-based vaccine targeting multiple selected neoepitopes in melanoma. Responses were detected against 60% of the predicted neoepitopes, with 57% of isolated CD4+, 17% of CD8+ and 26% of combined CD4+ and CD8+ responses. Ott *et al*. [36] also demonstrated the immunogenicity and feasibility of a vaccine that targets up to 20 predicted neoantigens in patients with melanoma. Using long synthetic 15–30 mer peptides with polyinosinic:polycytidylic acid and poly-L-lysine double-stranded RNA (poly-ICLC) (Hiltonol) as an adjuvant injected subcutaneously, the authors reported both CD4+ and CD8+ T-cell responses, respectively targeting 60% and 16% of the vaccine neoantigens. These results underline the importance of the CD4+ T cells in supporting CD8+ response and providing additional anti-tumoural effects. Of note, CD8+ T-cell response was evaluated after 2 weeks of *in vitro* restimulation in Ott's study. Thus, it is not possible to conclude that the vaccine really induced a specific response in patients because an *in vitro* priming against the tested antigens cannot be formally excluded. Still, four of six vaccinated patients had no recurrence 25 months after vaccination. The two others with recurrent diseases were successfully treated with anti-PD-1 therapy inducing a complete tumour regression.



Combination of cancer vaccine with ICI can also help improving T-cell response by promoting T-cell activation and epitope spreading [37,38]. In Sahin's study, the only complete response was observed in combination with PD1 blockade [27]. Several studies are currently underway based on this scheme; for instance, NCT02897765 study adds a personalised vaccine after 3 months of nivolumab in different types of cancers. The advantage in this context of advanced disease is the possibility of performing sequential tumour biopsies to evaluate changes in the immune infiltrate after vaccination, in addition to demonstrate the induction of a specific immune response detected in blood. Furthermore, neo-epitope-based vaccines could be an interesting approach in combination with ICIs for diseases with low or no response to ICIs alone despite a significant mutational burden. This is the case for microsatellite stable colorectal cancer, where response to immunotherapeutic approaches (eg T-cell recruiting bispecific antibody with atezolizumab) have shown clinical responses in favour of a functional immune system [39]. This may represent an interesting proof of concept of efficacy.

Depleting the immunosuppressive milieu (by using cyclophosphamide for regulatory T-cell depletion or gemcitabine for myeloid-derived suppressor cells) or promoting T-cell expansion by the use of cytokines such as IL-7 are other ways for improving T-cell response [40,41].

Finally, combination of different immunotherapies will probably be required to eradicate advanced tumours. Moynihan *et al.* evaluated in a murine syngeneic tumour model, a combination including a tumour-targeting antibody, recombinant IL-2, anti-PD-1 antibody and specific vaccine [42]. This combination induced tumour infiltration of both innate and adaptive immune cells, mediating a strong anti-tumoural effect with eradication of large established tumours. Although the tolerability of such combination must be confirmed in humans, it makes no doubt that enhancing the vaccinal response by different simultaneous ways is a promising approach.

### 3.3. Optimisation of the timing of vaccination

Most of studies have used cancer vaccines in an advanced or metastatic setting. Considering the risk of higher immunosuppression and of higher resistance associated to clonal heterogeneity in case of high tumour burden, it has been suggested that cancer vaccination should be used in an adjuvant setting in a context of low residual disease [1]. Vaccination could be also used in a neoadjuvant setting. In this context, it has been shown in murine models that T-cell stimulation by ICI was more efficient in preventing metastasis when applied before tumour resection [43]. In fact, more and more studies are currently being performed in an adjuvant situation (Table 1). Triple negative breast cancer, in

Table 1  
Ongoing recruiting studies using neopeptides for personalised therapeutic vaccine.

Study number	Promotor	Cancer	Vaccine type	Position
NCT02301611	Cancer Insight, LLC	Melanoma	DC	Adjuvant
NCT02348320	Washington University School of Medicine	Breast (TN)	DNA	Adjuvant
NCT02600949	M.D. Anderson Cancer Center	Pancreatic, colorectal	Long peptide	Advanced/metastatic
NCT02721043	Nina Bhardwaj, Icahn School of Medicine at Mount Sinai	Solid tumours	Long peptide	Adjuvant
NCT02808364	Guangdong 999 Brain Hospital	Glioblastoma	DC	Adjuvant
NCT02808416	Guangdong 999 Brain Hospital	Solid tumours with brain metastases	DC	Adjuvant
NCT02933073	UConn Health	Ovarian	Long peptide	Neoadjuvant or adjuvant
NCT02956551	Sichuan University	Carcinoma, non-small cell lung	DC	Advanced/metastatic
NCT02933073	UConn Health	Ovarian	Long peptide	Adjuvant
NCT03122106	Washington University School of Medicine	Pancreatic	DNA	Adjuvant
NCT03289962	Genentech, Inc.	Melanoma, non-small cell lung, bladder, colorectal, breast (TN), renal, head and neck	RNA	Advanced/metastatic
NCT03480152	National Cancer Institute (NCI)	Melanoma, colon, gastrointestinal, genitourinary, hepatocellular	RNA	Advanced/metastatic
NCT03552718	NantBioScience, Inc.	Colorectal, breast, head and neck, melanoma, non-small cell lung, pancreatic, liver	DC	Adjuvant
NCT03633110	Genocea Biosciences, Inc.	Melanoma, non-small cell lung, head and neck, urothelial carcinoma, renal cell carcinoma	Long peptide	Adjuvant
NCT03558945	Changhai Hospital	Pancreatic	NA	Adjuvant
NCT03645148	Zhejiang Provincial People's Hospital and Hangzhou Neoantigen Therapeutics Co., Ltd	Pancreatic	Long peptide	Advanced/metastatic

DC: dendritic cells, NA: not available, TN: triple negative

non-complete pathological response following neoadjuvant chemotherapy, is well suited to this type of approach. The risk of relapse is high, and the vaccine is set up in a context of low tumoural mass after conventional treatment combining neoadjuvant chemotherapy, surgery and radiotherapy. Several phase I trials are underway in this context, such as NCT02348320 study. It will be necessary to set up control groups in this setting to demonstrate a benefit in terms of survival. Furthermore, the absence of biopsy-accessible tumours restrains the immunological analyses to the blood level. Nevertheless, it is likely that the adjuvant setting in pathologies at high risk of relapse represents a positioning of choice for vaccine approaches, used alone or in combination with an anti-PD-1/-L1.

#### 4. Conclusion

Vaccination based on specific neoantigens in a personalised approach opens new therapeutic perspectives in oncology. Recent studies have indeed provided a clinical proof of concept in melanoma [25,36]. Immunogenicity is, however, too often considered as a surrogate marker of efficacy, and objective responses remain low. In advanced metastatic diseases, vaccination may either enhance the response to PD-1/-L1 antagonists by increasing the number of effectors within the tumour or induce an anti-tumoural T response in immunologically ‘cold’ tumours (characterised by the absence of T infiltration), a prerequisite for the activity of an anti-PD-1/-L1. The lack of HLA class I expression by tumour cells represent an important limitation of this approach. Nevertheless, CD4<sup>+</sup> T cells induced by vaccination may promote an effective anti-tumoural response by interferon- $\gamma$  secretion and stimulation of other cells from the immune environment. This aspect is barely known and needs to be clarified given the importance of the CD4<sup>+</sup> T-cell response generated during the first neoepitope-based vaccination trials [17]. Finally, adjuvant therapy represents a privileged positioning for vaccination, so as to allow long-term control of residual disease following the generation of an immune response.

#### Conflict of interest statement

S.D. is an employee for Collectis and reports personal fees from AstraZeneca, Elsalys, Erytech Pharma and Netris Pharma. The other authors declare no potential conflict of interest.

#### Acknowledgements

This work was supported by the ‘Institut National du Cancer INCa’ (PRT-K2017-072), the ‘Ligue contre le Cancer’, the SIRC project (LYRICAN, INCa-DGOS-Inserm\_12563) and the LABEX DEVweCAN (ANR-10-LABX-0061) of the University of Lyon.

#### References

- [1] Romero P, Banchereau J, Bhardwaj N, Cockett M, Disis ML, Dranoff G, et al. The human vaccines project: a roadmap for cancer vaccine development. *Sci Transl Med* 2016;8. <https://doi.org/10.1126/scitranslmed.aaf0685>. 334ps9-334ps9.
- [2] Hu Z, Ott PA, Wu CJ. Towards personalized, tumour-specific, therapeutic vaccines for cancer. *Nat Rev Immunol* 2018;18: 168–82. <https://doi.org/10.1038/nri.2017.131>.
- [3] Chan TA, Wolchok JD, Snyder A. Genetic basis for clinical response to CTLA-4 blockade in melanoma. *N Engl J Med* 2015; 373. <https://doi.org/10.1056/NEJMc1508163>. 1984–1984.
- [4] Rizvi NA, Mazières J, Planchard D, Stinchcombe TE, Dy GK, Antonia SJ, et al. Activity and safety of nivolumab, an anti-PD-1 immune checkpoint inhibitor, for patients with advanced, refractory squamous non-small-cell lung cancer (CheckMate 063): a phase 2, single-arm trial. *Lancet Oncol* 2015;16:257–65. [https://doi.org/10.1016/S1470-2045\(15\)70054-9](https://doi.org/10.1016/S1470-2045(15)70054-9).
- [5] Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, Eyring AD, et al. PD-1 blockade in tumors with mismatch-repair deficiency. *N Engl J Med* 2015;372:2509–20. <https://doi.org/10.1056/NEJMoa1500596>.
- [6] Yarchoan M, Hopkins A, Jaffee EM. Tumor mutational burden and response rate to PD-1 inhibition. *N Engl J Med* 2017;377: 2500–1. <https://doi.org/10.1056/NEJMc1713444>.
- [7] Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, et al. Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. *Cell Rep* 2017;18: 248–62. <https://doi.org/10.1016/j.celrep.2016.12.019>.
- [8] Matsushita H, Vesely MD, Koboldt DC, Rickert CG, Uppaluri R, Magrini VJ, et al. Cancer exome analysis reveals a T-cell-dependent mechanism of cancer immunoevasion. *Nature* 2012; 482:400–4. <https://doi.org/10.1038/nature10755>.
- [9] Nielsen M, Andreatta M. NetMHCpan-3.0; improved prediction of binding to MHC class I molecules integrating information from multiple receptor and peptide length datasets. *Genome Med* 2016; 8:33. <https://doi.org/10.1186/s13073-016-0288-x>.
- [10] Sette A, Vitiello A, Reherman B, Fowler P, Nayarsina R, Kast WM, et al. The relationship between class I binding affinity and immunogenicity of potential cytotoxic T cell epitopes. *J Immunol* 1994;153:5586–92.
- [11] van Rooij N, van Buuren MM, Philips D, Velds A, Toebes M, Heemskerk B, et al. Tumor exome analysis reveals neoantigen-specific T-cell reactivity in an ipilimumab-responsive melanoma. *J Clin Oncol* 2013;31:e439–42. <https://doi.org/10.1200/JCO.2012.47.7521>.
- [12] Calis JJA, Maybeno M, Greenbaum JA, Weiskopf D, De Silva AD, Sette A, et al. Properties of MHC class I presented peptides that enhance immunogenicity. *PLoS Comput Biol* 2013; 9, e1003266. <https://doi.org/10.1371/journal.pcbi.1003266>.
- [13] Kim S, Kim HS, Kim E, Lee MG, Shin E, Paik S, et al. Neo-pepsee: accurate genome-level prediction of neoantigens by harnessing sequence and amino acid immunogenicity information. *Ann Oncol* 2018. <https://doi.org/10.1093/annonc/mdy022>.
- [14] Łuksza M, Riaz N, Makarov V, Balachandran VP, Hellmann MD, Solovyyov A, et al. A neoantigen fitness model predicts tumour response to checkpoint blockade immunotherapy. *Nature* 2017. <https://doi.org/10.1038/nature24473>.
- [15] Yadav M, Jhunjhunwala S, Phung QT, Lupardus P, Tanguay J, Bumbaca S, et al. Predicting immunogenic tumour mutations by combining mass spectrometry and exome sequencing. *Nature* 2014;515:572–6. <https://doi.org/10.1038/nature14001>.
- [16] Bassani-Sternberg M, Bräunlein E, Klar R, Engleitner T, Sinitcyn P, Audehm S, et al. Direct identification of clinically relevant neoepitopes presented on native human melanoma tissue by mass spectrometry. *Nat Commun* 2016;7:13404. <https://doi.org/10.1038/ncomms13404>.

- [17] Kreiter S, Vormehr M, van de Roemer N, Diken M, Löwer M, Diekmann J, et al. Mutant MHC class II epitopes drive therapeutic immune responses to cancer. *Nature* 2015;520:692–6. <https://doi.org/10.1038/nature14426>.
- [18] Yossef R, Tran E, Deniger DC, Gros A, Pasetto A, Parkhurst MR, et al. Enhanced detection of neoantigen-reactive T cells targeting unique and shared oncogenes for personalized cancer immunotherapy. *JCI Insight* 2018;3. <https://doi.org/10.1172/jci.insight.122467>.
- [19] Yang B, Jeang J, Yang A, Wu TC, Hung C-F. DNA vaccine for cancer immunotherapy. *Hum Vaccines Immunother* 2014;10:3153–64. <https://doi.org/10.4161/21645515.2014.980686>.
- [20] Heil F, Hemmi H, Hochrein H, Ampenberger F, Kirschning C, Akira S, et al. Species-specific recognition of single-stranded RNA via toll-like receptor 7 and 8. *Science* 2004;303:1526–9. <https://doi.org/10.1126/science.1093620>.
- [21] Bowen WS, Srivastava AK, Batra L, Barsoumian H, Shirwan H. Current challenges for cancer vaccine adjuvant development. *Expert Rev Vaccines* 2018;17:207–15. <https://doi.org/10.1080/14760584.2018.1434000>.
- [22] Walter S, Weinschenk T, Stenzl A, Zdrojowy R, Pluzanska A, Szczylik C, et al. Multipptide immune response to cancer vaccine IMA901 after single-dose cyclophosphamide associates with longer patient survival. *Nat Med* 2012;18:1254–61. <https://doi.org/10.1038/nm.2883>.
- [23] Melief CJM, van der Burg SH. Immunotherapy of established (pre)malignant disease by synthetic long peptide vaccines. *Nat Rev Cancer* 2008;8:351–60. <https://doi.org/10.1038/nrc2373>.
- [24] Draper SJ, Heeney JL. Viruses as vaccine vectors for infectious diseases and cancer. *Nat Rev Microbiol* 2010;8:62–73. <https://doi.org/10.1038/nrmicro2240>.
- [25] Carreno BM, Magrini V, Becker-Hapak M, Kaabinejadian S, Hundal J, Petti AA, et al. A dendritic cell vaccine increases the breadth and diversity of melanoma neoantigen-specific T cells. *Science* 2015;348:803–8. <https://doi.org/10.1126/science.aaa3828>.
- [26] McGranahan N, Furness AJS, Rosenthal R, Ramskov S, Lyngaa R, Saini SK, et al. Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. *Science* 2016;351:1463–9. <https://doi.org/10.1126/science.aaf1490>.
- [27] Sahin U, Derhovanessian E, Miller M, Kloke B-P, Simon P, Löwer M, et al. Personalized RNA mutanome vaccines mobilize poly-specific therapeutic immunity against cancer. *Nature* 2017;547:222–6. <https://doi.org/10.1038/nature23003>.
- [28] Stronen E, Toebes M, Kelderman S, van Buuren MM, Yang W, van Rooij N, et al. Targeting of cancer neoantigens with donor-derived T cell receptor repertoires. *Science* 2016;352:1337–41. <https://doi.org/10.1126/science.aaf2288>.
- [29] Zaretsky JM, Garcia-Diaz A, Shin DS, Escuin-Ordinas H, Hugo W, Hu-Lieskovan S, et al. Mutations associated with acquired resistance to PD-1 blockade in melanoma. *N Engl J Med* 2016;375:819–29. <https://doi.org/10.1056/NEJMoa1604958>.
- [30] McGranahan N, Rosenthal R, Hiley CT, Rowan AJ, Watkins TBK, Wilson GA, et al. Allele-specific HLA loss and immune escape in lung cancer evolution. *Cell* 2017;171:1259–71. <https://doi.org/10.1016/j.cell.2017.10.001>. e11.
- [31] Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. *Nat Rev Cancer* 2012;12:252–64. <https://doi.org/10.1038/nrc3239>.
- [32] Hui E, Cheung J, Zhu J, Su X, Taylor MJ, Wallweber HA, et al. T cell costimulatory receptor CD28 is a primary target for PD-1-mediated inhibition. *Science* 2017;355:1428–33. <https://doi.org/10.1126/science.aaf1292>.
- [33] Kahles A, Lehmann K-V, Toussaint NC, Hüser M, Stark SG, Sachsenberg T, et al. Comprehensive analysis of alternative splicing across tumors from 8,705 patients. *Cancer Cell* 2018;34:211–24. <https://doi.org/10.1016/j.ccell.2018.07.001>. e6.
- [34] Turajlic S, Litchfield K, Xu H, Rosenthal R, McGranahan N, Reading JL, et al. Insertion-and-deletion-derived tumour-specific neoantigens and the immunogenic phenotype: a pan-cancer analysis. *Lancet Oncol* 2017;18:1009–21. [https://doi.org/10.1016/S1470-2045\(17\)30516-8](https://doi.org/10.1016/S1470-2045(17)30516-8).
- [35] Finn OJ, Rammensee H-G. Is it possible to develop cancer vaccines to neoantigens, what are the major challenges, and how can these be overcome?: neoantigens: nothing new in spite of the name. *Cold Spring Harb Perspect Biol* 2017;a028829. <https://doi.org/10.1101/cshperspect.a028829>.
- [36] Ott PA, Hu Z, Keskin DB, Shukla SA, Sun J, Bozym DJ, et al. An immunogenic personal neoantigen vaccine for patients with melanoma. *Nature* 2017;547:217–21. <https://doi.org/10.1038/nature22991>.
- [37] Gulley JL, Madan RA, Pachynski R, Mulders P, Sheikh NA, Trager J, et al. Role of antigen spread and distinctive characteristics of immunotherapy in cancer treatment. *JNCI J Natl Cancer Inst* 2017;109. <https://doi.org/10.1093/jnci/djw261>.
- [38] Melero I, Berman DM, Aznar MA, Korman AJ, Pérez Gracia JL, Haanen J. Evolving synergistic combinations of targeted immunotherapies to combat cancer. *Nat Rev Cancer* 2015;15:457–72. <https://doi.org/10.1038/nrc3973>.
- [39] Argilés G, Saro J, Segal NH, Melero I, Ros W, Marabelle A, et al. LBA-004Novel carcinoembryonic antigen T-cell bispecific (CEA-TCB) antibody: preliminary clinical data as a single agent and in combination with atezolizumab in patients with metastatic colorectal cancer (mCRC). *Ann Oncol* 2017;28. <https://doi.org/10.1093/annonc/mdx302.003>.
- [40] Le DT, Jaffee EM. Regulatory T cell modulation using cyclophosphamide in vaccine approaches: a current perspective. *Cancer Res* 2012;72:3439–44. <https://doi.org/10.1158/0008-5472.CAN-11-3912>.
- [41] Lee S, Margolin K. Cytokines in cancer immunotherapy. *Cancers* 2011;3:3856–93. <https://doi.org/10.3390/cancers3043856>.
- [42] Moynihan KD, Opel CF, Szeto GL, Tzeng A, Zhu EF, Engreitz JM, et al. Eradication of large established tumors in mice by combination immunotherapy that engages innate and adaptive immune responses. *Nat Med* 2016;22:1402–10. <https://doi.org/10.1038/nm.4200>.
- [43] Liu J, Blake SJ, Yong MCR, Harjunpää H, Ngiew SF, Takeda K, et al. Improved efficacy of neoadjuvant compared to adjuvant immunotherapy to eradicate metastatic disease. *Cancer Discov* 2016;6:1382–99. <https://doi.org/10.1158/2159-8290.CD-16-0577>.



## REVUE 4 - Développement des CAR-T dans les tumeurs solides

**Vincent Alcazer**<sup>1,2</sup>, *Christophe Delenda*<sup>3</sup>, *Laurent Poirot*<sup>3</sup>, *Stéphane Depil*<sup>1,3,4,5,\*</sup>

1 INSERM U1052, Centre de Recherche en Cancérologie, 69008 Lyon, France

2 Hospices civils de Lyon, Service d'hématologie, Centre Hospitalier Lyon Sud, 69310 Pierre-Bénite, France

3 Cellectis, 75013 Paris, France

4 Centre Léon Bérard, 69008 Lyon, France

5 Université Claude Bernard Lyon 1, 69100 Villeurbanne, France

\* Corresponding author: Stéphane Depil, Centre Léon Bérard, 28 Prom. Léa et Napoléon Bullukian, 69008 Lyon, France;

Tel: +33 (0) 4 78 78 28 28; Mail: [stephane.depil@lyon.unicancer.fr](mailto:stephane.depil@lyon.unicancer.fr)

Bull Cancer. 2018 Dec;105 Suppl 2:S178-S187. French. doi: 10.1016/S0007-4551(19)30048-7. PMID: 30686356.

# Développement des CAR-T dans les tumeurs solides

Vincent Alcazer<sup>1,2</sup>, Christophe Delenda<sup>3</sup>, Laurent Poirot<sup>3</sup>, Stéphane Depil<sup>1,3,4,5,\*</sup>

1. INSERM U1052, Centre de Recherche en Cancérologie, 69008 Lyon, France
2. Hospices civils de Lyon, Service d'hématologie, Centre Hospitalier Lyon Sud, 69310 Pierre-Bénite, France
3. Cellectis, 75013 Paris, France
4. Centre Léon Bérard, 69008 Lyon, France
5. Université Claude Bernard Lyon 1, 69100 Villeurbanne, France

\*Auteur correspondant.

Stéphane Depil  
stephane.depil@lyon.unicancer.fr

## Mots clés

CAR  
Tumeurs solides  
Cible  
Immunosuppression  
Combinaison  
Optimisation

## ■ Résumé

Les lymphocytes T à récepteur antigénique chimérique (CAR-T) ont démontré une efficacité remarquable dans certaines hémopathies malignes, notamment la leucémie aiguë lymphoblastique B et les lymphomes B pour les CAR anti-CD19. Les données dans les tumeurs solides restent quant à elles plutôt décevantes avec tout au plus quelques stabilisations et de rares réponses objectives observées. Ce manque d'efficacité peut être expliqué par différents facteurs : absence de cible spécifique (et donc risque accru de toxicité) et hétérogénéité d'expression, microenvironnement immunosuppresseur, homing insuffisant et faible pénétration intra-tumorale, défaut de persistance des CAR-T. De nombreuses approches sont actuellement envisagées pour contrecarrer ces différents mécanismes de résistance : construction de CAR bispécifiques ou utilisation de *gates* logiques, combinaison avec des inhibiteurs d'*immune checkpoint* ou CAR-T rendus insensibles aux molécules immunosuppressives, ajout de récepteurs de chémokines ou d'enzyme, combinaison avec un virus oncolytique, administration intra-tumorale, sélection de sous-populations mémoires et développement de CAR-T « armés » sécrétant des cytokines type IL-12, -15 ou -18. L'élaboration de CAR-T de dernière génération optimisés devrait à terme permettre de répondre à un besoin thérapeutique majeur notamment dans les tumeurs immunologiquement froides et/ou n'exprimant pas les molécules de CMH de classe I.

## Keywords

CAR  
Solid tumors  
Target  
Immunosuppression  
Combination  
Optimization

## ■ Abstract

### Development of CAR T-cells in solid tumors: challenges and perspectives

*While Chimeric Antigen Receptor (CAR) T-cells have shown outstanding results in some hematologic malignancies, studies in solid tumors are less encouraging with poor response rates. Several factors can account for this lack of efficiency in solid tumors: heterogeneous expression or absence of specific target antigen (and so higher risk of toxicity), immunosuppressive microenvironment, homing and tumoral trafficking issues or lack of CAR T-cell persistence. Different approaches can be considered to overcome these resistance mechanisms: bispecific CARs, use of logic gates, combination with immune checkpoint inhibitors, engineered CAR T-cells resistant to*



*immunosuppressive molecules, addition of chemokines or enzymes, combination with oncolytic virus, intra-tumoral administration, selection of memory T cell subpopulations and development of armored CAR T-cells secreting cytokines such as IL-12, -15 or -18. Last generation optimized CAR T-cell design should thus improve therapeutic efficiency. CAR-T cells may represent in a near future a therapeutic breakthrough also in solid tumors, especially in cold tumors and/or tumors lacking MHC class I expression.*

Cet article fait partie du numéro supplément *CAR-T Cells* réalisé avec le soutien institutionnel des partenaires Gilead et Celgene.

## Introduction

Les CAR-T représentent sans aucun doute une avancée thérapeutique majeure dans les hémopathies malignes. Les CAR-T de seconde génération (Figure 1) dirigés contre le CD19 ont ainsi démontré une activité remarquable dans la leucémie aiguë lymphoblastique B (LAL) et les lymphomes B. Le tisagenlecleucel (Kymriah, Novartis) a été approuvé par la FDA en août 2017 dans la LAL B en rechute de l'enfant et sujet jeune avec des taux de réponses complètes (RC) de 70-90 % [1]. L'axicabtagene ciloleucel (Yescarta, Kite Pharma) a été approuvé dans les lymphomes

malins non hodgkiniens à grandes cellules de phénotype B en rechute en octobre 2017 avec des taux de RC de l'ordre de 55 % [2]. Les CAR-T anti-BCMA donnent également des résultats très prometteurs dans le myélome multiple [3]. En revanche, les premiers résultats obtenus avec des CAR-T dans les tumeurs solides apparaissent nettement moins spectaculaires. Nous présenterons ici les principaux résultats obtenus dans les tumeurs solides. Nous analyserons les principaux obstacles expliquant les limites d'activité des CAR-T dans ce contexte et les différentes solutions possibles pour les contourner.

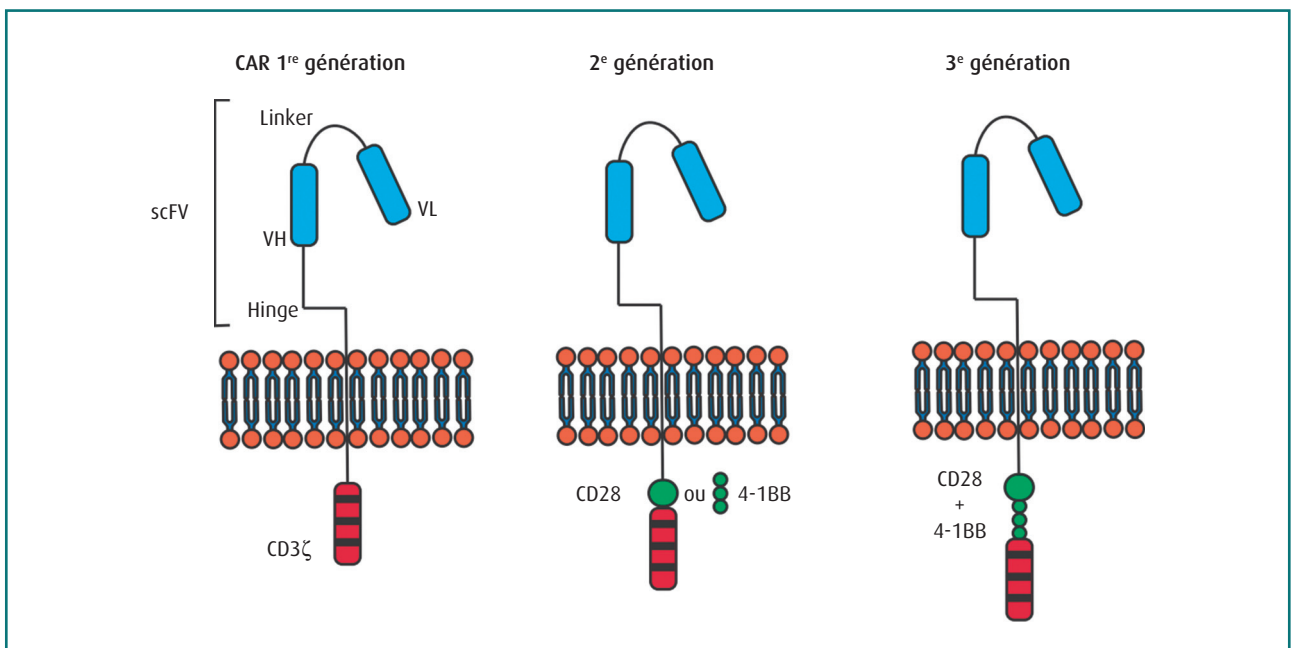


FIGURE 1

### Principes de construction des CAR.

Les CAR (« Chimeric Antigen Receptors ») correspondent à des protéines de fusion comprenant un domaine extracellulaire consistant en un fragment d'anticorps (scFV) reconnaissant un antigène membranaire cible présent sur la cellule tumorale, associé à un domaine intracellulaire provenant de protéines de signalisation des lymphocytes T.

La 1<sup>re</sup> génération de CAR comprenait uniquement CD3ζ. Un domaine de co-stimulation (CD28 ou 4-1BB) a été fusionné au CD3ζ dans les CAR de 2<sup>de</sup> génération.

La 3<sup>e</sup> génération comprend deux domaines de co-stimulation liés à CD3ζ.

# Résultats cliniques obtenus avec les CAR-T dans les tumeurs solides (Tableau 1)

**TABLEAU 1**  
**Synthèse des résultats des essais cliniques évaluant des CAR-T autologues dans les tumeurs solides.**

Cible	Type de cancer	Essai clinique	Vecteur Domaine de signalisation	Modalités d'administration	Patients (n)	Réponses cliniques	Références
<b>EGFR* &amp; CD133**</b>	Cérébral, Hépatique, ovaire, pancréas et sein, colorectal, LAM, LAL	NCT01869166/NCT02541370	Vecteur gammarétroviral CD3ζ & 4-1BB	IV : 3X (J0 à J2) 2.2x10E6* ou 1.22x10E6** cellules/kg	1	1/1 RP (8.5 mois) après EGFR CAR* 1/1 RP (4.5 mois) après CD133 CAR**	Feng et al. (2017a)
<b>CEA</b>	Hépatique	NCT01373047	Vecteur gammarétroviral CD3ζ & CD28	IA : 3X (sur 6 semaines) 1x10E8 à 1x10E10 cellules totales/dose	6	1/6 SD (25.5 mois)	Katz et al. (2015)
<b>CEA</b>	Hépatique	NCT02416466	Non spécifique	IA : 3X (toutes les 2 semaines) 1x10E10 cellules totales/dose	6	3/6 SD (moy. 7.5 mois, 3.7 à 12.5 mois)	Katz et al. (2017)
<b>CEA</b>	Colorectal et pancréatique	NCT02850536	Non spécifique	IA : 3X (1 fois/semaine) 1x10E10 cellules totales/dose	3	1/3 RC (11 mois)	Katz et al. (2018)
<b>CEA</b>	Pancréas, poumon et sein, colorectal et gastrique	NCT02349724	Vecteur lentiviral Domaine(s) non spécifié(s)	IV : 1X ou 2X 1x10E5 à 1x10E8 cellules/kg	10	7/10 SD (moy. 5.8 mois, 3.2 à 8.5 mois)	Zhang et al. (2017)
<b>CEA</b>	Ovaire, pancréas, poumon et sein, colorectal et gastrique	NCT01212887	Vecteur gammarétroviral CD3ζ	IV : 1X (J0) Max 5x10E10 cellules totales	14	7/14 SD (moy. 3.7 mois, 2.7 à 5.2 mois)	Thistlethwaite et al. (2017)
<b>cMET</b>	Sein	NCT01837602	mRNA CD3ζ & 4-1BB	IT : 1X (J0) 3x10E7 or 3x10E8 cellules totales	6	1/6 SD (1+ mois)	Tchou et al. (2017)
<b>EGFR</b>	Tumeurs solides (non spécifiées)	NCT01869166	Vecteur lentiviral CD3ζ & 4-1BB	1X à 3X 0.8x10E6 à 4.1x10E6 cellules/kg/dose	17	1/17 RC (22+ mois) 10/17 SD (moy. 6 mois, 2.5 à 15 mois)	Guo et al. (2018)
<b>EGFRvIII</b>	Gliome	NCT02209376	Vecteur lentiviral CD3ζ & 4-1BB	IV : 1X (J0) 1x10E7 cellules/kg	10	1/10 SD (18+ mois)	O'Rourke et al. (2015)
<b>GD2</b>	Mélanome	ACTRN12613000198729	Vecteur gammarétroviral CD3ζ, CD28 & OX40	IV : 1X (J0) 1x10E7 or 2x10E7 cellules/m2	4	4/4 progressions	Gargett et al. (2016)
<b>GD2</b>	Neuroblastome	NCT00085930	Vecteur gammarétroviral Domaine(s) non spécifié(s)	IV : 1X (J0) 1.2x10E7 à 1x10E8 cellules/m2	19	3/19 RC (moy. 28.3 mois, 6 à 58 mois) 1/19 RP (32 mois), 1/19 SD (6 mois)	Louis et al. (2011)
<b>GD2</b>	Neuroblastome	NCT02761915	Vecteur gammarétroviral CD3ζ & CD28	IV 1X (J0) 1x10E7 ou 1x10E8 CAR+ cellules/m2	4	1/4 RP (1.5 mois)	Straathof et al. (2018)
<b>GPC3</b>	Hépatique	NCT02395250	Vecteur lentiviral CD3ζ, CD28 & 4-1BB	IV : 1X au moins (J0) 1.3x10E5 to 1.5x10E7 cellules/kg	6	1/6 RP (12.8 mois) 3/6 SD (moy. 11.7 mois, 3.6 à 18.7 mois)	Zhai et al. (2017)

TABLEAU 1  
Synthèse des résultats des essais cliniques évaluant des CAR-T autologues dans les tumeurs solides.

Cible	Type de cancer	Essai clinique	Vecteur Domaine de signalisation	Modalités d'administration	Patients (n)	Réponses cliniques	Références
<b>HER2</b>	Glioblastome	NCT01109095	Vecteur gammarrétroviral CD3ζ & CD28	IV : 1X au moins (J0) 1x10E6 à 1x10E8 cellules/m2	16	1/16 RP (9 mois) 7/16 SD (1.5 à 30+ mois)	Ahmed <i>et al.</i> (2017)
<b>HER2</b>	Sarcome	NCT00902044	Vecteur gammarrétroviral CD3ζ & CD28	IV : 1X à 3X (toutes les 6/12 semaines) 1x10E8 cellules/m2/dose	6	1/6 RC, 2/6 SD (max 14.2 mois)	Hegde <i>et al.</i> (2017)
<b>HER2</b>	Tumeurs solides (non spécifiées)	NCT01935843	Vecteur non spécifié CD3ζ & (CD3ζ & 4-1BB)	IV : 1X ou 2X (J0) Moy. 2.1x10E6 cellules/kg/dose	11	1/11 RP (4.5 mois) 5/11 SD (moy. 4.8 mois, 1.5 à 8.3 mois)	Feng <i>et al.</i> (2017b)
<b>IL13RA2</b>	Gliome	NCT02208362	Vecteur lentiviral CD3ζ & 4-1BB	IT : 16X (sur 220 jours) 2x10E6 (1X) à 1x10E7 (2X à 16X) cellules totales/dose	1	1/1 RP (7.5 mois)	Brown <i>et al.</i> (2016)
<b>Mésothéline</b>	Poumon	NCT01355965	mRNA CD3ζ & 4-1BB	IV : 3X (1fois/semaine) à 10X (1 à 3 fois/semaine) IT : 9X-10X 1x10E8 à 1x10E9 cellules totales/ dose	2	1/1 RP (6 mois) 1/1 SD (2.5 mois)	Beatty <i>et al.</i> (2014)
<b>Mésothéline</b>	Pancréas	NCT01897415	mRNA CD3ζ & 4-1BB	IV : 3X (1 fois/semaine) 1x10E8 à 3x10E8 cellules/m2/dose	6	2/6 SD (max 4 mois)	Beatty <i>et al.</i> (2015)
<b>Mésothéline</b>	Ovaire, pancréas et poumon	NCT02159716	Vecteur lentiviral CD3ζ & 4-1BB	IV : 1X (J0) 3x10E7 ou 3x10E8 cellules/m2	6	6/6 SD (à 1 mois)	Tanyi <i>et al.</i> (2016)
<b>PSMA</b>	Prostate	NCT01140373	Vecteur gammarrétroviral CD3ζ & CD28	IV : 1X (J0) 1x10E7 à 3x10E7 cellules/kg	4	2/4 SD (6+ et 16+ mois)	Slovin <i>et al.</i> (2013)
<b>PSMA</b>	Prostate	NCT00664196	Vecteur gammarrétroviral Domains unspecified	IV : 1X (J0) 1x10E9 ou 1x10E10 cellules totales	18	2/5 RP (moy. 5+ mois)	Junghans <i>et al.</i> (2013)
<b>TAG-72</b>	Colorectal	C-9701	Vecteur gammarrétroviral CD3ζ	IV : 3X or 5X (toutes les 2 semaines) 1x10E8 à 1x10E10 cellules totales/ dose	9	9/9 progressions	Hege <i>et al.</i> (2017)
<b>TAG-72</b>	Colorectal	C-9702	Vecteur gammarrétroviral CD3ζ	IA : 4X 1x10E9 ou 1x10E10 cellules totales/ dose	3	3/3 progressions	Hege <i>et al.</i> (2017)

LAM : Leucémie Aiguë Myéloïde, LAL : Leucémie aigüe Lymphoïde, IA : Intra-artérielle, IT : Intra-tumorale, IV : Intra-veineuse, RC : réponse complète, RP : réponse partielle, SD : Maladie stable

### CAR dirigé contre la mésothéline

La mésothéline est un antigène tumoral présent au niveau du mésothéliome pleural, des cancers du pancréas et ovaire. Compte tenu d'une expression également partagée par le péricard, la plèvre et le péricarde, le premier CAR-T évalué dans ce contexte a été construit pour avoir une expression transitoire par transfection ARN sans lymphodéplétion préalable [4]. Le profil de tolérance était satisfaisant en dehors d'une réaction anaphylactique à l'encontre du scFv murin utilisé dans la construction, favorisée ici par un schéma d'administration répété. D'autres essais ont ensuite été mis en place dans les cancers du pancréas et de l'ovaire utilisant un CAR classique de seconde génération (toujours avec scFv murin). Seules des stabilisations ont pu être observées, sans aucune réponse objective [5].

### CAR dirigé contre CEA

Quatre essais de phase I évaluant un CAR dirigé contre la protéine CEA dans les tumeurs hépatiques et digestives ont été rapportés [6,7]. Le profil de tolérance apparaît satisfaisant mais on note l'absence d'expansion durable des CAR-T injectés (CAR de seconde génération CD3 $\zeta$  / CD28 dans l'étude de Zhang). Des stabilisations de la maladie (SD) ont pu être observées, sans réponse objective.

### CAR dirigé contre EGFR et variants

Deux essais chinois ont évalués un CAR dirigé contre le récepteur à l'EGF (EGFR) natif dans les cancers des voies biliaires [8] et le cancer du poumon non à petites cellules [9], montrant dans le premier cas une RC et 10 SD sur 17 patients évalués, et dans le second 2 réponses partielles (RP) et 5 SD sur 11 patients évalués. Le profil de tolérance semble acceptable avec une toxicité cutanéomuqueuse n'excédant pas le grade 1/2.

EGFRvIII est une forme mutée d'EGFR communément associée au glioblastome. Une étude a ainsi évalué un CAR de seconde génération (CD3 $\zeta$  / 4-1BB) anti-EGFRvIII dans le glioblastome en rechute [10]. Le profil de tolérance était satisfaisant. Une maladie résiduelle stable pendant 18 mois a été rapportée chez un patient sur les 10 traités, avec une expansion transitoire des CAR-T détectable dans le sang. La présence de CAR-T dans la tumeur a été mise en évidence chez sept patients ayant bénéficié d'une chirurgie après l'infusion. Cette infiltration était associée à une perte d'expression de la cible chez cinq patients, soulignant le problème d'hétérogénéité d'expression antigénique dans la tumeur.

### CAR dirigé contre HER2

Le premier essai mené par le NCI évaluant un CAR de troisième génération (CD28/4-1BB/CD3 $\zeta$ ) avec un scFv de forte affinité dérivé de l'Herceptine a dû être stoppé suite au décès d'un patient atteint d'un cancer colique ayant présenté une détresse respiratoire associée à la présence du CAR-T dans le poumon, en lien avec un faible niveau d'expression de HER2 dans le tissu pulmonaire [11]. Un autre essai a évalué des doses plus faibles de CAR-T possédant un scFv de plus faible affinité dans le sarcome, montrant un profil de tolérance satisfaisant mais une activité clinique modeste [12]. Un essai a également évalué des CAR-T

anti-HER2 administrés IV sans lymphodéplétion préalable chez 17 patients présentant un glioblastome HER2-positif en rechute. Sur 16 patients évaluables, un patient a présenté une RP et 7 une SD avec une bonne tolérance globale [13].

### CAR dirigé contre PSMA

PSMA est une protéine membranaire surexprimée dans le cancer de la prostate par la plupart des cellules tumorales et par les néo-vaisseaux, suggérant un effet anti-angiogénique ajouté à l'effet anti-tumoral direct d'un CAR-T dirigé contre PSMA. Deux CAR-T de seconde génération anti-PSMA (CD3 $\zeta$  / CD28) ont été évalués en phase I. Dans l'essai du MSKCC une SD de 6 mois a pu être obtenue chez 2 patients sur 4 [14]. L'étude du Roger William Medical Center a mis en évidence 2 RP sur 5 patients évalués. La réponse était associée à un taux plasmatique d'IL-2 plus élevé [15]. Ces résultats justifient la poursuite d'essais cliniques évaluant des CAR optimisés ciblant PSMA (essai NCT03089203 par exemple).

### CAR dirigé contre IL-13R $\alpha$

IL-13R $\alpha$  est surexprimé dans la moitié des glioblastomes avec une expression faible dans le tissu cérébral normal. Une première publication montrait la tolérance satisfaisante de l'injection intracrânienne de CAR-T dirigés contre IL-13R $\alpha$  avec 2 patients présentant des signes de réponse à l'IRM [16]. Un patient en rémission après injections intracrâniennes répétées de CAR-T dirigés contre IL-13R $\alpha$  a été rapporté dans le *New England Journal of Medicine*. À noter qu'une augmentation concordante des cellules immunitaires et de différentes cytokines était observée dans le liquide céphalorachidien [17].

### CAR dirigé contre GD2

GD2 est un disialoganglioside surexprimé par les tumeurs d'origine neuro-ectodermique notamment le neuroblastome. Un CAR de première génération a été évalué chez 19 patients atteints de neuroblastome et a permis l'obtention de 3 RC et 1 RP [18]. Une étude de phase I d'escalade de doses évaluant un CAR de dernière génération a été rapportée à l'AAO 2018. Les premiers paliers montraient l'absence de réponse, en lien avec l'absence d'expansion des CAR-T injectés. Au dernier palier testé ( $1 \times 10^8$  cellules/m<sup>2</sup>), une réponse transitoire a pu être observée [19].

### CAR dirigé contre glypican 3

Glypican-3 est un protéoglycane de surface surexprimé dans le carcinome hépatocellulaire. Un CAR de troisième génération anti-glypican 3 a été évalué au Renji Hospital. Les patients qui n'ont pas reçu de lymphodéplétion préalable ont progressé alors qu'une réponse objective et 3 SD prolongées ont été observées parmi les 6 patients qui ont reçu une lymphodéplétion [20].

### Autres CAR

D'autres essais sont en cours avec des CAR dirigés contre CD70, CD 171 (L-1 CAM) dans le neuroblastome, claudine 18 dans les cancers du pancréas et de l'estomac, EPCAM, EPHA2, FAP, LMP1 dans le carcinome indifférencié du nasopharynx lié à l'EBV,

MUC1, MUC16, PD-L1, PSCA dans le pancréas, ROR1, TAG-72 dans le cancer colorectal, etc. Les résultats cliniques restent attendus.

### Facteurs limitant l'efficacité des CAR-T dans les tumeurs solides

Comme le montrent les différentes études mentionnées plus haut, les CAR-T ont pour l'instant donné des résultats beaucoup moins satisfaisants dans les tumeurs solides que dans les hémopathies, avec quelques stabilisations et très peu de réponses objectives. Différents facteurs, détaillés ci-dessous, se combinent dans les tumeurs solides pour expliquer cette relative inefficacité des CAR-T « classiques ».

### Absence d'antigènes membranaires spécifiques et hétérogénéité d'expression

L'origine myélo-sanguine permet d'utiliser dans les hémopathies des antigènes d'expression restreinte tels que CD19, CD20, CD22, CD123, CLL1, BCMA, etc. Les cibles membranaires dans les tumeurs solides sont souvent surexprimées par la tumeur mais retrouvées également dans différents tissus sains parfois critiques. Si le différentiel d'expression permet un niveau de sélectivité tumorale suffisant pour des approches anticorps classiques ou même armés, un niveau d'expression même faible par un tissu sain peut être à l'origine d'une toxicité rédhibitoire avec un CAR optimisé. L'expression faible de HER2 par le tissu pulmonaire explique ainsi la toxicité observée avec un CAR de forte affinité dérivé de l'Herceptine [11].

Une autre problématique liée à la cible est l'hétérogénéité d'expression avec la sélection des clones tumoraux n'exprimant pas l'antigène d'intérêt (mécanisme de résistance mis en évidence en hématologie avec CD19 malgré une expression forte et homogène au départ). Il paraît ainsi nécessaire de combiner différentes spécificités. De telles combinaisons sont ainsi envisagées dans le glioblastome compte tenu des limitations observées avec un CAR ciblant uniquement EGFRvIII [10].

### Immunosuppression associée au microenvironnement tumoral

Le microenvironnement tumoral est caractérisé par la présence de cellules immunosuppressives comprenant notamment des cellules T régulatrices, des cellules suppressives d'origine myéloïde (MDSCs) et des macrophages M2. Ces différentes populations cellulaires et les cytokines sécrétées, comme le TGF- $\beta$  et l'IL-10, contribuent à réduire l'activité des CAR-T qui arriveraient au sein de cet environnement hostile du point de vue immunitaire [21].

À cela s'ajoute l'expression de différents checkpoints négatifs (PD-L1, IDO, TIM3, LAG3, etc.) présents de manière constitutive ou induite [22]. Une éventuelle activation des CAR-T avec production d'IFN- $\gamma$  favorisera ainsi l'expression de PD-L1 ou IDO au sein de la tumeur, ce qui viendra limiter l'efficacité dans le temps des CAR-T injectés.

D'autres facteurs inhibiteurs sont présents dans les tumeurs solides, associés à l'hypoxie, le manque de nutriments, un pH acide ou la présence d'acide lactique [23].

### Problèmes de homing et d'accès au site tumoral des CAR-T

Alors que la pénétration des CAR-T dans la moelle osseuse et les ganglions semble satisfaisante, leur pénétration dans les tumeurs solides semble plus compliquée en raison des barrières physiques et biochimiques mises en place. Le trafic des lymphocytes T jusqu'au site tumoral correspond en fait à un processus dynamique mettant en jeu un réseau complexe d'interactions. Les récepteurs aux chémokines exprimés par les CAR-T doivent correspondre aux chémokines sécrétées par la tumeur. Il a ainsi été montré qu'une inadéquation entre les chémokines et les récepteurs était impliquée dans le faible recrutement des CAR-T au niveau du site tumoral. De nombreuses tumeurs sécrètent en effet soit peu de chémokines soit des chémokines pour lesquelles les lymphocytes T effecteurs n'expriment pas le récepteur correspondant [21]. L'endothélium tumoral représente également une barrière à l'infiltration des cellules T. Une bonne pénétration dans la tumeur nécessite donc que les cellules T et l'endothélium expriment les molécules et les récepteurs d'adhésion adéquats. Durant le processus d'extravasation, les lymphocytes T dégradent des composantes de la membrane basale sous-endothéliale et de la matrice extra-cellulaire, dont les héparanes sulfates. Les CAR-T doivent ainsi libérer l'héparanase pour accéder aux cellules tumorales. Un déficit en héparanase a été associé à une diminution de l'activité anti-tumorale des CAR-T dans des modèles tumoraux possédant un stroma dense [24].

### Absence de persistance des CAR-T

Un faible niveau de persistance des CAR-T a été observé dans les tumeurs solides comparativement à ce qui a pu être décrit avec les CAR-T autologues dirigés contre CD19 dans les hémopathies lymphoïdes. La persistance de CAR-T à des niveaux significatifs dépend des capacités d'expansion initiale d'une part et de la présence de cellules de phénotype mémoire central d'autre part. Ainsi, une faible présence de cellules mémoires de type « *central memory* » ( $T_{CM}$ ) ou « *stem cell-like memory* » ( $T_{SCM}$ ) au sein des CAR-T injectés, un environnement immunosuppresseur peu propice à la survie des CAR-T, la difficulté d'accès aux cellules tumorales exprimant l'antigène nécessaire à l'activation et l'expansion des CAR-T, ou encore l'immunogénicité de la construction du CAR, sont autant de facteurs pouvant limiter la persistance [25].

## Approches destinées à augmenter l'efficacité des CAR-T dans les tumeurs solides (Tableau 2)

### Stratégies destinées à pallier à l'absence de cible spécifique exprimée de manière homogène

Compte tenu de l'absence de cible réellement spécifique et de l'hétérogénéité d'expression des antigènes dans les tumeurs solides, une possibilité théorique est de discriminer les cellules tumorales des cellules saines à partir d'une combinatoire d'antigènes. Ceci a amené à développer des constructions de

TABLEAU 2  
Facteurs limitant l'efficacité des CAR-T dans les tumeurs solides et solutions proposées.

Facteurs limitants	Stratégies proposées
Absence de cible spécifique / Hétérogénéité d'expression des cibles	- CAR bispécifiques - « Gates » logiques (AND, NOT et OR)
Immunosuppression	- Combinaisons avec des inhibiteurs d'immune checkpoint - CAR-T avec KO de PD-1, PD-1 dominant négatif/chimérique (PD-1/CD28) - CAR-T TGFβ-R II KO ou dominant négatif
Problème de homing et d'accès au site tumoral	- Ajout de récepteurs de chémokines (CXCR2, CCR2b, CCR4) - Expression d'enzymes type héparanase - Combinaison avec virus oncolytique - Administration intra-tumorale
Manque de persistance	- Sélection et enrichissement de T <sub>scm</sub> et T <sub>cm</sub> - Utilisation de lymphocytes mémoires spécifiques de virus (EBV, CMV) - CAR-T « armés » sécrétant des cytokines (IL-12, -15, -18)

T<sub>cm</sub> : Lymphocyte T « central memory », T<sub>scm</sub> : Lymphocyte T « stem cell-like memory »

CAR s'articulant autour de circuits logiques booléens, avec le développement de « gates » AND, NOT et OR (Figure 2). Dans le cas de la « gate AND », l'activation n'a lieu qu'en cas de reconnaissance de 2 cibles A et B de manière simultanée. Une première approche a consisté à séparer le signal CD3ζ et le signal de co-stimulation sur 2 récepteurs chimériques différents exprimés dans la même cellule T [26]. Une autre approche repose sur l'utilisation d'un récepteur Notch synthétique (synNotch). La fixation à un premier antigène A induit le clivage du domaine Notch contenant un activateur transcriptionnel qui est transloqué au noyau. Ce signal active la transcription du CAR se liant à l'antigène B pour induire l'activité cytotoxique du CAR [27]. Dans le cas d'une « gate NOT », il y a combinaison entre un récepteur CAR activateur reconnaissant A et un récepteur CAR inhibiteur (dont la région intracellulaire est dérivée d'un récepteur inhibiteur comme PD-1 ou CTLA-4) reconnaissant B (antigène exprimé par une

cellule normale). Le lymphocyte T s'activera uniquement en cas de fixation à l'antigène A, protégeant ainsi les cellules normales exprimant B. Enfin, dans une « gate OR », le lymphocyte T possède 2 récepteurs CAR reconnaissant A et B, permettant à la cellule T de s'activer en cas de reconnaissance de A et/ou de B, ce qui permet de contrecarrer la résistance liée à l'absence d'expression d'une cible en cas d'hétérogénéité tumorale.

**Stratégies pour contrecarrer les mécanismes immunosuppresseurs**

**Combinaison avec un immune checkpoint inhibiteur et construction de CAR-T insensibles à PD-L1**

Il existe un rationnel solide pour associer les CAR-T avec un inhibiteur de checkpoint type anti-PD1/-L1. Les tumeurs solides vont, en effet, (i) soit exprimer d'emblée PD-L1 au niveau des

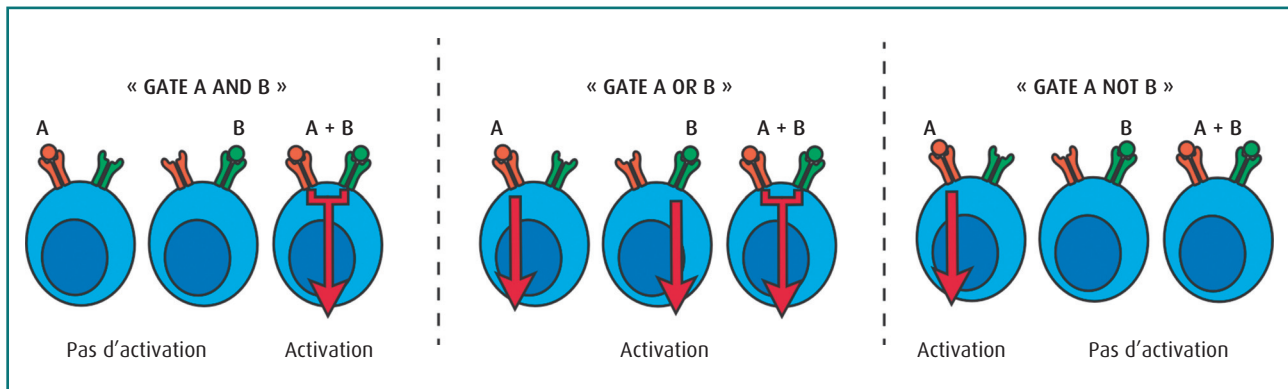


FIGURE 2  
Principe de fonctionnement des « gates » logiques.

Le lymphocyte T exprime deux récepteurs différents : l'un spécifique d'un antigène A, l'autre d'un antigène B.  
 Gate « AND » : Le signal d'activation n'est déclenché qu'en cas de reconnaissance concomitante des deux antigènes A et B ;  
 Gate « OR » : Le signal d'activation est transmis en cas de reconnaissance de l'antigène A, de B ou des deux.  
 Gate « NOT » : Le signal d'activation est transmis en cas de reconnaissance de l'antigène A, mais pas en cas de reconnaissance de B (signal inhibiteur).



cellules tumorales ou de leur microenvironnement, (ii) soit l'activité des CAR-T va induire l'expression de PD-L1 dans la tumeur en réponse à la production d'IFN- $\gamma$ , ce qui va venir moduler son efficacité [22]. Différentes données précliniques confirment ce rationnel de combinaison [28,29]. Une autre possibilité est d'utiliser un CAR-T avec un KO de PD-1, un dominant négatif de PD-1 ou un PD-1 chimérique PD1/CD28 utilisant la partie cytoplasmique de CD28 pour transformer un signal inhibiteur en signal activateur [30,31].

### **Développement de CAR-T TGF $\beta$ -RII négatif ou dominant négatif**

Le TGF $\beta$  est produit par de nombreux types tumoraux. Le TGF $\beta$  inhibe les fonctions des lymphocytes T et semble jouer un rôle majeur dans les phénomènes d'exclusion immunitaire des cellules T [32]. Différentes stratégies peuvent être utilisées pour contrecarrer le TGF $\beta$  : association avec un anticorps anti-TGF $\beta$ , utilisation de CAR-T KO pour le TGF $\beta$  Récepteur II ou CAR-T exprimant un dominant négatif du TGF $\beta$  Récepteur II [21]. Suite à la preuve de concept réalisée en préclinique dans un modèle de prostate [33], un essai de phase I est actuellement en cours dans le cancer de la prostate avec un CAR-T PSMA- TGF $\beta$  Récepteur II dominant négatif (NCT03089203).

### **Stratégies destinées à favoriser l'accès à la tumeur**

#### **Ajout de récepteurs de chémokines**

Plusieurs récepteurs de chémokines ont été évalués en préclinique pour favoriser l'accès à la tumeur des lymphocytes T. Il a ainsi été montré que la surexpression de CXCR2 augmentait la migration des lymphocytes T via la reconnaissance de CXCL1 produit par la tumeur [34]. L'équipe de Moon a évalué un CAR anti-mésothéline surexprimant CCR2b dans un modèle de mésothéliome et a pu mettre en évidence une augmentation de la migration des CAR-T dans la tumeur associée à une activité anti-tumorale accrue [35]. Une observation similaire a pu être faite avec un CAR-T anti-GD2 exprimant CCR2b dans un modèle de neuroblastome sécrétant CCL2 [36]. Enfin, il a été montré que l'expression de CCR4 par des CAR-T anti-CD30 augmentait leur migration dans un modèle de xénotransplante de lymphome de Hodgkin sécrétant CCL17 [37].

#### **Approches favorisant la pénétration dans le stroma tumoral**

La surexpression de l'héparanase (enzyme dégradant l'héparane sulfate) est susceptible de favoriser l'infiltration des CAR-T au sein du stroma tumoral, notamment si celui-ci est dense. Il a ainsi été montré qu'un CAR-T exprimant l'héparanase possédait une activité anti-tumorale accrue, en rapport avec une meilleure pénétration au sein de la tumeur [24].

#### **Combinaison avec des virus oncolytiques**

L'inflammation induite par un virus oncolytique est susceptible de favoriser le recrutement de CAR-T et de moduler l'environnement immunosuppresseur. Il a ainsi été montré qu'un virus oncolytique produisant la chémokine RANTES et l'IL-15 possédait

non seulement un effet lytique direct sur les cellules tumorales, mais en plus favorisait le recrutement et la survie de CAR-T GD2 dans un modèle de neuroblastome [38]. Les virus oncolytiques vont, d'une manière générale, induire un microenvironnement pro-inflammatoire favorisant l'infiltration des cellules T, avec production des chémokines CXCL9, CXCL10 et CXCL11, et rompre la barrière physique à la pénétration des cellules T. L'inflammation va aussi entraîner une augmentation des sélectines sur les cellules endothéliales favorisant l'infiltration T. Enfin, certains virus comme le virus de la stomatite vésiculaire (VSV) peuvent se répliquer au sein des fibroblastes associés aux tumeurs (CAFs) et les détruire. Ces fibroblastes ayant un rôle majeur dans la mise en place d'un stroma desmoplasique et immunosuppresseur, leur déplétion par un virus oncolytique va favoriser la pénétration des cellules T au sein de la tumeur. Il existe donc un rationnel solide pour associer CAR-T et virus oncolytique. De plus, la possibilité d'une utilisation systémique de certains virus oncolytiques permet d'envisager une distribution des CAR-T aux différentes lésions tumorales [39].

#### **Administration intra-tumorale des CAR-T**

L'administration de CAR-T ciblant la mésothéline en intra-pleural a été comparée avec la voie IV dans un modèle orthotopique de mésothéliome. Une efficacité supérieure était observée après injection intra-pleurale associée à une persistance prolongée. L'élimination de lésions extra-thoraciques était également observée suggérant une circulation des CAR-T à partir du site d'injection [40]. Des injections locales ont été réalisées en clinique dans le glioblastome, suggérant un effet positif de l'injection de CAR-T IL-13R $\alpha$  dans la cavité tumorale réséquée et dans les ventricules, avec un effet à la fois sur les lésions cérébrales et médullaires [17].

#### **Approches destinées à renforcer la persistance et l'activité des CAR-T**

Compte tenu de l'importance des cellules mémoires T<sub>CM</sub> ou T<sub>SCM</sub> pour assurer une persistance au long cours des CAR-T injectés, certains auteurs ont proposé de transduire des T<sub>SCM</sub> et T<sub>CM</sub> sélectionnés [41]. Des lymphocytes T mémoires spécifiques de virus comme EBV et CMV (caractérisés par une persistance à long terme) et transduits avec le récepteur CAR d'intérêt ont également été utilisés [18].

D'autres équipes ont modifié les CAR-T de manière à leur faire sécréter des cytokines favorisant leur survie ou leur activité. Il a ainsi été montré qu'un CAR-T sécrétant de l'IL-12 (de manière constitutive ou après activation) possédait une activité cytotoxique accrue tout en induisant une repolarisation Th1 attirant les cellules T endogènes et des cellules de l'immunité innée [42]. Une publication récente montre qu'un CAR-T sécrétant de l'IL-18 présente une expansion et une persistance augmentées [43]. De même, l'expression d'IL-15 est susceptible de favoriser la prolifération et la persistance des CAR-T [44].

## **Conclusion et perspectives**

Même si les résultats observés jusqu'à ce jour restent assez décevants dans les tumeurs solides, de nombreuses approches

sont envisagées pour permettre la construction de CAR-T de nouvelle génération optimisés, possédant une meilleure sélectivité tumorale, de meilleures capacités d'accès à la tumeur et une activité accrue dans un contexte immunosuppresseur. Les approches de *gene editing* aujourd'hui disponibles permettent de cumuler de nombreuses modifications, qui devront néanmoins satisfaire à un processus rigoureux de contrôle qualité et de qualification réglementaire. Les approches de combinaison ouvrent également des perspectives tout à fait intéressantes. À terme, les approches CAR-T pourraient représenter une thérapeutique de choix, « supra-physiologique », capable de venir modifier

l'histoire naturelle de tumeurs immunologiquement froides, et ce de manière indépendante de l'expression du CMH de classe I par la tumeur. Le réel potentiel de « rupture thérapeutique » associée justifie donc la poursuite des efforts accomplis dans ce domaine.

#### Déclaration de liens d'intérêts :

S. Depil déclare être employé de Cellectis et avoir réalisé des activités de consultance ou des prestations pour : AstraZeneca, BMS, Servier, Eisai, Erytech, PDC<sup>®</sup>line Pharma, Epivax, Netris Pharma, et Celyad. C. Delenda et L. Poirot déclarent être employés de Cellectis. Cet article fait partie du numéro supplément *CAR-T Cells* réalisé avec le soutien institutionnel des partenaires Gilead et Celgene.

## Références

- [1] Maude SL, Laetsch TW, Buechner J, Rives S, Boyer M, Bittencourt H, et al. Tisagenlecleucel in Children and Young Adults with B-Cell Lymphoblastic Leukemia. *New England Journal of Medicine* 2018;378:439-48.
- [2] Neelapu SS, Locke FL, Bartlett NL, Lekakis LJ, Miklos DB, Jacobson CA, et al. Axicabtagene Ciloleucel CAR T-Cell Therapy in Refractory Large B-Cell Lymphoma. *New England Journal of Medicine* 2017;377:2531-44.
- [3] Brudno JN, Maric I, Hartman SD, Rose JJ, Wang M, Lam N, et al. T Cells Genetically Modified to Express an Anti - B-Cell Maturation Antigen Chimeric Antigen Receptor Cause Remissions of Poor-Prognosis Relapsed Multiple Myeloma. *JCO* 2018;36:2267-80.
- [4] Beatty GL, Haas AR, Maus MV, Torigan DA, Soulen MC, Plesa G, et al. Mesothelin-specific chimeric antigen receptor mRNA-engineered T cells induce anti-tumor activity in solid malignancies. *Cancer Immunol Res* 2014;2:112-20.
- [5] Tanyi JL, Haas AR, Beatty GL, Stashwick CJ, O'Hara MH, Morgan MA, et al. Anti-mesothelin chimeric antigen receptor T cells in patients with epithelial ovarian cancer. *JCO* 2016;34:5511-5511.
- [6] Katz SC, Burga RA, McCormack E, Wang LJ, Mooring W, Point GR, et al. Phase I Hepatic Immunotherapy for Metastases Study of Intra-Arterial Chimeric Antigen Receptor - Modified T-cell Therapy for CEA+ Liver Metastases. *Clin Cancer Res* 2015;21:3149-59.
- [7] Zhang C, Wang Z, Yang Z, Wang M, Li S, Li Y, et al. Phase I Escalating-Dose Trial of CAR-T Therapy Targeting CEA+ Metastatic Colorectal Cancers. *Molecular Therapy* 2017;25:1248-58.
- [8] Guo Y, Feng K, Liu Y, Wu Z, Dai H, Yang Q, et al. Phase I Study of Chimeric Antigen Receptor - Modified T Cells in Patients with EGFR-Positive Advanced Biliary Tract Cancers. *Clin Cancer Res* 2018;24:1277-86.
- [9] Feng K, Guo Y, Dai H, Wang Y, Li X, Jia H, et al. Chimeric antigen receptor-modified T cells for the immunotherapy of patients with EGFR-expressing advanced relapsed/refractory non-small cell lung cancer. *Sci China Life Sci* 2016;59:468-79.
- [10] O'Rourke DM, Nasrallah MP, Desai A, Melenhorst JJ, Mansfield K, Morrissette JJD, et al. A single dose of peripherally infused EGFRVIII-directed CAR T cells mediates antigen loss and induces adaptive resistance in patients with recurrent glioblastoma. *Sci Transl Med* 2017;9.
- [11] Morgan RA, Yang JC, Kitano M, Dudley ME, Laurencot CM, Rosenberg SA. Case Report of a Serious Adverse Event Following the Administration of T Cells Transduced With a Chimeric Antigen Receptor Recognizing ERBB2. *Molecular Therapy* 2010;18:843-51.
- [12] Ahmed N, Brawley VS, Hegde M, Robertson C, Ghazi A, Gerken C, et al. Human Epidermal Growth Factor Receptor 2 (HER2) - Specific Chimeric Antigen Receptor - Modified T Cells for the Immunotherapy of HER2-Positive Sarcoma. *JCO* 2015;33:1688-96.
- [13] Ahmed N, Brawley V, Hegde M, Bielamowicz K, Kalra M, Landi D, et al. HER2-Specific Chimeric Antigen Receptor-Modified Virus-Specific T Cells for Progressive Glioblastoma: A Phase 1 Dose-Escalation Trial. *JAMA Oncol* 2017;3:1094-101.
- [14] Slovin SF, Wang X, Hullings M, Arauz G, Bartido S, Lewis JS, et al. Chimeric antigen receptor (CAR+) modified T cells targeting prostate-specific membrane antigen (PSMA) in patients (pts) with castrate metastatic prostate cancer (CMPC). *JCO* 2013;31:72-72.
- [15] Junghans RP, Ma Q, Rathore R, Gomes EM, Bais AJ, Lo ASY, et al. Phase I Trial of Anti-PSMA Designer CAR-T Cells in Prostate Cancer: Possible Role for Interacting Interleukin 2-T Cell Pharmacodynamics as a Determinant of Clinical Response. *Prostate* 2016;76:1257-70.
- [16] Brown CE, Badie B, Barish ME, Weng L, Ostberg JR, Chang W-C, et al. Bioactivity and Safety of IL13 $\alpha$ 2-Redirected Chimeric Antigen Receptor CD8+ T Cells in Patients with Recurrent Glioblastoma. *Clin Cancer Res* 2015;21:4062-72..
- [17] Brown CE, Alizadeh D, Starr R, Weng L, Wagner JR, Naranjo A, et al. Regression of Glioblastoma after Chimeric Antigen Receptor T-Cell Therapy. *New England Journal of Medicine* 2016;375:2561-9..
- [18] Louis CU, Savoldo B, Dotti G, Pule M, Yvon E, Myers GD, et al. Antitumor activity and long-term fate of chimeric antigen receptor-positive T cells in patients with neuroblastoma. *Blood* 2011;118:6050-6.49.
- [19] Straathof K, Flutter B, Wallace R, Thomas S, Cheung G, Collura A, et al. A Cancer Research UK phase I trial of anti-GD2 chimeric antigen receptor (CAR) transduced T-cells (1RG-CART) in patients with relapsed or refractory neuroblastoma. *Proceedings: AACR Annual Meeting 2018; April 14-18, 2018; Chicago, IL.*
- [20] Zhai B, Shi D, Gao H, Qi X, Jiang H, Zhang Y, et al. A phase I study of anti-GPC3 chimeric antigen receptor modified T cells (GPC3 CAR-T) in Chinese patients with refractory or relapsed GPC3+ hepatocellular carcinoma (r/r GPC3+ HCC). *JCO* 2017;35:3049-3049.
- [21] Mirzaei HR, Rodriguez A, Shepphird J, Brown CE, Badie B. Chimeric Antigen Receptors T Cell Therapy in Solid Tumor: Challenges and Clinical Applications. *Front Immunol* 2017;8.
- [22] Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. *Nature Reviews Cancer* 2012;12:252-64.
- [23] Zhang E, Gu J, Xu H. Prospects for chimeric antigen receptor-modified T cell therapy for solid tumors. *Molecular Cancer* 2018;17:7.
- [24] Caruana I, Savoldo B, Hoyos V, Weber G, Liu H, Kim ES, et al. Heparanase promotes tumor infiltration and antitumor activity of CAR-redirected T lymphocytes. *Nat Med* 2015;21:524-9.
- [25] Jackson HJ, Rafiq S, Brentjens RJ. Driving CAR T-cells forward. *Nat Rev Clin Oncol* 2016;13:370-83.

- [26] Kloss CC, Condomines M, Cartellieri M, Bachmann M, Sadelain M. Combinatorial antigen recognition with balanced signaling promotes selective tumor eradication by engineered T cells. *Nat Biotechnol* 2013;31:71–5.
- [27] Roybal KT, Rupp LJ, Morsut L, Walker WJ, McNally KA, Park JS, et al. Precision Tumor Recognition by T Cells With Combinatorial Antigen-Sensing Circuits. *Cell* 2016;164:770–9.
- [28] Moon EK, Wang L-C, Dolfi DV, Wilson CB, Ranganathan R, Sun J, et al. Multifactorial T-cell Hypofunction That Is Reversible Can Limit the Efficacy of Chimeric Antigen Receptor – Transduced Human T cells in Solid Tumors. *Clin Cancer Res* 2014;20:4262–73.
- [29] John LB, Devaud C, Duong CPM, Yong CS, Beavis PA, Haynes NM, et al. Anti-PD-1 Antibody Therapy Potently Enhances the Eradication of Established Tumors By Gene-Modified T Cells. *Clin Cancer Res* 2013;19:5636–46.
- [30] Cherkassky L, Morello A, Villena-Vargas J, Feng Y, Dimitrov DS, Jones DR, et al. Human CAR T cells with cell-intrinsic PD-1 checkpoint blockade resist tumor-mediated inhibition. *J Clin Invest* 2016;126:3130–44.
- [31] Liu X, Ranganathan R, Jiang S, Fang C, Sun J, Kim S, et al. A chimeric switch-receptor targeting PD1 augments the efficacy of second generation CART-Cells in advanced solid tumors. *Cancer Res* 2016;76:1578–90.
- [32] Mariathasan S, Turley SJ, Nickles D, Castiglioni A, Yuen K, Wang Y, et al. TGF $\beta$  attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. *Nature* 2018;554:544–8.
- [33] Kloss CC, Lee J, Zhang A, Chen F, Melenhorst JJ, Lacey SF, et al. Dominant-Negative TGF- $\beta$  Receptor Enhances PSMA-Targeted Human CAR T Cell Proliferation And Augments Prostate Cancer Eradication. *Molecular Therapy* 2018.
- [34] Peng W, Ye Y, Rabinovich BA, Liu C, Lou Y, Zhang M, et al. Transduction of Tumor-Specific T Cells with CXCR2 Chemokine Receptor Improves Migration to Tumor and Antitumor Immune Responses. *Clin Cancer Res* 2010;16:5458–68.
- [35] Moon EK, Carpenito C, Sun J, Wang L-CS, Kapoor V, Predina J, et al. Expression of a Functional CCR2 Receptor Enhances Tumor Localization and Tumor Eradication by Retargeted Human T cells Expressing a Mesothelin-Specific Chimeric Antibody Receptor. *Clin Cancer Res* 2011;17:4719–30. doi:10.1158/1078-0432.CCR-11-0351.
- [36] Craddock JA, Lu A, Bear A, Pule M, Brenner MK, Rooney CM, et al. Enhanced tumor trafficking of GD2 chimeric antigen receptor T cells by expression of the chemokine receptor CCR2b. *J Immunother* 2010;33:780–8.
- [37] Di Stasi A, De Angelis B, Rooney CM, Zhang L, Mahendravada A, Foster AE, et al. T lymphocytes coexpressing CCR4 and a chimeric antigen receptor targeting CD30 have improved homing and antitumor activity in a Hodgkin tumor model. *Blood* 2009;113:6392–402.
- [38] Nishio N, Diaconu I, Liu H, Cerullo V, Caruana I, Hoyos V, et al. Armed oncolytic virus enhances immune functions of chimeric antigen receptor-modified T cells in solid tumors. *Cancer Res* 2014;74:5195–205.
- [39] Twumasi-Boateng K, Pettigrew JL, Kwok YYE, Bell JC, Nelson BH. Oncolytic viruses as engineering platforms for combination immunotherapy. *Nat Rev Cancer* 2018;18:419–32.
- [40] Adusumilli PS, Cherkassky L, Villena-Vargas J, Colovos C, Servais E, Plotkin J, et al. Regional delivery of mesothelin-targeted CAR T cell therapy generates potent and long-lasting CD4-dependent tumor immunity. *Sci Transl Med* 2014;6:261ra151.
- [41] Wang X, Popplewell LL, Wagner JR, Naranjo A, Blanchard MS, Mott MR, et al. Phase 1 studies of central memory-derived CD19 CAR T-cell therapy following autologous HSCT in patients with B-cell NHL. *Blood* 2016;127:2980–90.
- [42] Zhang L, Kerkar SP, Yu Z, Zheng Z, Yang S, Restifo NP, et al. Improving Adoptive T Cell Therapy by Targeting and Controlling IL-12 Expression to the Tumor Environment. *Molecular Therapy* 2011;19:751–9.
- [43] Avanzi MP, Yeku O, Li X, Wijewarnasuriya DP, van Leeuwen DG, Cheung K, et al. Engineered Tumor-Targeted T Cells Mediate Enhanced Anti-Tumor Efficacy Both Directly and through Activation of the Endogenous Immune System. *Cell Rep* 2018;23:2130–41.
- [44] Krenciute G, Prinzing BL, Yi Z, Wu M-F, Liu H, Dotti G, et al. Transgenic Expression of IL15 Improves Antiglioma Activity of IL13R $\alpha$ 2-CAR T Cells but Results in Antigen Loss Variants. *Cancer Immunol Res* 2017;5:571–81.



## **REVUE 5 - Les LT à récepteur antigénique chimérique (CAR-T) : principes physiopathologiques, actualités et perspectives**

***Vincent Alcazer***<sup>1,2</sup>, *Stéphane Depil*<sup>2,3</sup>

1 Service d'hématologie clinique, Centre hospitalier Lyon Sud, Pierre-Bénite

2 Centre de recherche en cancérologie de Lyon, INSERM U1052, Lyon

3 Centre Léon Bérard, Lyon

\* Correspondance :

Vincent Alcazer, Centre de Recherche en Cancérologie de Lyon, Lyon, France.

Email : [vincent.alcazer@lyon.unicancer.fr](mailto:vincent.alcazer@lyon.unicancer.fr)

Onko+. 29 sept 2019;11(89). French.



# CAR-T cells en hémato-oncologie

## LES LYMPHOCYTES T À RÉCEPTEUR ANTIGÉNIQUE CHIMÉRIQUE (CAR-T) : PRINCIPES PHYSIOPATHOLOGIQUES, ACTUALITÉS ET PERSPECTIVES



**Dr Vincent Alcazer**<sup>1,2</sup>,  
**Pr Stéphane Depil**<sup>2,3</sup>

<sup>1</sup>Service d'hématologie clinique, Centre hospitalier Lyon Sud, Pierre-Bénite

<sup>2</sup>Centre de recherche en cancérologie de Lyon, Inserm U1052, Lyon

<sup>3</sup>Centre Léon Bérard, Lyon

Les lymphocytes T à récepteur chimérique antigénique (CAR-T) ont révolutionné la prise en charge des hémopathies lymphoïdes B CD19+. Beaucoup d'efforts restent à faire pour généraliser les résultats à d'autres cancers et notamment aux tumeurs solides.

### Résumé

Les lymphocytes T à récepteur antigénique chimérique (CAR-T) ont bouleversé la prise en charge des hémopathies B CD19+. Ces récepteurs chimériques permettent de moduler la reconnaissance de la cible par le lymphocyte T ainsi que son activation, en surpassant différents mécanismes de résistance habituels tels que la perte du CMH de type I. Cet article fait le point sur ce nouveau type de thérapie cellulaire. Après un bref rappel du fonctionnement des lymphocytes T en situation normale et tumorale, le principe de construction des lymphocytes CAR-T y est abordé. Les données actuelles d'efficacité sont ensuite passées en revue, dans les hémopathies lymphoïdes B CD19+ d'abord, puis dans les autres hémopathies et les tumeurs solides. Les principales toxicités sont abordées, à la lumière des recommandations américaines actuelles. Enfin, les principales limites de cette approche sont discutées avant d'ouvrir sur les perspectives actuelles.

### Abstract

#### CAR-T cells in hematology-oncology

Chimeric Antigenic Receptor T cells (CAR T-cells) have demonstrated an outstanding efficacy in CD19+ lymphoid malignancies. Modulation of T-cells target recognition and activation can be modulated by these chimeric receptors, surpassing different tumor escape mechanisms such as MHC class I loss. In this article, we review the current knowledge on CAR T-cells therapy. Starting with a brief recall on normal T lymphocyte's functions in normal and oncologic contexts, we expose the main CAR T-cells construction principles. Current efficacy and safety data are then reviewed, in CD19+ B cells malignancies but also in other hematological and solid malignancies. Current challenges and perspectives are finally discussed.

### Introduction

Parmi les nombreuses thérapies innovantes en oncologie, les lymphocytes à récepteur antigénique chimérique (lymphocytes CAR-T) sont actuellement l'objet d'un intérêt particulier. Cet engouement se reflète dans l'abondance de

la littérature d'une part, les multiples études venant chaque semaine augmenter les connaissances actuelles sur le sujet, et sur le nombre croissant d'essais thérapeutiques d'autre part, avec plus de 200 essais recensés en 2017, en grande majorité dans les hémopathies lymphoïdes B (1).

Cet article fait le point sur ce qui est considéré actuellement comme une révolution thérapeutique dans les hémopathies. Après un bref rappel indispensable sur le fonctionnement des lymphocytes T normaux en contexte sain et tumoral, nous aborderons le fonctionnement même des



lymphocytes CAR-T et les principales données d'efficacité et de tolérance avant de finir sur un bref aperçu des perspectives dans le domaine.

## Principes

### Les lymphocytes T normaux

Pour bien comprendre le réel changement de paradigme introduit par les lymphocytes CAR-T, il convient de faire un bref rappel sur le fonctionnement des lymphocytes T (LT) normaux.

Cellule clé de l'immunité adaptative, le LT a besoin d'être activé par une cellule présentatrice d'antigène (CPA) avant de pouvoir exercer toute fonction. Les LT naïfs, présents dans les organes lymphoïdes secondaires (OLS), ont ainsi besoin de trois signaux pour être pleinement activés et pouvoir se différencier en cellules effectrices pouvant migrer en périphérie et exercer leur activité.

#### > Le premier signal

Le premier signal, constitué par l'antigène en lui-même, est le signal d'activation à proprement parler. Les LT ne peuvent pas reconnaître un antigène directement : celui-ci doit être présenté par la CPA *via* une molécule du complexe-majeur d'histocompatibilité (CMH). Le complexe antigène-CMH est alors reconnu par le récepteur du LT (TCR pour *T cell receptor*). Les LT CD8<sup>+</sup> reconnaissent les antigènes présentés par le CMH de type I, tandis que les LT CD4<sup>+</sup> reconnaissent les antigènes présentés par le CMH de type II.

#### > Le deuxième signal

Le deuxième signal, dit de costimulation, permet la survie et l'expansion des LT : les récepteurs de costimulation CD28 et 4-1BB situés à la surface des LT sont stimulés par les ligands

activateurs présents sur les cellules présentatrices d'antigènes, tels que ceux de la famille B7. À noter qu'une fois activé par le premier signal, le LT exprime de nombreux autres récepteurs de costimulation (comme ICOS par exemple), mais également des récepteurs inhibiteurs tels que CTLA4, reconnaissant les molécules de la famille B7 avec jusqu'à 20 fois plus d'affinité que CD28 et permettant le rétrocontrôle de l'activation des LT.

#### > Le troisième signal

Le troisième signal est un signal cytokinique permettant la différenciation des LT en cellules effectrices. Contrairement aux LT CD8<sup>+</sup> qui se différencient principalement en LT effecteurs cytotoxiques, les LT CD4<sup>+</sup> peuvent se différencier en de nombreuses sous-populations présentant des fonctions bien différentes. Une fois activés et différenciés en cellules effectrices, les LT n'ont plus besoin de signaux de costimulation pour agir. Ils restent toutefois restreints par le CMH, ne pouvant reconnaître que les cibles présentées par le CMH.

### Les réponses lymphocytaires T en contexte tumoral

Toutes ces étapes indispensables à l'activation des LT se retrouvent dans le contexte tumoral.

#### > Cycle de reconnaissance immunitaire

Daniel Chen et Ira Mellman ont ainsi décrit sept étapes clés indispensables

à la reconnaissance et la destruction d'une cellule tumorale par un LT CD8<sup>+</sup> cytotoxique, constituant le cycle de l'immunité anti-tumorale (2). Ainsi, pour être reconnue par un LT, une cellule tumorale doit :

1. émettre des antigènes tumoraux (= cibles),
2. qui seront captés par les CPA,
3. migrant alors vers les organes lymphoïdes secondaires pour pouvoir y activer les LT naïfs en y assurant la présence des trois signaux.
4. Les LT activés doivent ensuite migrer vers la tumeur par la circulation sanguine,
5. puis infiltrer la tumeur et son microenvironnement,
6. pour enfin reconnaître la cible présentée à la surface de la cellule tumorale sur une molécule de CMH,
7. et exercer leur activité cytotoxique.

#### > Phénotype de la tumeur

Ce cycle de reconnaissance immunitaire peut être défaillant à chacune de ces étapes, définissant ainsi le phénotype de la tumeur.

• **Les tumeurs dites froides**, non infiltrées, peuvent être des tumeurs « *immune-desert* » par défaillance des premières étapes (pas de cible, pas de capture par les CPA ou pas d'activation des LT dans les OLS) ou des tumeurs « *immune-excluded* » (défaut de pénétration des LT par restriction stromale).

• **Les tumeurs dites chaudes**, infiltrées, présentent des défaillances tardives du cycle : microenvironnement immunosuppresseur,

expression de *checkpoint* inhibiteurs inactivant les LT, etc.

Ce phénotype constitue un point majeur garant de l'efficacité de certaines thérapies immunologiques, une tumeur non infiltrée ayant peu de chance de répondre aux inhibiteurs de *checkpoint* immunitaire par exemple.

Au total, nous pouvons voir que la reconnaissance de la cellule tumorale par un LT est un processus long et complexe, sujet à de nombreuses possibilités d'échappement tumoral.

### Les lymphocytes T modifiés et CAR

#### > Les lymphocytes T modifiés

Afin de contourner ces limitations présentées par les LT "classiques", des LT avec des TCR modifiés ont été mis au point, permettant de ne plus être limité par les étapes précoces d'activation ainsi que la faible expression de la cible. Des résultats prometteurs ont été obtenus dans les tumeurs solides malgré des problèmes importants de toxicité rencontrés lors des premiers essais du fait de la *cross* réactivité avec les tissus normaux (3). La problématique de l'échappement tumoral par perte du CMH de type I reste néanmoins une limitation importante de ce type d'approche (4).

#### > Les CAR

La mise au point de récepteur antigénique chimérique (CAR pour *chimeric antigen receptor*) a permis également de surpasser les limitations liées à l'activation et à la présentation par le CMH.

#### Constitution

Les CAR sont constitués d'un domaine de reconnaissance de

### Les lymphocytes CAR-T combinent ainsi l'avantage de la reconnaissance antigénique forte et non restreinte par le CMH d'un anticorps avec les propriétés effectrices d'un lymphocyte T.

l'antigène correspondant à un fragment d'anticorps scFv (*single-chain variable fragment*) associé au domaine de signalisation du TCR (domaine CD3) (5). Contrairement au TCR reconnaissant uniquement des structures protéiques présentées par le CMH, les immunoglobulines peuvent reconnaître tout type de structure chimique et ne sont par ailleurs pas restreintes par le CMH. Des domaines de costimulation additionnels tels que CD28, CD137 (4-1BB) et OX40 ont par la suite été ajoutés dans les CAR de deuxième et troisième générations, permettant une amélioration de la survie et de la persistance des lymphocytes CAR-T activés (Fig. 1).

#### Principe

Le principe des lymphocytes CAR-T autologues consiste dans un premier temps à récupérer les LT du patient par aphérèse. Ces LT sont ensuite soumis à une expansion polyclonale (le plus souvent *via* des billes de stimulation CD3-CD28) avant d'être transduits avec le CAR d'intérêt (différentes méthodes pouvant être employées, la transfection lentivirale ou l'utilisation de rétrotransposon constituant les deux techniques principales utilisées actuellement). Les lymphocytes CAR-T obtenus sont ensuite réinjectés au patient, après lymphodéplétion préalable par chimiothérapie (en général fludarabine et cyclophosphamide), permettant une expansion et un maintien optimal des lymphocytes CAR-T par promotion de l'expansion homéostatique.

#### Avantages et inconvénients

Les lymphocytes CAR-T combinent ainsi l'avantage de la reconnaissance antigénique forte et non restreinte par le CMH d'un anticorps avec les propriétés effectrices d'un lymphocyte T. De nombreuses voies d'échappement tumoral, avec en premier lieu la perte du CMH, mécanisme majeur d'évasion immunitaire, ne sont donc plus un obstacle avec une telle approche (6). Ces avantages sont tout de même accompagnés de certains inconvénients, les CAR se limitant à la reconnaissance d'antigènes de surface extracellulaires et ne pouvant plus reconnaître de cible antigénique intracellulaire.

#### Efficacité

##### Les CAR CD19+ dans les hémopathies lymphoïdes B

Si les lymphocytes CAR-T présentent de nombreux avantages théoriques, le véritable engouement actuel fait suite aux données d'efficacité observée chez les patients traités pour des hémopathies lymphoïdes B CD19+.

#### > Approbation par la FDA et l'EMA

Après des résultats décevants observés avec la première génération de CAR ne possédant pas de domaine de costimulation, la deuxième génération de CAR aura finalement tenu ses promesses, menant à l'approbation par la FDA en 2017 puis par l'EMA de deux produits :

- d'un côté, le **tisagenlecleucel**

(« *Tisa-cel* », Kymriah®, Novartis), possédant un domaine de stimulation 4-1BB et approuvé dans les leucémies aiguës lymphoblastiques (LAL) B en rechute ou réfractaire (R/R) jusqu'à 25 ans et dans les lymphomes malins non hodgkiniens (LNH) B R/R ;

• de l'autre côté, l'**axicabtagene ciloleucel** (« *Axi-cel* », Yescarta®, Kite Pharma), possédant un domaine de costimulation CD28 et approuvé pour les LNH B R/R.

### > Résultats

C'est avant tout dans les LAL B CD19+ que les résultats se sont révélés très prometteurs, avec des taux de réponses complètes de l'ordre de 80 % et une survie globale de 76 % à 1 an pour le *Tisa-cel* (7). Au-delà des taux de réponse en eux-mêmes, c'est surtout la population ayant bénéficié du traitement, comprenant des patients multi-traités et/ou réfractaires, qui souligne le réel apport des CAR dirigés contre CD19.

### > Les études princeps

L'approbation des CAR CD19+ a par la suite été confortée par deux études princeps, **ZUMA-1** pour l'*Axi-cel* et **JULIET** pour le *Tisa-cel*, qui ont évalué l'efficacité des CAR-T chez des patients atteints de LNH B R/R. Les taux de réponse observés ont alors été confirmés récemment par des études de suivi à long terme, avec 83 % de réponses objectives, dont 58 % de réponses complètes et une médiane de survie globale non atteinte pour le suivi de l'essai **ZUMA-1**, et 52 % de réponses objectives, dont 40 % de réponses complètes pour l'essai **JULIET** (8, 9).

Outre les différences concernant le domaine de costimulation, une différence majeure de design dans ces deux études pourrait expliquer ces différences d'efficacité : les patients

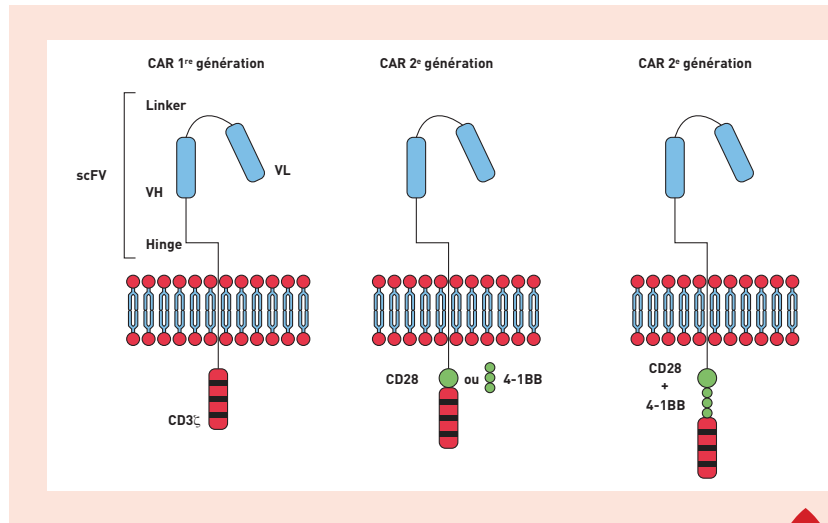


Figure 1

### Structure des récepteurs antigéniques chimériques CAR.

Les CAR correspondent à des protéines de fusion comprenant un domaine extracellulaire consistant en un fragment d'anticorps (scFV) reconnaissant un antigène membranaire cible présent sur la cellule tumorale, associé à un domaine intracellulaire provenant de protéines de signalisation des lymphocytes T. La première génération de CAR comprenait uniquement un domaine CD3ζ. Un domaine de costimulation (CD28 ou 4-1BB) a été fusionné au CD3ζ dans les CAR de deuxième génération. La troisième génération comprend deux domaines de costimulation liés à CD3ζ.

de l'essai **JULIET** pouvaient recevoir une chimiothérapie d'attente (*bridge*), avec un délai médian de délivrance des CAR-T de 54 jours. Cette chimiothérapie n'était pas permise dans l'essai **ZUMA-1**, avec un délai médian de délivrance des CAR-T de 17 jours permettant une infusion relativement rapide des CAR-T. Les patients ayant besoin d'une telle thérapie étaient ainsi exclus de **ZUMA-1**, rendant les deux cohortes difficilement comparables.

### > Données de vie réelle

Point important, ces données d'efficacité ont été confirmées sur une étude de vie réelle présentée au congrès 2018 de l'association américaine d'hématologie, avec des chiffres très proches de ceux observés dans les essais (10). Outre le taux de réponse, le point majeur est la durée de réponse chez les patients

en rémission complète, ce qui n'a jamais été observé auparavant.

### Les autres cibles dans les hémopathies

Devant les réponses obtenues chez des patients réfractaires ou lourdement prétraités, de nouvelles cibles ont été recherchées pour étendre l'utilisation des CAR à d'autres pathologies.

### > Les CAR anti-BCMA dans le myélome

Les CAR anti-BCMA (*B cell maturation antigen*) sont ainsi en cours de développement dans le myélome, avec des résultats préliminaires encourageants (11). Deux des patients traités à la plus forte dose de  $9 \times 10^6$  cellules/kg présentaient notamment une maladie réfractaire avec une infiltration médullaire par plus de 80 % de plasmocytes malins. Une infusion

de lymphocytes CAR-T anti-BCMA a permis l'obtention d'une rémission complète stringente dans les deux cas. Si l'un des patients est toujours en très bonne réponse partielle après 28 semaines de traitement, l'autre a maintenu sa réponse pendant 17 semaines avant de rechuter. À la différence des LMNH, des rechutes sont observées de manière plus fréquente, suggérant un contexte et un environnement immunitaire différents dans le myélome.

**> Les CAR anti-CD22 dans les leucémies aiguës lymphoblastiques**

Des CAR dirigés contre CD22 ont également été évalués dans la LAL avec 73 % de rémissions complètes observées chez 15 patients (12).

**> Le cas des hémopathies myéloïdes**

La situation semble plus complexe dans le cas des pathologies myéloïdes, notamment des leucémies aiguës myéloïdes (LAM). La sélection d'une cible pertinente, suffisamment exprimée par les cellules leucémiques (et notamment les cellules souches leucémiques) et non exprimée par le tissu sain pour limiter les toxicités dites *on-target/off-tumor*, reste pour le moment problématique.

Plusieurs essais sont actuellement en cours concernant les cibles classiques que sont CD33 et CD123. Le cas d'un patient présentant une LAM R/R ayant reçu un traitement par lymphocytes CAR-T dirigés contre le CD33 a été rapporté : l'infusion s'est compliquée d'un syndrome de relargage cytokinique majeur ainsi que d'une pancytopenie prolongée. Malgré une nette diminution du pourcentage de blastes médullaires après infusion, une augmentation progressive a par la suite été observée avec une rechute de la maladie à 9 semaines (13).

D'autres cibles non exprimées ou exprimées à plus faible taux par la

cellule souche hématopoïétique normale telles que CLL1 (CLEC12A) ou CD44v6 ont démontré des résultats intéressants dans des modèles pré-cliniques et sont toujours en cours d'investigation dans des essais de phase I (14).

**Les tumeurs solides**

**> Problématiques**

Le succès rencontré dans les hémopathies lymphoïdes B ne s'est pour l'instant pas exporté aux tumeurs solides. Outre la problématique de la cible, de nombreux facteurs contribuent à limiter l'efficacité des lymphocytes CAR-T dans ces situations :

- difficulté à trouver une cible spécifique (risque de toxicité *on-target/off-tumor* limitante (voir toxicités et limites)),
- hétérogénéité d'expression,
- environnement immunosuppresseur,
- difficultés de *homing* et de pénétration des CAR-T dans les tissus tumoraux.

**> Principaux résultats**

Nous avons récemment passé en revue les différents résultats obtenus à ce jour avec les CAR dans les tumeurs solides (15).

**Les CAR anti-EGFR dans le cancer du poumon non à petites cellules et les cancers des voies biliaires**

Parmi les principaux essais, les CAR dirigés contre le récepteur à l'EGF (EGFR) dans le cancer du poumon non à petites cellules et les cancers des voies biliaires ont montré des premiers résultats :

- une rémission complète et 10 maladies stables (SD) sur 17 patients  **dans les cancers des voies biliaires,**
- deux réponses partielles et cinq SD sur 11 patients évalués  **dans le cancer du poumon non à petites cellules** (16, 17).

Le profil de tolérance était par ailleurs acceptable avec une toxicité cutanéomuqueuse n'excédant pas le grade 1/2.

**Les CAR anti-CEA dans les tumeurs hépatiques et digestives**

Des CAR dirigés contre CEA ont été évalués dans les tumeurs hépatiques et digestives, avec des stabilisations sans réponses objectives observées. Il faut noter cependant l'absence d'expansion durable des lymphocytes CAR-T injectés dans ce contexte (18).

**Les CAR anti-PSMA**

Deux CAR de seconde génération anti-PSMA (domaine de costimulation CD28) ont été évalués en phase I. Dans l'essai du Memorial Sloan-Kettering Cancer Center (MSKCC), une SD de 6 mois a pu être obtenue chez deux patients sur quatre (19). L'étude du Roger William Medical Center a quant à elle mis en évidence deux réponses partielles sur cinq patients évalués (20).

**> Autres cibles en cours d'évaluation**

De nombreuses autres cibles telles que GD2, glypican 3, mésothéline, WT1 sont actuellement en cours d'essai dans des phases précliniques ou cliniques de phase I (14, 15). Des résultats prometteurs obtenus avec l'administration intra-tumorale de CAR-T anti-mésothéline dans le mésothéliome ont également été présentés cette année à l'ASCO (21).

**Toxicités**

**Le syndrome de relargage cytokinique (CRS pour cytokine release syndrome)**

**> Synthèse des cytokines**

L'engagement du CAR entraîne la

**Tableau 1 - Échelle de cotation du syndrome de relargage cytokinique d'après (25).**

À noter qu'il n'existe pas de seuil consensuel pour l'hypoxémie, les auteurs considérant pour le CRS le « besoin d'oxygène pour supplémenter un défaut d'oxygénation et non par simple confort ».

Grade	Grade 1	Grade 2	Grade 3	Grade 4
<b>Fièvre</b>	≥ 38°C	≥ 38°C	≥ 38°C	≥ 38°C
<b>Hypotension (pression artérielle systolique &lt; 90 mmHg)</b>	Absence	Ne nécessitant pas de vasopresseurs	Nécessitant vasopresseurs à fortes doses	Nécessitant plusieurs vasopresseurs
<b>Hypoxémie</b>	Absence	Nécessitant oxygène ≤ 6 l/min	Nécessitant oxygène > 6 l/min	Nécessitant ventilation mécanique

synthèse de multiples cytokines par les lymphocytes T activés, telles que l'IL-2, IL-6, IFN $\gamma$  et GM-CSF, ainsi que par les cellules immunitaires de l'environnement avec notamment la synthèse d'IL-1 et d'IL-6 par les monocytes et les macrophages, qui semblent jouer un rôle majeur dans l'initiation du CRS (22, 23).

#### > Symptômes

Ce relargage cytokinique s'accompagne de symptômes généraux non spécifiques associant à des degrés divers :

- fièvre,
- frissons,
- myalgies,
- tachycardie.

Dans ses formes les plus sévères, le CRS peut toucher n'importe quel organe, allant jusqu'à **la défaillance multiviscérale**. Des phénomènes de **fuite capillaire** sont également observés, liés à une atteinte de l'endothélium, les cellules endothéliales pouvant là encore être source d'IL-6 (24).

#### > Gradation du CRS

Des critères consensuels pour grader le CRS ont été récemment établis par

la société américaine de transplantation et de thérapie cellulaire et sont repris dans le **tableau 1** (25). À noter que dans cette classification, il est précisé que les toxicités organiques associées peuvent accompagner le CRS, mais ne participent pas à sa cotation.

#### > Épidémiologie

Le CRS s'installe généralement rapidement après l'infusion des CAR, entre 1 et 14 jours après. Les fortes masses tumorales, les comorbidités, et l'utilisation de CAR à domaine de costimulation CD28 tendraient à favoriser le CRS, sans que cela soit pour le moment clairement démontré (22). L'incidence du CRS est variable, de l'ordre de 93 % dans l'essai **ZUMA-1** (11 % de grade 3-4) et 58 % dans l'essai **JULIET** (22 % de grade 3-4) (8, 9). Ces taux étaient retrouvés dans les différents essais concernant les tumeurs solides (15).

#### > Traitements

En cas de CRS grade 1, des soins de support seuls sont préconisés. Ce n'est qu'en cas de fièvre persistante ou de CRS grade 2 qu'un traitement est recommandé. Dans la majorité

des cas, le CRS répond à l'injection de **tocilizumab**, anticorps dirigé contre le récepteur de l'IL-6, et les symptômes sont complètement réversibles. En ciblant directement l'IL-6 soluble avec une affinité extrêmement forte, le **siltuximab** pourrait constituer une alternative intéressante au tocilizumab (22). Les **glucocorticoïdes** constituent également une option thérapeutique du fait de leur effet anti-inflammatoire, et sont réservés aux cas réfractaires au traitement de première ligne par tocilizumab en raison de leurs potentiels effets immunosuppresseurs sur les lymphocytes CAR-T (bien que les résultats obtenus ne suggèrent pas d'effet délétère de leur utilisation sur l'efficacité).

#### La toxicité neurologique

Si le CRS était un événement plutôt attendu, les complications neurologiques observées par la suite l'étaient moins.

#### > Symptômes

Ces toxicités, bien que réversibles dans la plupart des cas, peuvent entraîner des symptômes plus ou moins sévères (agitation, aphasie,



**Tableau 2 - Échelle de cotation des symptômes neurologiques liés aux lymphocytes CAR-T (ICANS) d'après (25).**

Le grade d'ICAN est déterminé par l'événement le plus sévère.

	Grade 1	Grade 2	Grade 3	Grade 4
Score ICE	7-9	3-6	0-2	0
Conscience	Spontanément réveillable	Réveillable à la voix	Réveillable aux stimuli tactiles	Non réveillable, stupeur, coma
Convulsions			Focalisée ou généralisée avec résolution rapide Tracé EEG sans convulsion, résolutif après traitement	Convulsions prolongées > 5 min menaçant le pronostic vital
Moteur				Hémi-parésie ou paraparésie
Pression intra-crânienne élevée ou œdème cérébral			Œdème focalisé sur l'imagerie	Œdème cérébral diffus, œdème papillaire, posture de décortication

encéphalopathie...) menant dans certains cas au décès. Elles sont regroupées sous le terme de **syndrome de neurotoxicité associé aux cellules immunitaires effectrices** (ou ICANS pour *immune effector cells associated neurotoxicity syndrome*) et peuvent survenir de façon concomitante ou, plus fréquemment, après le CRS (22, 25). Les symptômes les plus précoces regroupent une dysgraphie, puis une aphasie, avec en particulier une difficulté à nommer les objets, mais les cas les plus graves peuvent également survenir sans aucun signe avant-coureur. Il semblerait que les complications neurologiques soient plus fréquentes en cas de LAL, de forte infiltration médullaire, de présence d'un CRS et/ou d'antécédents neurologiques (26). Les atteintes diffèrent également selon les produits (25).

**> Physiopathologie des ICANS**

La physiopathologie des ICANS est encore mal comprise. Une activation endothéliale menant à une rupture

de la barrière hémato-encéphalique et au passage des lymphocytes CAR-T a pu être démontrée par plusieurs équipes, retrouvant par ailleurs la présence des lymphocytes CAR-T dans le liquide céphalo-rachidien en l'absence de toute cellule tumorale (26). Une équipe italienne a récemment mis au point un modèle murin récapitulant la physiopathologie du CRS et de l'atteinte neurologique en cas d'infusion de CAR dirigé contre CD19 (23). Norelli *et al.* ont ainsi montré que les monocytes étaient la principale source d'IL-6 et d'IL-1 pendant le CRS, et que si l'utilisation de tocilizumab prévenait le CRS, celui-ci ne prévenait pas la neurotoxicité. L'utilisation d'anakinra en revanche, un anti-IL-1, a permis une réversibilité du CRS et de la neurotoxicité, ouvrant une nouvelle voie thérapeutique prometteuse.

**> Évolution et gradation**

Une fois installées, les atteintes neurologiques durent généralement de 2 à 4 jours, mais peuvent

durer de quelques heures à quelques semaines (22). Dans la grande majorité des cas, ces symptômes sont spontanément réversibles. Une harmonisation de la gradation des symptômes a été proposée par la société américaine de transplantation et de thérapie cellulaire en se fondant sur les critères CARTOX (Tab. 2) (25).

**> Suivi et traitements**

Un outil de suivi et dépistage des atteintes neurologiques est également proposé (CARTOX-10), reprenant de nombreux éléments du MMSE. Un examen neurologique quotidien de dépistage est ainsi indispensable, avec un examen approfondi (si possible par un neurologue) en cas d'apparition de tout symptôme neurologique.

Dès l'apparition de symptômes d'ICANS de grade 1, un **traitement symptomatique** sera de mise en évitant les traitements pouvant favoriser une dépression du système nerveux central. La réalisation



d'une IRM cérébrale avec et sans injection de produit de contraste, d'un fond d'œil à la recherche d'un œdème papillaire, d'une ponction lombaire avec mesure de la pression d'ouverture et d'un électroencéphalogramme (EEG) de 30 minutes sont recommandés (22). Un traitement par **lorazépam** ou **lévétiracétam** est préconisé en cas de tracé épileptique sans convulsions associées. Un traitement par **tocilizumab** sera débuté en cas de CRS associé. Un traitement par **dexaméthasone** ou **méthylprednisolone** sera proposé en l'absence de CRS ou en cas de CRS réfractaire au tocilizumab.

### Les toxicités on-target/off-tumor

En l'absence de cible entièrement spécifique de la tumeur, les toxicités dites *on-target/off-tumor* liées à l'expression de la cible en dehors de la tumeur restent une préoccupation.

#### > Cas de CD19

Dans le cas de CD19, l'infusion et la persistance des CAR s'accompagnent principalement d'une lymphopénie B prolongée. L'absence de CD19 sur les plasmocytes permet de conserver les cellules mémoires matures productrices d'immunoglobulines, permettant de garder une mémoire humorale couvrant notamment les cibles vaccinales telles que le tétanos. Ces propriétés d'expression de l'antigène CD19 participent en grande partie au bon profil de tolérance des CAR anti-CD19.

#### > Expression de la cible en dehors de la tumeur

La vigilance s'impose toutefois face à une expression *a priori* restreinte d'une cible : le premier essai évaluant un CAR de troisième génération dirigé contre ERBB2/HER2 a

**Il semblerait que les complications neurologiques soient plus fréquentes en cas de LAL, de forte infiltration médullaire, de présence d'un CRS et/ou d'antécédents neurologiques.**

dû être stoppé suite au décès d'un patient traité pour cancer colique et ayant présenté une détresse respiratoire. La mise en évidence des lymphocytes CAR-T dans les poumons a permis de montrer une faible expression de ERBB2 dans ce tissu (27). D'une façon similaire, l'essai de phase I évaluant un CAR dirigé contre CAIX (*carbonic anhydrase IX*) dans le carcinome rénal s'est accompagné d'une hépatotoxicité inattendue, liée à une faible expression de CAIX au niveau de l'épithélium biliaire normal (28).

### Limites

Quelle que soit l'efficacité des lymphocytes CAR-T, de nombreux obstacles demeurent.

#### La cible

La problématique de la cible est au premier plan : pour être reconnu, l'antigène doit être présent à la surface de la cellule tumorale. Impossible donc de cibler les antigènes intracellulaires. Par ailleurs, le succès rencontré dans les hémopathies lymphoïdes B est en grande partie due à l'excellente cible représentée par CD19, à la fois partagée par l'ensemble du contingent tumoral, mais également restreinte à la lignée B, limitant les toxicités *on-target/off-tumor*. Néanmoins, le fait d'augmenter la pression immunitaire sur une cible unique expose aux rechutes avec perte de la cible, ouvrant la voie à des CARs ciblant plusieurs cibles.

#### Les rechutes et résistances induites

Malgré l'efficacité souvent mise au premier plan, les rechutes restent un problème important.

#### > Perte de la cible

La perte de la cible, comme cela est bien décrit maintenant dans le cas des LAL CD19+, est un des premiers mécanismes d'échappement mis en place par la tumeur, soulignant la nécessité de développer de nouvelles approches complémentaires et de nouvelles cibles. Plusieurs mécanismes ont été décrits pouvant expliquer cette perte :

- présence initiale et sélection d'un sous-clone CD19 négatif,
- mutations additionnelles et émergence d'un variant issu d'un épissage alternatif de CD19,
- changement de lignée du lymphoblaste vers la lignée myéloïde.

Plus récemment, des phénomènes de trogocytose de CD19 par les lymphocytes CAR-T ont été décrits (capture de la cible par les cellules T), pouvant expliquer les rechutes avec un clone exprimant faiblement CD19 (29).

#### > Absence d'expansion ou de persistance des lymphocytes CAR-T

L'absence d'expansion ou de persistance des lymphocytes CAR-T pose également problème dans certains cas. De nombreux travaux ont permis de faire la lumière récente sur le sujet, permettant de définir des différences d'expansion et

de persistance selon le domaine de costimulation (30). Le phénotype des lymphocytes modifiés pourrait jouer également un rôle important, avec des approches sélectionnant les sous-types de lymphocytes (ratio CD4/CD8 contrôlé) ou réalisant un enrichissement en lymphocytes T de type *stem cell memory* pour une persistance au long cours (31).

**> Perte du CAR**

La perte du CAR par absence d'expansion ou destruction *via* notamment anticorps anti-scFV murin peut également consister un des facteurs limitant l'efficacité thérapeutique (32).

**Le délai d'infusion**

Une des limitations actuelles majeures des lymphocytes CAR-T est la durée de fabrication. Les usines de production étant pour le moment encore localisées aux États-Unis, le temps écoulé entre le prélèvement du patient par aphérèse et l'infusion du produit fini varie entre 4 et 8 semaines. Ce délai devrait être réduit avec l'arrivée d'usines de production directement en Europe.

Cet effet est d'autant plus problématique que les indications actuelles de ces produits sont limitées aux patients en rechute ou réfractaires, pour lesquels une thérapie d'attente n'est pas toujours évidente à maintenir.

**Perspectives**

**L'affinité du CAR et les cibles**

**> Modification de l'affinité du CAR pour sa cible**

La modification de l'affinité du CAR pour sa cible pourrait permettre de réintroduire les antigènes associés à

la tumeur, avec leur avantage d'être partagés. Ainsi, une équipe américaine a mis au point un CAR ciblant EGF-R avec une faible affinité. Le CAR ciblait le tissu tumoral possédant une forte expression d'EGF-R, épargnant ainsi le tissu sain où l'expression y était plus faible (33).

**> Autres paramètres à moduler**

Outre l'affinité, de nombreux paramètres tels que la structure transmembranaire ou le domaine de signalisation peuvent moduler la fonction du CAR et sont actuellement source de multiples études (34).

**> Cibles multiples**

Le risque lié à l'utilisation d'une cible tumorale unique peut être surpassé par l'emploi de cibles multiples. Les CAR bispécifiques émergent ainsi, permettant de réduire la pression immunitaire sur cible unique. Les CAR CD19-CD20 ou CD19-CD22 ont ainsi montré un intérêt pour répondre à la problématique des rechutes CD19 négatives dans les LAL B (35).

**Les CAR-T avec gates logiques**

**> Les CAR bispécifiques**

La nécessité d'un signal de costimulation pour activer les LT normaux est une sécurité qui n'existe plus sur les lymphocytes CAR-T. L'ajout d'un second CAR sur un même lymphocyte T peut permettre de restaurer un point de contrôle. Des lymphocytes CAR-T ne s'activant qu'en cas de reconnaissance de la double cible ont pu ainsi être mis au point, permettant de gagner en spécificité. Wilkie *et al.* ont ainsi mis au point un CAR bispécifique ciblant HER2 et Muc1, deux antigènes exprimés dans plusieurs tumeurs solides. La partie HER2 contient le domaine

CD3 epsilon, et la partie Muc1, le signal CD28 de costimulation. Ainsi, le CAR ne peut être pleinement activé qu'en cas de coexistence des deux cibles (36).

**> Utilisation d'un récepteur**

**Notch synthétique**

Une autre approche repose sur l'utilisation d'un récepteur Notch synthétique (synNotch) (37, 38). La fixation à un premier antigène A induit le clivage du domaine Notch contenant un activateur transcriptionnel qui est transloqué au noyau. Ce signal active la transcription du CAR se liant à l'antigène B, induisant, lui, l'activité cytotoxique.

**La gate NOT**

Dans le cas d'une « gate NOT », il y a combinaison entre un récepteur CAR activateur reconnaissant A et un récepteur CAR inhibiteur (dont la région intracellulaire est dérivée d'un récepteur inhibiteur comme PD-1 ou CTLA-4) reconnaissant B (antigène exprimé par une cellule normale). Le lymphocyte T s'activera uniquement en cas de fixation à l'antigène A, protégeant ainsi les cellules normales exprimant B.

**La gate OR**

Enfin, dans une « gate OR », le lymphocyte T possède deux récepteurs CAR reconnaissant A et B, permettant à la cellule T de s'activer en cas de reconnaissance de A et/ou de B, ce qui permet de contrecarrer la résistance liée à l'absence d'expression d'une cible en cas d'hétérogénéité tumorale.

**Les CAR allogéniques**

**> Avantages**

Bien que le développement des CAR selon une modalité auto-logue permette de s'affranchir des

problématiques d'alloréactivité, la possibilité de produire des lymphocytes CAR-T allogéniques prêts à l'emploi pourrait permettre de bénéficier plus rapidement du traitement. Le délai actuel d'infusion reste en effet une problématique importante, particulièrement pour les populations déjà pré-traitées présentant une maladie active. Produits à partir de lymphocytes de donneurs sains, les lymphocytes CAR-T allogéniques promettent également de passer outre de nombreux mécanismes d'échappement tumoral liés notamment à l'épuisement de la réponse T chez le patient.

#### > Les approches envisagées

Plusieurs approches sont envisagées :

- utilisation de cellules T mémoires EBV-spécifiques pour éviter le risque de réaction du greffon contre l'hôte,
- utilisation de CAR NK ou iNKT, de lymphocytes gamma delta, ou de T alpha bêta délétés pour leur TCR.

Une preuve de concept a été apportée avec un lymphocyte CAR-T allogénique TCR KO dirigé contre CD19 chez deux enfants présentant une LAL R/R (39). L'obtention d'une rémission moléculaire chez les deux patients a permis la réalisation d'une greffe allogénique par la suite.

Les résultats préliminaires avec un CAR T anti-CD19 TCR KO (UCART19, Cellectis/Servier) montrent un effet anti-leucémique, mais la durée de réponse et la persistance des CAR T allogéniques administrés restent une question majeure. Il a été proposé de réaliser un KO de la bêta 2 microglobuline pour inhiber l'expression des molécules HLA de classe I et ainsi réduire la réactivité des cellules T du receveur à l'encontre des CAR-T

### La possibilité de produire des lymphocytes CAR-T allogéniques prêts à l'emploi pourrait permettre de bénéficier plus rapidement du traitement.

injectés. Néanmoins, la reconnaissance des CAR-T allogéniques n'exprimant pas les molécules de classe I, par les lymphocytes NK du patient, pose dans ce cas un problème (activité des NK à l'encontre de cellules HLA I négatives selon le modèle de *self-missing*). Il a été proposé d'insérer en plus un inhibiteur de NK comme HLA-E afin de prévenir la lyse par les NK (40).

#### CAR-T modifiés, gène suicide et inactivation

L'avènement des modifications génétiques rendues possibles notamment par les technologies telles que CRISPR-CAS9 ou TALEN élargit encore les domaines du possible concernant le développement des CAR.

#### > Gène suicide

Les lymphocytes CAR-T peuvent être transduits avec un gène suicide inductible tel que celui de la caspase 9 (iCasp9), qui est activé par l'administration de l'AP1903 induisant l'apoptose des lymphocytes modifiés (34).

#### > Modification du tissu sain

Une autre approche plus originale dans les hémopathies myéloïdes réside dans la modification du tissu sain. Une équipe a ainsi montré la faisabilité dans un modèle murin de la suppression de l'expression du CD33 par la moelle saine afin de rendre la cible spécifique de la LAM (41). L'équipe de Michel Sadelain au MSKCC a également proposé l'introduction du CAR directement

dans le locus du TCR, permettant un contrôle physiologique de l'expression du CAR et ainsi une amélioration de sa fonction (42).

### Conclusion

Les lymphocytes CAR-T ont révolutionné la prise en charge des hémapathies CD19+, expliquant l'entrainement actuel pour ce type d'approche. Si beaucoup d'avancées ont été réalisées depuis le début de leur développement, de nombreux obstacles persistent actuellement. La généralisation de ce type d'approche à d'autres cancers et notamment aux tumeurs solides demandera ainsi de multiples efforts de recherche à différents niveaux, depuis la définition de cibles spécifiques limitant les toxicités *off-tumor*, jusqu'à la production de CAR de plus en plus performants. ■

✳ Vincent Alcazer déclare ne pas avoir de lien d'intérêt. Stéphane Depil a été employé de Cellectis et a assuré des activités de consultance ou des prestations pour : AstraZeneca, BMS, Servier, Elsalys, Erytech, PDC\*line Pharma, Epivax Oncology, Netris Pharma, Celyad.

### Mots-clés

Lymphocytes, CAR-T, Récepteur antigénique chimérique, Thérapie cellulaire, Hémopathies, Tumeurs solides

### Keywords

T cells, CAR-T, Chimeric antigenic receptor, Cells therapy, Hematological malignancies, Solid malignancies



**Bibliographie**

1. Hartmann J, Schüßler-Lenz M, Bondanza A, Buchholz CJ. Clinical development of CAR T cells-challenges and opportunities in translating innovative treatment concepts. *EMBO Mol Med* 2017 ; 9 : 1183-97.
2. Chen DS, Mellman I. Oncology meets immunology: the cancer-immunity cycle. *Immunity* 2013 ; 39 : 1-10.
3. Linette GP, Stadtmauer EA, Maus MV et al. Cardiovascular toxicity and titin cross-reactivity of affinity-enhanced T cells in myeloma and melanoma. *Blood* 2013 ; 122 : 863-71.
4. Cameron BJ, Gerry AB, Dukes et al. Identification of a Titin-derived HLA-A1-presented peptide as a cross-reactive target for engineered MAGE A3-directed T cells. *Sci Transl Med* 2013 ; 5 : 197ra103.
5. June CH, O'Connor RS, Kawalekar OU et al. CAR T cell immunotherapy for human cancer. *Science* 2018 ; 359 : 1361-5.
6. Garrido F, Aptsiauri N, Doorduijn EM et al. The urgent need to recover MHC class I in cancers for effective immunotherapy. *Curr Opin Immunol* 2016 ; 39 : 44-51.
7. Maude SL, Laetsch TW, Buechner J et al. Tisagenlecleucel in children and young adults with B-cell lymphoblastic leukemia. *N Engl J Med* 2018 ; 378 : 439-48.
8. Locke FL, Ghobadi A, Jacobson CA et al. Long-term safety and activity of axicabtagene ciloleucel in refractory large B-cell lymphoma (ZUMA-1): a single-arm, multi-centre, phase 1-2 trial. *Lancet Oncol* 2019 ; 20 : 31-42.
9. Schuster SJ, Bishop MR, Tam CS et al. Tisagenlecleucel in adult relapsed or refractory diffuse large B-cell lymphoma. *N Engl J Med* 2019 ; 380 : 45-56.
10. Jacobson CA, Hunter B, Armand P et al. Axicabtagene ciloleucel in the real world: outcomes and predictors of response, resistance and toxicity. *Blood* 2018 ; 132 : 92-2.
11. Ali SA, Shi V, Maric I et al. T cells expressing an anti-B-cell maturation antigen chimeric antigen receptor cause remissions of multiple myeloma. *Blood* 2016 ; 128 : 1688-700.
12. Fry TJ, Shah NN, Orentas RJ et al. CD22-targeted CAR T cells induce remission in B-ALL that is naive or resistant to CD19-targeted CAR immunotherapy. *Nat Med* 2018 ; 24 : 20-8.
13. Wang Q, Wang Y, Lv H et al. Treatment of CD33-directed chimeric antigen receptor-modified T cells in one patient with relapsed and refractory acute myeloid leukemia. *Mol Ther* 2015 ; 23 : 184-91.
14. Hofmann S, Schubert M-L, Wang L et al. Chimeric antigen receptor (CAR) T cell therapy in acute myeloid leukemia (AML). *J Clin Med* 2019 ; 8 : 200.
15. Alcazer V, Delenda C, Poirrot L, Depil S. Développement des CAR-T dans les tumeurs solides. *Bull Cancer* 2018 ; 105 : S15-24.
16. Guo Y, Feng K, Liu Y et al. Phase I study of chimeric antigen receptor-modified T cells in patients with EGFR-positive advanced biliary tract cancers. *Clin Cancer Res* 2018 ; 24 : 1277-86.
17. Feng K, Guo Y, Dai H et al. Chimeric antigen receptor-modified T cells for the immunotherapy of patients with EGFR-expressing advanced relapsed/refractory non-small cell lung cancer. *Sci China Life Sci* 2016 ; 59 : 468-79.
18. Katz SC, Burga RA, McCormack E et al. Phase I Hepatic immunotherapy for metastases study of intra-arterial chimeric antigen receptor-modified T-cell therapy for CEA+ liver metastases. *Clin Cancer Res* 2015 ; 21 : 3149-59.
19. Slovin SF, Wang X, Hullings M et al. Chimeric antigen receptor (CAR+) modified T cells targeting prostate-specific membrane antigen (PSMA) in patients (pts) with castrate metastatic prostate cancer (CMPC). *J Clin Oncol* 2013 ; 31 : 72.
20. Junghans RP, Ma Q, Rathore R et al. Phase I trial of anti-PSMA designer CAR-T cells in prostate cancer: possible role for interacting interleukin 2-T cell pharmacodynamics as a determinant of clinical response. *Prostate* 2016 ; 76 : 1257-70.
21. Adusumilli PS, Zauderer mg, Rusch VW et al. Abstract CT036: A phase I clinical trial of malignant pleural disease treated with regionally delivered autologous mesothelin-targeted CAR T cells: Safety and efficacy. *Cancer Res* 2019 ; 79 : CT036.
22. Neelapu SS, Tummala S, Kebriaei P et al. Chimeric antigen receptor T-cell therapy - assessment and management of toxicities. *Nat Rev Clin Oncol* 2018 ; 15 : 47-62.
23. Norelli M, Camisa B, Barbiera G et al. Monocyte-derived IL-1 and IL-6 are differentially required for cytokine-release syndrome and neurotoxicity due to CAR T cells. *Nat Med* 2018 ; 24 : 739-48.
24. Obstfeld AE, Frey NV, Mansfield K et al. Cytokine release syndrome associated with chimeric-antigen receptor T-cell therapy: clinicopathological insights. *Blood* 2017 ; 130 : 2569-72.
25. Lee DW, Santomasso BD, Locke FL et al. ASTCT consensus grading for cytokine release syndrome and neurologic toxicity associated with immune effector cells. *Biol Blood Marrow Transplant* 2019 ; 25 : 625-38.
26. Gust J, Hay KA, Hanafi L-A et al. Endothelial activation and blood-brain barrier disruption in neurotoxicity after adoptive immunotherapy with CD19 CAR-T cells. *Cancer Discov* 2017 ; 7 : 1404-19.
27. Morgan RA, Yang JC, Kitano M et al. Case report of a serious adverse event following the administration of T cells transduced with a chimeric antigen receptor recognizing ERBB2. *Mol Ther* 2010 ; 18 : 843-51.
28. Lamers CH, Sleijfer S, van Steenbergen S et al. Treatment of metastatic renal cell carcinoma with CAIX CAR-engineered T cells: clinical evaluation and management of on-target toxicity. *Mol Ther* 2013 ; 21 : 904-12.
29. Hamieh M, Dobrin A, Cabriolu A et al. CAR T cell trogocytosis and cooperative killing regulate tumour antigen escape. *Nature* 2019 ; 568 : 112.
30. Milone MC, Fish JD, Carpenito C et al. Chimeric receptors containing CD137 signal transduction domains mediate enhanced survival of T cells and increased antileukemic efficacy in vivo. *Mol Ther* 2009 ; 17 : 1453-64.
31. Wang X, Popplewell LL, Wagner JR et al. Phase 1 studies of central memory-derived CD19 CAR T-cell therapy following autologous HSCT in patients with B-cell NHL. *Blood* 2016 ; 127 : 2980-90.
32. Dotti G, Gottschalk S, Savoldo B, Brenner MK. Design and development of therapies using chimeric antigen receptor-expressing T cells. *Immunol Rev* 2014 ; 257 : 107-26.
33. Caruso HG, Hurton LV, Najjar A et al. Tuning sensitivity of CAR to EGFR density limits recognition of normal tissue while maintaining potent antitumor activity. *Cancer Res* 2015 ; 75 : 3505-18.
34. Srivastava S, Riddell SR. Engineering CAR-T cells: Design concepts. *Trends Immunol* 2015 ; 36 : 494-502.
35. Zah E, Lin M-Y, Silva-Benedict A et al. T Cells expressing CD19/CD20 bispecific chimeric antigen receptors prevent antigen escape by malignant B cells. *Cancer Immunol Res* 2016 ; 4 : 498-508.
36. Wilkie S, van Schalkwyk MCI, Hobbs S et al. Dual targeting of ErbB2 and MUC1 in breast cancer using chimeric antigen receptors engineered to provide complementary signaling. *J Clin Immunol* 2012 ; 32 : 1059-70.
37. Roybal KT, Rupp LJ, Morsut L et al. Precision tumor recognition by T cells with combinatorial antigen-sensing circuits. *Cell* 2016 ; 164 : 770-9.
38. Srivastava S, Salter AI, Liggitt D et al. Logic-gated ROR1 chimeric antigen receptor expression rescues T cell-mediated toxicity to normal tissues and enables selective tumor targeting. *Cancer Cell* 2019 ; 35 : 489-503.e8.
39. Qasim W, Zhan H, Samarasinghe S et al. Molecular remission of infant B-ALL after infusion of universal TALEN gene-edited CAR T cells. *Sci Transl Med* 2017 ; 9 : eaaj2013.
40. Gornalusse GG, Hirata RK, Funk SE et al. HLA-E-expressing pluripotent stem cells escape allogeneic responses and lysis by NK cells. *Nat Biotechnol* 2017 ; 35 : 765-72.
41. Kim MY, Yu K-R, Kenderian SS et al. Genetic inactivation of CD33 in hematopoietic stem cells to enable CAR T Cell immunotherapy for acute myeloid leukemia. *Cell* 2018 ; 173 : 1439-53.e19.
42. Eyquem J, Mansilla-Soto J, Giavridis T et al. Targeting a CAR to the TRAC locus with CRISPR/Cas9 enhances tumour rejection. *Nature* 2017 ; 543 : 113-7.

## Vous recherchez un médecin ou un remplaçant pour compléter votre équipe médicale ?

Contactez nos services pour une diffusion maximale de votre petite annonce dans la revue et sur les sites associés

Votre contact : Claire Voncken - Tél. : 01 49 29 29 20 - Mail : cvoncken@expressiongroupe.fr  
ou connectez-vous sur la rubrique "petites annonces" de nos sites :

[WWW.ONKO.FR](http://WWW.ONKO.FR)

OFFRES-SANTÉ 

## REVUE 6 - Toward “off-the-shelf” allogeneic CAR T cells

*Blake T. Aftab*<sup>1\*</sup>, *Barbra Sasu*<sup>2</sup>, *Janani Krishnamurthy*<sup>1</sup>, *Eric Gschweng*<sup>2</sup>, ***Vincent Alcazer***<sup>3</sup>,  
*Stéphane Depil*<sup>3,4,5\*</sup>

1 Atara Biotherapeutics, Thousand Oaks, CA, USA

2 Allogene Therapeutics, South San Francisco, CA, USA

3 Centre de Recherche en Cancérologie de Lyon, Lyon, France

4 Centre Léon Bérard, Lyon, France

5 Université, Claude Bernard Lyon 1, Lyon, France

\* Corresponding authors:

Stéphane Depil, Centre de Recherche en Cancérologie de Lyon, Lyon, France.

Email: [stephane.depil@lyon.unicancer.fr](mailto:stephane.depil@lyon.unicancer.fr)

Blake T. Aftab, Atara Biotherapeutics, Thousand Oaks, CA, USA.

Email: [baftab@atarabio.com](mailto:baftab@atarabio.com)

Adv Cell Gene Ther. 2020;3(3):e86. doi: 10.1002/acg2.86.



# Toward “off-the-shelf” allogeneic CAR T cells

Blake T. Aftab<sup>1</sup> | Barbra Sasu<sup>2</sup> | Janani Krishnamurthy<sup>1</sup> | Eric Gschweng<sup>2</sup> |  
Vincent Alcazer<sup>3</sup> | Stéphane Depil<sup>3,4,5</sup> 

<sup>1</sup>Atara Biotherapeutics, Thousand Oaks, CA, USA

<sup>2</sup>Allogene Therapeutics, South San Francisco, CA, USA

<sup>3</sup>Centre de Recherche en Cancérologie de Lyon, Lyon, France

<sup>4</sup>Centre Léon Bérard, Lyon, France

<sup>5</sup>Université, Claude Bernard Lyon 1, Lyon, France

## Correspondence

Stéphane Depil, Centre de Recherche en Cancérologie de Lyon, Lyon, France.  
Email: stephane.depil@lyon.unicancer.fr

Blake T. Aftab, Atara Biotherapeutics, Thousand Oaks, CA, USA.  
Email: baftab@atarabio.com

## Abstract

Chimeric antigen receptor (CAR) T cell therapy represents a major breakthrough in the field of immuno-oncology. Many potential issues are apparent for autologous CAR T cell therapy, such as time for manufacturing and need for interim therapies in progressing patients, wide variations in terms of quality and quantity of T cells, and difficulty to obtain enough cells for redosing. “Off-the-shelf” allogeneic CAR T cells premanufactured from third-party donors may theoretically provide solutions to these different problems. However, allogeneic T cells possess foreign immunological identities that can lead to histocompatibility considerations such as graft-versus-host disease and rejection of allogeneic cells. This review outlines the major recent advances for off-the-shelf T cell therapies currently in clinical trials or in preclinical development and describes strategies for reengineering or selecting specific T cell immune identities to create safe and efficient immunotherapies for patients.

## KEYWORDS

allogeneic T cell, chimeric antigen receptor T Cell, gene editing, immunotherapy, virus-specific T cell

## 1 | INTRODUCTION

Chimeric antigen receptor (CAR) T cell therapy is considered as a therapeutic breakthrough in some hematological malignancies, with recent approvals of Kymriah (tisagenlecleucel) and Yescarta (axicabtagene ciloleucel), and compelling pivotal results supporting liso-cel (lisocabtagene maraleucel).<sup>1,2</sup> However, the logistics involved in manufacturing and treatment with autologous CAR T cells adds a layer of complexity for clinicians and patients. Furthermore, there are several difficulties for which opportunities for improvement exist. The latter include:

- **Time to manufacturing:** Manufacturing of autologous CAR T cells today is largely centralized to comply with quality control and reproducibility rules. On-site or point-of-care production schemes propose to reduce logistic complexity and overall vein-to-vein time, but this strategy of production has been limited to clinical

trials thus far and still takes several days. CAR T cell manufacturing may thus create undesirable wait times for patients, especially for those with rapid progression of the disease. Furthermore, manufacturing failure may prevent patients from receiving CAR T cell therapy.<sup>3-5</sup> Complexities of cross-departmental and cross-institutional coordination of manufacturing and treatment may delay cell infusion.<sup>6</sup>

- **Need for interim therapies:** Bridging chemotherapies may be necessary to control disease progression while the patient is waiting for infusion of CAR T cells, posing issues for refractory malignancies.<sup>3,5</sup>
- **Risk of malignant contamination:** The leukapheresis product may contain malignant cells that become transduced with the CAR protein that binds cell surface antigen, protecting the contaminating cells from CAR T cell-mediated elimination.<sup>7</sup> This risk may depend on the primary disease and timing of cell harvest.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. *Advances in Cell and Gene Therapy* published by John Wiley & Sons Ltd



- *T cell variability*: Patient-specific variation in T cell phenotype and prevalence, like the presence of more or less 'early' or 'stem cell-like' T cell subsets, may yield different expansion and persistence of cell products.<sup>8</sup> Many factors may be involved in this T cell variability: age, disease, previous lines of treatment, explaining the interpatient heterogeneity observed with the use of autologous CAR T cells.<sup>8</sup>
- *Risk of insufficient T cell expansion*: Starting material is limited to T cells collected from leukapheresis and may not sufficiently expand to high numbers during manufacturing, especially in patients who are heavily pretreated and lymphopenic.<sup>3,5</sup>
- *T cell dysfunction*: A cancer patient's T cells may be dysfunctional as a result of disease burden or prior lines of therapy, and may therefore be a poor drug delivery vehicle for CAR T cell therapies.<sup>9-12</sup> Collecting T cells during the first lines of cancer therapy may reduce the risk associated with the cancer treatments but not the effects of tumor-associated immunosuppression.
- *Limited opportunity for redosing*: CAR T cells are prepared in a single batch and may be of limited quantity, such that patients may not have the opportunity to quickly and easily receive a new infusion of CAR T cells.

Unlike autologous T cell therapies, "off-the-shelf" T cells from third-party donors would be expanded in high numbers *prior* to treatment and made quickly available to patients.<sup>13,14</sup> This strategy could overcome the different limitations discussed above. A manufacturing protocol performed on T cells from a single donor leukapheresis has the potential to supply enough cells to produce hundreds of doses, and a single donor may be able to support multiple rounds of manufacturing.<sup>15,16</sup> This high-volume manufacturing allows for an opportunity to quickly and easily re-dose patients with a new off-the-shelf cell therapy without delays due to manufacturing or scale of production. Of note, in the potential instance of an immunogenic response to an off-the-shelf therapy, the use of cells from another donor further facilitates re-dosing options. Furthermore, a pretreatment manufacturing protocol allows for featuring multiple edits, manipulations, or TCR selection strategies that otherwise would be difficult to accommodate in the autologous setting. As CAR T design gains in complexity, this may become even more important, with the weight and complexity of safety assays making testing on individual products prohibitive.

A T cell therapy sourced from healthy allogeneic donors, however, introduces potential risks for issues related to incomplete histocompatibility. Due to the highly polymorphic nature of human leukocyte antigen (HLA) genes, allogeneic donor lymphocyte infusions are difficult to achieve without including consideration for risk of graft-versus-host disease (GvHD)<sup>17,18</sup> and durability of donor cells.<sup>19</sup> Complete donor-recipient matching of HLA haplotype is limited by low compatible donor availability, and healthcare providers must assess risk factors based on the extent of HLA matches for unmodified, unselected, or unmanipulated allogeneic T cells.<sup>20</sup>

In this Review, we will present the main approaches that are currently in development to deliver allogeneic T cell therapy and to

address these potentials for GvHD and cell rejection. In this context, we will focus on TCR editing-based approaches and virus-specific memory T cells as examples of clinically advanced nonalloreactive T cells. We will discuss the challenge of persistence and the perspectives of optimization.

## 2 | STRATEGIES TO AVOID GvHD

GvHD is believed to be largely driven by donor T cell recognition of host peptide-HLA complexes through the  $\alpha\beta$  T cell receptor complex ( $\alpha\beta$ TCR). To eliminate  $\alpha\beta$ TCR-mediated GvHD, two approaches can be employed:

- *Use of  $\alpha\beta$ TCR negative cells*: Editing cells to remove the TCR or using alternate cell types that lack  $\alpha\beta$ TCR.
- *Use of nonalloreactive T cells*: Utilizing T cells with  $\alpha\beta$ TCR specificity directed toward nonalloreactive targets.

## 3 | USE OF $\alpha\beta$ TCR NEGATIVE CELLS

Cell therapy has broadly validated the use of  $\alpha\beta$  T cells as potent and efficient killers.<sup>21-23</sup> Given their relative abundance in peripheral blood and their ability to proliferate rapidly upon stimulation, they are attractive for large scale manufacturing of allogeneic cells to treat a substantial number of patients per batch.<sup>24,25</sup>

Several strategies have been used to remove or decrease expression or function of the TCR complexes on these cells. In one approach, a truncated dominant-negative CD3 $\zeta$  protein (termed TCR inhibitor molecule, or TIM) was expressed in T cells, resulting in reduced TCR signaling and protection from xenogeneic GvHD.<sup>26</sup> In another approach, an antibody-derived single-chain variable fragment specific for CD3 $\epsilon$  was combined with amino acid sequences retaining it intracellularly. These protein expression blockers (PEBLs) colocalized intracellularly with CD3 $\epsilon$  and inhibited surface expression of CD3 and  $\alpha\beta$ TCR.<sup>27</sup> Small hairpin RNA (shRNA) has also been proposed to decrease expression of TCR chains.<sup>28</sup> Since the above approaches have the possibility to retain a low but functional amount of TCR, most entrants into the allogeneic space have opted for cellular gene editing approaches to ensure complete removal of the TCR. Several editing technologies are being used and all have in common their ability to cut DNA in a sequence-specific manner, creating a double strand break that then is patched using the cell's natural DNA repair machinery via nonhomologous end-joining (NHEJ). This restores the integrity of the chromosome but creates insertions or deletions (INDELS), which lead to disruption of the locus (most often *TRAC*, the TCR $\alpha$  common chain). With no functional *TRAC* gene, the TCR complex cannot form.<sup>29</sup>

No approaches for  $\alpha\beta$ TCR removal result in 100% of cells having  $\alpha\beta$ TCR reduced or eliminated, necessitating a purification step to remove residual  $\alpha\beta$ TCR positive cells. The efficiency of this process is

critical to reduce the risk of any contaminating  $\alpha\beta$ TCR positive cells that may induce GvHD.

Transcription activator-like effector nucleases (TALEN) were the first technology to be used in patients for  $\alpha\beta$ TCR removal,<sup>13</sup> entering the clinic for compassionate use of a CD19 CAR for pediatric acute lymphoblastic leukemia (ALL) in 2015<sup>30</sup> and in two phase 1 trials for pediatric and adult ALL in 2016.<sup>31</sup> Grade 2 skin GvHD was observed in one of the two compassionate use cases,<sup>30</sup> and grade 1 skin GvHD in the 2 of 21 patients in the two UCART19 clinical trials. No higher grades of GvHD were observed. Other gene editing technologies currently being employed preclinically and in the clinic for allogeneic CARs are clustered regularly interspaced short palindromic repeat (CRISPR), ARC editing megaTALs and Zinc fingers.<sup>32</sup>

Gene edited allogeneic CAR T cells are classed as gene therapies by the FDA, necessitating long-term follow-up of patients to monitor for safety.<sup>33</sup> Since gene editing may have a low frequency of off-target cutting, monitoring of predicted off-target cut sites and nonbiased approaches to identify off-target events are necessary. Assays to look for any consequences of off-target cutting (such as eg inactivation of tumor suppressors) are also required, which includes demonstrating lack of growth factor independence. If more than one edit is made simultaneously, consequences of translocation must also be investigated and translocation frequencies must be monitored.<sup>34</sup> Although no instances of viral integration-related transformation have been reported with CAR T therapy to date, there is also a potential risk that this might occur given the random integration of both retroviral and lentiviral vectors. In addition to using gene editing to knock out  $\alpha\beta$ TCR expression, it can also be used as a tool to add the CAR into the disrupted locus, thereby producing site specific integration.<sup>35,36</sup> Site specific integration may possess other advantages, for instance potentially lowering risk of translocations with two or more edits by favoring homology-directed repair rather than NHEJ.<sup>37</sup>

Although  $\alpha\beta$  T cells represent the 'gold standard' of cell sources for therapy, they are not the only option available. Several effector cell types exist which lack an  $\alpha\beta$ TCR or express an invariant version and are not expected to cause GvHD.<sup>38</sup> NK cells present in the peripheral blood for example can be expanded *ex vivo* to create an engineered product. CD19 CAR modified NK cells have been described,<sup>39</sup> showing activity in patients in a published phase I/II.<sup>40</sup> Eight 11 treated patients had a clinical response, including seven complete remissions (CRs) and one remission of a Richter's transformation with persistent chronic lymphocytic leukemia. Other cell types such as iNKT or  $\gamma\delta$ T cells may be as potent as  $\alpha\beta$  T cells and may have advantages in terms of their innate receptor activity, their ability to be expanded readily to large numbers using their natural ligands<sup>41</sup> and ability to penetrate solid tumors.<sup>42</sup> Their relative frequencies in peripheral blood are, however, very low and any contamination with alloreactive  $\alpha\beta$  T cells would require similar purification to avoid the risk of GvHD.

If a renewable cell source could be used to generate gene edited T cells or other cell types, clones with ideal characteristics and defined composition could be generated; for instance, absence of

translocations, successful gene editing to result in a completely  $\alpha\beta$ TCR negative cell source. Work has begun to derive T cells engineered from induced pluripotent stem cells (iPSCs).<sup>43-45</sup> These cells can be engineered, then genetically screened clones can be expanded and differentiated. Proof of concept for this approach has been generated for NK and T cells,<sup>46</sup> with at least one T cell product heading into the clinic (Table 1). Nonetheless, technical challenges still remain for alternatives to  $\alpha\beta$  T cells, including demonstrating functional equivalence to conventional T cells.

Most of the work described above has been carried out in the context of CD19 CARs since CD19 still represents the most clinically validated cell therapy target. Allogeneic cell products are now starting to be created against targets for other hematologic malignancies, such as BCMA,<sup>47,48</sup> and for some solid tumor targets. Although more challenges are expected for CAR therapies moving into solid tumor indications, these may be the areas where allogeneic approaches give the most benefit. Multiple gene edits to overcome solid tumor challenges may be possible, and the larger patient populations requiring treatment will be much more feasible with an allogeneic cell source.

### 3.1 | Selection of nonalloreactive T cells: A nongene edited approach

Alternative approaches to editing of  $\alpha\beta$ TCR have also demonstrated significant clinical progress for off-the-shelf cell therapies. One example of this has been the use of virus-specific T (VST) cells derived from healthy donors. Exposure of lymphocyte collections to viral antigens *ex vivo* expand T cells expressing virus-specific  $\alpha\beta$ TCR. Virally-infected antigen presenting cells are commonly used for this stimulation and expansion of T cells expressing virus-specific  $\alpha\beta$ TCR<sup>49,50</sup> and may be genetically engineered<sup>51</sup> or be directly loaded with viral peptides.<sup>14</sup> Similarly, virus-specific  $\alpha\beta$ TCR can be directly isolated using MHC-bound antigen peptide.<sup>52,53</sup> The resulting virus-specific  $\alpha\beta$ TCR repertoire functions via formation of a classical immune synapse and is thus activated via specific recognition of virus and not host peptide in complex with a specific HLA. As a primary therapeutic composition, VST cells can be manufactured from multiple donors, each associated with different HLA alleles, and banked in large numbers based on HLA restriction.<sup>14,53</sup>

One immediate application for these off-the-shelf banks of VST cells is in the treatment of diseases originating from viral infection. For instance, Epstein-Barr virus (EBV) is a known driver of cell transformation and progression of several diseases including lymphoproliferative disorders,<sup>50,54-56</sup> nasopharyngeal carcinoma,<sup>49</sup> leiomyosarcoma,<sup>57</sup> and multiple sclerosis.<sup>58</sup> The first evidence of successful use of off-the-shelf VST therapy was performed with minimally HLA matched EBV-specific T cells to treat patients with EBV-positive lymphoma following solid organ transplant.<sup>56</sup> In this multicenter phase II trial, 14 of 33 patients achieved CR and seven patients achieved partial remissions. Similarly, Prockop et al has described the treatment of 46 patients with rituximab relapsed/

TABLE 1 Clinical trials evaluating allogeneic T cells

Sponsor	Therapy name	Therapy target	Development stage	Clinical Trial ID	Disease	Trial start date	Gene edits used
Servier Allogene Therapeutics	UCART19	CD19	Phase 1	NCT02808442 (PALL)	Pediatric R/R ALL	6/3/16	CD52KO TRACKO CD19 CAR (Lenti)
Servier/Allogene Therapeutics	UCART19	CD19	Phase 1	NCT02746952 (CALM)	R/R ALL	8/1/16	CD52KO TRACKO CD19 CAR (Lenti)
Collectis	UCART123	CD123	Phase 1	NCT03190278 (AMELI-01)	R/R AML	6/19/17	TRACKO CD123 CAR (Lenti)
Collectis	UCART123	CD123	Phase 1 <sup>a</sup>	NCT03203369 (ABC123)	BPDCN	6/28/17	TRACKO CD123 CAR (Lenti)
Poseida	P-BCMA-101	BCMA	Phase 1	NCT03288493	Multiple Myeloma	9/20/17	BCMA Centyrin (piggyBac)
Celyad	CYAD-101	NKG2D Ligands	Phase 1	NCT03692429 (alloSHRINK)	CRC	11/28/18	TIM NKG2D CAR (Retro)
Precision Biosciences	PBCAR0191	CD19	Phase 1/2	NCT03666000	R/R B-ALL and R/R NHL	3/11/19	CD19 CAR --> TRAC locus (ARCUS)
Allogene Therapeutics	ALLO-501	CD19	Phase 1/2	NCT03939026 (ALPHA)	R/R DLBCL and R/R FL	5/1/19	CD52KO TRACKO CD19 CAR (Lenti)
CRISPR Therapeutics	CTX110	CD19	Phase 1	NCT04035434	NHL	7/22/19	MHC-I KO TRACKO CD19 CAR SSI
Allogene Therapeutics	ALLO-715	BCMA	Phase 1	NCT04093596 (UNIVERSAL)	Multiple myeloma	9/23/19	CD52KO TRACKO BCMA CAR (Lenti)
Collectis	UCART123	CD123	Phase 1	NCT04106076	High risk AML	7/11/19	TRACKO CD123 CAR (Lenti)
Collectis	UCART22	CD22	Phase 1	NCT04150497 (BALLI-01)	R/R B-ALL and R/R NHL	10/17/19	TRACKO CD22 CAR (Lenti)
Collectis	UCARTCS1	CS1	Phase 1	NCT04142619 (MELANI-01)	Multiple Myeloma	11/21/19	CS1KO TRACKO CS1 CAR (Lenti)
CRISPR Therapeutics	CTX120	BCMA	Phase 1	NCT04244656	Multiple myeloma	1/28/20	MHC-I KO TRACKO BCMA CAR SSI
Virus-specific T cells							
Atara Biotherapeutics	Tabelecleucel	EBV antigens Phase 3	NCT03394365 (ALLELE)	R/R EBV+PTLD	After SOT or HCT	12/29/17	No Edits
Atara Biotherapeutics	Tabelecleucel	EBV antigens	Phase 1/2	NCT03769467 (EBV+)	NPC) EBV+ NPC	11/28/18	No edits
Atara Biotherapeutics	Tabelecleucel	EBV antigens	Phase 1/2	NCT00002663	EBV-associated malignancies	3/1/95	No edits
AlloVir	Viralym-M	Multiviral antigens	(N/A)	NCT02108522 (CHARMS)	Viral infections after SCT	6/1/14	No edits

(Continues)

TABLE 1 (Continued)

Sponsor	Therapy name	Therapy target	Development stage	Clinical Trial ID	Disease	Trial start date	Gene edits used
AlloVir	Viralym-M	Multiviral antigens	Phase 1/2	NCT01570283 (ARMS)	Viral infections	9/1/12	No edits
AlloVir	Viralym-C	Multiviral antigens	Phase 1	NCT02313857	CMV infections	9/1/15	No edits
MSKCC	[none]	EBV antigens and CD19	Phase 1	NCT01430390	B cell malignancies	9/1/11	+ CD19 CAR
Fred-Hutch	[none]	EBV antigens and WT1	Phase 1/2	NCT01640301	AML	7/12/12	+ Affinity-enhanced WT1 TCR
Atara Biotherapeutics	ATA188	EBV antigens	Phase 1	NCT03283826	Multiple sclerosis	10/19/17	No edits
Atara Biotherapeutics	ATA520	EBV antigens and WT1	Phase 1	NCT00620633	Leukemia	2/1/08	No edits
Atara Biotherapeutics	ATA320	CMV antigens	Phase 2	NCT02136797	CMV infection after HSCT	5/1/14	No edits
Baylor College of Medicine	[none]	EBV antigens (EBNA1, LMP, BARF1)	Phase 1	NCT02287311 (MABEL)	EBV+lymphoma	2/1/15	No edits

Start dates as reported in ClinicalTrials.gov

Abbreviations: ALL, acute lymphoblastic leukemia; AML, acute myeloid leukemia; BPDCN, blastic plasmacytoid dendritic cell neoplasm; CMV, cytomegalovirus; DLBCL, diffuse large B-cell lymphoma; EBV, Epstein Barr virus; FL, follicular lymphoma; HCT, allogeneic hematopoietic cell transplant; NHL, non-Hodgkin lymphoma; NPC, nasopharyngeal carcinoma; PTLD, post-transplant lymphoproliferative disease; R/R, relapsed/refractory; SOT, solid organ transplant.

<sup>a</sup>Study terminated.

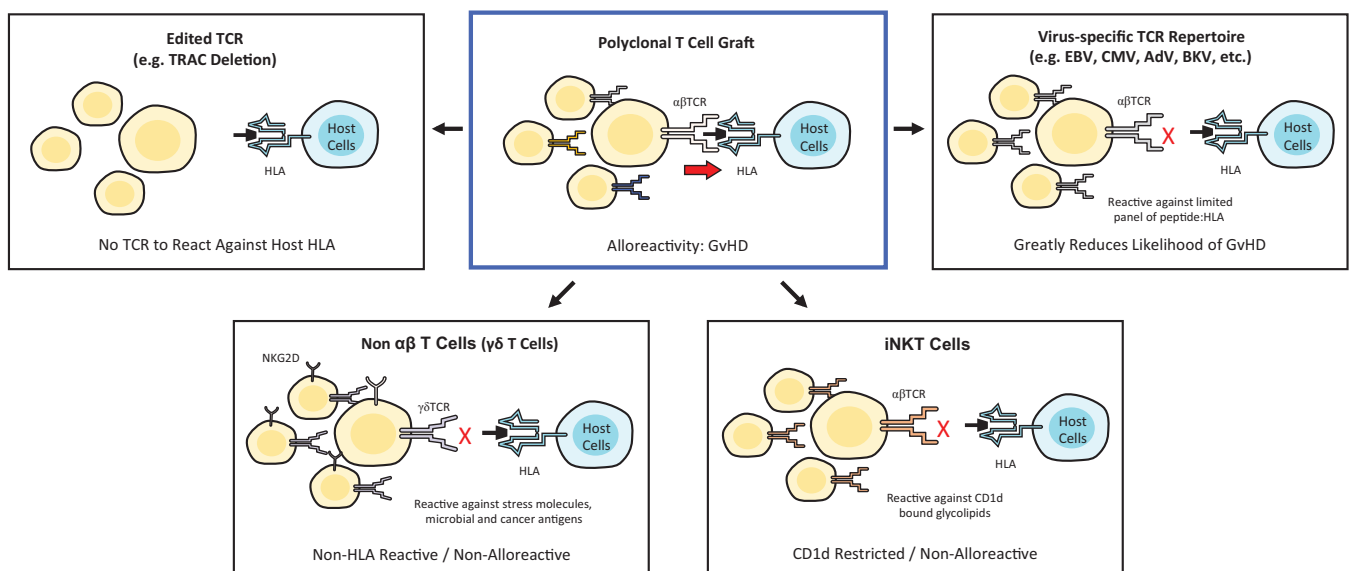
refractory EBV-positive lymphoma using a bank of partially HLA matched off-the-shelf EBV-specific T cells.<sup>59</sup> In this experience, response rates in patients developing lymphoma following a previous stem cell transplant or solid organ transplant were 68% and 54% respectively. Of the 46 patients treated, only one developed a *de novo* low grade GvHD, which resolved with topical therapy. Despite partial HLA-matching, allogeneic EBV-specific T cells expanded in patients without lymphodepletion pretreatment and were detected in circulation for up to 27.3 months after the date of infusion.<sup>60,61</sup>

Beyond EBV-specific T cells, other off-the-shelf VST therapies have been deployed similarly in the clinic, including those specific for adenovirus, BK virus, and human herpesvirus 6 (HHV-6).<sup>62-64</sup> In these examples, low incidence of GvHD is attributed to enrichment of virus-specific  $\alpha\beta$ TCR repertoire and the concurrent depletion of cells expressing functionally alloreactive  $\alpha\beta$ TCRs. These studies demonstrate that a low degree of HLA matching provides sufficient persistence and patient exposure to facilitate complete and durable clinical response. Together, examples of allogeneic VST cell therapies have clinical history establishing benchmarks for efficacy and safety in the context of partial HLA allele match, including with as few as one HLA allele match.<sup>54</sup> This limited requirement for HLA matching for this class of allogeneic cell therapy significantly decreases the required size of an off-the-shelf bank containing suitable HLA coverage to address a general population.

Similar to previously described engineered T cell platforms, VSTs can also be transduced to additionally express a CAR, functionally redirecting their effector function to new, nonviral tumor targets

through CAR-specific and HLA-independent mechanisms.<sup>65-67</sup> EBV antigens may also be present along with other B-cell surface markers, such as CD19, commonly expressed in B cell malignancies.<sup>68</sup> In this case, EBV-specific T cells transduced to express a CD19-targeting CAR can recognize and kill lymphoma cells via either the endogenous virus-specific TCR, which targets EBV antigen through HLA-dependent mechanisms, or via engagement of the CAR against CD19-antigen in an HLA-independent mechanism. Preclinical studies with off-the-shelf allogeneic EBV-specific CD19-CAR T cells demonstrate this capability with CARs containing either CD28 or 4-1BB co-stimulatory domains, while maintaining the central memory VST phenotype when stimulated with antigen.<sup>66,67</sup> Preliminary results from a phase 1 evaluation of allogeneic EBV-specific CD19-CAR T cells recently described 83% response rate and no reported incidence of confirmed GvHD or CRS or ICANS above grade 2 across six patients with relapsed/refractory B cell malignancies with 26.9 months median follow-up.<sup>69</sup>

Similar strategies have also been extended to solid tumor targets, as demonstrated with autologous EBV-specific VST expressing CAR targeting disialoganglioside (GD2) for treatment of neuroblastoma.<sup>70,71</sup> In this phase I trial, investigators reported 3 of 11 patients with active disease achieved CR with EBV-specific GD2-targeted CAR T cells. In this study, the GD2 CAR-VST demonstrated superior cell persistence (up to 192 weeks), twice as long as standard  $\alpha\beta$ T cells bearing CAR alone (96 weeks). This comparable persistence benefit may provide further support for transitioning these approaches to the allogeneic setting.



**FIGURE 1** Methods to prevent alloreactivity. Top Center: The introduction of a polyclonal  $\alpha\beta$ T cell product brings the risk of graft recognition via HLA, known as GvHD. Top Left: T cell mediated reactivity against host tissues can be completely abrogated by deletion of the TCR, typically by targeting the TCR $\alpha$  common chain (TRAC) with a nuclease, leaving the sole reactivity of the allogeneic product through the defined specificity of the CAR. Top Right: An oligoclonal  $\alpha\beta$  T cell product such as viral specific T cells (VST) has a lower likelihood of reacting against host HLA, thus reducing the chance of GVHD. Bottom Left: The  $\gamma\delta$  T cell receptor is not restricted on HLA, thus does not cause GvHD via HLA mismatch. The ligands bound by the  $\gamma\delta$  TCR are still being discovered, but include nonself molecules such as microbial pathogens and cancer antigens. Additionally,  $\gamma\delta$  T cells express NKG2D which functionalizes them against cancer cells which may express stress ligands. Bottom Right: iNKT cells express an invariant  $\alpha\beta$  TCR which recognizes lipid antigens through the nonclassical HLA molecule CD1d

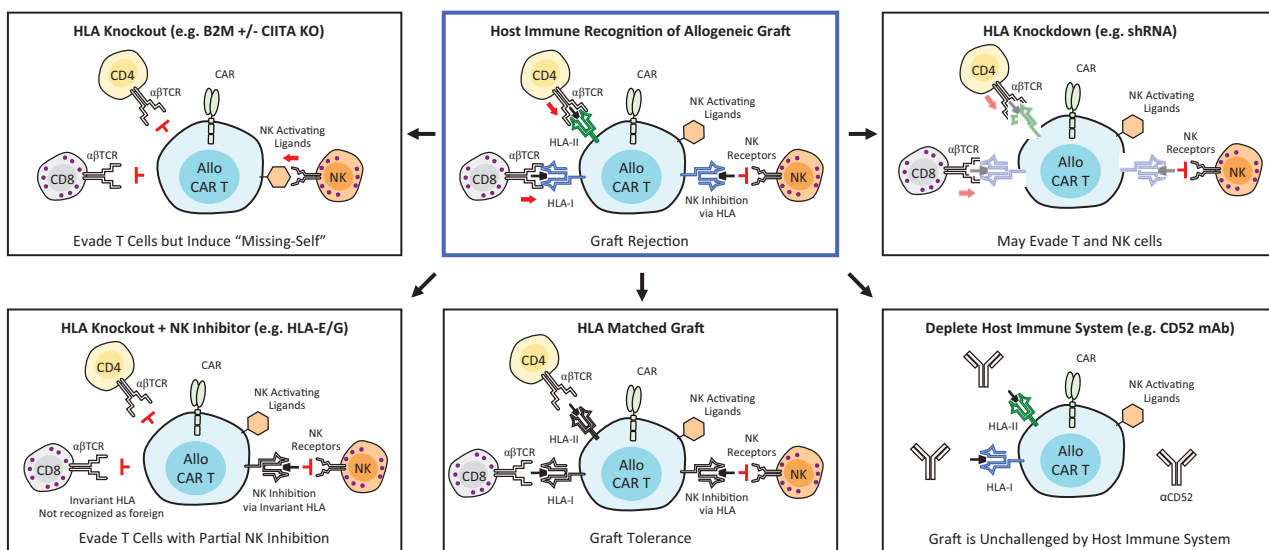
Beyond CAR, allogeneic VSTs have also been combined with engineered TCR approaches. Similar to redirection of VST function via a CAR, redirection may also be achieved via transduction with engineered TCRs directed toward nonviral tumor-specific antigens. In this way, allogeneic donor-derived EBV-specific VSTs further transduced with an engineered TCR specific to Wilm's tumor 1 antigen (WT1) have also been explored clinically.<sup>72-74</sup> This therapy was recently tested in patients with acute myeloid leukemia (AML) who received an allogeneic hematopoietic cell transplantation (HCT). These WT1 TCR-redirection EBV-specific T cells were associated with increased persistence of greater than 10 months, while similarly WT1-TCR transduced CMV-specific T cells persistence was sporadic and declined after 2 months.<sup>72,74</sup> This differential persistence was associated with an enriched central memory T cell profile commonly seen with the EBV subset of VSTs. At 44 months after infusion, 100% relapse-free survival was reported for AML patients receiving EBV-specific WT1-targeting T cells, and incidence of GvHD remained low and attributable to prior HCT.<sup>73</sup>

As a specific example of a nongene edited allogeneic  $\alpha\beta$ T cell, VSTs may provide advantages and feasible alternatives for creating off-the-shelf cell therapies that do not require elimination of either endogenous  $\alpha\beta$ TCR expression or the expression of HLA. These advantages potentially impact both the ease of GMP manufacturing and risk of gene edited T cells inducing extraneous effects, such as

p53-mediated DNA damage response.<sup>75</sup> EBV-specific T cells may also provide additional potential for approaching other solid tumor targets, as EBV-specific T cells have been shown to traffic to and penetrate solid tumor tissues, where they may facilitate broad anti-tumor immune response upon stimulation.<sup>76</sup>

#### 4 | EXPANSION AND PERSISTENCE OF ALLOGENEIC CAR T CELLS

The clinical efficacy of CAR T cells is associated with their initial expansion to achieve a sufficient effector-to-target ratio. Post-infusion CAR T cell expansion has been reported to be a major factor for clinical response in the case of autologous CD19 CAR T cells in B-ALL.<sup>77</sup> This expansion is enhanced by homeostatic cytokines such as IL-7 and IL-15, whose levels are elevated after lymphodepletion.<sup>78,79</sup> CAR T cell persistence can result in long-term immunosurveillance. The persistence required to achieve a prolonged remission may vary according to the disease. One school of thought proposes that prolonged persistence may be required for a long-term control of diseases such as B-ALL.<sup>5</sup> In contrast, most of DLBCL patients achieving a CR at 6 months remain disease-free.<sup>4</sup> This difference probably reflects the biology of the disease and the persistence or not of a tumor reservoir after a first step of tumor killing.



**FIGURE 2** Methods to prevent graft rejection. Top Center: Allogeneic CAR T cells can be recognized by host CD8 and CD4 cells via TCR-HLA-I and -II interactions respectively. Recognition of activating ligands such as MICA/B by NK cells results in cell lysis, which may be inhibited by KIR-HLA. Bottom Center: Selecting a partially HLA matched graft, as found with virus specific banked T cells, reduces recognition and supports graft tolerance. Top Left: Knockout of the conserved B2M subunit abrogates surface HLA-I expression, rendering cells invisible to CD8 T cells. Knockout of the transcription factor CIITA reduces HLA-II expression, thus avoiding detection by CD4 T cells. However, eliminating HLA also removes the inhibition of NK cells, inducing potential “missing-self” cytotoxicity. Bottom Left: An invariant HLA molecule lacking in sufficient polymorphism to be recognized by T cells, such as HLA-E or HLA-G, may be added to HLA knockout cells to restore NK inhibition, thus avoiding both T and NK cells. Top Right: Knockdown of HLA-I and -II by shRNA may be employed to reduce recognition by T cells while maintaining enough HLA to inhibit NK. Bottom Right: Host immune cells such as T, B, and NK express CD52. By knocking out CD52 on allogeneic CAR T cells, a CD52 monoclonal antibody can be given systemically to deplete host immune cells while sparing the engineered graft



Costimulatory domains significantly affect CAR T cell expansion and persistence. It has been shown that CD19.28.z-CAR T cells expand more rapidly and to higher levels than CD19.BB.z-CAR T cells.<sup>80,81</sup> This accelerated and higher expansion of T cells transduced with a CAR with CD28 versus 4-1BB costimulatory domain may be explained by a greater signal strength downstream of CD28 versus 4-1BB.<sup>82</sup> However, autologous T cells expressing a CAR with a CD28 costimulatory domain rarely persist more than 1 or 2 months, whereas those with 4-1BB can persist for months or years.<sup>5,77</sup> Because the half-life of allogeneic CAR T cells will be reduced in most of cases due to rejection by the host immune system, it may be suggested in this context to use preferentially CD28 as a costimulatory domain, in order to obtain a rapid expansion of the CAR T cells and thus optimize the antitumor effect on a limited period of time. The use of CD28 may be, however, associated with a higher risk of cytokine release syndrome compared to 4-1BB.<sup>83</sup> Of note, there is no data available to date in the clinic to compare allogeneic CAR T cells using CD28 versus 4-1BB as a costimulatory domain.

Once the risk of GvHD has been controlled, a potential issue remaining for allogeneic CAR T cell therapy is their limited persistence in the recipient. Different approaches can be proposed to allow initial expansion and improve persistence of allogeneic CAR T cells: HLA-matching, reinforcement of the lymphodepletion, HLA class I and II deletion, and use of immunomodulatory factors.

To increase persistence, a bank of cells could be prepared from donors with different HLA alleles, providing partial HLA coverage for the majority of the targeted patient population (as performed for VST cells derived from third party donors). Considering that the most important alleles to match for are HLA-A, -B, -C and potentially -DR, the size of the bank could be relatively limited. However, the persistence of the injected cells will be still reduced in absence of full HLA matching compared to autologous T cells. Indeed, the results obtained with third party donor-derived VST establish a persistence of the cells of 14-90 days and in the case of partially matched EBV specific T cells persistence for up to 27.3 months after the date of infusion has been reported<sup>60,61</sup> whereas VST from HCT donors were particularly long (up to 10 years).<sup>84</sup>

The first clinical trials evaluating gene-edited allogeneic CAR T cells have used a reinforced lymphodepletion. An anti-CD52 antibody was added to the classical fludarabine and cyclophosphamide combination, with the aim of inducing transient tolerance to the transplant by increasing the depth and duration of the lymphopenia. CAR T cells are in this case protected by deletion of CD52.<sup>18</sup> This approach has been shown to lead to expansion of the allogeneic CAR T cells associated with efficacy. Two compassionate use patients first treated with UCART19 showed durable CR<sup>30</sup> and 82% of patients in the two phase 1 trials who were treated with the anti-CD52-containing lymphodepletion regimen achieved CR.<sup>31</sup> Furthermore, in two of three patients who failed to show benefit from the initial infusion, reinfusion gave minimal residual disease negative-CR. In a similar way, deletion of the dCK protein may provide resistance against purine nucleotide analogue chemotherapies such as fludarabine used for lymphodepletion.<sup>85</sup> These approaches have the advantage

of suppressing all immune cell classes which may mediate rejection such as T, B, NK and potentially monocytes. However, the flip side of this benefit is that a prolonged immunosuppression is associated with a higher risk of opportunistic infections and viral reactivation.

Another approach consists in modifying allogeneic cells in order to obtain "universal cells" that will evade host immune detection. In the case of CD8 T cell rejection, the priority is to eliminate the expression of HLA class I molecules on CAR T cells. Deletion of the conserved gene beta 2-microglobulin completely removes surface expression of HLA class I.<sup>86</sup> However, while immunogenic recognition by CD8 T cells will be reduced by this approach, the complete loss of HLA class I molecules will increase the risk of recognition of the allogeneic CAR T cells by NK cells, the so called 'missing self' mechanism. NK cell mediated destruction of HLA-edited T cells has been proposed to be prevented by expression on CAR T cells of non-polymorphic HLA molecules such as HLA-E that will bind inhibitory receptors on NK cells.<sup>87</sup> In addition to CD8 T cell-mediated rejection, CD4 T cells can contribute to rejection of allogeneic T cells through recognition of HLA class II molecules. For this reason, limiting class II expression by deleting one of the transcription factors required for all HLA class II genes, such as *CIITA* or *RFXANK*, has been proposed.<sup>88</sup>

Finally, expression of immunomodulatory molecules, such as PD-L1 or FASL, may suppress or kill the recipient's activated T lymphocytes respectively.<sup>88</sup> This approach may be, however, limited by the expression of PD-1 and FAS molecules on the activated CAR T cells themselves, inducing thus their inhibition for PD-L1/PD-1 and fratricide killing for FASL/FAS interaction. How inhibition of phagocytosis by monocytes or macrophages may increase the persistence of allogeneic T cells is less clear. Adding molecules such as CD47 and CD200 may allow modified allogeneic T cells to avoid macrophage-mediated clearance.<sup>88,89</sup> Currently it is unclear if some of the above will be dominant, or if all modifications will be needed to avoid rejection.

The different possibilities of T cell engineering open new perspectives in the field of immune cell therapy. However, it remains to be elucidated whether the production of universal cells that completely evade the host immune system may also induce safety issues, since the immune system will not be able to recognize transformed cells in case of an oncogenic event. The initial screening of the administered allogeneic T cells for oncogenic mutations would therefore be of crucial importance in the context of extensive editing of the T cell genome. Given that late oncogenic events can still occur, and because rare oncogenic cells may have not been detected during the process, an optimized suicide system may be advantageous in modified T cells to allow their efficient elimination.<sup>90</sup> Linking the suicide gene to a cell-division gene, as described by Liang et al,<sup>91</sup> may be particularly useful in this context.

## 5 | CONCLUSIONS

Engineered adoptive cell immunotherapies are quickly showing promise for patients. As logistical and scientific insights into the manufacturing of these therapies evolve, patients could greatly

benefit from readily available off-the-shelf cell therapies tailored to the patient and their disease. By engineering T cell immune identities, off-the-shelf allogeneic T cell therapies currently in clinical development (Table 1) could provide broader immunotherapeutic options for patients. Current studies with nonengineered allogeneic T cells, such as the phase 3 ALLELE (NCT03394365) trial for patients with relapsed/refractory EBV+ post-transplant lymphoproliferative disease, and engineered allogeneic CAR T cell studies like UCART19 for ALL (NCT02746952), could begin to lay the groundwork for off-the-shelf therapies as standard-of-care treatments for patients.

### CONFLICT OF INTEREST

SD has been employee of Collectis. He is the founder of ErVaccine Technologies and member of the management team of Swiss Rockets. Consulting or scientific advisory boards: Servier, Celyad, PDC\*line Pharma, Erytech, AstraZeneca, Elsalys, Netris Pharma, Vaximm. BTA and JK are employees of Atara Biotherapeutics. BS and EG are employees of Allogene Therapeutics

### ETHICAL STATEMENT

The authors confirm that the ethical policies of the journal, as noted on the journal's author guidelines page, have been adhered to. No ethical approval was required as this is a review article with no original research data.

### ORCID

Stéphane Depil  <https://orcid.org/0000-0001-6011-3945>

### REFERENCES

- Leyfman Y. Chimeric antigen receptors: unleashing a new age of anti-cancer therapy. *Cancer Cell Int.* 2018;18:182.
- Abramson JS, Palomba ML, Gordon LI, et al. Pivotal safety and efficacy results from Transcend NHL 001, a multicenter phase 1 study of Lisocabtagene Maraleucel (liso-cel) in relapsed/refractory (R/R) large B cell lymphomas. *Blood.* 2019;134(Suppl. 1):241.
- Schuster SJ, Svoboda J, Chong EA, et al. Chimeric antigen receptor T cells in refractory B-cell lymphomas. *N Engl J Med.* 2017;377(26):2545–2554.
- Neelapu SS, Locke FL, Bartlett NL, et al. Axicabtagene Ciloleucel CAR T-cell therapy in refractory large B-cell lymphoma. *N Engl J Med.* 2017;377(26):2531–2544.
- Maude SL, Laetsch TW, Buechner J, et al. Tisagenlecleucel in children and young adults with B-cell lymphoblastic leukemia. *N Engl J Med.* 2018;378(5):439–448.
- McGuirk J, Waller EK, Qayed M, et al. Building blocks for institutional preparation of CTL019 delivery. *Cytotherapy.* 2017;19(9):1015–1024.
- Ruella M, Xu J, Barrett DM, et al. Induction of resistance to chimeric antigen receptor T cell therapy by transduction of a single leukemic B cell. *Nat Med.* 2018;24(10):1499–1503.
- Hoffmann J-M, Schubert M-L, Wang L, et al. Differences in expansion potential of naive chimeric antigen receptor T cells from healthy donors and untreated chronic lymphocytic leukemia patients. *Front Immunol.* 2017;8:1956.
- Das RK, Vernau L, Grupp SA, et al. Naive T-cell deficits at diagnosis and after chemotherapy impair cell therapy potential in pediatric cancers. *Cancer Discov.* 2019;9(4):492–499.
- Azzaoui I, Uhel F, Rossille D, et al. T-cell defect in diffuse large B-cell lymphomas involves expansion of myeloid-derived suppressor cells. *Blood.* 2016;128(8):1081–1092.
- Schietinger A, Philip M, Krisnawan VE, et al. Tumor-specific T cell dysfunction is a dynamic antigen-driven differentiation program initiated early during tumorigenesis. *Immunity.* 2016;45(2):389–401.
- Totterman TH, Carlsson M, Simonsson B, Bengtsson M, Nilsson K. T-cell activation and subset patterns are altered in B-CLL and correlate with the stage of the disease. *Blood.* 1989;74(2):786–792.
- Poirot L, Philip B, Schiffer-Mannioui C, et al. Multiplex genome-edited T-cell manufacturing platform for "off-the-shelf" adoptive T-cell immunotherapies. *Cancer Res.* 2015;75(18):3853–3864.
- Withers B, Clancy L, Burgess J, et al. Establishment and operation of a third-party virus-specific T cell bank within an allogeneic stem cell transplant program. *Biol Blood Marrow Transplant.* 2018;24(12):2433–2442.
- Atara Biotherapeutics Corporate Presentation. In the 38th Annual J.P. Morgan Healthcare Conference. 2020. San Francisco, CA.
- Allogene Corporate Presentation. In The 38th Annual J.P. Morgan Healthcare Conference. 2020. San Francisco, CA.
- Lee SJ, Klein J, Haagenson M, et al. High-resolution donor-recipient HLA matching contributes to the success of unrelated donor marrow transplantation. *Blood.* 2007;110(13):4576–4583.
- Loiseau P, Busson M, Balere M-L, et al. HLA association with hematopoietic stem cell transplantation outcome: the number of mismatches at HLA-A, -B, -C, -DRB1, or -DQB1 is strongly associated with overall survival. *Biol Blood Marrow Transplant.* 2007;13(8):965–974.
- Petersdorf EW, Hansen JA, Martin PJ, et al. Major-histocompatibility-complex class I alleles and antigens in hematopoietic-cell transplantation. *N Engl J Med.* 2001;345(25):1794–1800.
- Howard CA, Fernandez-Vina MA, Appelbaum FR, et al. Recommendations for donor human leukocyte antigen assessment and matching for allogeneic stem cell transplantation: consensus opinion of the Blood and Marrow Transplant Clinical Trials Network (BMT CTN). *Biol Blood Marrow Transplant.* 2015;21(1):4–7.
- Chandran SS, Klebanoff CA. T cell receptor-based cancer immunotherapy: Emerging efficacy and pathways of resistance. *Immunol Rev.* 2019;290(1):127–147.
- June CH, Sadelain M. Chimeric antigen receptor therapy. *N Engl J Med.* 2018;379(1):64–73.
- Rosenberg SA, Restifo NP. Adoptive cell transfer as personalized immunotherapy for human cancer. *Science.* 2015;348(6230):62–68.
- Graham C, Jozwik A, Pepper A, Benjamin R. Allogeneic CAR-T cells: more than ease of access? *Cells.* 2018;7(10):155.
- Zhao J, Lin Q, Song Y, Liu D. Universal CARs, universal T cells, and universal CAR T cells. *J Hematol Oncol.* 2018;11(1):132.
- Gilham DE, Michaux A, Breman E, et al. TCR inhibitory molecule as a promising allogeneic NKG2D CAR-t cell approach. *J Clin Oncol.* 2018;36(15 suppl):e15042–e15042.
- Kamiya T, Wong D, Png YT, Campana D. A novel method to generate T-cell receptor-deficient chimeric antigen receptor T cells. *Blood Adv.* 2018;2(5):517–528.
- Bunse M, Bendle GM, Linnemann C, et al. RNAi-mediated TCR knockdown prevents autoimmunity in mice caused by mixed TCR dimers following TCR gene transfer. *Mol Ther.* 2014;22(11):1983–1991.
- Hall C, Berkhout B, Alarcon B, Sancho J, Wileman T, Terhorst C. Requirements for cell surface expression of the human TCR/CD3 complex in non-T cells. *Int Immunol.* 1991;3(4):359–368.
- Qasim W, Zhan H, Samarasinghe S, et al. Molecular remission of infant B-ALL after infusion of universal TALEN gene-edited CAR T cells. *Sci Transl Med.* 2017;9:374.

31. Benjamin R, Graham C, Yallop D, et al. Preliminary data on safety, cellular kinetics and anti-leukemic activity of UCART19, an allogeneic anti-CD19 CAR T-cell product, in a pool of adult and pediatric patients with high-risk CD19<sup>+</sup> relapsed/refractory B-cell acute lymphoblastic leukemia. *Am Soc Hematol*. 2018;132(Suppl 1):896.
32. Porteus M. Genome editing: A new approach to human therapeutics. *Annu Rev Pharmacol Toxicol*. 2016;56:163–190.
33. Long CBER. Term follow-up after administration of human gene therapy products, 2018. Retrieved from <https://www.fda.gov/media/113768/download>
34. Stadtmauer EA, Fraietta JA, Davis MM, et al. CRISPR-engineered T cells in patients with refractory cancer. *Science*. 2020;367:6481.
35. Eyquem J, Mansilla-Soto J, Giavridis T, et al. Targeting a CAR to the TRAC locus with CRISPR/Cas9 enhances tumour rejection. *Nature*. 2017;543(7643):113–117.
36. MacLeod DT, Antony J, Martin AJ, et al. Integration of a CD19 CAR into the TCR alpha chain locus streamlines production of allogeneic gene-edited CAR T cells. *Mol Ther*. 2017;25(4):949–961.
37. Myer VE (2019) Controlling rearrangement frequencies in the context of multigene genome editing. In *Genome Engineering: From Mechanisms to Therapies (B4)*. Victoria, BC: Keystone Symposia.
38. Godfrey DI, Le Nours J, Andrews DM, Uldrich AP, Rossjohn J. Unconventional T cell targets for cancer immunotherapy. *Immunity*. 2018;48(3):453–473.
39. Liu E, Tong Y, Dotti G, et al. Cord blood NK cells engineered to express IL-15 and a CD19-targeted CAR show long-term persistence and potent antitumor activity. *Leukemia*. 2018;32(2):520–531.
40. Liu E, Marin D, Banerjee P, et al. Use of CAR-transduced natural killer cells in CD19-positive lymphoid tumors. *N Engl J Med*. 2020;382(6):545–553.
41. Capsomidis A, Benthall G, Van Acker HH, et al. Chimeric antigen receptor-engineered human gamma delta T cells: enhanced cytotoxicity with retention of cross presentation. *Mol Ther*. 2018;26(2):354–365.
42. Gentles AJ, Newman AM, Liu CL, et al. The prognostic landscape of genes and infiltrating immune cells across human cancers. *Nat Med*. 2015;21(8):938–945.
43. Clarke RL, van der Stegen S, Lee T, et al. Abstract LB-108: Generation of off-the-shelf TCR-less CAR-targeted cytotoxic T cells from renewable pluripotent cells for cancer immunotherapy. *Cancer Res*. 2018;78(13 suppl.):LB-108-LB-108.
44. Montel-Hagen A, Seet CS, Li S, et al. Organoid-induced differentiation of conventional T cells from human pluripotent stem cells. *Cell Stem Cell*. 2019;24(3), 376–389.e8.
45. Schmitt TM, de Pooter RF, Gronski MA, Cho SK, Ohashi PS, Zúñiga-Pflücker JC. Induction of T cell development and establishment of T cell competence from embryonic stem cells differentiated in vitro. *Nat Immunol*. 2004;5(4):410–417.
46. Themeli M, Kloss CC, Ciriello G, et al. Generation of tumor-targeted human T lymphocytes from induced pluripotent stem cells for cancer therapy. *Nat Biotechnol*. 2013;31(10):928–933.
47. Dar H, Henderson D, Padalia Z, et al. Preclinical development of CTX120, an allogeneic CAR-T cell targeting BCMA. *Blood*. 2018;132(Suppl. 1):1921.
48. Sommer C, Boldajipour B, Kuo TC, et al. Preclinical evaluation of allogeneic CAR T cells targeting BCMA for the treatment of multiple myeloma. *Mol Ther*. 2019;27(6):1126–1138.
49. Prockop S, Doubrovina E, Baxi SS, et al. Treatment of EBV<sup>+</sup> nasopharyngeal carcinoma with banked EBV-specific cytotoxic T cells. *J Clin Oncol*. 2016;34(15 suppl):3012.
50. Prockop SE, Li AI, Baiocchi R, et al. Efficacy and safety of ATA129, partially matched allogeneic third-party Epstein-Barr virus-targeted cytotoxic T lymphocytes in a multicenter study for post-transplant lymphoproliferative disorder. *Biol Blood Marrow Transplant*. 2018;24(3):S41–S42.
51. Pender MP, Csurhes PA, Smith C, et al. Epstein-Barr virus-specific T cell therapy for progressive multiple sclerosis. *JCI Insight*. 2018;3(22):e124714.
52. Cobbold M, Khan N, Pourghesari B, et al. Adoptive transfer of cytomegalovirus-specific CTL to stem cell transplant patients after selection by HLA-peptide tetramers. *J Exp Med*. 2005;202(3):379–386.
53. Sukdolak C, Tischer S, Dieks D, et al. CMV-, EBV- and ADV-specific T cell immunity: screening and monitoring of potential third-party donors to improve post-transplantation outcome. *Biol Blood Marrow Transplant*. 2013;19(10):1480–1492.
54. Barker JN, Doubrovina E, Sauter C, et al. Successful treatment of EBV-associated posttransplantation lymphoma after cord blood transplantation using third-party EBV-specific cytotoxic T lymphocytes. *Blood*. 2010;116(23):5045–5049.
55. Haque T, Wilkie GM, Jones MM, et al. Allogeneic cytotoxic T-cell therapy for EBV-positive posttransplantation lymphoproliferative disease: results of a phase 2 multicenter clinical trial. *Blood*. 2007;110(4):1123–1131.
56. Haque T, Wilkie GM, Taylor C, et al. Treatment of Epstein-Barr virus-positive post-transplantation lymphoproliferative disease with partly HLA-matched allogeneic cytotoxic T cells. *Lancet*. 2002;360(9331):436–442.
57. Kurlander LS, Srinivasan A, Ghobadi A, et al. Efficacy and safety of tabellecleucel in patients with Epstein-Barr virus-associated leiomyosarcomas (EBV+ LMS). *Ann Oncol*. 2018;29(Suppl. 10):x11–x16.
58. Bar-Or A, Pender MP, Khanna R, et al. Epstein-Barr virus in multiple sclerosis: theory and emerging immunotherapies. *Trends Mol Med*. 2020;26(3):296–310.
59. Prockop S, Doubrovina E, Suser S, et al. Off-the-shelf EBV-specific T cell immunotherapy for rituximab-refractory EBV-associated lymphoma following transplant. *J Clin Invest*. 2020;130(2):733–747.
60. Aftab BT, Shen RR, Guzman-Becerra N, et al. Correlation of Epstein-Barr virus-specific cytotoxic T lymphocyte precursors (EBV-CTLp) after tabellecleucel with response and survival in rituximab relapsing or refractory EBV<sup>+</sup> post-transplant lymphoproliferative disease (PTLD). *J Clin Oncol*. 2020;38(5 suppl):38–38.
61. Prockop S, Doubrovina E, Hasan A, et al. Durable responses to 3rd party viral specific T cells mediated by different patterns of engraftment. *Cytotherapy*. 2019;21(5 Suppl.):S26.
62. Leen AM, Bollard CM, Mendizabal AM, et al. Multicenter study of banked third-party virus-specific T cells to treat severe viral infections after hematopoietic stem cell transplantation. *Blood*. 2013;121(26):5113–5123.
63. Melenhorst JJ, Leen AM, Bollard CM, et al. Allogeneic virus-specific T cells with HLA alloreactivity do not produce GVHD in human subjects. *Blood*. 2010;116(22):4700–4702.
64. Tzannou I, Papadopoulou A, Naik S, et al. Off-the-shelf virus-specific T cells to treat BK virus, human herpesvirus 6, cytomegalovirus, Epstein-Barr virus, and adenovirus infections after allogeneic hematopoietic stem-cell transplantation. *J Clin Oncol*. 2017;35(31):3547–3557.
65. Cruz CRY, Micklethwaite KP, Savoldo B, et al. Infusion of donor-derived CD19-redirected virus-specific T cells for B-cell malignancies relapsed after allogeneic stem cell transplant: a phase 1 study. *Blood*. 2013;122(17):2965–2973.
66. Curran KJ, Taylor C, Doubrovina E, et al. Validation of donor derived virus specific T-lymphocytes genetically modified to target the CD19 antigen for the treatment of relapsed leukemia. *Mol Ther*. 2011;19:S90.
67. Shen RR, Pham CD, Wu M, Munson DJ, Aftab BT. CD19 chimeric antigen receptor (CAR) engineered Epstein-Barr virus (EBV) specific T cells – an off-the-shelf, allogeneic CAR T-cell immunotherapy platform. *Cytotherapy*. 2019;21(5 Suppl.):S11.
68. Tse E, Kwong YL. Epstein Barr virus-associated lymphoproliferative diseases: the virus as a therapeutic target. *Exp Mol Med*. 2015;47:e136.

69. Curran KJ, Sauter CS, Kernan NA, et al. Durable remission following "Off-the-Shelf" Chimeric Antigen Receptor (CAR) T-cells in patients with relapse/refractory (R/R) B-cell malignancies. *Biol Blood Marrow Transplant*. 2020;26(3):S89.
70. Pule MA, Savoldo B, Myers GD, et al. Virus-specific T cells engineered to coexpress tumor-specific receptors: persistence and antitumor activity in individuals with neuroblastoma. *Nat Med*. 2008;14(11):1264–1270.
71. Louis CU, Savoldo B, Dotti G, et al. Antitumor activity and long-term fate of chimeric antigen receptor-positive T cells in patients with neuroblastoma. *Blood*. 2011;118(23):6050–6056.
72. Chapuis A, Egan DN, Bar M, et al. EBV-specific donor cells transduced to express a high-affinity WT1 TCR can prevent recurrence in post-HCT patients with high-risk AML. *Blood*. 2016;128(22):1001.
73. Chapuis AG, Egan DN, Bar M, et al. T cell receptor gene therapy targeting WT1 prevents acute myeloid leukemia relapse post-transplant. *Nat Med*. 2019;25(7):1064–1072.
74. Greenberg PD, Anderson KG, Egan D, et al. Abstract IA01: Engineering T cell responses to tumors: taking the immune system where no responses have gone before. *Cancer Immunol Res*. 2016;4(11 Suppl.):IA01–IA01.
75. Haapaniemi E, Botla S, Persson J, Schmierer B, Taipale J. CRISPR-Cas9 genome editing induces a p53-mediated DNA damage response. *Nat Med*. 2018;24(7):927–930.
76. Rosato PC, Wijeyesinghe S, Stolley JM, et al. Virus-specific memory T cells populate tumors and can be repurposed for tumor immunotherapy. *Nat Commun*. 2019;10(1):567.
77. Park JH, Rivière I, Gonen M, et al. Long-term follow-up of CD19 CAR therapy in acute lymphoblastic leukemia. *N Engl J Med*. 2018;378(5):449–459.
78. Fry TJ, Connick E, Falloon J, et al. A potential role for interleukin-7 in T-cell homeostasis. *Blood*. 2001;97(10):2983–2990.
79. Sandau MM, Winstead CJ, Jameson SC. IL-15 is required for sustained lymphopenia-driven proliferation and accumulation of CD8 T cells. *J Immunol*. 2007;179(1):120–125.
80. Lee DW, Kochenderfer JN, Stetler-Stevenson M, et al. T cells expressing CD19 chimeric antigen receptors for acute lymphoblastic leukaemia in children and young adults: a phase 1 dose-escalation trial. *Lancet*. 2015;385(9967):517–528.
81. Mueller KT, Maude SL, Porter DL, et al. Cellular kinetics of CTL019 in relapsed/refractory B-cell acute lymphoblastic leukemia and chronic lymphocytic leukemia. *Blood*. 2017;130:2317–2325.
82. Salter AI, Ivey RG, Kennedy JJ, et al. Phosphoproteomic analysis of chimeric antigen receptor signaling reveals kinetic and quantitative differences that affect cell function. *Sci Signal*. 2018;11(544).
83. Majzner RG, Mackall CL. Clinical lessons learned from the first leg of the CAR T cell journey. *Nat Med*. 2019;25(9):1341–1355.
84. O'Reilly RJ, Prockop S, Hasan AN, Koehne G, Doubrovina E. Virus-specific T-cell banks for 'off the shelf' adoptive therapy of refractory infections. *Bone Marrow Transpl*. 2016;51(9):1163–1172.
85. Valton J, Guyot V, Marechal A, et al. A multidrug-resistant engineered CAR T cell for allogeneic combination immunotherapy. *Mol Ther*. 2015;23(9):1507–1518.
86. Wang D, Quan Y, Yan Q, Morales JE, Wetsel RA. Targeted disruption of the beta2-microglobulin gene minimizes the immunogenicity of human embryonic stem cells. *Stem Cells Transl Med*. 2015;4(10):1234–1245.
87. Gornalusse GG, Hirata RK, Funk SE, et al. HLA-E-expressing pluripotent stem cells escape allogeneic responses and lysis by NK cells. *Nat Biotechnol*. 2017;35(8):765–772.
88. Lanza R, Russell DW, Nagy A. Engineering universal cells that evade immune detection. *Nat Rev Immunol*. 2019;19(12):723–733.
89. Harding J, Vintersten-Nagy K, Shutova M, et al. Induction of long-term allogeneic cell acceptance and formation of immune privileged tissue in immunocompetent hosts. *bioRxiv*. 2019;716571.
90. DiStasi A, Tey S-K, Dotti G, et al. Inducible apoptosis as a safety switch for adoptive cell therapy. *N Engl J Med*. 2011;365(18):1673–1683.
91. Liang Q, Monetti C, Shutova MV, et al. Linking a cell-division gene and a suicide gene to define and improve cell therapy safety. *Nature*. 2018;563(7733):701–704.

**How to cite this article:** Aftab BT, Sasu B, Krishnamurthy J, Gschweng E, Alcazer V, Depil S. Toward "off-the-shelf" allogeneic CAR T cells. *Adv Cell Gene Ther*. 2020;3:e86. <https://doi.org/10.1002/acg2.86>

## REVUE 7 - Human Endogenous Retroviruses (HERVs): Shaping the Innate Immune Response in Cancers

***Vincent Alcazer***<sup>1,2\*</sup>, *Paola Bonaventura*<sup>1,3</sup>, *Stéphane Depil*<sup>1,3,4,5\*</sup>

1 Cancer Research Center of Lyon, 69008 Lyon, France

2 Department of Clinical Hematology, Centre Hospitalier Lyon Sud, Hospices Civils de Lyon, 69310 Pierre-Bénite, France

3 Centre Léon Bérard, 69008 Lyon, France

4 Université Claude Bernard Lyon 1, 69008 Lyon, France

5 ErVaccine Technologies, 69008 Lyon, France

\* Corresponding authors:

Vincent Alcazer, Cancer Research Center of Lyon, 69008 Lyon, France

Email: [vincent.alcazer@lyon.unicancer.fr](mailto:vincent.alcazer@lyon.unicancer.fr)

Stéphane Depil, Cancer Research Center of Lyon, 69008 Lyon, France

Email: [stephane.depil@lyon.unicancer.fr](mailto:stephane.depil@lyon.unicancer.fr)

Cancers. 2020 Mar 6;12(3):610. doi: 10.3390/cancers12030610. PMID: 32155827; PMCID: PMC7139688.



Review

# Human Endogenous Retroviruses (HERVs): Shaping the Innate Immune Response in Cancers

Vincent Alcazer <sup>1,2,\*</sup>, Paola Bonaventura <sup>1,3</sup> and Stephane Depil <sup>1,3,4,5,\*</sup><sup>1</sup> Cancer Research Center of Lyon, 69008 Lyon, France; paola.bonaventura@lyon.unicancer.fr<sup>2</sup> Department of Clinical Hematology, Centre Hospitalier Lyon Sud, Hospices Civils de Lyon, 69310 Pierre-Bénite, France<sup>3</sup> Centre Léon Bérard, 69008 Lyon, France<sup>4</sup> Université Claude Bernard Lyon 1, 69008 Lyon, France<sup>5</sup> ErVaccine Technologies, 69008 Lyon, France

\* Correspondence: vincent.alcazer@chu-lyon.fr (V.A.); stephane.depil@lyon.unicancer.fr (S.D.)

Received: 10 February 2020; Accepted: 5 March 2020; Published: 6 March 2020



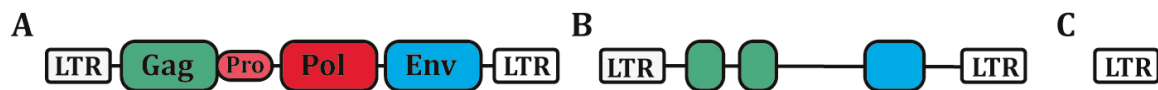
**Abstract:** Human Endogenous Retroviruses (HERVs) are accounting for 8% of the human genome. These sequences are remnants from ancient germline infections by exogenous retroviruses. After million years of evolution and multiple integrations, HERVs have acquired many damages rendering them defective. At steady state, HERVs are mostly localized in the heterochromatin and silenced by methylation. Multiple conditions have been described to induce their reactivation, including auto-immune diseases and cancers. HERVs re-expression leads to RNA (simple and double-stranded) and DNA production (by reverse transcription), modulating the innate immune response. Some studies also argue for a role of HERVs in shaping the evolution of innate immunity, notably in the development of the interferon response. However, their exact role in the innate immune response, particularly in cancer, remains to be defined. In this review, we see how HERVs could be key-players in mounting an antitumor immune response. After a brief introduction on HERVs characteristics and biology, we review the different mechanisms by which HERVs can interact with the immune system, with a focus on the innate response. We then discuss the potential impact of HERVs expression on the innate immune response in cancer.

**Keywords:** human endogenous retroviruses; innate immunity; cancer

## 1. Introduction

About half of the human genome is composed of transposable elements (TEs) [1,2]. Initially considered as “junk DNA”, these mobile sequences have been shown to provide a major source of structural variation in the genome with an impact on oncogenesis [1]. Among them, Human Endogenous Retroviruses (HERVs) represent 8% of the human genome. These DNA sequences of retroviral origin are the legacy of ancient germ line infections by exogenous retroviruses, persisting as integrated retroviruses. HERVs complete structure is the same as that of exogenous retroviruses, consisting of four genes (*gag*, *pro*, *pol* and *env*) flanked by two long-terminal repeats (LTRs) (Figure 1A) [3]. These LTRs function as promoters for HERV expression, have strong RNA regulatory sequences and contain transcription factor binding sites. Successive recombination between the two LTR led to internal deletion in the vast majority of HERVs, leading to the excision of most of the coding regions [2]. The hundred thousand copies of HERVs present in the human genome is thus mainly constituted of solitary LTR, only few thousands among the most recent integrated ones (such as the HML family) keeping partially conserved open reading frame (ORF) (Figure 1B,C). It is important to note that humans differ from other mammals in this point, and particularly from mice which, unlike humans, possess complete replication-competent ERVs sequences [4].





**Figure 1.** Structure of HERVs. (A) The complete proviral sequence of HERVs is composed of four protein-coding sequences (Gag, Pro, Pol and Env) surrounded by two long-terminal repeat (LTR) promoter sequences. (B) Most HERVs have acquired many damages on their ORFs, rendering them defective. Still, protein-coding sequences remain in some HERVs. (C) The vast majority of HERVs is now composed of solitary LTR. HERVs: Human Endogenous Retroviruses, LTR: Long-terminal Repeat, ORFs: Open-reading frames.

HERVs are mainly localized in heterochromatin and repressed by epigenetic silencing at steady state [5]. Only few physiological functions have been linked to HERVs expression so far. The best known physiological example are Syncitins 1 and 2, two HERV-W/HERV-FRD derived envelope proteins expressed in the placenta playing a role in trophoblast formation and foeto-maternal tolerance [6]. Different studies also suggest a role for HERVs in the complex human brain development: In human neural progenitor cells, HERVs act as a docking platform for the epigenetic co-repressor protein TRIM28, inducing a local state of heterochromatin around these HERVs [7]. It has been shown that the transfer from motoneuron to muscles of Arc/Arg3.1, a crucial protein for synaptic plasticity, is regulated by a Gag-like protein contained in its 3' UTR region [8]. Conversely, increased HERVs transcription and activity has been described in several pathological conditions, including auto-immune diseases and cancer [1,4]. Indeed, the loss of CpG methylation in tumors seems to preferentially affect retrotransposon elements and particularly HERVs [4].

Due to their retrotransposition ability, HERVs can be involved in oncogenesis in several ways such as insertional mutagenesis or chromosomal instability [1]. LTR-induced oncogene transcription has also been reported recently [9]. Furthermore, HERV-derived proteins can exert direct oncogenic activity, as it is the case for instance with Np9, an HERV-K derived protein that can activate Wnt/ $\beta$ -catenin, Ras-ERK, cMyc/AKT and Notch1 signaling pathways in chronic lymphocytic leukemia [10].

HERVs may have different impacts on the immune response. On one hand, the transmembrane subunit of the Envelope (Env) protein from different HERVs contains a conserved immunosuppressive domain (ISD) [6]. Of note, the mechanisms and targets of Env-mediated immunosuppression are not fully defined, and its role in cancer immune escape is not clearly established. On the other hand, HERVs' replication intermediates can stimulate innate sensors promoting antitumoral immune response by triggering a type I and III Interferon (IFN) response [11,12]. HERVs expression is thus a double-edged sword for antitumor immunity, favoring either immune escape or immune activation.

In this review, we will present the main data linking HERVs expression with the innate immune response in cancer. After a brief review of the mechanisms regulating HERVs expression, we will discuss the interactions between HERVs and the innate immune response, before considering the implication of HERVs expression in the antitumor immune response.

## 2. HERVs Expression: Interaction with the Innate Immune System

### 2.1. HERVs Are Silenced by Epigenetic Mechanisms at Steady State

Epigenetics plays a key role in embryonic development, allowing a fine regulation of cell fate and pluripotency and contributing to HERVs repression [5]. Histone H3 trimethylation, at the lysine 9 position (H3K9me3) by the histone-lysine N-methyltransferase SETDB1, is involved in early HERVs' silencing during embryogenesis, while DNA methylation is thought to be the main mechanism in differentiated cells. As more than 90% of methylated CpGs are located in part of the genome harboring TEs, some authors even suggested that cytosine methylation was developed as a defense mechanism against TEs spreading in human [11,13]. The exact role of each mechanism is, however, still a matter of debate, different studies suggesting a more generalized role for SETDB1 in differentiated cells such

as lymphocytes [14,15]. In mice, SETDB1 was shown to ensure Th2 cell stability by repressing ERVs that control the Th1 gene network [15], and also to repress the expression of Fc $\gamma$ R-IIb in T cells, an inhibitory Fc receptor regulating the signaling through the TCR complex [16]. Accordingly, SETDB1 knock out induces an increased phosphorylation of ZAP70 in response to CD3 agonism, resulting in an increased cell activation and death at early thymic selection stage. In B cells, loss of SETDB1 induces a significant expression of HERVs, incompatible with B cell development and survival [14,17].

TRIM28 may also play an important role as shown in human neural progenitor cells by inducing a local heterochromatin state to suppress HERVs expression [7]. Derepression of this local network by knockdown of TRIM28 induces the transcription of both HERVs and their surrounding genes, providing another layer of epigenetic control in these cells. Another key player is the retinoblastoma protein (pRB) which recruits EZH2 to establish trimethylation of the 27th lysin of histone H3 (H3K27me<sub>3</sub>, a marker of heterochromatin) on tandem and interspersed genomic repeats in somatic cells [18]. Multiple factors are thus participating in the control of HERVs expression through epigenetic modifications, and there is still a need to better understand and identify all the actors involved in this regulation.

## 2.2. HERVs Expression Is Modulated by the Immune Response, Microbiota and Viral Infections

Multiple immune stimuli have been shown to induce ERVs expression. First evidences of Murine Leukemia Virus (MLVs) induction following murine B cells activation were reported in the 80's by Stoye J. and Moroni C [19]. The authors showed that both B cell mitogens (lipopolysaccharide (LPS) and lipoprotein (LP)) and 5-bromo-2'-deoxyuridine (BrdU) were able to induce ERVs expression in mice, and that their combination was synergistic, BrdU preventing the methylation of newly replicated DNA. More recently, Young G. and Kassiotis G. assessed ERVs expression by microarray in murine and human immune cells under different stimulatory conditions [20]. ERVs assessment in gut tissues from wild type or MYD88 knock out mice bred in two different conditions (specific pathogen-free versus germ-free) revealed that expression of certain ERVs was dependent on the presence of the gut microbiota. Consequently, ERVs expression was largely decreased in MYD88<sup>-/-</sup> mice, supporting the role of the microbiota in inducing a basal ERVs expression in mice's gut through MYD88 signaling. Still in mice, deficiency in Toll-like Receptors (TLR)-3 and -9 (recognizing double stranded (ds) RNA) or in TLR-7 (recognizing single stranded (ss) RNA) was shown to allow endogenous MLVs reactivation [21].

In human, analysis of blood samples from HIV-1 infected patients show an increase in HERV-K RNA expression [22]. HERV-K protein production is also increased after HIV-1 infection of CD4 T cell lines [23]. Similar results were found in gut biopsies from ulcerative colitis patients compared to healthy individuals, confirming that not only external pathogens but also commensal bacteria can modulate HERVs expression [20]. Peripheral blood mononuclear cells (PBMCs) treated by type I IFN express high levels of Alu elements, and IFN- $\alpha$  stimulation following a viral infection is known to induce transcription of certain HERVs [24,25]. In this latter case, it was clearly shown that the expression of HERV-K18 Env-derived superantigen IDDMK1,222, which is known to induce the activation of autoreactive T cells associated with type I diabetes development, is strongly induced by IFN- $\alpha$  [25]. Type II IFN can also promote the expression of a particular class of HERVs, located in the 3' untranslated region of specific genes enriched for STAT1 and EZH2 regulation motifs [26]. IFN- $\gamma$  exposure induces the bidirectional transcription from these sequences, leading to dsRNA generation with 3' antisense retroviral coding sequences (labelled SPARCS which correspond to antisense ERVs) transcription.

Other mechanisms can participate to ERVs transcription in case of viral infection. Small ubiquitin-like modifiers (SUMO) are a family of proteins involved in the regulation of multiple cellular proteins and pathways by inducing post-translational modifications [27]. SUMOylation has been shown to be a key process during the host response to viral infection by suppressing the polyubiquitination and degradation of PRRs such as RIG-1 and MDA-5 [28]. There is also evidence for TRIM28 SUMOylation, but not phosphorylation, to be implicated in ERVs repression [29]. Recently, it has been shown that influenza virus infections induce loss of a SUMO-modified form of TRIM28, leading to a derepression of ERVs [30]. The resulting dsRNA-mediated type I IFN response enhances

the host antiviral response, illustrating how an aberrant transcription of self-RNA derived from ERVs can be sensed as non-self by PRRs during an infection. Nevertheless, the IFN-antagonist protein NS1 produced by the wild-type influenza virus can oligomerize around dsRNAs to mitigate the antiviral effects of ERVs transcription [30].

Adaptive immunity could also play a role in the control of ERVs expression. Antibody-deficient mice present the spontaneous expression of fully infectious MLVs copies, leading to the development of lymphomas [31]. Evidences of specific CD8<sup>+</sup> and CD4<sup>+</sup> T cells directed against HERV-derived peptides have also been reported, as discussed later.

### 2.3. Role of PRRs, TLR Signaling and B Cell Response in the Control of ERVs

While studies have provided a better understanding of the different mechanisms underlying HERVs expression, the link between HERVs and the immune system still remains poorly defined, with most evidences coming from exogenous retroviruses or mice models ([2] for review). As discussed above, TLR deficiency can lead to ERVs reactivation in mice [21]. Yu et al. showed that while TLR-7 deficiency alone leads to spontaneous viremia, additional TLR-3 and -9 deficiencies lead to the development of acute lymphoblastic T cell lymphoma. Interestingly, TLR-7 deficient mice expressed ERVs but lacked ERVs-specific antibody response, confirming a previous study showing that TLR-7 is essential for initiating a B cell response directed against ERVs, and that CD4<sup>+</sup> T cell responses are also attenuated in the absence of TLR-7, whereas CD8<sup>+</sup> T cell responses are conserved [32]. In this latter study, the authors showed that intrinsic TLR signaling in B cells is mediating the response, while TLR signaling in DCs is less important. [33]. Various levels of antibody specificities and efficiencies in TLR-3 and 9 simple or double deficient mice also suggest a modulatory role for these two TLRs in ERV-specific antibody response. Of note, both studies are describing responses against a fully infectious murine ERV, the exogenous Friend-MuLV retrovirus, and may not be extrapolated to human ERVs, since most of them are defective.

Altogether these data illustrate how the immune system modulates and controls ERVs expression, leading to malignancy in case of immune deficiency. Recently, it has been shown that interferon- $\gamma$ -inducible protein 16 (IFI16), an AIM2-like receptor for DNA, is able to sense HERV-K-derived ssDNA in somatic tissues, while being absent in stem cells [34]. It seems likely that different mechanisms evolved to provide a better control of an aberrant transcription of these sequences in somatic tissues while allowing expression in stem cells where they may contribute to cell identity.

### 2.4. HERVs as Network Regulator: the Example of the IFN- $\gamma$ Network

With their ability to move and replicate within the genome, TEs have provided a major source of gene regulation. Indeed, once inserted, TEs can provide an alternative promoter sequence and introduce an alternative transcription start site, interrupt a pre-existing promoter, introduce a transcription factor binding site, drive an antisense transcription or silence a whole region by inducing a local state of heterochromatin [35]. TEs can also act post-transcriptionally by providing an alternative polyadenylation site, an alternative splicing or a target for micro RNA. Among the multiple examples of gene network regulated by TEs, it is now clear that HERVs participated in the shaping and the development of the IFN- $\gamma$  network [36]. By providing already functional cis-regulatory elements to surrounding genes, TEs insertion promotes emergence of complex gene regulatory networks that would take much longer time to develop by independent de novo mutations [37]. Chuong et al. thus found 27 families of TEs enriched within IFN- $\gamma$ -responsive cis regulatory elements, among which 20 were LTRs from the HERV family with STAT1 binding motifs [36]. The authors also showed that the only STAT1 binding site within the 50 kb of the Absent in Melanoma 2 gene (AIM2, an IFN-stimulated gene encoding a sensor of foreign cytosolic DNA promoting the assembly of the inflammasome) is located in a LTR sequence, derepressed by H3K27 acetylation after IFN- $\gamma$  exposure. Suppression of this LTR element by Crispr-Cas9 completely abrogated AIM2 expression upon IFN- $\gamma$  treatment. Activated caspase-1 levels, a surrogate marker of pyroptotic cell death, were also markedly reduced after

vaccinia-virus infection, demonstrating that this LTR element is indispensable for the inflammatory response to infection. In line with this, an impaired activity of IFN-stimulated genes is observed in cell lines deleted of these HERV sequences [36]. HERVs may thus have been key actors in the development and the shaping of our immune system, by providing ready-to-use IFN- $\gamma$  inducible enhancers.

### 3. HERVs as Modulators of the Innate Immune Response

#### 3.1. Activation of Innate Immunity by HERV Nucleic Acids

Pattern recognition receptors (PRRs) are key receptors of the innate immune system. Their ability to directly recognize conserved pathogen-associated molecular patterns (PAMPs) allow them to quickly trigger the innate immune response, acting in first line after an infection. Among them, two main families are specialized in the recognition of nucleic acids: endosomal PRRs (the TLR family), and cytosolic PRRs (including the RIG-I-Like Receptors (RLRs), the NOD-like receptors (NLRs), the C-type lectin receptors (CLRs) families and analogous DNA sensing receptors including oligoadenylate synthase (OAS) proteins and cyclic guanosine monophosphate–adenosine monophosphate (cGAMP) synthase (cGAS)) [38]. Once engaged, PRRs trigger different signaling cascades leading to inflammatory cytokines production and converging to type I IFN production. The pathways involved depends on the type of PRR. Schematically, TLRs signal through MYD88 or TRIF, while RIG-I (which recognizes short dsRNA and RNA with 5' triphosphate ends) and MDA5 (which recognizes long dsRNA) both signal through MAVS, and DNA sensors signal through STING [38]. All together these effectors give rise to the innate immune response and the “danger signal” required for inducing a specific adaptive response [39].

The recognition of nucleic acids is at the time a very efficient strategy to fight viral infections and paradoxically a risky strategy due to the possibility of recognizing self-nucleic acids, promoting autoimmunity. With their ability to produce both RNA and DNA intermediates (by reverse-transcription), ERVs are constituting perfect candidates for stimulating PRRs, as complementary sense and anti-sense ERVs transcripts are forming both ssRNA and dsRNA. SsRNA can be sensed by TLR-7 and 8 resulting in IFN- $\alpha$  secretion by stimulated dendritic cells (DCs) and macrophages, as shown with HIV [40]. DsRNA, which are not found in normal cells, could be one of the most immunogenic nucleic acid PAMPs. HERV dsRNA can be recognized by TLR-3, RIG-I and MDA5, the two latters signaling through MAVS to induce a type I IFN response [11].

Once retrotranscribed into DNA, retroviruses can be sensed by cGAS to produce cGAMP, which binds and activates STING to trigger a type I IFN response through nuclear factor-kappa B (NF- $\kappa$ B) and IFN regulatory factor 3 (IRF3) activation [41]. DsDNA could also be sensed by DNA-dependent activator of IFN-regulatory factors (DAI), and DNA:RNA hybrids by TLR-9, even if this has not been demonstrated for the particular case of HERVs [42,43].

Regulation mechanisms exist, such as 3' repair exonuclease 1 (TREX1), a 3'→5' DNA exonuclease that degrades retroviral DNA, thus preventing their accumulation [44]. TREX1 knock down increases type I IFN production after HIV infection in mucosal cells [45]. Albeit performed in mice, these experiments suggest that similar mechanisms could exist for modulating immune activation from nucleic acid sensing.

#### 3.2. Activation of Innate Immunity by HERV-Derived Proteins

Not only nucleic acids but also proteins patterns can stimulate the innate immune response. Transmembrane TLR-2 and 4 are able to recognize retroviral envelope glycoproteins (at least of exogenous origin as shown for HIV-1), triggering a pro-inflammatory response through NF- $\kappa$ B activation [46]. Specific data regarding HERVs are still lacking to further extrapolate these results. Envelope subunit of multiple sclerosis-associated retroviral element (MSRV) from the HERV-W family was shown to specifically activate immune cells through CD14 and TLR-4, leading to production of pro-inflammatory cytokines such as IL-1 $\beta$ , IL-6 and TNF- $\alpha$ . It can activate DCs to promote the



development of a Th1-response [47]. This could be an important immune activation pathway in the pathogenesis of multiple sclerosis, as suggested by a multiple sclerosis mouse model where HERV-W Env (Syncitin) overexpression in astrocytes led to neuroinflammation and oligodendrocytes' death [48].

### 3.3. Suppression of the Immune Response

Most exogenous retroviruses have been described to suppress the host's immune system to maintain the infection [49]. Regarding ERVs, this effect has been co-opted for the foeto-maternal tolerance, promoted by the expression in the placenta of Syncitin 2, a HERV-FRD-derived Env protein [50]. This shared property between exogenous and endogenous retroviruses led to the identification of an immunosuppressive domain (ISD) in the transmembrane unit of the Env protein (reviewed in [51]). Exposure of human PBMCs to ISD-derived peptides from different retroviruses, including HIV and HERV-K, increases the expression of different cytokines, including IL-10, IL-6, IL-8, RANTES, MCP-1, MCP-2, TNF- $\alpha$ , MIP-1 $\alpha$ , MIP-1 $\beta$ , MIP-3, IL-1 $\beta$  and Gro( $\alpha,\beta,\gamma$ ), and decreases the expression of IL-2 and CXCL9 [52,53]. Although all studies converge so far to a direct effect on cellular immune effectors, the exact impact of ISD remains unclear, and no interaction with the innate immune response has been clearly highlighted. An interesting study was performed in zebra-fish embryos, which have only innate immunity during the first month of development [54]. Embryos injected with HERV-K-containing extra-cellular vesicles derived from two colorectal adenocarcinoma cell lines treated with decitabine, showed a reduced expression of IL-1 $\beta$  and myeloperoxidase. This effect was however modest and needs to be investigated in other models. Of note, it has been also suggested that HERV-derived Env proteins could target T cell activation indirectly by modulating the stimulatory activity of dendritic cells [55].

## 4. HERVs as Innate “Adjuvants” in Cancer

### 4.1. Turning Cancerous Cells into Virus-Infected Cells

While the effects of 5-Azacytidine (AZA) in promoting ERVs expression had been described since 1984, the link with its anti-tumoral effect in human was only recently established [11,56]. Studies first showed the link between the use of AZA and the modulation of multiple immune pathways, including upregulation of immunomodulatory pathways and cancer testis antigens [57]. The classification of primary human tumors based on the union of these pathways confirmed the enrichment of this AZA-induced Immune Genes (AIM) after AZA therapy in breast and colorectal cancer trials, suggesting that patients with low AIM signature at diagnosis might benefit most from receiving epigenetic therapy prior to immunotherapy.

Chiappinelli K. et al. showed that AZA-induced response was mediated by cytosolic sensing of dsRNA via TLR-3 following expression of demethylated HERVs, leading to a type I IFN production and apoptosis of tumor cells. This activation could stimulate anti-tumor immunity by activating innate lymphoid cells and promoting the infiltration and clonal expansion of tumor-specific cytotoxic T cells (“viral mimicry mechanism”). The induction of HERVs expression could also lead to the exposure of several HERV-derived tumor-associated antigens, leading to new targets on the tumor cells [58]. These results were confirmed by Roulois D. et al. with the use of a slightly different DNA methyltransferase inhibitor (DNMTi), 5-aza-2'-deoxycytidine (AZA-CdR) in enriched sphere cultures derived from human colorectal cancer [12]. A low dose of 5-AZA-CdR also induced type I IFN and reduced the frequency of colorectal cancer initiating cells (CICs) independently of the cancer's methylation profile. Knockdown of MDA5, MAVS, or IRF7 was sufficient to render these cells insensitive to 5-AZA-CdR treatment, in contrast with RIG1 knockdown, suggesting a major role for MDA5 sensor in 5-AZA-CdR efficiency. Interestingly, while 5-AZA-CdR also induced the expression of IFN-responsive genes, the MDA5-mediated apoptosis of tumor cells seemed to be independent of the IFN response, as already suggested in melanoma [59].

Other groups found similar results in different cancer types. In myelodysplastic syndromes, decrease in DNA methylation following AZA treatment was shown to be enriched in heterochromatin zones and not in coding genes, suggesting that HERVs activation is one of the mechanisms contributing to AZA clinical effects in these diseases [60]. High-grade glioblastoma, which exhibits repressive chromatin marks (by loss of H3K27 trimethylation) at some sites, contains reciprocal active chromatin marks (by gain of H3K27 acetylation) enriched at repeated elements sites, leading to their increased expression [61]. This acetylation can in fact be seen as a therapeutic vulnerability, rendering TEs and HERVs more susceptible to activation by epigenetic therapies such as AZA or histone deacetylase inhibitors (HDACi).

HERVs expression was also shown to overcome the resistance to anti-PD1 therapy in a melanoma mouse model: the ablation of the histone H3K4 demethylase LSD1 (KDM1A) enhances tumor immunogenicity and T cell infiltration by inducing dsRNA production through HERVs expression, leading to a type I IFN response [62]. This could be a way to turn a so-called “cold” into a “hot” tumor.

In a different approach, Cyclin-dependent kinases 4 and 6 inhibitors (CDK4/6is) have been shown to increase the anti-tumoral immune response by activating the expression of HERVs in tumor cells [63]. By inhibiting the phosphorylation of the pRB, they reduce DNMTs activity, promoting the expression of HERVs and other genes regulating immune functions. The increased intracellular levels of dsRNA stimulate the production of type III IFN, enhancing tumor antigen presentation and promoting a specific cytotoxic T cell response. CDK4/6is also lead to a selective suppression of regulatory T cell proliferation, potentially related to the higher expression of pRB by these latter. As with DNMTs, CDK4/6is might enhance the efficacy of immunotherapy such as immune checkpoints blockade.

Altogether these results explain the mechanism behind the anti-tumoral effects of DNMTs, characterized by a delayed response time and paradoxically an absence of predictive methylation markers. By mimicking viral infection into the tumor cell through HERVs expression, epigenetic therapies could thus represent one approach to turn “cold” into “hot” tumors by promoting immune-cells infiltration and antigen presentation through type I IFN production while triggering cancerous-cell apoptosis and providing new targets for adaptive response [58,64]. All these effects put DNMTs as privileged candidates for association with immune therapies.

#### *4.2. Impact of HERVs Expression in the Response to Immune Checkpoint Inhibitors*

By activating both innate and adaptive immune response, HERVs may represent key factors in the response to immunotherapy. HERV expression has recently been associated with response to anti-PD1 in clear cell renal cell carcinoma [65]. The “viral defense” gene expression signature associated with HERVs expression is also predictive of immune checkpoint responses in melanoma patients, and DNMTs treatment sensitize to anti-CTLA-4 therapy in a mouse melanoma model [11]. In the same way, Heidegger S. and Wintges A. et al. showed that RIG-I activation is indispensable for responsiveness to CTLA-4 checkpoint blockade [66].

HERVs could thus be involved in the response to checkpoint blockade in tumors characterized by a low mutational burden. It has recently been shown that in rhabdoid tumors, a pediatric cancer exhibiting one of the lowest mutational burden, the driver mutation characterized by a biallelic loss of SMARCB1 (a core member of a chromatin remodeler complex) triggers the expression of HERVs, explaining thus the unusual immune infiltration of these tumors, which is driven by the subsequent type III IFN response [67].

HERVs expression may also modulate the antitumor immune response once the latter has been initiated. For instance, SPARCS expression following IFN- $\gamma$  exposure could be a positive feedback mechanism fostering the antitumoral immune response [26].

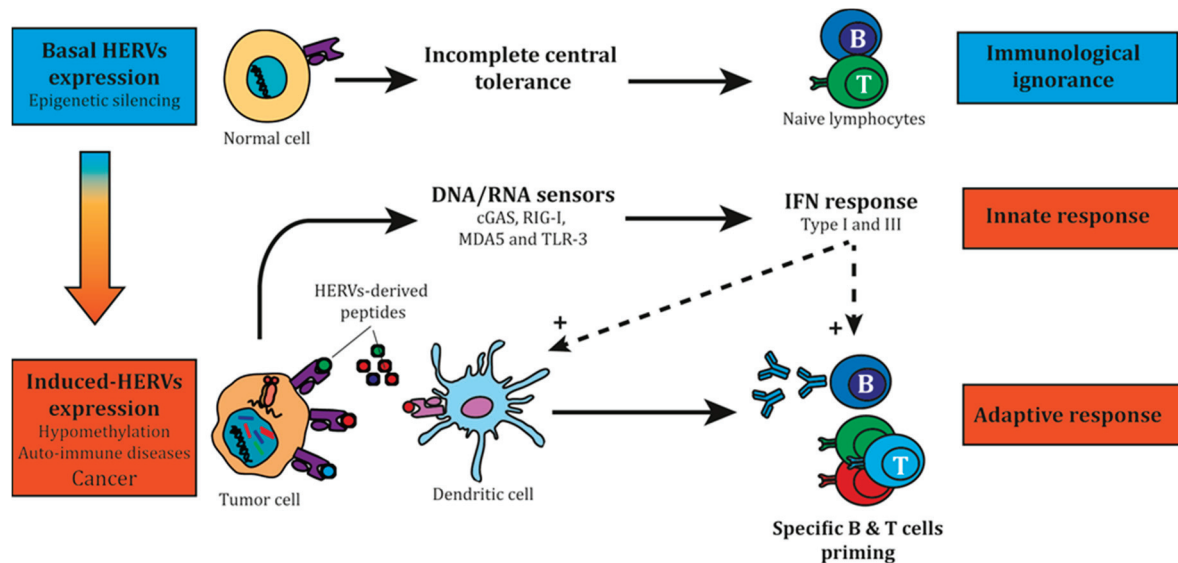
## **5. Conclusions**

The integration of ERVs in the genome represents a good example of biological exaptation. ERVs have been co-opted by the host during evolution to modulate key functions and complex gene



regulatory networks (such as IFN- $\gamma$ ). ERVs have also shaped the immune system, which, in turn, may control their expression. Since ERVs represent major factors of activation of the innate immune response, the induction of their expression in tumor cells is of great therapeutic interest. Indeed, many epigenetic drugs may act through ERVs derepression. This could represent a key mechanism to transform cold tumors into hot ones, justifying the rationale of combination of epigenetic drugs with immune checkpoint inhibitors.

With the potential to stimulate both the innate response through DNA/RNA sensing and the adaptive immune response through antigens recognition [68–70], ERVs provide unique and promising targets for cancer immunotherapy (Figure 2).



**Figure 2.** HERVs as relevant and unique targets for cancer immunotherapy. At steady state, most HERVs are not expressed, mainly silenced by epigenetic mechanisms. Basal HERV expression can be found in some tissues and in the thymus, with an incomplete central tolerance. Different circumstances such as auto-immune diseases or cancer can foster HERVs expression, leading to the production of both HERVs-derived nucleic acids and proteins. These intermediates can promote an innate immune response by inducing type I and III IFN production through DNA/RNA sensors, and also an adaptive immune response through B and T cell epitopes. IFN: Interferon.

**Author Contributions:** Conceptualization, V.A., P.B. and S.D.; writing—original draft preparation, V.A.; writing—review and editing, V.A., P.B. and S.D. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was financially supported by the PRT-K grant for research in translational medicine for oncology, the Ligue Nationale contre le Cancer, the LabEx DEVweCAN and LYriCAN and the Agence de Biomédecine.

**Conflicts of Interest:** S.D. is founder of ErVaccine Technologies and member of the management team of Swiss Rockets. Consulting or scientific advisory boards: Netris Pharma, Vaximm. The other authors have no conflict of interest to declare.

## References

1. Burns, K.H. Transposable elements in cancer. *Nat. Rev. Cancer* **2017**, *17*, 415–424. [[CrossRef](#)]
2. Kassiotis, G.; Stoye, J.P. Immune responses to endogenous retroelements: Taking the bad with the good. *Nat. Rev. Immunol.* **2016**, *16*, 207–219. [[CrossRef](#)]
3. Vargiu, L.; Rodriguez-Tomé, P.; Sperber, G.O.; Cadeddu, M.; Grandi, N.; Blikstad, V.; Tramontano, E.; Blomberg, J. Classification and characterization of human endogenous retroviruses; mosaic forms are common. *Retrovirology* **2016**, *13*, 7. [[CrossRef](#)]

4. Kassiotis, G. Endogenous retroviruses and the development of cancer. *J. Immunol.* **2014**, *192*, 1343–1349. [[CrossRef](#)] [[PubMed](#)]
5. Rowe, H.M.; Trono, D. Dynamic control of endogenous retroviruses during development. *Virology* **2011**, *411*, 273–287. [[CrossRef](#)] [[PubMed](#)]
6. Dupressoir, A.; Laviaille, C.; Heidmann, T. From ancestral infectious retroviruses to bona fide cellular genes: Role of the captured syncytins in placentation. *Placenta* **2012**, *33*, 663–671. [[CrossRef](#)] [[PubMed](#)]
7. Brattås, P.L.; Jönsson, M.E.; Fasching, L.; Nelander Wahlestedt, J.; Shahsavani, M.; Falk, R.; Falk, A.; Jern, P.; Parmar, M.; Jakobsson, J. TRIM28 controls a gene regulatory network based on endogenous retroviruses in human neural progenitor cells. *Cell Rep.* **2017**, *18*, 1–11. [[CrossRef](#)]
8. Ashley, J.; Cordy, B.; Lucia, D.; Fradkin, L.G.; Budnik, V.; Thomson, T. Retrovirus-like gag protein arc1 binds RNA and traffics across synaptic boutons. *Cell* **2018**, *172*, 262–274.e11. [[CrossRef](#)]
9. Jang, H.S.; Shah, N.M.; Du, A.Y.; Dailey, Z.Z.; Pehrsson, E.C.; Godoy, P.M.; Zhang, D.; Li, D.; Xing, X.; Kim, S.; et al. Transposable elements drive widespread expression of oncogenes in human cancers. *Nat. Genet.* **2019**, *51*, 611–617. [[CrossRef](#)]
10. Chen, T.; Meng, Z.; Gan, Y.; Wang, X.; Xu, F.; Gu, Y.; Xu, X.; Tang, J.; Zhou, H.; Zhang, X.; et al. The viral oncogene Np9 acts as a critical molecular switch for co-activating  $\beta$ -catenin, ERK, Akt and Notch1 and promoting the growth of human leukemia stem/progenitor cells. *Leukemia* **2013**, *27*, 1469–1478. [[CrossRef](#)]
11. Chiappinelli, K.B.; Strissel, P.L.; Desrichard, A.; Li, H.; Henke, C.; Akman, B.; Hein, A.; Rote, N.S.; Cope, L.M.; Snyder, A.; et al. Inhibiting DNA methylation causes an interferon response in cancer via dsrna including endogenous retroviruses. *Cell* **2015**, *162*, 974–986. [[CrossRef](#)] [[PubMed](#)]
12. Roulois, D.; Loo Yau, H.; Singhanian, R.; Wang, Y.; Danesh, A.; Shen, S.Y.; Han, H.; Liang, G.; Jones, P.A.; Pugh, T.J.; et al. DNA-demethylating agents target colorectal cancer cells by inducing viral mimicry by endogenous transcripts. *Cell* **2015**, *162*, 961–973. [[CrossRef](#)] [[PubMed](#)]
13. Bestor, T.H.; Tycko, B. Creation of genomic methylation patterns. *Nat. Genet.* **1996**, *12*, 363–367. [[CrossRef](#)] [[PubMed](#)]
14. Collins, P.L.; Kyle, K.E.; Egawa, T.; Shinkai, Y.; Oltz, E.M. The histone methyltransferase SETDB1 represses endogenous and exogenous retroviruses in B lymphocytes. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 8367–8372. [[CrossRef](#)] [[PubMed](#)]
15. Adoue, V.; Binet, B.; Malbec, A.; Fourquet, J.; Romagnoli, P.; van Meerwijk, J.P.M.; Amigorena, S.; Joffre, O.P. The histone methyltransferase setdb1 controls the helper Cell lineage integrity by repressing endogenous retroviruses. *Immunity* **2019**, *50*, 629–644.e8. [[CrossRef](#)] [[PubMed](#)]
16. Martin, F.J.; Xu, Y.; Lohmann, F.; Ciccone, D.N.; Nicholson, T.B.; Loureiro, J.J.; Chen, T.; Huang, Q. KMT1E-mediated chromatin modifications at the Fc $\gamma$ RIIb promoter regulate thymocyte development. *Genes Immun.* **2015**, *16*, 162–169. [[CrossRef](#)]
17. Kato, M.; Takemoto, K.; Shinkai, Y. A somatic role for the histone methyltransferase Setdb1 in endogenous retrovirus silencing. *Nat. Commun.* **2018**, *9*, 1683. [[CrossRef](#)]
18. Ishak, C.A.; Marshall, A.E.; Passos, D.T.; White, C.R.; Kim, S.J.; Cecchini, M.J.; Ferwati, S.; MacDonald, W.A.; Howlett, C.J.; Welch, I.D.; et al. An RB-EZH2 complex mediates silencing of repetitive DNA sequences. *Mol. Cell* **2016**, *64*, 1074–1087. [[CrossRef](#)]
19. Stoye, J.P.; Moroni, C. Endogenous retrovirus expression in stimulated murine lymphocytes. Identification of a new locus controlling mitogen induction of a defective virus. *J. Exp. Med.* **1983**, *157*, 1660–1674. [[CrossRef](#)]
20. Young, G.R.; Mavrommatis, B.; Kassiotis, G. Microarray analysis reveals global modulation of endogenous retroelement transcription by microbes. *Retrovirology* **2014**, *11*, 59. [[CrossRef](#)]
21. Yu, P.; Lübben, W.; Slomka, H.; Gebler, J.; Konert, M.; Cai, C.; Neubrandt, L.; Prazeres da Costa, O.; Paul, S.; Dehnert, S.; et al. Nucleic acid-sensing Toll-like receptors are essential for the control of endogenous retrovirus viremia and ERV-induced tumors. *Immunity* **2012**, *37*, 867–879. [[CrossRef](#)] [[PubMed](#)]
22. Contreras-Galindo, R.; Kaplan, M.H.; Markovitz, D.M.; Lorenzo, E.; Yamamura, Y. Detection of HERV-K(HML-2) viral RNA in plasma of HIV type 1-infected individuals. *AIDS Res. Hum. Retroviruses* **2006**, *22*, 979–984. [[CrossRef](#)] [[PubMed](#)]
23. Contreras-Galindo, R.; López, P.; Vélez, R.; Yamamura, Y. HIV-1 infection increases the expression of human endogenous retroviruses type K (HERV-K) in vitro. *AIDS Res. Hum. Retroviruses* **2007**, *23*, 116–122. [[CrossRef](#)]

24. Hung, T.; Pratt, G.A.; Sundararaman, B.; Townsend, M.J.; Chaivorapol, C.; Bhangale, T.; Graham, R.R.; Ortmann, W.; Criswell, L.A.; Yeo, G.W.; et al. The Ro60 autoantigen binds endogenous retroelements and regulates inflammatory gene expression. *Science* **2015**, *350*, 455–459. [[CrossRef](#)] [[PubMed](#)]
25. Stauffer, Y.; Marguerat, S.; Meylan, F.; Ucla, C.; Sutkowski, N.; Huber, B.; Pelet, T.; Conrad, B. Interferon- $\alpha$ -induced endogenous superantigen: A model linking environment and autoimmunity. *Immunity* **2001**, *15*, 591–601. [[CrossRef](#)]
26. Cañadas, I.; Thummalapalli, R.; Kim, J.W.; Kitajima, S.; Jenkins, R.W.; Christensen, C.L.; Campisi, M.; Kuang, Y.; Zhang, Y.; Gjini, E.; et al. Tumor innate immunity primed by specific interferon-stimulated endogenous retroviruses. *Nat. Med.* **2018**, *24*, 1143–1150. [[CrossRef](#)] [[PubMed](#)]
27. Everett, R.D.; Boutell, C.; Hale, B.G. Interplay between viruses and host sumoylation pathways. *Nat. Rev. Microbiol.* **2013**, *11*, 400–411. [[CrossRef](#)] [[PubMed](#)]
28. Hu, M.-M.; Liao, C.-Y.; Yang, Q.; Xie, X.-Q.; Shu, H.-B. Innate immunity to RNA virus is regulated by temporal and reversible sumoylation of RIG-I and MDA5. *J. Exp. Med.* **2017**, *214*, 973–989. [[CrossRef](#)]
29. Lee, A.; Cingöz, O.; Sabo, Y.; Goff, S.P. Characterization of interaction between Trim28 and YY1 in silencing proviral DNA of Moloney murine leukemia virus. *Virology* **2018**, *516*, 165–175. [[CrossRef](#)]
30. Schmidt, N.; Domingues, P.; Golebiowski, F.; Patzina, C.; Tatham, M.H.; Hay, R.T.; Hale, B.G. An influenza virus-triggered SUMO switch orchestrates co-opted endogenous retroviruses to stimulate host antiviral immunity. *PNAS* **2019**, *116*, 17399–17408. [[CrossRef](#)]
31. Young, G.R.; Eksmond, U.; Salcedo, R.; Alexopoulou, L.; Stoye, J.P.; Kassiotis, G. Resurrection of endogenous retroviruses in antibody-deficient mice. *Nature* **2012**, *491*, 774–778. [[CrossRef](#)] [[PubMed](#)]
32. Browne, E.P. Toll-like receptor 7 controls the anti-retroviral germinal center response. *PLoS Pathog.* **2011**, *7*, e1002293. [[CrossRef](#)] [[PubMed](#)]
33. Kane, M.; Case, L.K.; Wang, C.; Yurkovetskiy, L.; Dikiy, S.; Golovkina, T.V. Innate immune sensing of retroviral infection via Toll-like receptor 7 occurs upon viral entry. *Immunity* **2011**, *35*, 135–145. [[CrossRef](#)]
34. Hurst, T.P.; Aswad, A.; Karamitros, T.; Katzourakis, A.; Smith, A.L.; Magiorkinis, G. Interferon-inducible protein 16 (IFI16) has a broad-spectrum binding ability against ssDNA targets: An evolutionary hypothesis for antiretroviral checkpoint. *Front. Microbiol.* **2019**, *10*. [[CrossRef](#)] [[PubMed](#)]
35. Feschotte, C. Transposable elements and the evolution of regulatory networks. *Nat. Rev. Genet.* **2008**, *9*, 397–405. [[CrossRef](#)] [[PubMed](#)]
36. Chuong, E.B.; Elde, N.C.; Feschotte, C. Regulatory evolution of innate immunity through co-option of endogenous retroviruses. *Science* **2016**, *351*, 1083–1087. [[CrossRef](#)]
37. Lynch, V.J. A copy-and-paste gene regulatory network. *Science* **2016**, *351*, 1029–1030. [[CrossRef](#)]
38. Barbalat, R.; Ewald, S.E.; Mouchess, M.L.; Barton, G.M. Nucleic acid recognition by the innate immune system. *Annu. Rev. Immunol.* **2011**, *29*, 185–214. [[CrossRef](#)]
39. Iwasaki, A.; Medzhitov, R. Control of adaptive immunity by the innate immune system. *Nat. Immunol.* **2015**, *16*, 343–353. [[CrossRef](#)]
40. Heil, F.; Hemmi, H.; Hochrein, H.; Ampenberger, F.; Kirschning, C.; Akira, S.; Lipford, G.; Wagner, H.; Bauer, S. Species-specific recognition of single-stranded RNA via toll-like receptor 7 and 8. *Science* **2004**, *303*, 1526–1529. [[CrossRef](#)]
41. Gao, D.; Wu, J.; Wu, Y.-T.; Du, F.; Aroh, C.; Yan, N.; Sun, L.; Chen, Z.J. Cyclic GMP-AMP synthase is an innate immune sensor of HIV and other retroviruses. *Science* **2013**, *341*, 903–906. [[CrossRef](#)] [[PubMed](#)]
42. Takaoka, A.; Wang, Z.; Choi, M.K.; Yanai, H.; Negishi, H.; Ban, T.; Lu, Y.; Miyagishi, M.; Kodama, T.; Honda, K.; et al. DAI (DLM-1/ZBP1) is a cytosolic DNA sensor and an activator of innate immune response. *Nature* **2007**, *448*, 501–505. [[CrossRef](#)] [[PubMed](#)]
43. Rigby, R.E.; Webb, L.M.; Mackenzie, K.J.; Li, Y.; Leitch, A.; Reijns, M.A.M.; Lundie, R.J.; Revuelta, A.; Davidson, D.J.; Diebold, S.; et al. RNA:DNA hybrids are a novel molecular pattern sensed by TLR9. *EMBO J.* **2014**, *33*, 542–558. [[CrossRef](#)] [[PubMed](#)]
44. Stetson, D.B.; Ko, J.S.; Heidmann, T.; Medzhitov, R. Trex1 prevents cell-intrinsic initiation of autoimmunity. *Cell* **2008**, *134*, 587–598. [[CrossRef](#)]
45. Wheeler, L.A.; Trifonova, R.T.; Vrbanc, V.; Barteneva, N.S.; Liu, X.; Bollman, B.; Onofrey, L.; Mulik, S.; Ranjbar, S.; Luster, A.D.; et al. TREX1 knockdown induces an interferon response to HIV that delays viral infection in humanized mice. *Cell Rep.* **2016**, *15*, 1715–1727. [[CrossRef](#)]

46. Nazli, A.; Kafka, J.K.; Ferreira, V.H.; Anipindi, V.; Mueller, K.; Osborne, B.J.; Dizzell, S.; Chauvin, S.; Mian, M.F.; Ouellet, M.; et al. HIV-1 gp120 induces TLR2-and TLR4-mediated innate immune activation in human female genital epithelium. *J. Immunol.* **2013**, *191*, 4246–4258. [[CrossRef](#)]
47. Rolland, A.; Jouvin-Marche, E.; Viret, C.; Faure, M.; Perron, H.; Marche, P.N. The envelope protein of a human endogenous retrovirus-W family activates innate immunity through CD14/TLR4 and promotes th1-like responses. *J. Immunol.* **2006**, *176*, 7636–7644. [[CrossRef](#)]
48. Antony, J.M.; van Marle, G.; Opii, W.; Butterfield, D.A.; Mallet, F.; Yong, V.W.; Wallace, J.L.; Deacon, R.M.; Warren, K.; Power, C. Human endogenous retrovirus glycoprotein-mediated induction of redox reactants causes oligodendrocyte death and demyelination. *Nat. Neurosci.* **2004**, *7*, 1088–1095. [[CrossRef](#)]
49. Rosenberg, N.; Jolicoeur, P. Retroviral pathogenesis. In *Retroviruses*; Coffin, J.M., Hughes, S.H., Varmus, H.E., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, USA, 1997; ISBN 978-0-87969-571-2.
50. Mangeney, M.; Renard, M.; Schlecht-Louf, G.; Bouallaga, I.; Heidmann, O.; Letzelter, C.; Richaud, A.; Ducos, B.; Heidmann, T. Placental syncytins: Genetic disjunction between the fusogenic and immunosuppressive activity of retroviral envelope proteins. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 20534–20539. [[CrossRef](#)]
51. Denner, J. The transmembrane proteins contribute to immunodeficiencies induced by HIV-1 and other retroviruses. *AIDS* **2014**, *28*, 1081–1090. [[CrossRef](#)]
52. Denner, J. Expression and function of endogenous retroviruses in the placenta. *APMIS* **2016**, *124*, 31–43. [[CrossRef](#)] [[PubMed](#)]
53. Morozov, V.A.; Thi, V.L.D.; Denner, J. The transmembrane protein of the human endogenous retrovirus—K (HERV-K) modulates cytokine release and gene expression. *PLoS ONE* **2013**, *8*, e70399. [[CrossRef](#)] [[PubMed](#)]
54. Ferrari, L.; Cafora, M.; Rota, F.; Hoxha, M.; Iodice, S.; Tarantini, L.; Dolci, M.; Delbue, S.; Pistocchi, A.; Bollati, V. Extracellular vesicles released by colorectal cancer cell lines modulate innate immune response in zebrafish model: The possible role of human endogenous retroviruses. *Int. J. Mol. Sci.* **2019**, *20*, 3669. [[CrossRef](#)]
55. Hummel, J.; Kämmerer, U.; Müller, N.; Avota, E.; Schneider-Schaulies, S. Human endogenous retrovirus envelope proteins target dendritic cells to suppress T-cell activation. *Eur. J. Immunol.* **2015**, *45*, 1748–1759. [[CrossRef](#)]
56. Jaenisch, R.; Schnieke, A.; Harbers, K. Treatment of mice with 5-azacytidine efficiently activates silent retroviral genomes in different tissues. *Proc. Natl. Acad. Sci. USA* **1985**, *82*, 1451–1455. [[CrossRef](#)] [[PubMed](#)]
57. Li, H.; Chiappinelli, K.B.; Guzzetta, A.A.; Easwaran, H.; Yen, R.-W.C.; Vatapalli, R.; Topper, M.J.; Luo, J.; Connolly, R.M.; Azad, N.S.; et al. Immune regulation by low doses of the DNA methyltransferase inhibitor 5-azacytidine in common human epithelial cancers. *Oncotarget* **2014**, *5*, 587–598. [[CrossRef](#)] [[PubMed](#)]
58. Chiappinelli, K.B.; Zahnow, C.A.; Ahuja, N.; Baylin, S.B. Combining epigenetic and immunotherapy to combat cancer. *Cancer Res.* **2016**, *76*, 1683–1689. [[CrossRef](#)] [[PubMed](#)]
59. Besch, R.; Poeck, H.; Hohenauer, T.; Senft, D.; Häcker, G.; Berking, C.; Hornung, V.; Endres, S.; Ruzicka, T.; Rothenfusser, S.; et al. Proapoptotic signaling induced by RIG-I and MDA-5 results in type I interferon-independent apoptosis in human melanoma cells. *J. Clin. Investig.* **2009**, *119*, 2399–2411. [[CrossRef](#)]
60. Tobiaasson, M.; Abdulkadir, H.; Lennartsson, A.; Katayama, S.; Marabita, F.; De Paepe, A.; Karimi, M.; Krjutskov, K.; Einarsdottir, E.; Grövdal, M.; et al. Comprehensive mapping of the effects of azacitidine on DNA methylation, repressive/permmissive histone marks and gene expression in primary cells from patients with MDS and MDS-related disease. *Oncotarget* **2017**, *8*, 28812–28825. [[CrossRef](#)]
61. Krug, B.; De Jay, N.; Harutyunyan, A.S.; Deshmukh, S.; Marchione, D.M.; Guilhamon, P.; Bertrand, K.C.; Mikael, L.G.; McConechy, M.K.; Chen, C.C.L.; et al. Pervasive H3K27 acetylation leads to ERV expression and a therapeutic vulnerability in H3K27M gliomas. *Cancer Cell* **2019**, *36*, 338–339. [[CrossRef](#)]
62. Sheng, W.; LaFleur, M.W.; Nguyen, T.H.; Chen, S.; Chakravarthy, A.; Conway, J.R.; Li, Y.; Chen, H.; Yang, H.; Hsu, P.-H.; et al. Lsd1 ablation stimulates anti-tumor immunity and enables checkpoint blockade. *Cell* **2018**, *174*, 549–563. [[CrossRef](#)] [[PubMed](#)]
63. Goel, S.; DeCristo, M.J.; Watt, A.C.; BrinJones, H.; Sceneay, J.; Li, B.B.; Khan, N.; Ubellacker, J.M.; Xie, S.; Metzger-Filho, O.; et al. CDK4/6 inhibition triggers anti-tumour immunity. *Nature* **2017**, *548*, 471–475. [[CrossRef](#)] [[PubMed](#)]

64. Bonaventura, P.; Shekarian, T.; Alcazer, V.; Valladeau-Guilemond, J.; Valsesia-Wittmann, S.; Amigorena, S.; Caux, C.; Depil, S. Cold tumors: A therapeutic challenge for immunotherapy. *Front. Immunol.* **2019**, *10*, 168. [[CrossRef](#)] [[PubMed](#)]
65. Smith, C.C.; Beckermann, K.E.; Bortone, D.S.; De Cubas, A.A.; Bixby, L.M.; Lee, S.J.; Panda, A.; Ganesan, S.; Bhanot, G.; Wallen, E.M.; et al. Endogenous retroviral signatures predict immunotherapy response in clear cell renal cell carcinoma. *J. Clin. Investig.* **2018**, *128*, 4804–4820. [[CrossRef](#)] [[PubMed](#)]
66. Heidegger, S.; Wintges, A.; Stritzke, F.; Bek, S.; Steiger, K.; Koenig, P.-A.; Göttert, S.; Engleitner, T.; Öllinger, R.; Nedelko, T.; et al. RIG-I activation is critical for responsiveness to checkpoint blockade. *Sci. Immunol.* **2019**, *4*, eaau8943. [[CrossRef](#)]
67. Leruste, A.; Tosello, J.; Ramos, R.N.; Tauziède-Espariat, A.; Brohard, S.; Han, Z.-Y.; Beccaria, K.; Andrianteranagna, M.; Caudana, P.; Nikolic, J.; et al. Clonally expanded T cells reveal immunogenicity of rhabdoid tumors. *Cancer Cell* **2019**, *36*, 597–612. [[CrossRef](#)]
68. Wang-Johanning, F.; Frost, A.R.; Johanning, G.L.; Khazaeli, M.B.; LoBuglio, A.F.; Shaw, D.R.; Strong, T.V. Expression of human endogenous retrovirus k envelope transcripts in human breast cancer. *Clin. Cancer Res.* **2001**, *7*, 1553–1560.
69. Rycaj, K.; Plummer, J.B.; Yin, B.; Li, M.; Garza, J.; Radvanyi, L.; Ramondetta, L.M.; Lin, K.; Johanning, G.L.; Tang, D.G.; et al. Cytotoxicity of human endogenous retrovirus K-specific T cells toward autologous ovarian cancer cells. *Clin. Cancer Res.* **2015**, *21*, 471–483. [[CrossRef](#)]
70. Krishnamurthy, J.; Rabinovich, B.A.; Mi, T.; Switzer, K.C.; Olivares, S.; Maiti, S.N.; Plummer, J.B.; Singh, H.; Kumaresan, P.R.; Huls, H.M.; et al. Genetic engineering of T cells to target herv-K, an ancient retrovirus on melanoma. *Clin. Cancer Res.* **2015**, *21*, 3241–3251. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).





## ARTICLES – RESULTATS

---

## **ARTICLE 1 - Identification of shared tumor epitopes from endogenous retroviruses inducing high avidity cytotoxic T cells for cancer immunotherapy**

*Paola Bonaventura*<sup>1,2\*</sup>, ***Vincent Alcazer***<sup>1\*</sup>, *Virginie Mutez*<sup>3</sup>, *Laurie Tonon*<sup>4</sup>, *Juliette Martin*<sup>5</sup>, *Nicolas Chuvin*<sup>3</sup>, *Emilie Michel*<sup>6</sup>, *Rasha Boulos*<sup>3</sup>, *Yann Estornes*<sup>3</sup>, *Jenny Valladeau-Guilemond*<sup>1</sup>, *Alain Viari*<sup>4</sup>, *Christophe Caux*<sup>1,2</sup>, *Stéphane Depil*<sup>1,2,3,5</sup>†

1 Centre de Recherche en Cancérologie de Lyon (CRCL), UMR INSERM U1052 CNRS 5286, Lyon, France.

2 Centre Léon Bérard, Lyon, France.

3 ErVaccine Technologies, Lyon, France.

4 Synergie Lyon Cancer, Plateforme de bioinformatique « Gilles Thomas », Lyon, France.

5 CNRS- Institut de Biologie et Chimie des Proteines UMR 5086, Lyon, France

6 Université Claude Bernard Lyon 1, Lyon, France.

\* These authors contributed equally to this work

†**To whom correspondence should be addressed:** Stéphane Depil, MD, PhD, Centre Léon Bérard, 28 rue Laennec, 69 008 Lyon; e-mail: [stephane.depil@lyon.unicancer.fr](mailto:stephane.depil@lyon.unicancer.fr) ; Tel : + 33 4 72 68 19 30

==== Article actuellement en revue dans Science Advances =====

## Résumé

Dans ce premier travail, nous proposons une méthode reproductible et généralisable à grande échelle pour identifier et prioriser les épitopes T CD8<sup>+</sup> issus de HERVs exprimés dans un cancer donné. A partir de la base TCGA, nous avons tout d'abord réalisé une analyse préliminaire de l'expression des HERVs dans le cancer du sein selon une approche classique type « Rooney ». Ces premières analyses nous ont révélé une importante corrélation entre plusieurs HERVs de la famille HERV-K fortement exprimés dans le sous-type basal et des signatures lymphocytaires T CD8<sup>+</sup> ainsi que d'activation de *checkpoints* immunitaires. Basé sur ces résultats, nous avons aligné les différents HERVs corrélés aux réponses LT CD8<sup>+</sup> afin de définir des zones partagées riches en épitopes présentant une forte affinité prédite pour le HLA-A\*02. Parmi les différents candidats, nous avons choisi les 6 peptides (P1 à P6) les plus partagés pour validation biologique. Nous montrons qu'il existe des évidences de traduction pour 5 de ces peptides dans des données de spectrométrie de masse de cancer du sein.

Afin de généraliser cette approche à plus large échelle, nous avons ensuite analysé la base de données pancancer du TCGA (8 893 échantillons de 29 types de cancer solides différents) via l'approche *hervQuant*, permettant une quantification plus exhaustive élargie à 3 713 loci HERVs enrichis en séquences contenant des ORFs ainsi qu'une meilleure prise en charge des *multimaps*. Afin de limiter le nombre de peptides candidats, nous avons établi une méthode d'analyse contenant plusieurs étapes originales permettant de limiter le nombre de faux positifs. Parmi les différents ajouts, l'annotation de l'immunogénicité potentielle des HERVs via une approche d'apprentissage automatique, le filtre des candidats exprimés par les LT ou les lymphocytes NK ainsi que l'alignement des ORFs prédits sur des protéines connues de Gag et Pol nous permettent de limiter drastiquement le nombre final de peptides à tester.

D'un point de vue biologique, nous montrons que ces peptides peuvent induire des réponses LT CD8<sup>+</sup> spécifiques et polyfonctionnelles chez le donneur sain après stimulation de PBMCs par un protocole classique de *priming* de LT par cellules dendritiques autologues à partir de monocytes, ainsi que par un protocole de *priming* accéléré via des agonistes TLR. Ces réponses sont particulièrement importantes pour les peptides P1, P2, P4 et P6. Nous avons ensuite généré des clones dirigés contre ces 4 peptides via le tri de LT CD8<sup>+</sup> spécifique obtenus chez le donneur sain et expandus pendant 14 jours, et caractérisé leur TCR par séquençage des chaînes  $\alpha$  et  $\beta$ . Nous montrons que ces clones présentent une forte avidité pour leur cible, ainsi qu'une cytotoxicité face à la lignée MDA-MB-231 de cancer du sein triple négatif exprimant nos cibles. Enfin, nous détectons la présence de LT CD8<sup>+</sup> spécifique dirigés contre P1, P4 et P6 parmi les lymphocytes infiltrant la tumeur de patientes présentant un cancer du sein triple négatif au diagnostic, et montrons que ces clones sont cytotoxiques dans un modèle d'organoïde humain.

Ainsi, nous proposons une méthode applicable à grande échelle permettant d'identifier des épitopes LT CD8<sup>+</sup> à partir des HERVs exprimés dans un cancer donné en réduisant grandement le nombre de candidats et de faux positifs. Nous démontrons également le caractère immunogène des peptides sélectionnés ainsi que la cytotoxicité de leurs clones LT CD8<sup>+</sup> respectifs, ouvrant la voie à l'utilisation des peptides issus de HERVs pour des approches d'immunothérapie ciblée.

## Identification of shared tumor epitopes from endogenous retroviruses inducing high avidity cytotoxic T cells for cancer immunotherapy

Paola Bonaventura<sup>1,2,†</sup>, Vincent Alcazer<sup>1,†</sup>, Virginie Mutez<sup>3</sup>, Laurie Tonon<sup>4</sup>, Juliette Martin<sup>5</sup>, Nicolas Chuvin<sup>3</sup>, Emilie Michel<sup>3</sup>, Rasha Boulos<sup>3</sup>, Yann Estornes<sup>3</sup>, Jenny Valladeau-Guilemond<sup>1</sup>, Alain Viari<sup>4</sup>, Christophe Caux<sup>1,2</sup>, Stéphane Depil<sup>1,2,3,6,\*</sup>

### Affiliations:

1 Centre de Recherche en Cancérologie de Lyon (CRCL), UMR INSERM U1052 CNRS 5286, Lyon, France.

2 Centre Léon Bérard, Lyon, France.

3 ErVaccine Technologies, Lyon, France.

4 Synergie Lyon Cancer, Plateforme de bioinformatique « Gilles Thomas », Lyon, France.

5 CNRS- Institut de Biologie et Chimie des Protéines UMR 5086, Lyon, France

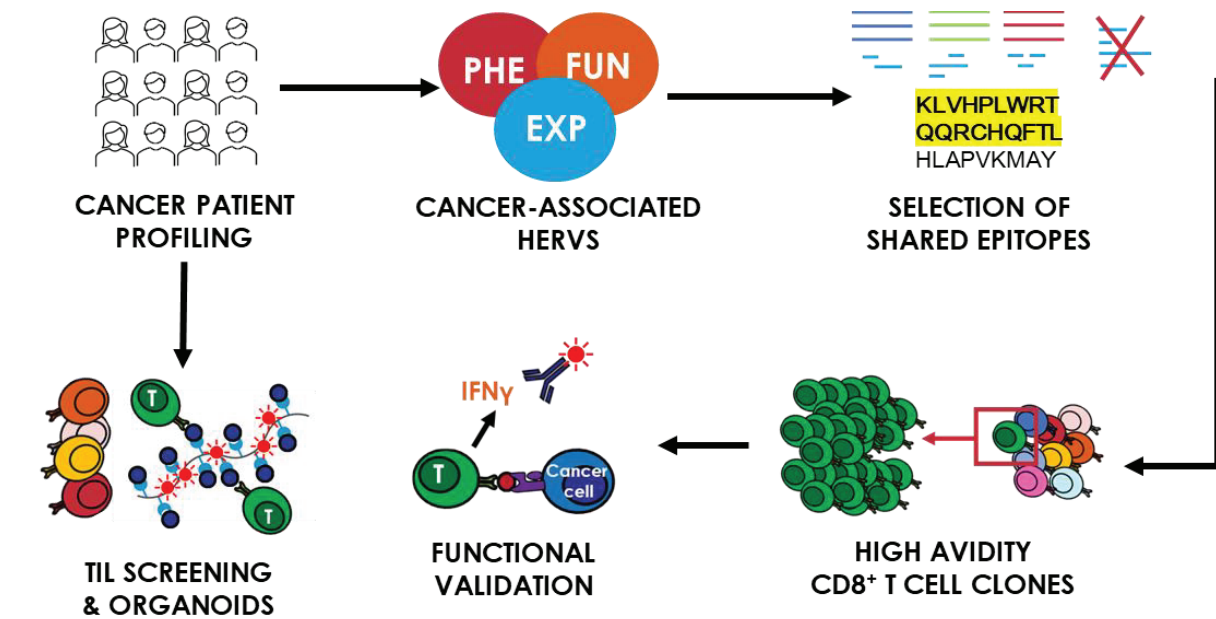
6 Université Claude Bernard Lyon 1, Lyon, France.

† These authors contributed equally to this work

\* **Correspondance:** Stéphane Depil, MD, PhD, Centre Léon Bérard, 28 rue Laennec, 69 008 Lyon; e-mail: [stephane.depil@lyon.unicancer.fr](mailto:stephane.depil@lyon.unicancer.fr) ; Tel : + 33 4 72 68 19 30

**One Sentence Summary:** A bioinformatic-based method was developed to identify shared tumor epitopes derived from HERVs and inducing high avidity CD8+ T cells cytotoxic against tumor cells.

## Graphical abstract



## Abstract

Human endogenous retroviruses (HERVs) represent 8% of the human genome. Given their viral origin, HERV products may represent tumor antigens relevant for cancer immunotherapy. We developed a systematic bioinformatics-based approach to identify shared CD8<sup>+</sup> T cell epitopes derived from cancer-associated HERVs in solid tumors. Six HLA-A2 epitope candidates among the most commonly shared epitopes with evidence of translation were selected for further immunological evaluation. *In vitro* priming assays confirmed the immunogenicity of these epitopes that induced high avidity CD8<sup>+</sup> T cell clones able to specifically recognize and kill HLA-A2<sup>+</sup> tumor cells. Furthermore, epitope-specific CD8<sup>+</sup> T cells were identified by dextramer-staining among tumor infiltrating lymphocytes from HLA-A2<sup>+</sup> breast cancer patients. Finally, we showed that HERV-specific T cells can lyse patient-derived organoids. These shared virus-like epitopes could be of major interest for the development of cancer vaccine approaches or T cell-based immune therapies, especially in tumors with low or intermediate mutational burden.

## **Main Text:**

### **Introduction**

The adaptive T cell immune response in cancer relies on the recognition of tumor epitopes specifically expressed by tumor cells (1). The role of neoantigens, generated by non-synonymous mutations specific to the tumor genome, has been extensively studied in the last decade (2) and many clinical trials testing combinations of neoantigens in personalized cancer vaccines have been initiated, with encouraging preliminary results (3). However, determining the optimal combination of neoepitopes for each patient remains challenging. Furthermore, many tumors are characterized by a low or moderate tumor mutational burden. Therefore, unveiling other families of tumor antigens, such as those derived from splice variants, fusion proteins or endogenous retroelements, possibly shared among different cancer subtypes, is of utmost importance for the development of off-the-shelf therapies in solid tumors (4).

Human endogenous retroviruses (HERVs) represent 8% of the human genome (5). Although most HERV genes are non-functional due to DNA recombination, mutations, and deletions, some produce functional proteins including the group-specific antigen (Gag), polymerase (Pol) with reverse transcriptase, and the envelope (Env) surface unit (6). Most HERVs are silenced by epigenetic mechanisms in normal cells (7). HERVs are unmethylated and aberrantly expressed in some tumors, including breast cancer, ovarian cancer (8), prostate cancer (9) and melanoma (10). HERVs were reported to be possible pathogenic agents in carcinogenesis, through their involvement in insertional mutagenesis, chromosomal aberrations or LTR-induced oncogene-activation (6, 11). Furthermore, some HERV proteins, like HERV-K Rec and Np9, are also putative oncogenes (6).

HERVs may represent an interesting source of shared tumor antigens (5). Due to their homology with “non-self” viral antigens and their limited expression in normal tissues, some HERVs may induce efficient cytotoxic T cell (CTL) responses in the absence of negative selection of reactive T cells in the thymus (12). A study reporting a CD8+ T cell response against an ERVK3-derived epitope in a melanoma patient was published by Schiavetti et al. in 2002 (13). Since this first report, other studies have shown that HERVs can induce T cell responses in cancer patients, notably in renal cell carcinoma (14, 15), colorectal cancer (16), seminoma (17) and breast cancer (18). However, very few studies have thus far identified HERV-derived T cell epitopes and characterized the corresponding T cell clones. The large number of HERVs integrated in the genome and the associated polymorphism of these repeated sequences render target discovery and validation very difficult, hampering the development of immune therapy



targeting HERV-derived epitopes. We present herein an original bioinformatics-based method to identify shared T cell epitopes derived from HERV viral antigens specifically overexpressed in different cancer subtypes. We show that these HERV-derived epitopes efficiently prime antitumor CTL clones of high avidity and could thus represent relevant targets for cancer immunotherapy.

## Results

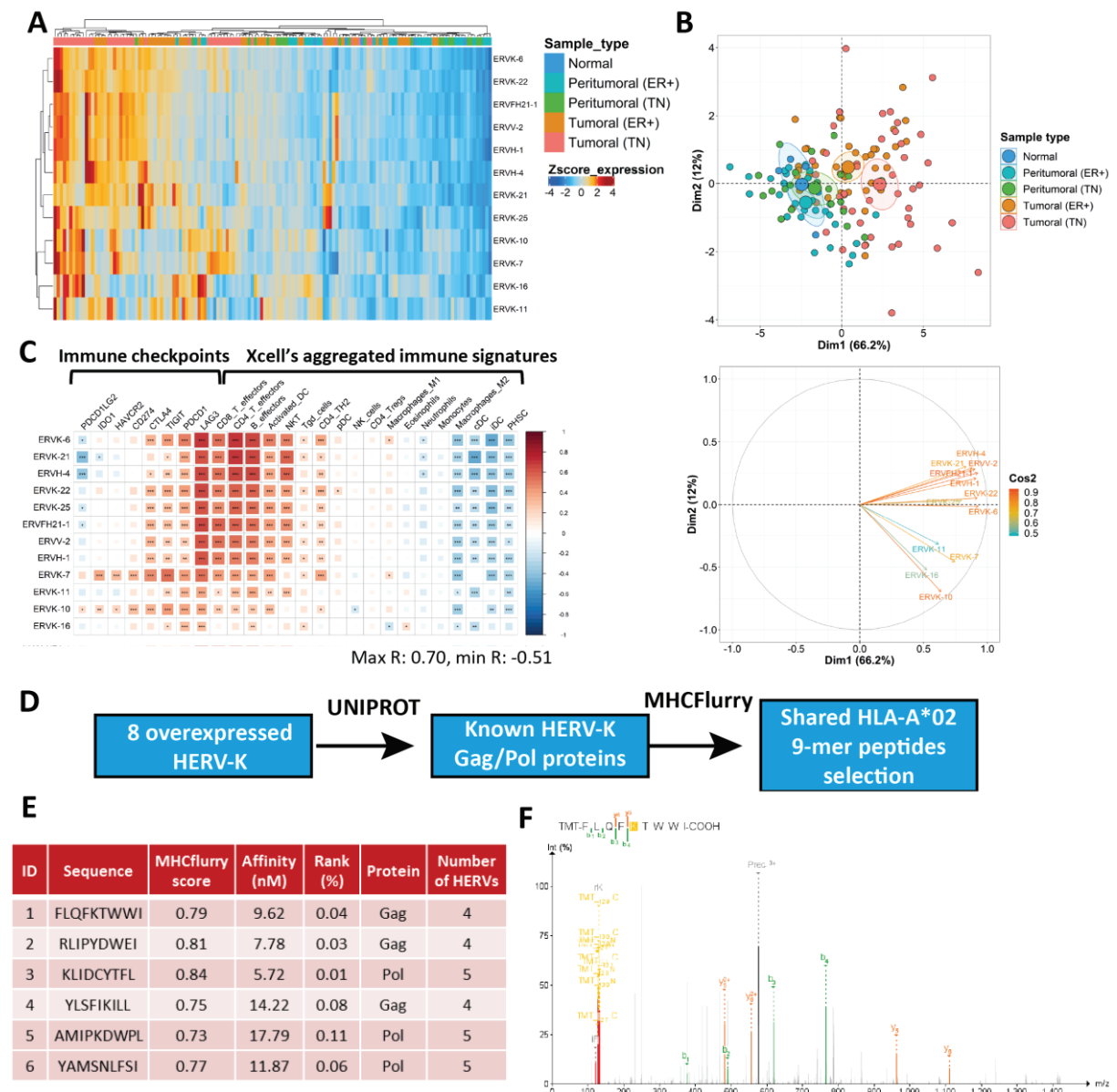
### *HERVs overexpressed in triple negative breast cancers (TNBCs) contain shared T cell epitopes*

To establish the proof-of-concept that HERVs overexpressed in a specific disease contain shared T cell epitopes, we first assessed HERV expression in breast cancer, owing to the well-characterized expression of HERVs in this disease (19) and the availability of a RNA sequencing (RNAseq) database comprising normal samples. For this initial approach, we used the genomic annotation of 66 HERVs established by Mayer et al (20), which contains only HERV sequences with a known corresponding protein entry in Uniprot (21). Differential HERV expression between tumor samples and pooled normal samples (peritumoral and normal breast tissues) revealed 9 and 12 HERVs significantly overexpressed in the estrogen receptor-positive (ER+) and TNBC subgroups, respectively (Fig. 1A, Supplementary Fig. S1A and S1B). Principal component analysis performed on the 12 HERVs overexpressed in the TNBC subgroup highlighted a good discrimination based on the first and second axes with 66% and 12% of variance, respectively, due mainly to HERVs from the HERV-K/HML-2 family (Fig. 1B). These HERVs were highly correlated with Xcell's T and B cell immune signatures (22) (max Spearman's R: 0.71, corrected P-value < 0.001) and negatively correlated with myeloid signatures such as M2 macrophages (Fig. 1C). HERV expression was also correlated with that of negative immune checkpoints such as PDCD1, CTLA4, LAG3 and TIGIT (Fig. 1C).

Based on the specific expression of a limited number of HERVs characterizing the TNBC subtype, we speculated that it may be possible to identify shared domains containing T cell epitopes within Gag, Pol and Env proteins. Indeed, common regions in sequences from the HERV-K/HML-2 family members overexpressed in TNBCs (HERV-K 6, 7, 10, 11, 21 and 25) were uncovered. Retrieval of the corresponding Gag, Pol and Env proteins from Uniprot followed by strong binder prediction via MHCflurry analysis led to the identification of several regions enriched in T cell epitope candidates for the most common HLA class I and II alleles (Fig. 1D). For further biological validation and immunological assays, we selected six 9-mer strong binders for HLA-A\*0201 from the most conserved sites: three from Gag (P1, P2 and P4)

and three from Pol (P3, P5 and P6) (Fig. 1E). Analysis of mass spectrometry data from The Cancer Genome Atlas (TCGA) and Clinical Proteomic Tumor Analysis Consortium (CPTAC) (23) showed evidence of translation for P1, P2, P3, P5 and P6 peptides in breast cancer samples (Fig. 1F and supplementary Fig. 1C). P4 was also selected as it had been described among HLA-I eluted peptides from tumors (patent WO2019/162110A1). Importantly, alignment against the human proteome revealed that the sequences of these epitope candidates did not match any self-protein sequence (Supplementary Table S1).

**Figure 1**



**Fig. 1. TNBC is characterized by the overexpression of HERVs mainly from the HERV-K family**

- (A) Heatmap of HERVs overexpressed in the TNBC subtype. Z-score of  $\log_2(\text{CPM} + 1)$  expression is represented.
- (B) Principal component analysis based on the 12 HERVs overexpressed in the TNBC subtype. Individuals (up) and variables (down) are represented.
- (C) Correlation between immune signatures and the expression of the 12 HERVs overexpressed in the TNBC subtype. Spearman's correlation coefficient is represented. (FDR-adjusted p-value: \*: <0.05, \*\*: < 0.01, \*\*\*: <0.001)
- (D) Workflow of HLA-A\*02 epitope selection from the 8 overexpressed HERV-K sequences with known Gag or Pol proteins.
- (E) Characteristics of the 6 selected HLA-A2 epitopes. 9-mer peptides were selected according to their predicted HLA-A\*02 affinity (considering strong binders for percentile ranks  $\leq 0.5$ ) and the number of HERVs containing their sequences.
- (F) MS/MS detection of P1 epitope in a sample from the CPTAC breast cancer prospective dataset. MS/MS spectrum is identified by Pepquery analysis (Peptide Spectrum Match, p value 0.0063)

*PDCD1LG2: Programmed cell death 1 ligand 2 (PD-L2), IDO1: Indoleamine 2,3-dioxygenase, HAVCR2: Hepatitis A virus cellular receptor 2 (TIM-3), CTLA4: Cytotoxic T-lymphocyte-associated protein 4, TIGIT: T cell immunoreceptor with Ig and ITIM domains, PDCD1: Programmed cell death 1, LAG3: Lymphocyte-activation gene 3, DC: Dendritic Cell, pDC: plasmacytoid DC, cDC: conventional DC, iDC: immature DC, NKT: Natural Killer T cell, Tgd\_cells: Gamma-Delta T cells, PHSC: Progenitor and Hematopoietic Stem Cells, ER+: Estrogen Receptor positive, TNBC: Triple Negative Breast Cancer.*

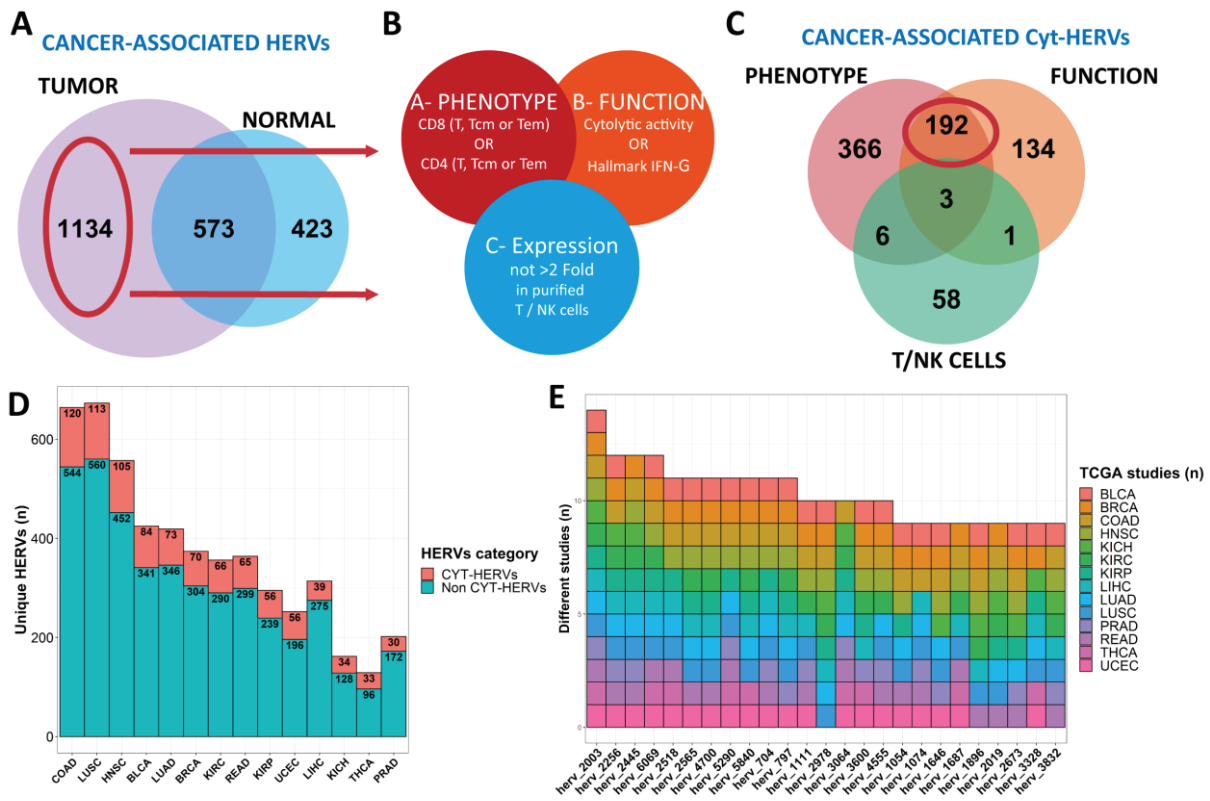
### ***A machine learning-based approach enhances the identification of HERVs associated with CTL response***

To improve the epitope detection process and generalize it to other cancers, we developed a new pipeline for annotating HERVs. For this, we reviewed multiple HERV databases and selected a more recent and complete reference of 3,173 HERVs, mostly composed of complete proviral sequences, thus having a higher probability of containing translated peptides (24). We assessed HERV expression in 8,893 primary tumor samples from 29 different cancer types from TCGA pancancer RNAseq database using HervQuant (15) (Supplementary Fig. S2A and Supplementary Table S2). We selected cancers with at least 10 available matched peritumoral samples ( $n = 14$ ; Supplementary Table S2) to filter HERVs highly expressed in tumors and not in normal tissue

(cancer-associated HERVs or CAH). Differential HERV expression analysis unveiled 1,134 CAH candidates (Fig. 2A). To reduce the number of candidates to test, we next applied a second filter to retain only HERVs associated with a CTL response among CAHs (cyt-HERVs). Cyt-HERV annotation was based on two inclusion criteria, namely the association of each HERV with at least one CD4 or CD8 T cell phenotype (A) and function (B) signature, and one exclusion criterion, namely its expression by purified T or NK cells (C) (A and B not C) (Fig. 2B). To reduce the risk of false-positive association and control for the high collinearity encountered with HERV expression, these associations were evaluated by L1 penalized regression (25) to retain only HERVs highly associated with CTL responses, controlling for cancer subtypes. A machine learning-based approach was used to test the associations independently for each cancer type (see Methods for full details) leading to the final identification of 192 cyt-HERVs (Fig. 2C). Sub-cancer analysis revealed that colon adenocarcinoma (COAD), lung squamous cell carcinoma (LUSC), head and neck squamous cell carcinoma (HNSC), bladder urothelial carcinoma (BLCA) and lung adenocarcinoma (LUAD) were the top 5 cancers with the highest total number of cyt-HERVs (Fig. 2D).

Overall, cyt-HERVs constituted around 15% of total CAHs, greatly reducing the number of potential candidates. Among the most shared cyt-HERVs, 11 were overexpressed in more than 10 different types of cancers, including 3 HERVs (*herv\_2256*, *herv\_6069* and *herv\_4700*) formerly reported to induce CD8<sup>+</sup> T cell responses (10, 14, 15) (Fig. 2E). Analysis of the mean beta-value of the 10 nearest surrounding probes from TCGA Illumina 450k methylation data revealed more cyt-HERVs significantly correlated with local demethylation ( $n = 37$ ) than methylation ( $n = 15$ ), suggesting a partial epigenetic control of these HERVs (Supplementary Fig. S2B).

**Figure 2**



**Fig. 2. Pancancer identification of HERVs associated with CTL responses**

(A) Venn diagram representing the total number of HERVs overexpressed in a tumor versus its normal counterpart (peritumoral tissue), and the total number of HERVs overexpressed in a normal peritumoral tissue versus its tumoral counterpart. HERVs overexpressed in at least 1 tumor and never overexpressed in any peritumoral tissue are considered cancer-associated.

(B) Venn diagram of the selection criteria for a HERV to be annotated as associated with CTL response (cyt-HERV). Each HERV had to be associated with both a phenotype (CD8 or CD4 T cell signatures) and a function (cytolytic activity (Granzyme B & Perforin 1) or IFN- $\gamma$  signature) criteria (A and B) and not overexpressed in normal purified T/ NK cells (A AND B NOT C).

(C) Venn diagram of cancer-associated HERVs' association with CTL responses criteria defined in A. A total of 192 HERVs are annotated as cyt-HERVs.

(D) Proportion of cancer-associated HERVs annotated as cyt-HERVs per cancer subtype. Cyt-HERVs are colored in red.

**(E)** Top 25 most shared cyt-HERVs with their respective tumors. The number of different cancers is represented on the y-axis.

*CTL: Cytotoxic T cell response, Cyt-HERVs: HERVs associated with CTL response in cancer, TNBC: Triple-Negative Breast Cancer*

***Selection of conserved Gag and Pol HERV-K/HML-2 motifs among cyt-HERVs leads to the identification of shared CD8<sup>+</sup> T cell epitopes***

We next assessed the presence of shared T cell epitopes among these cyt-HERVs. We translated our 192 cyt-HERV sequences into the 6 possible frames and retained predicted open-reading frames (ORFs) of at least 10 amino acids. To reduce the number of false-positives (non-translated sequences), we aligned these ORFs against known HERV-K/HML-2 Gag and Pol proteins referenced in UniProt and kept only ORFs with 90% homology with known existing HML-2 proteins (26). This conservative approach led to the identification of 57 HML-2 HLA-A\*0201 epitope candidates from 27 distinct ORFs (Supplementary Fig. S3A), with *herv\_2410* and *herv\_6069* showing the highest number of conserved HML-2-derived ORFs (Supplementary Fig. S3B). To better appreciate the distribution of these epitopes, we relocated each peptide among all the CAHs. The top 25 most shared epitopes included our 6 previously identified peptides (P1 to P6), with 13 unique epitopes present in at least 10 different HERVs (Figure 3A). YAMSNLFSI (P6), in particular, was shared by 27 different CAHs.

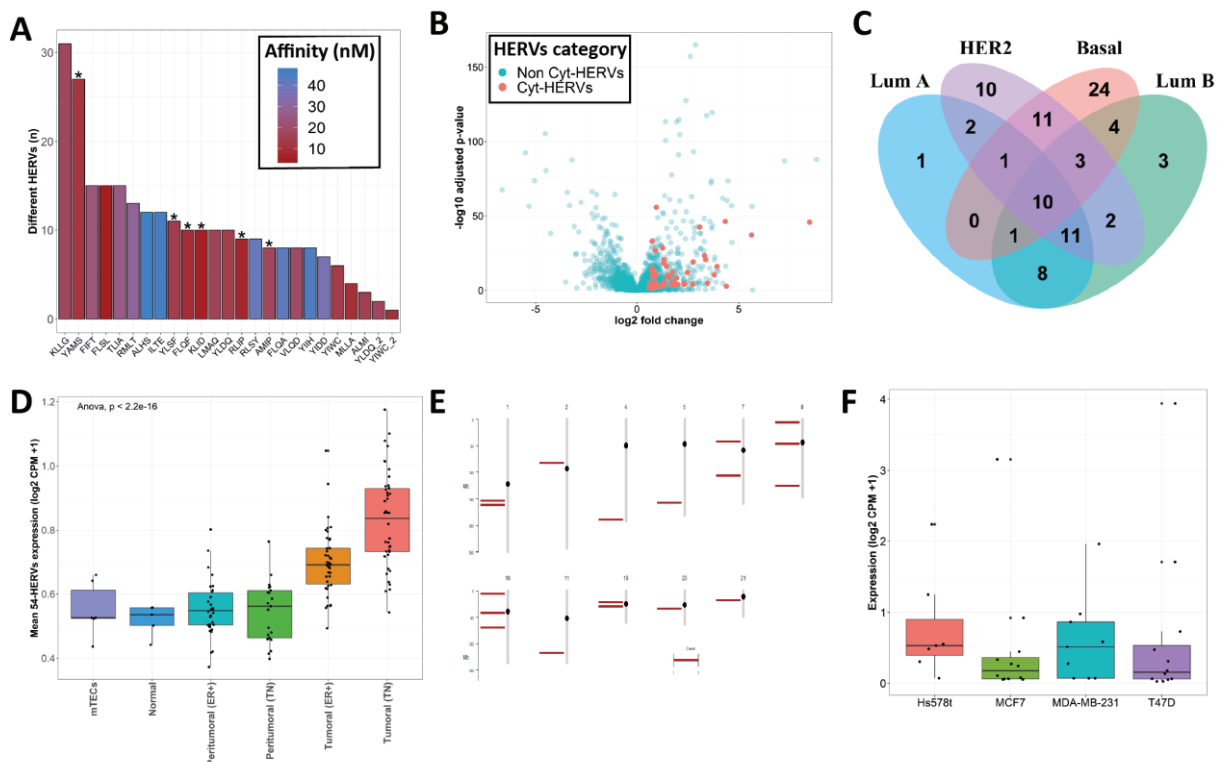
To extend our previous findings to TNBCs and reinforce our approach, we then focused on breast cancer. Differential HERV expression analysis uncovered a total of 497 CAHs expressed across different breast cancer subtypes, among which 91 were annotated as cyt-HERVs (Fig. 3B and 3C). Fifty-four of these 91 cyt-HERVs were expressed in the basal subtype (Fig. 3C). The mean expression of these 54 cyt-HERVs was significantly higher in TNBC and ER+ samples compared to peritumoral or normal breast tissues from an independent dataset (Fig. 3D). Interestingly, analysis of the mean expression level of these 54 HERVs in medullary thymic epithelial cells (mTECs) revealed a similar profile as in normal tissues (Fig. 3D), suggesting that the thymus may not induce a negative selection of the corresponding HERV antigen-specific T cells. We confirmed the high expression of these 54 cyt-HERVs in breast cancer cell lines sequenced in Varley et al.'s study (Supplementary Fig. S3C) and in cell lines from the Broad Institute Cancer Cell Line Encyclopedia (Supplementary Fig. S3D). The top 25 most shared epitope candidates among all CAHs expressed in the basal subtype contained the 6 previously identified peptides P1-P6 (Supplementary Fig. S3E).



We next selected HERVs containing the sequences of the 6 previously identified epitope candidates P1-P6 among the CAHs expressed in the basal subtype. Eighteen different HERVs contained at least one of the 6 peptides in their ORFs (Table 1). Genomic mapping of the corresponding loci showed a diffuse location for these 18 HERVs on chromosomes (Fig. 3E). Analysis of ribosome profiling (riboseq) data from a previously published study (27) revealed evidence of translation for these 18 HERVs in 4 different breast tumor cell lines including 2 basal and 2 luminal A subtypes (Fig. 3F).

Overall, our bioinformatics approach allowed us to select a limited number of HERV-derived T cell epitopes specifically overexpressed by tumor cells and most likely to be immunogenic among a large number of HERV candidates.

**Figure 3**



**Fig. 3. Shared CD8<sup>+</sup> T cell epitopes derived from conserved Gag and Pol HERV-K/HML-2 motifs are expressed in TNBCs**

**(A)** Bar chart of the top 25 most shared peptides predicted as strong HLA-A\*02 binders among the 192 cyt-HERVs. Previously identified peptides (P1 to P6) are shown in bold and marked with a star.

- (B) Volcano plot of HERVs overexpressed in basal breast cancer subtype compared to peritumoral samples in TCGA database. cyt-HERVs are represented in red.
- (C) Venn diagram of total number of cyt-HERVs overexpressed in each subtype of breast cancer in TCGA database.
- (D) Mean expression of the 54 cyt-HERVs overexpressed in TCGA basal subtype, in the independent database of Varley et al (28) and in medullary thymic epithelial cells.
- (E) Chromosomal location of the 18 peptide-containing HERVs overexpressed in the basal breast cancer subtype. Each bar corresponds to a HERV locus.
- (F) Expression of the 18 peptide-containing HERVs in the breast cancer basal (Hs578t & MDA-MB-231) and luminal A (MCF7 and T47D) cell lines analyzed by Riboseq.

*mTECs: medullary thymic epithelial cells, TNBC: Triple-Negative Breast Cancer*

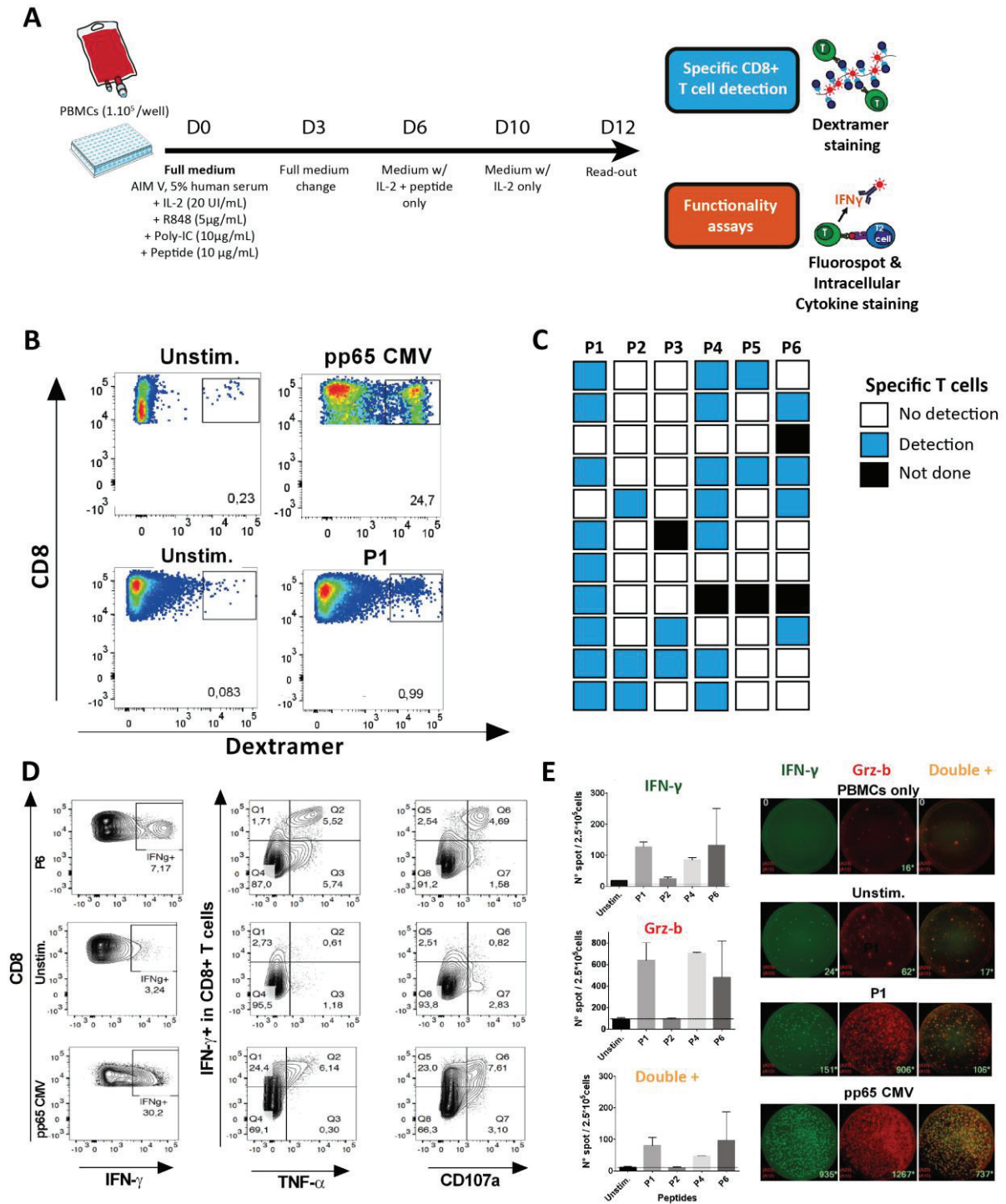
### ***HERV-derived epitopes induce strong and polyfunctional T cell responses***

We then evaluated the capacity of the selected epitope candidates to induce efficient T cell responses. The HLA-A2 affinity of the 6 selected peptides was first confirmed using an in vitro binding assay on purified HLA-A\*02:01 molecules (supplementary Fig. S4A). To assess the immunogenicity of these 6 peptides, we developed an optimized in vitro priming assay performed on peripheral blood mononuclear cells (PBMCs) from HLA-A2-positive donors (Fig. 4A and Methods for details). The dextramer-based quantification of peptide-specific CD8<sup>+</sup> T cells (gating strategy supplementary Fig. S4B) revealed the presence of specific T cells for all peptides, with variations among donors (Fig. 4B and 4C and Supplementary Fig. S4C). P1 appeared to be the most immunogenic peptide with significant T cell responses in 9/11 donors, followed by P4 (7/10), P6 (4/9) and P2 (3/11) (Fig. 4C). The immunogenicity of these peptides was further confirmed by a classical assay using monocyte-derived dendritic cells prepared from 5 HLA-A2-positive healthy donors (Supplementary Fig S4D). Flow cytometry analysis showed a CD8<sup>+</sup> T cell IFN- $\gamma$  production when peptide-stimulated PBMCs were co-cultured with T2 cells pulsed with the cognate epitopes (Supplementary Fig. S4B and S4E). Of note, P1 also induced the highest IFN- $\gamma$  response compared to the other peptides.

Based on these results, we selected P1, P2, P4 and P6 for further experiments. A polyfunctional IFN- $\gamma$ <sup>+</sup> TNF- $\alpha$ <sup>+</sup>-specific CD8<sup>+</sup> T cell response was observed upon co-culture with peptide-pulsed T2 cells, associated with the presence of the degranulation marker CD107a (Fig. 4D).

Fluorospot assay in the same co-culture conditions confirmed the secretion of IFN- $\gamma$  and granzyme B with the presence of double-positive cells (Fig. 4E).

**Figure 4**



**Fig. 4. HERV-derived epitopes induce polyfunctional CD8<sup>+</sup> T cell responses**

**(A)** Schematic representation of the *in vitro* priming protocol. PBMCs from HLA-A2-positive healthy donors are cultured in 96 well plates with AIM V Medium with 5% human serum

enriched with IL-2, TLR ligands (R848 and poly-IC) and the peptide of interest. Medium is changed every 3-4 days, without TLR ligands on day 6 and without peptide on day 10. Cells are collected and counted for analysis (Fluorospot, dextramer and intracellular staining) on day 12.

**(B)** Representative example of dextramer staining of peptide-stimulated (lower right panel) and unstimulated (lower left panel) CD8<sup>+</sup> T cells. CMV pp65 stimulated (upper right) and unstimulated (upper left) PBMCs are used as controls.

**(C)** Summary of the results obtained with PBMCs from 11 HLA-A2-positive healthy donors: blue squares indicate detection of epitope-specific CD8<sup>+</sup> T cells; white squares indicate a lack of detection; black squares are depicted when the test was not performed.

**(D)** Plots of IFN- $\gamma$  (left panels), IFN- $\gamma$  and TNF- $\alpha$  (center panels) or IFN- $\gamma$  and CD107a (right panels) staining gated on CD8<sup>+</sup> T cells. Healthy donor PBMCs were stimulated with peptide (here P6, upper line), no peptide (central line) or CMV pp65 peptide (bottom line). Plots are representative of the data summarized in figure S4E (n=11).

**(E)** Mean number and standard deviation of IFN- $\gamma$ +, Grz-b+ and double IFN- $\gamma$ + Grz-b+ spots counted on Fluorospot after stimulation of HLA-A2-positive healthy donor PBMCs with different peptides. Experiments were performed in duplicates for each represented condition. Representative wells are shown on the right panel: IFN- $\gamma$  (green, left), Grz-b (red, center) and double positive (yellow, right) spots for PBMCs alone and 10:1 PBMCs:T2 co-cultures using either unstimulated PBMCs, P1-stimulated PBMCs or CMV pp65 positive control PBMCs.

*CMV: Cytomegalovirus, IFN: Interferon, pp65: phosphoprotein 65, TLR: Toll Like Receptor, TNF: Tumor Necrosis Factor, Unstim: Unstimulated.*

### ***Epitope-specific CD8<sup>+</sup> T cell clones are characterized by T cell receptors (TCR) of high predicted affinity***

P1, P2, P4 and P6-specific CD8<sup>+</sup> T cells were sorted by flow cytometry using dextramer staining and expanded on feeder cells (see Methods). More than 90% (90-99%) of the CD8<sup>+</sup> T cells were dextramer-positive after one (P1) or two steps (P2, P4 and P6) of selection-expansion (Supplementary Fig. S5A). TCR $\beta$  immunosequencing confirmed the presence of dominant clones with a unique V $\beta$  rearrangement representing 90.8%, 90.7%, 99.6% and 76% of the expanded T

cells for P1, P2, P4 and P6, respectively (Fig. 5A and 5B). Of note, the V/D/J recombination sequences of TCR $\beta$  characterizing these clones were not present in the T cell bulk before peptide stimulation (threshold sensitivity:  $3 \times 10^{-6}$ ). TCR $\alpha$  chains were also sequenced and confirmed the presence of a unique major clone for P1, P4 and P6, enabling TCR pairing and modeling. Because 2 major V $\alpha$  rearrangements were obtained for P2, the predominant rearrangement occurring at a 60% frequency was used for TCR modeling.

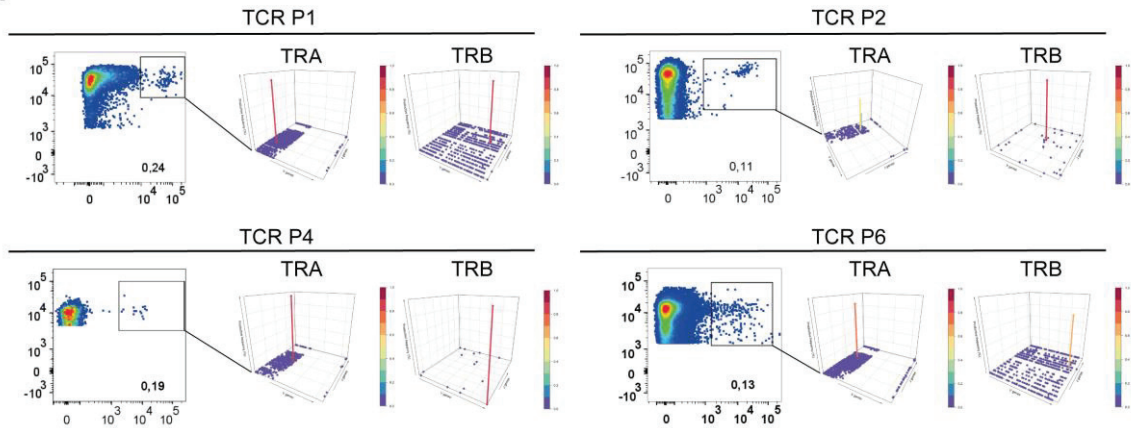
The affinity of the T cell clones specific for the peptides P1, P2, P4 and P6 was then characterized by considering 3D models of the TCR-peptide-MHC (pMHC) complexes (see Methods). The stability of macro-molecular complexes is due to the formation of favorable interactions at the interface such as hydrogen bonds, salt bridges and hydrophobic interactions. Figure 5C displays these favorable interactions in each of the 3D models of TCR-pMHC complexes. These interactions involve specific side-chains of the peptides, that are exposed at the TCR-pMHC interface. In the TCR P1 complex, Phe1, Phe4 and Trp8 side chains of the peptide form several hydrophobic interactions, and the backbone atoms of Phe 4 and Ile 9 are involved in H-bonds. In the TCR P2 complex, several hydrophobic interactions are mediated by peptide residues Pro4, Tyr5 and Trp7. In TCR P4, peptide residues Ile5, Ile7 and Leu8 form several hydrophobic interactions, while Tyr1 and Lys6 side chains, as well as Phe4 and Leu8 backbone atoms are involved in H-bonds. In TCR P6, Tyr1, Ser4, Asn5, Leu6 and Phe7 form several hydrophobic interactions, while Ser4/Leu6/Ser8 backbones and Tyr1/Ser4/Asn5/Ser8 side-chains form 8 H-bonds.

Overall, this analysis of the predicted 3D models suggests that the TCR-pMHC complexes are stabilized by several favorable non-covalent interactions, supporting that the TCR identified after clonal expansion of HERV-specific T cells form a stable complex with the peptides presented by HLA-A2 molecules. To gain further insight, we submitted the 3D models to binding affinity prediction (Fig. 5D). When compared to reference TCR-pMHC complexes available in the Protein Data Bank and obtained from crystallography data, the predicted affinities of the identified TCRs match clinically relevant TCR affinities, such as TCRs targeting MAGE-A3, NY-ESO-1, MART-1, HTLV or CMV (Fig. 5E). Hence, the HERV-specific TCRs identified are predicted to stably interact with their respective pMHC complexes, reminiscent of high affinity TCRs.



**Figure 5**

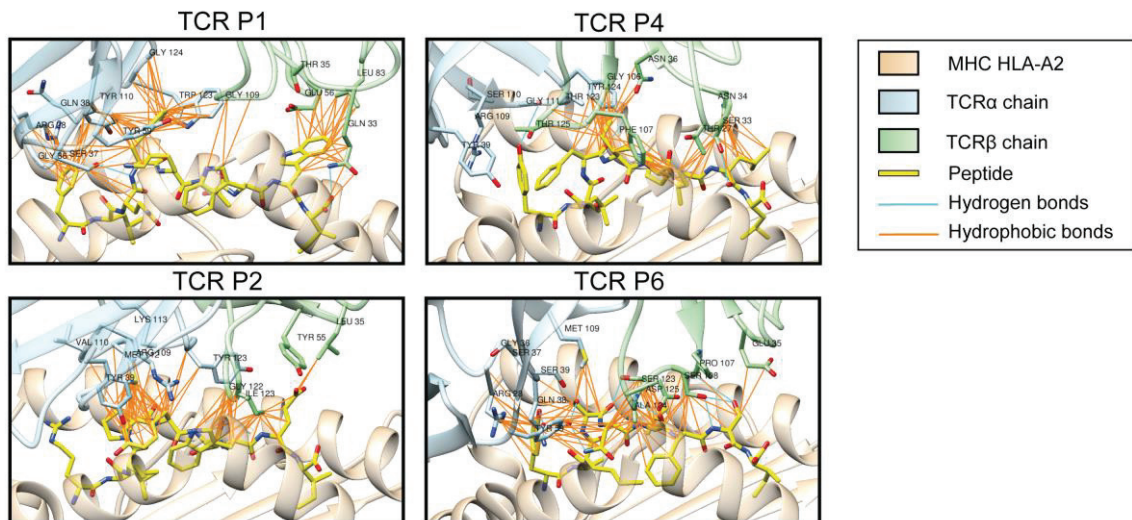
**A**



**B**

Antigen	TCR $\alpha$			TCR $\beta$			
	Frequency	V allele	J allele	Frequency	V allele	D allele	J allele
P1	91%	TRAV12-02*01	TRAJ24-01*01	91%	TRBV20-01*01	TRBD02-01*01	TRBJ02-03*01
P2	60%	TRAV26-01*02	TRAJ50-01*01	91%	TRBV09-01*01	TRBD02-01*01	TRBJ02-02*01
P4	98%	TRAV38-02*01	TRAJ40-01*01	99.60%	TRBV30-01*05	ND	TRBJ01-01*01
P6	89%	TRAV12-02*01	TRAJ49-01*01	76%	TRBV28-01*01	TRBD02-01*01	TRBJ02-03*01

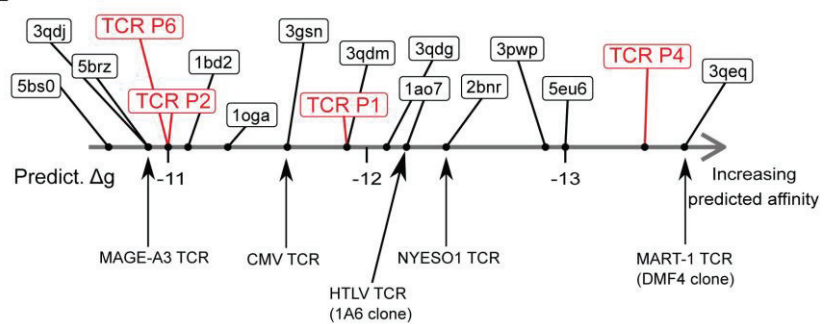
**C**



**D**

	Predict. $\Delta g$
*** TCR P4	-13.4
*** TCR P1	-11.9
*** TCR P2	-11.0
*** TCR P6	-11.0

**E**



**Figure 5: Visualization of CDR Loops and CDR Loop interactions with peptides**



- (A)** Epitope-specific CD8<sup>+</sup> T cells from HLA-A2-positive healthy donors were sorted by FACS using dextramer staining. Sorted cells (left in all four panels, gated) underwent TCR sequencing and results are shown in Manhattan plot reports of the V/J recombination of the TCR- $\alpha$  (center in all panels) and TCR- $\beta$  (right in all panels). V and J segments are represented according to chromosomal location on the X and Y-axis, respectively. Productive frequencies of clones are represented on the Z-axis.
- (B)** Productive frequency of the TCR $\alpha$  and TCR $\beta$  CDR3 sequences for the top clones specific to each peptide (P1, P2, P4, P6) and the corresponding resolved V, D and J alleles.
- (C)** 3D representative models of the TCR-pMHC interface for each identified TCR. The MHC chains (HLA-A2) are colored in beige, TCR alpha chains in light blue, TCR beta chains in light green, and peptides in yellow. Peptide and TCR residues involved in the TCR-pMHC interaction are shown as sticks, with O atoms in red, N atoms in dark blue and S atoms in yellow. Hydrogen bonds are represented as cyan lines, salt bridge as red lines, and hydrophobic contacts as orange lines.
- (D)** Predicted binding affinities (Predict.  $\Delta g$ ) are expressed in kcal/mol. Average values are reported. Stars indicate significant statistical test (Welch two-sample t-test) at the 5% level.
- (E)** Diagram ranking of modeled HERV-specific TCR-pMHC and reference TCR-pMHC complexes available in the Protein Data Bank and obtained from crystallography data, according to their predicted binding affinity.

*CDR: complementarity-determining region, TCR: T cell receptor; MHC: major histocompatibility complex; pMHC: peptide-MHC, TRA: alpha chain of TCR; TRB: beta chain of TCR.*

### ***High avidity HERV-specific T cell clones recognize and kill tumor cells***

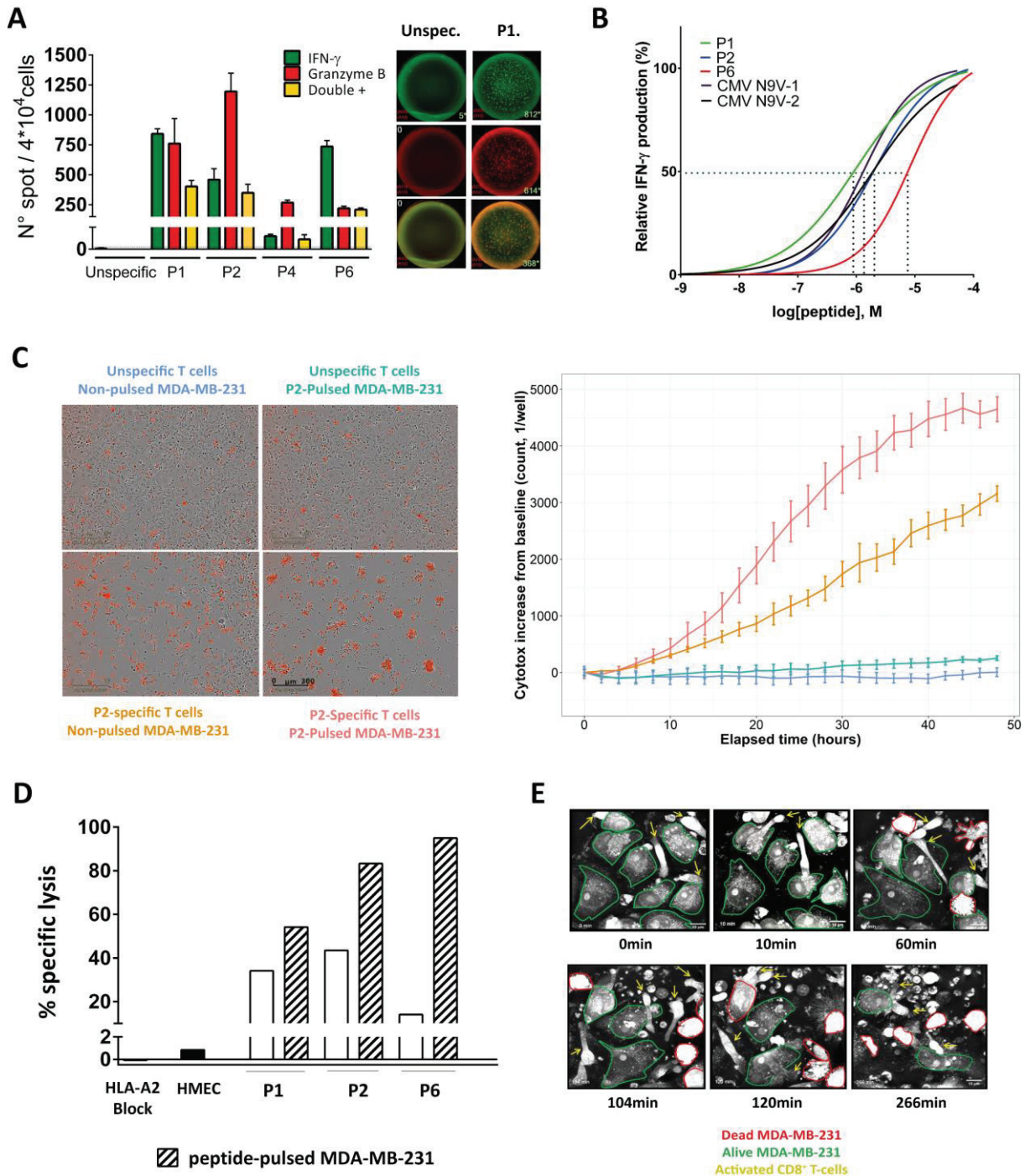
The functionality of the sorted and expanded epitope-specific CD8<sup>+</sup> T cells was confirmed by Fluorospot using peptide-pulsed T2 cells. A dual IFN- $\gamma$  and granzyme B secretion was shown, except for the P4-specific T cell clone that appeared poorly functional (notably in terms of IFN- $\gamma$  production) and was then excluded from further analyses (Fig. 6A). The functional avidity was subsequently assessed by loading T2 cells with decreasing concentrations of the cognate peptide (ranging from 10<sup>-4</sup> to 10<sup>-9</sup> M) and measuring the lowest peptide concentration necessary to provoke IFN- $\gamma$  responses in 50% of cells (defined as half maximal effective concentration, i.e. EC50). The EC50 values, estimated at 6.6 x 10<sup>-7</sup> M, 1.9 x 10<sup>-6</sup> M and 6.8 x 10<sup>-6</sup> M for P1, P2 and P6-specific T cells, respectively, are in the same order of magnitude as neoepitope-specific T

cell clones (29) and CMV-specific T cells ( $1.2 \times 10^{-6}$  and  $1.9 \times 10^{-6}$  for N9V-1 and N9V-2, respectively), suggesting an absence of negative selection in the thymus of the most reactive T cell clones (Fig. 6B).

We next assessed the capacity of these HERV epitope-specific CD8<sup>+</sup> T cells to recognize and kill tumor cells, using as a target the HLA-A2-positive MDA-MB-231 basal BRCA tumor cell line, previously shown to express HERVs containing epitope sequences (Fig. 3F and Supplementary Fig. S3D). Tumor cells were co-cultured with the epitope-specific T cells or with the dextramer-negative CD8<sup>+</sup> T cell fraction sorted and expanded in the same conditions (negative controls). Flow cytometry analysis highlighted IFN- $\gamma$  production by approximately 25% of epitope-specific T cells in contact with MDA-MB-231, with a significant increase ( $> 6$ -fold) compared to the background observed with non-specific T cells. This IFN- $\gamma$  production was inhibited by a HLA-A2 blocking monoclonal antibody, demonstrating that the T cell clones specifically recognized the tumor cells in a HLA-A2 restricted manner (Supplementary Fig. S6A and S6B).

In order to monitor tumor cell death in real-time, we performed an immune-cell killing assay using the IncuCyte technology. T cell clones induced a significant and HLA-A2-restricted killing of MDA-MB-231 cells, as shown by the time-dependent increase in the amount of Cytotox fluorescent reagent of target cells. In contrast, the dextramer-negative fraction of T cells did not induce significant cell death of MDA-MB-231 cells (pulsed or not with the peptide) (Fig. 6C). Specific lysis (at effector to target ratio E:T = 2:1) was calculated based on the quantification of target cell death after 48 hours after subtracting the alloreactive background (assessed by the target cell death induced with the corresponding dextramer-negative T cell fraction) (see Methods). A particularly high specific lysis of the tumor cells was achieved with P1 and P2-specific T cells (35% and 44%, respectively), with a more moderate lysis (15%) with P6-specific T cells. The specific lysis was further increased when the target tumor cells were pulsed with the cognate epitope, reaching 55%, 80% and even 95% for P1, P2 and P6-specific T cells, respectively. Of note, epitope-specific T cell clones did not kill HLA-A2-positive human mammary epithelial cells (HMECs) used here as a negative, normal cell, control (Fig. 6D). These results were further validated using the 3D microscopy Nanolive technology, showing morphological signs of activation of specific T cells associated with killing of the majority of the tumor cells after 4.5 hours (E:T=10:1). Again, no cell death was observed when the specific T cells were co-cultured with HMECs and when MDA-MB-231 cells were co-cultured with non-specific T cells (Fig. 6E, Supplementary videos 1, 2 and 3). Altogether, these data show that the selected epitopes elicit high avidity CD8<sup>+</sup> T cell clones that selectively recognize and kill HERV-expressing tumor cells.

**Figure 6**



**Fig. 6. HERV-specific T cell clones are functional, recognize and kill tumor cells**

(A) Left: Mean number and standard derivation of IFN- $\gamma$  (green), Grz-b (red) and double positive (yellow) spots counted on Fluorospot. Experiments were performed in duplicates for each condition represented for all healthy donors. Right: Representative wells of IFN- $\gamma$  (green, upper quadrants), Grz-b (red, center) and double positive (yellow, lower

quadrants) spots for unspecific (dextramer-negative, left) and peptide-specific (here P1) T cells (right) co-cultured with T2 cells in a ratio 10:1 for 24 hours.

**(B)** Representative analysis of the functional avidity of CD8<sup>+</sup> T cell clones calculated as nonlinear fit of normalized IFN- $\gamma$  production. N9-V1 and N9V-2 are 2 different CMV-specific T cell clones (see Methods). EC<sub>50</sub> are represented for each clone by the interpolation of the dashed lines with the X-axis.

**(C)** Left: Representative 10x images of co-cultures of specific T cells (here P2-specific) with MDA-MB-231 cell line after 48h cell killing assay using IncuCyte technology. Unspecific (upper panels) and specific (lower panels) T cells were co-cultured with non-pulsed (left) and peptide-pulsed (here P2) (right) MDA-MB231 at E:T =2:1. Dead MDA-MB-231 are depicted in red on images. Right: Representative plot of 48-hour cell death quantification represented as fluorescence intensity increase from the baseline (Y-axis) in function of the time (hours, X-axis). Colors indicating the combinations of effectors (P2-specific or unspecific T cells) and targets (MDA-MB-231 pulsed or not with P2) are the same as those used on the left.

**(D)** Specific tumor cell lysis at 48h. The mean percentage of specific lysis of technical triplicates is plotted for each condition. Dashed-pattern bars represent tumor cells pulsed with the cognate peptide. Data are representative of at least 2 independent experiments.

**(E)** 60X-pictures of MDA-MB-231 co-cultured with P1-specific CD8<sup>+</sup> T cells (E:T = 10:1) acquired at different time points using Nanolive technology. T cells are shown by yellow arrows. Living and dead cells are contoured in green and in red, respectively.

*FACS: Fluorescence-activated Cell Sorting, Grz-b: Granzyme B, IFN: Interferon, TCR: T Cell Receptor: E:T : effector to target ratio*

### ***HERV-specific T cells are present among tumor infiltrating T cells***

In order to test our hypothesis that an adaptive immune response against HERVs may exist in cancer patients, we assessed by dextramer-staining the presence of HERV epitope-specific T cells among polyclonally expanded tumor infiltrating lymphocytes (TILs) from TNBC HLA-A2 patients. HERV-specific TILs were observed for at least one epitope in 7 of the 11 analyzed tumor samples, with variations in terms of epitope specificity and frequency from one patient to another. P1, P4 and P6 were the most frequently recognized peptides, with a dextramer-based identification in 4/11, 4/11 and 5/11 cases, respectively, whereas no significant staining was seen for P3 (Fig. 7A and 7B).

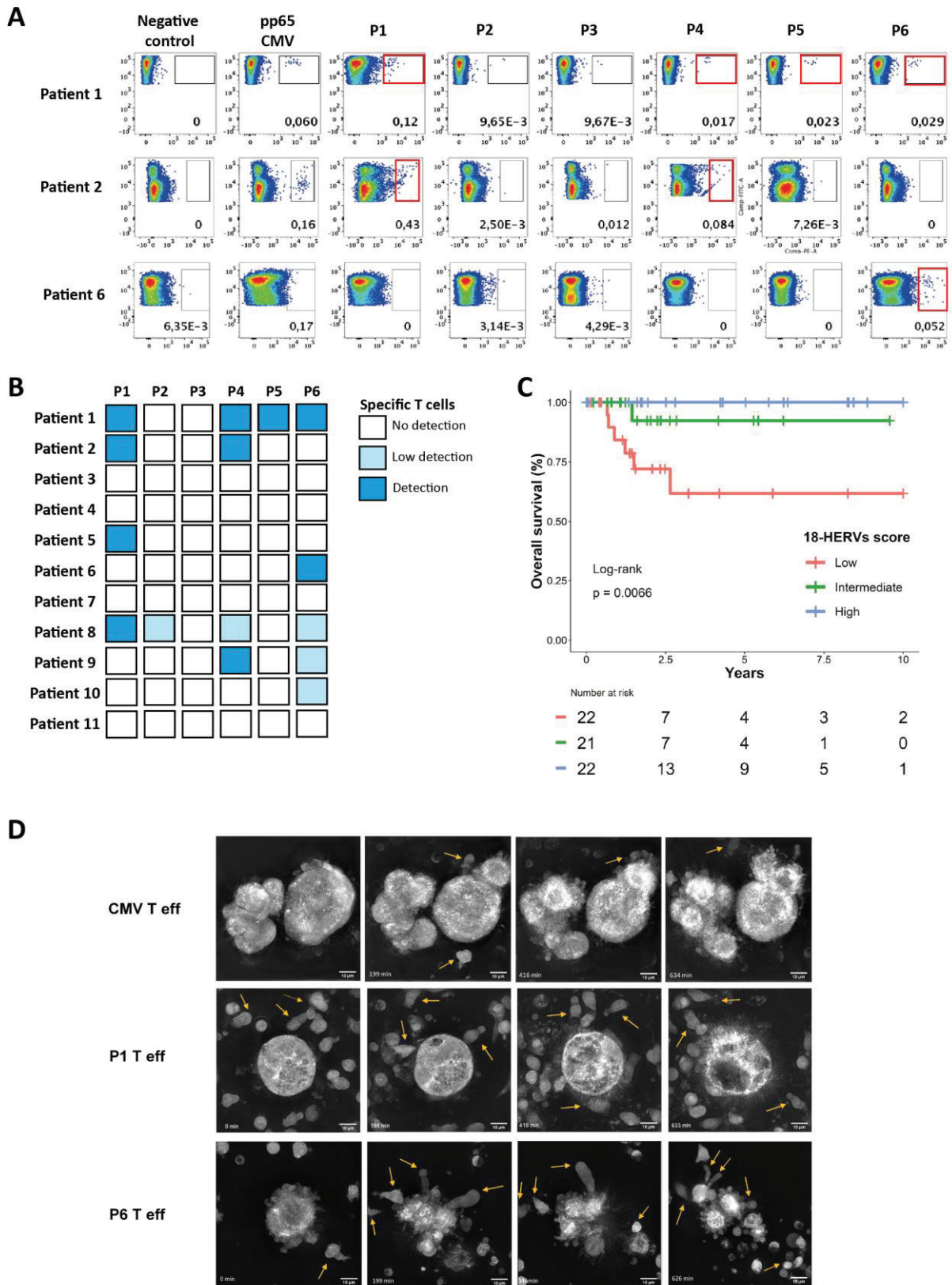
These results prompted us to investigate the potential link between the outcome of TNBC patients and the expression of the 18 HERVs containing these HLA-A2 epitopes. We established a score based on the mean expression of these 18 HERVs in HLA-A2 patients with basal breast cancer from TCGA cohort. Interestingly, HLA-A2-positive patients with a high or intermediate 18-HERVs score had a significantly better overall survival than those with a low score ( $P = 0.0066$ ) (Fig. 7C). This prognostic impact was not observed in the overall population (Supplementary Fig. S7A).

Finally, we evaluated the antitumor activity of HERV-specific T cells against primary tumor cells by using organoids derived from the tumor of patient 8 (Supplementary Fig. S7B, see Methods). RNAseq analysis confirmed the expression of the 18 epitopes-containing HERVs at early and late passage (Supplementary Fig. S7C). Tumor organoids were co-cultured with P1, P6 or CMV-specific CD8<sup>+</sup> T cell clones in a 3D microscopy Nanolive experiment (E:T=10:1). Whereas no activation of T cells was observed with CMV-specific T cells, P1 and P6-specific T cells exhibited signs of active proliferation associated with lysis of the organoids (Fig. 7D).

Altogether, these last results suggest that HERV-specific T cells are induced during tumor development and may participate in the antitumor immune response.



**Figure 7**



**Fig. 7. HERV-specific T cells are present among tumor infiltrating T cells**



- (A)** Representative panels of dextramer staining for the 6 HERV epitopes in TILs from 3 HLA-A2-positive TNBC patients. Dextramer positive cells were gated on CD8<sup>+</sup> T-cells. Negative control: dextramer complexed to a non-natural irrelevant peptide (ALIAPVHAV).
- (B)** Results summary of dextramer CD8<sup>+</sup> T cell detection among TILs from 11 HLA-A2-positive TNBC patients for the 6 HERV peptides. Blue squares indicate the detection of epitope-specific CD8<sup>+</sup> T cells; light blue squares indicate low detection and white squares indicate no detection.
- (C)** Overall survival according to 18-HERVs score in TCGA HLA-A2 TNBC patients (n=65). Patients were divided in three groups according to the score terciles: blue line, high expression (n=22); green line, intermediate expression (n=21); red line, low expression (n=22)
- (D)** 60X-pictures of TNBC organoids co-cultured with CMV, P1 or P6-specific CD8<sup>+</sup> T cell clones (top-down) acquired at different time points using Nanolive technology. T cells are shown by yellow arrows.

*TILs: Tumor Infiltrating Lymphocytes, TNBC: Triple-negative Breast Cancer*

## Discussion

The first study providing a comprehensive characterization of the expression of HERVs in cancer was published in 2015 by Rooney et al (30). Using the same annotations, we assessed their expression in an independent RNAseq database with different breast cancer subtypes (28). This approach demonstrated that a limited number of HERV-K/HML-2 members are specifically overexpressed in TNBCs and are associated with T cell infiltration and several negative immune checkpoints. These results suggest that HERVs not only trigger an innate response through RNA sensing (31, 32) but also induce an adaptive immune response. However, this preliminary approach presents several limitations, including a small reference set of only 66 HERVs, which pooled results from heterogeneous studies performed with different methodologies. Recently, Smith et al developed the computational workflow HervQuant for the quantification of HERVs from RNAseq data (15), based on Vargiu's reference that compiles 3,173 intact full-length HERV sequences (24). These sequences underwent a first round of selection to retain HERVs specifically overexpressed in tumors. In order to limit the number of candidates for further validation, we next developed a machine learning-based approach using

penalized regression, allowing a stringent selection by retaining only cancer-associated HERVs significantly associated with a CTL response.

This whole approach led to the selection of a limited number of cyt-HERVs per disease. We did not find a significant correlation with demethylation for all the HERVs studied. This may be due to the limitations of the methylation assay, which evaluates methylation in sites located at variable distances from the selected HERV locus. More recent techniques such as ATAC-seq could be used to more accurately address the epigenetic regulation of HERVs. Interestingly, the largest numbers of cyt-HERVs were in tumors that respond well to checkpoint inhibitors, such as lung, head and neck and renal cell carcinoma, but also in colon cancer, regardless of the subtype and mismatch repair status. The immune impact of HERVs and cyt-HERVs in this latter disease remains to be addressed in a dedicated study to better understand the interactions and the mechanisms underlying the observed immune resistance in mismatch-repair proficient colon cancer.

A limited number of shared T cell epitopes were identified by our approach. The fact that each epitope sequence is present in multiple HERVs located at different chromosomal loci may represent a major advantage compared to other tumor antigens by significantly decreasing the risk of selection of tumor clones losing the epitope. The selected epitopes induce functional T cells with cytolytic properties, in agreement with the hypothesis that HERV-reactive T cells are not eliminated in the thymus during the negative selection process. Priming assays do not rule out the possibility that memory T cells are also activated. However, the low frequency of specific T cells, which could be estimated at less than  $3 \times 10^{-6}$  based on the analysis of specific TCR $\beta$  rearrangements in the T cell bulk before peptide stimulation, suggests that a low number of naïve T cells have been primed (33). The variations in responses observed between the donors may be due to the availability of the right TCR according to each individual T cell repertoire.

Our main assumption was that HERV-derived epitopes, due to their homology with virus sequences, may elicit high affinity functional T cells. The selected epitopes induced CD8<sup>+</sup> T cell clones of high predicted affinity and high functional avidity, as observed with virus antigens. The epitope specificity of these T cells is highlighted by the HLA-A2 restriction, the lack of cytotoxicity against non-tumor HLA-A2-positive cells, the increased cytotoxicity observed with pulsed target cells and structural modeling.

Finally, we found pre-existing HERV-specific T cells among TILs from TNBC samples, with variable frequencies and specificities among patients. This observation suggests that the selected epitopes (at least P1, P4 and P6) are efficiently processed and presented to T cells during

tumor development. Based on the prognostic impact of epitope-containing HERVs in HLA-A2 patients, we will further assess the links between the expression levels of these HERVs, the frequency of HERV-specific T cells among TILs and patient outcome.

Our results show HERV-derived targets as a class of “virus-like” tumor antigens shared by specific tumor subtypes. HERV-derived epitopes could be used in a strategy of innovative cancer vaccines, especially in tumors with low or intermediate mutational burden that are poor candidates for mutation-associated neoepitope-based vaccines. Furthermore, the characterization of TCRs specific to HERV-derived epitopes may also lead to the development of an immune cell therapy based on TCR-engineered T cells.

## **MATERIALS AND METHODS**

### **Datasets**

For RNA-seq data, raw fastq files were accessed from the NCBI Gene Expression Omnibus (GEO) portal, under the accession number GSE58135 for Varley et al. independent breast cancer dataset (28), GSE74246 for the sorted PBMC dataset (34), GSE127825 and GSE127826 for the six mTECs samples (35). TCGA pancancer raw fastq files were accessed from the Genomic Data Commons (GDC) portal (<https://portal.gdc.cancer.gov/>). Cell line data were accessed from the Broad Institute Cancer Cell Line Encyclopedia (CCLE) portal (<https://portals.broadinstitute.org/ccle>).

### **HERV expression quantification**

Two different published methods were used to assess HERV expression (15, 30). In the first one, described by Rooney et al., RNAseq reads were mapped with bowtie2 (36) to an annotation of 66 known HERVs sequences (20) and normalized against the count of mapped reads, leading to counts per million (CPM). For the second one, the HervQuant pipeline was used (15). Briefly, RNAseq reads were mapped with STAR v2.7.3a (37) to the hg19 reference transcriptome compiled with the annotation of 3,173 HERV sequences (24). Multimaps  $\leq 10$  and mismatch  $\leq 7$  were allowed, as in the original publication. BAM outputs were filtered for reads that mapped HERV sequences using SAMtools v1.4 (38) and then quantified using Salmon v0.7.2 (39). Raw counts were normalized to counts per million total reads and then  $\log_2 + 1$  transformed. Results were comparable to the original data published by Smith et al. (15).

### **Quality check / sample filtering**

Only primary solid tumor samples (TCGA code 01) were included, regrouping 9,718 samples from 32 different cancer types, from which 9,492 were analyzable for HERV expression. Quality

check resulted in complete removal of ESCA and STAD samples due to a largely skewed HERV distribution, leading to the final analysis of 8,893 samples from 29 different cancer types (Supplementary Fig. S2 and Supplementary Table 1).

### **Immune signatures and genetic alterations**

Phenotypic immune signatures were calculated with the Xcell method (22). For the TCGA Pancancer samples, Xcell signatures were directly downloaded from the Xcell website ([https://xcell.ucsf.edu/xCell\\_TCGA\\_RSEM.txt](https://xcell.ucsf.edu/xCell_TCGA_RSEM.txt)). For the GSM1401648 dataset, signatures were calculated for the whole dataset, and immune signatures were filtered after. Interferon-gamma (IFN- $\gamma$ ) signature was calculated by single-sample gene set variation analysis (GSVA) (40) based on the HALLMARK\_INTERFERON\_GAMMA\_RESPONSE signature from the Molecular Signature Database (<http://software.broadinstitute.org/gsea/msigdb/index.jsp>). Enrichment scores were calculated for each sample per cancer type. The cytolytic activity (CYT\_score) was calculated as the geometric mean of granzyme-B (GRZB) and perforin (PRF1) expression, as previously described (30). TCGA Pancancer Genetic alterations were retrieved from Thorsson V. et al. (41).

### **Cancer-associated and cyt-HERV annotation**

To define cancer-specificity, differential HERV expression was performed between tumor samples and their respective normal peritumoral matched tissues. Only TCGA studies with at least 10 peritumoral samples were included (n=14 different cancer types). Differential HERV expression analysis was performed independently for each TCGA cancer type. Having filtered-out any HERV expressed more than 2-fold in any normal tissue compared to its matched tumor, remaining HERVs overexpressed more than 2 fold in at least one cancer compared to its normal counterpart (peritumoral tissue) was considered cancer-associated.

To be annotated as potentially immunogenic, each cancer-associated HERV had to be associated with at least one phenotype (A) criterion and one functionality (B) criterion and not be overexpressed by T/NK cells (C). Phenotype criteria included association with either CD4 or CD8+ T cell signatures as defined by the Xcell method (22). Function criteria included association with either IFN- $\gamma$  or the cytolytic activity, defined by the geometric mean of granzyme-A (GZMA) and perforin (PRF1) expression (30). Normal PBMC expression was assessed in an independent dataset of sorted-PBMCs from healthy donors (34). HERV expression was compared independently in T cells and NK cells to the rest of PBMCs.

### **L1-Penalized regression (LASSO)**

Associations were calculated by Lasso regression using the glmnet and the c060 packages (42, 43). Gaussian distribution was considered for the CYT score and the IFN- $\gamma$  signatures, and Poisson distribution was considered for the Xcell signatures. HERVs were analyzed as  $\log_2(\text{CPM}+1)$ , requiring no further standardization. For each cancer type, a model was built based on optimal parameters found with 10-fold cross validation. Each HERV with a positive coefficient in the final model (based on the lambda parameter minimizing the mean-squared error) was considered to be associated with the variable.

### **Epitope screening**

Open-reading frame (ORF) detection was performed using sixpack from EMBOSS v6.6.0.0 (44). Detected ORFs of more than 10 amino acids were then aligned to known HML-2 (HERV-K) Gag, Pro, Pol, Env, Rec and Np9 proteins referenced in UniProt (21). Blast with optimal parameters for retrovirus was used (Word size = 3, composition-based statistics, no “low-complexity-region” filter). Conserved sequences aligned with Gag and Pol proteins with more than 90% identity and an e-value < 0.05 were then screened for predicted HLA-A\*02 strong binders using MHCflurry v1.3 (45). Peptides with a rank  $\leq$  0.5 percentile were considered to be strong-binders. The human proteome was downloaded on Uniprot (ID: UP000005640) to validate the absence of match before peptide synthesis and in vitro validation.

### **Analysis of peptidome proteomic datasets**

Raw MS/MS datasets were downloaded from CPTAC (23) for breast cancer studies (46, 47). Retrieved MS/MS spectra were converted to MGF format using msconvert from proteowizard (48). Then the list of peptides was analysed using the standalone version of Pepquery (v.1.6.2.0) (49). The used command line was: `java -Xmx10G -jar pepquery.jar -fixMod 6,62,108 -varMod 117 -maxVar 3 -c 1 -tol 10 -tolu ppm -minScore 12 -e 1 -um -hc TRUE -n 1000 -itol 0.05 -m 1 cpu 12 -pep  $\{\text{peptides\_list}\}$  -db  $\{\text{Reference\_database}\}$  -ms  $\{\text{MS\_database}\}$  -o  $\{\text{output\_directory}\}$ . For the second data set the fixmod and varMod in the command line were adapted like the following: -fixmod 6,103,157 -varMod 101,117.`

### **Riboseq analysis**

Ribosome-profiling data were retrieved from a previously published study (27). Raw fastq files were pre-processed as described in the initial publication. Briefly, adapter sequences were trimmed from raw data using cutadapt 1.1 with parameters (`--quality-base=33 -O 12 -m 20 -q 5`) and mapped to our reference hg19-HERV transcriptome.

### **Statistical analysis**

All analyses were performed using R statistical software version 3.6.0. Differential HERV expression analysis was performed using DESEQ2 v1.24.0 (50) and logarithmic fold changes were shrunk with the apeglm package (51).

### **Biological samples**

Blood from healthy donors was obtained from the “Etablissement Français du Sang” (Lyon). Fresh TNBC (n=11) were provided by the tissue bank of Centre Léon Bérard (CLB) (BB-0033-00050, CRB - CLB, Lyon, France; French agreement number: AC-2013-1871), after approval from the institutional review board and ethics committee (L-06-36 and L-11-26) and patient written informed consent, in accordance with the Declaration of Helsinki.

### **Peptides synthesis**

Peptides were synthesized at JPT peptide Technologies (GE, EU) with a specification and a purity >90%. Lyophilized powder was resuspended in 1% DMSO distilled water.

### **Cell lines**

MDA-MB-231 basal breast cancer epithelial cells were obtained from American Type Culture Collection (ATCC catalog name: HTB-26) and cultured in 10% FBS DMEM (Gibco, FR, EU) 1% Penicillin/streptomycin 1% L-Glutamine. HMEC primary cells were obtained from Promocell (GE, EU) and cultured in mammary epithelial growth medium (Promo Cell, GE, EU)

### ***In vitro* priming assays**

PBMCs were obtained by Ficoll density gradient centrifugation (Eurobio, FR, EU). They were rapidly thawed at 37°C and extensively washed, let at room temperature or overnight at 37°C before assessing their viability.  $0.15 \times 10^6$  PBMCs per well were cultured in 96 well plates with AIM V Medium (Gibco, FR, EU) enriched with 5µg/mL R-848 (Resquimod), 10µg/mL HMW poly-IC (both Invivogen, FR, EU), 20IU/mL IL-2 (PROLEUKIN aldesleukine, Novartis Pharma, CH, EU) and 10µg/mL of the peptide of interest at day 0. After 3, 6 and 10 days 100µL of medium were replaced by enriched fresh medium (IL-2 and peptide only at day 6 and IL-2 only at day 10) and splitted if necessary. On day 12 cells were collected and counted for analysis.

### **Feeding protocol**

Dextramer single-cell sorted CD8<sup>+</sup> T cells were expanded on a feeder composed by 35 Gy-irradiated allogeneic PBMCs and B-lymphoblastoid cell lines in a ratio 10:1. Feeder cells were plated in a 96-well round bottom plate at a concentration of  $0.10 \times 10^6$  cells per well in RPMI 5% human serum with PHA-L 1.5µg/mL (Merck KgAa, GE, EU) and IL-2 150 IU/mL (Novartis



Pharma, CH, EU) and up to  $5 \times 10^3$  sorted cells were added per well. Cells were cultured for 14 days and medium was replaced when needed with fresh IL-2 enriched RPMI 5% human serum. This process was repeated if needed.

### **TCR immunosequencing**

DNA from specific CD8<sup>+</sup> T cells and the corresponding bulk PBMCs was extracted using the QIAGEN QIAmp® DNA Blood Micro kit (QIAGEN, GE, EU) and sent for TCR survey and deep analysis to Adaptive Biotechnologies (WA, US).

### **Generation and refinement of 3D models**

The full TCR sequences of both alpha and beta chains were reconstructed for each T cell clone from the results of the immunosequencing as previously published (52). For variable domains, TRA and TRB CDR3 nucleotide sequences were obtained from immunosequencing (Fig. 5A) and the 5' and 3' ends of the TRAV and TRBV regions were obtained from International Immunogenetics Information System (IMGT) online database. Human constant domains of TRA and TRB were added in 3' of the variable domains to reconstitute the full-length TCR. These full-length TCR sequences, together with the MHC and peptide sequences were submitted to the CBS TCRpMHCmodels-1.0 web server, specifically developed for the automatic structural modeling of TCR-pMHC complexes (53) using template-based modeling. TCR residues are renumbered using a standardized procedure (54, 55). The initial models generated by the web server were further refined in four rounds using a protocol adapted from Bobisse et al (29). Briefly, the CDR loops were refined by pairs using Modeller software version 9.25 (56, 57): CDR loops  $\alpha 1/\alpha 2$  at round 1,  $\alpha 1/\alpha 3$  at round 2,  $\beta 1/\beta 2$  at round 3 and  $\beta 1/\beta 3$  at round 4. Only the residues in coil conformations in the initial models were refined. In each round, 500 models were generated and the best model based on the Modeller internal DOPE score was selected and used as input for the next round.

At the end of the four rounds, a representative model was chosen based on the consensus of unweighted contacts as follows: contacts between residues at the TCR-pMHC interface (defined by a distance lower than 5 Angstrom between heavy atoms) were counted in the set of 4\*25 models with best DOPE scores of each refinement round. Then, among the 25 models with best DOPE scores in the final round, the model with the highest number of recurrent contacts (referred as un-normalized CONSRANK score (58, 59)) is elected as the representative model.

A quantitative view of potential stabilizing interactions between TCR and pMHC is provided by the frequencies of inter-residue contacts observed in the set of 4\*25 models with best DOPE scores of each refinement round (Supplementary Fig. S5D). These TCR-pMHC interactions concur with previous studies (60). Whereas CDR1 and CDR3 loops of both TCR $\alpha$  and TCR $\beta$  chains interact with both the MHC and the peptide, the CDR2 loops interact mostly with the MHC molecule.

### ***3D structure analysis***

The structural similarity between representative models of different complexes was assessed by the RMSD between backbone CDR loops computed with UCSF Chimera (61) (Supplementary Fig. S5B and S5C). As expected, structural variability was highest within CDR $\beta$ 3 loops. UCSF Chimera was also used for hydrogen bonds and hydrophobic contact detection and structure visualization.

### ***Binding affinity prediction***

The binding affinity was predicted using the prodigy method (62) which uses a linear model based on the number and types of contacts at the interface. For each complex, instead of running one prediction on the representative model, we averaged the predictions obtained for the 25 models with best DOPE scores obtained at round 4 of the refinement protocol.

### **Fluorospot**

After co-culture of PBMCs or CD8<sup>+</sup> T cells with T2 cells pulsed or not with the peptide in a ratio 10:1 in AIM V medium (Gibco, FR, EU) at 37°C 9% CO<sub>2</sub> for 24 hours, double-color Fluorospot (CTL GmbH, CA, US) with IFN- $\gamma$  AF488 and Grz-b CTL-red was performed according to manufacturer's instructions. Revelation plate was read on ImmunoSpot® S6 ULTIMATE UV Image Analyzer and analyzed with the ImmunoSpot® analysis software.

### **Functional avidity**

Dextramer-isolated specific CD8<sup>+</sup> T cells for the selected peptides (P1, P2, P6) were used in a functional avidity IFN- $\gamma$  production test by enzyme-linked immunosorbent assay (IFN- $\gamma$  ELISA, Thermo Fisher scientific, FR, EU). Two CMV T cell clones specific for the immuno-dominant epitope N9V (NLVPMVATV) were used. N9V-1 corresponds to pp65 dextramer-selected CD8<sup>+</sup> T cells and N9V-2 is a CD8<sup>+</sup> T cell clone kindly provided by Dr Henri Vie. Functional avidity of specific CD8<sup>+</sup> T-cell responses was assessed by performing limiting peptide dilutions from 10<sup>-4</sup> to 10<sup>-9</sup> M (log) charged on T2 cells pulsed for 5 hours. After wash, peptide-pulsed T2 cells were co-

cultured with specific CD8<sup>+</sup> T cell in a ratio 1:1 in AIM-V medium (Gibco, FR, EU) supplemented with 5% of human serum. After 18 hours, supernatants were collected and ELISA was performed. The peptide concentration required to achieve a half maximal cytokine response (EC<sub>50</sub>) was determined (graphpad prism, version 6.0 for Windows was used for the 50% EC (EC<sub>50</sub>) determinations, R>0.98)

### **Live imaging**

Cells were plated in DMEM (Gibco, FR, EU) medium, 10% SVF, 1% Penicillin/Streptomycin. For IncuCyte analysis medium was removed from the 96 wells plate after overnight cell adhesion. A blocking HLA-A2 antibody (GeneTex, clone BB7.2, GTX75806, CA, US) was added in AIM-V medium (Gibco, FR, EU) for 1 hour, according to conditions. T cells were then added in an effector to target (E:T) ratio 2:1 in the presence of IncuCyte Cytotox dye (Essen Bioscience, UK, EU) for cell death quantification. A 48-hour live imaging was performed at 37°C 5%CO<sub>2</sub> with Incucyte Zoom. Cell death was calculated as the total number of counted stained cells corrected by the number of counted stained cells at baseline. Maximum killing was established using DMSO. Specific lysis was calculated according to the following formula:

$$\% \text{ specific lysis} = \left( \frac{\text{((HERV-specific T cells induced target cell death - spontaneous target cell death) - (non-specific dextramer-negative T cells induced target cell death - spontaneous target cell death))}}{\text{(DMSO induced target cell death - spontaneous target cell death)}} \right) \times 100$$

For Nanolive imaging T cells were then added with an E:T 10:1 and phase imaging was performed every minute using Nanolive microscope 3D cell explorer.

### **Tumor dilacerations: organoids and TILs expansion**

Tumor tissues were dissected into fragments of approximately 1 mm<sup>3</sup> and dilacerated with collagenase IV and DNase for 45 minutes in 20% SVF enriched RPMI. The tumor lysate was centrifuged at 1500 rpm for 5 minutes and resuspended in 5% human serum enriched RPMI. Cells were counted and plated at a density of 5X10<sup>4</sup> cells per well in a flat bottom 96-well plate with anti-CD3 anti-CD28 Dynabeads (Dynabeads, Gibco, EU) and IL-2 at 100 IU/mL in a ratio beads to cells of 1:4.

For organoids, a part of the tumor lysate (3 to 10 million of cells) was resuspended in 10mL of Advanced DMEM/F12 medium. Cells were centrifuged at 500 rcf for 10 seconds and then resuspended in full medium. This protocol was repeated 3 to 5 times according to the cell number at the beginning to enrich the cell suspension in epithelial cells. These cells were then cultured according to the protocol previously described by Dreuhs et al. (63).

## Multi-parametric Flow Cytometry

T cells were counted and co-cultured with T2 cells loaded or not with the cognate peptide in a 5:1 ratio. After one hour CD107a antibody (BD, clone H4A3) was added in each well with Golgi plug (1/1000) (10 µg/mL, BD, FR, EU). After 5 hours, viability, surface and intra-cellular staining were performed. To assess cytokine expression in CD8<sup>+</sup> T cells an intracellular staining with the FoxP3 Fixation and Permeabilization kit (Thermo Fisher scientific, Life Technologies, CA, US) was used, according to manufacturer's instructions.

Dextramer staining was performed on PBMCs after a 12-day culture (priming protocol) or on TILs expanded for 14 days after tumor dilaceration. Cells were washed in 2 mL washing buffer (PBS + 2% FBS + 2mM EDTA (Sigma Alderich, MI, US)) and stained for 10 minutes with dextramers (Immudex ApS, DK, EU) at room temperature prior to viability and surface marker staining. Washing was performed 2 times to avoid non-specific dextramer staining. CMV pp65 NLVPMVATV was used as positive control. For TILs analysis, a dextramer complexed to a non-natural irrelevant peptide (ALIAPVHAV) was used as negative control.

All samples were analyzed on a LSR-Fortessa (BD Biosciences, FR, EU) with conserved settings throughout the entire study. Data were analyzed using FlowJo Software (Tree Star v10.4, NJ, USA).

## Supplementary Materials

### Materials and Methods

Fig. S1. Differential HERV expression in breast cancer subtypes (Varley's database) and peptide mass spectra

Fig. S2. TCGA pancancer flow-chart and methylation analysis

Fig. S3. TCGA basal breast-cancer cyt-HERVs and peptide analysis

Fig. S4 HERV-derived epitopes CD8<sup>+</sup> T cell responses and gating strategies

Fig. S5. Dextramer staining after feeder expansion and analysis of 3D structural models

Fig. S6 Flow cytometry assays on T cells co-cultured with tumor cells

Fig. S7 HERV expression in TNBC patients and organoids cell surface markers

Table S1. Blast results of the 6 HERV peptides aligned against the human proteome.

Table S2: Number of samples from each cancer type included in the study

Supplementary videos 1, 2 and 3: 3D microscopy imaging using Nanolive technology : P1-specific T cells co-cultured with MDA-MB-231 (video 1) or HMECs (video 2) and dextramer-negative non-specific T cells co-cultured with MDA-MB-231 (video 3).

## References and Notes

1. P. G. Coulie, B. J. Van den Eynde, P. van der Bruggen, T. Boon, Tumour antigens recognized by T lymphocytes: at the core of cancer immunotherapy, *Nat. Rev. Cancer* 14, 135–146 (2014).
2. M. Yarchoan, B. A. Johnson, E. R. Lutz, D. A. Laheru, E. M. Jaffee, Targeting neoantigens to augment antitumour immunity, *Nat. Rev. Cancer* 17, 209–222 (2017).
3. V. Alcazer, P. Bonaventura, L. Tonon, S. Wittmann, C. Caux, S. Depil, Neoepitopes-based vaccines: challenges and perspectives, *Eur. J. Cancer* 108, 55–60 (2019).
4. C. C. Smith, S. R. Selitsky, S. Chai, P. M. Armistead, B. G. Vincent, J. S. Serody, Alternative tumour-specific antigens, *Nat. Rev. Cancer* 19, 465–478 (2019).
5. A. S. Attermann, A.-M. Bjerregaard, S. K. Saini, K. Grønbaek, S. R. Hadrup, Human endogenous retroviruses and their implication for immunotherapeutics of cancer, *Ann. Oncol.* 29, 2183–2191 (2018).
6. R. F. Downey, F. J. Sullivan, F. Wang-Johanning, S. Ambs, F. J. Giles, S. A. Glynn, Human endogenous retrovirus K and cancer: Innocent bystander or tumorigenic accomplice?, *Int. J. Cancer* 137, 1249–1257 (2015).
7. H. M. Rowe, D. Trono, Dynamic control of endogenous retroviruses during development, *Virology* 411, 273–287 (2011).
8. K. Rycaj, J. B. Plummer, B. Yin, M. Li, J. Garza, L. Radvanyi, L. M. Ramondetta, K. Lin, G. L. Johanning, D. G. Tang, F. Wang-Johanning, Cytotoxicity of human endogenous retrovirus K-specific T cells toward autologous ovarian cancer cells, *Clin. Cancer Res.* 21, 471–483 (2015).
9. B. S. Reis, A. A. Jungbluth, D. Frosina, M. Holz, E. Ritter, E. Nakayama, T. Ishida, Y. Obata, B. Carver, H. Scher, P. T. Scardino, S. Slovin, S. K. Subudhi, V. E. Reuter, C. Savage, J. P. Allison, J. Melamed, E. Jäger, G. Ritter, L. J. Old, S. Gnjatic, Prostate cancer progression correlates with increased humoral immune response to a human endogenous retrovirus GAG protein, *Clin. Cancer Res.* 19, 6112–6125 (2013).
10. J. Krishnamurthy, B. A. Rabinovich, T. Mi, K. C. Switzer, S. Olivares, S. N. Maiti, J. B. Plummer, H. Singh, P. R. Kumaresan, H. M. Huls, F. Wang-Johanning, L. J. N. Cooper, Genetic Engineering of T Cells to Target HERV-K, an Ancient Retrovirus on Melanoma, *Clin. Cancer Res.* 21, 3241–3251 (2015).
11. H. S. Jang, N. M. Shah, A. Y. Du, Z. Z. Dailey, E. C. Pehrsson, P. M. Godoy, D. Zhang, D. Li, X. Xing, S. Kim, D. O'Donnell, J. I. Gordon, T. Wang, Transposable elements drive widespread expression of oncogenes in human cancers, *Nat. Genet.* 51, 611–617 (2019).
12. G. Kassiotis, J. P. Stoye, Immune responses to endogenous retroelements: taking the bad with the good, *Nat. Rev. Immunol.* 16, 207–219 (2016).

13. F. Schiavetti, J. Thonnard, D. Colau, T. Boon, P. G. Coulie, A Human Endogenous Retroviral Sequence Encoding an Antigen Recognized on Melanoma by Cytolytic T Lymphocytes, *Cancer Res.* 62, 5510–5516 (2002).
14. Y. Takahashi, N. Harashima, S. Kajigaya, H. Yokoyama, E. Cherkasova, J. P. McCoy, K. Hanada, O. Mena, R. Kurlander, T. Abdul, R. Srinivasan, A. Lundqvist, E. Malinzak, N. Geller, M. I. Lerman, R. W. Childs, Regression of human kidney cancer following allogeneic stem cell transplantation is associated with recognition of an HERV-E antigen by T cells, *J. Clin. Invest.* 118, 1099–1109 (2008).
15. C. C. Smith, K. E. Beckermann, D. S. Bortone, A. A. De Cubas, L. M. Bixby, S. J. Lee, A. Panda, S. Ganesan, G. Bhanot, E. M. Wallen, M. I. Milowsky, W. Y. Kim, W. K. Rathmell, R. Swanstrom, J. S. Parker, J. S. Serody, S. R. Selitsky, B. G. Vincent, Endogenous retroviral signatures predict immunotherapy response in clear cell renal cell carcinoma, *J. Clin. Invest.* 128, 4804–4820 (2018).
16. C. S. Mullins, M. Linnebacher, Endogenous retrovirus sequences as a novel class of tumor-specific antigens: an example of HERV-H env encoding strong CTL epitopes, *Cancer Immunol. Immunother.* 61, 1093–1100 (2012).
17. S. Rakoff-Nahoum, P. J. Kuebler, J. J. Heymann, M. E Sheehy, G. M. Ortiz, G. S Ogg, J. D. Barbour, J. Lenz, A. D. Steinfeld, D. F. Nixon, Detection of T lymphocytes specific for human endogenous retrovirus K (HERV-K) in patients with seminoma, *AIDS Res. Hum. Retroviruses* 22, 52–56 (2006).
18. F. Wang-Johanning, L. Radvanyi, K. Rycaj, J. B. Plummer, P. Yan, K. J. Sastry, C. J. Piyathilake, K. K. Hunt, G. L. Johanning, Human endogenous retrovirus K triggers an antigen-specific immune response in breast cancer patients, *Cancer Res.* 68, 5869–5877 (2008).
19. G. L. Johanning, G. G. Malouf, X. Zheng, F. J. Esteva, J. N. Weinstein, F. Wang-Johanning, X. Su, Expression of human endogenous retrovirus-K is strongly associated with the basal-like breast cancer phenotype, *Sci. Rep.* 7 (2017), doi:10.1038/srep41960.
20. J. Mayer, J. Blomberg, R. L. Seal, A revised nomenclature for transcribed human endogenous retroviral loci, *Mob. DNA* 2, 7 (2011).
21. UniProt: a worldwide hub of protein knowledge, *Nucleic Acids Res.* 47, D506–D515 (2019).
22. D. Aran, Z. Hu, A. J. Butte, xCell: digitally portraying the tissue cellular heterogeneity landscape, *Genome Biol.* 18, 220 (2017).
23. N. J. Edwards, M. Oberti, R. R. Thangudu, S. Cai, P. B. McGarvey, S. Jacob, S. Madhavan, K. A. Ketchum, The CPTAC Data Portal: A Resource for Cancer Proteomics Research, *J. Proteome Res.* 14, 2707–2713 (2015).
24. L. Vargiu, P. Rodriguez-Tomé, G. O. Sperber, M. Cadeddu, N. Grandi, V. Blikstad, E. Tramontano, J. Blomberg, Classification and characterization of human endogenous retroviruses; mosaic forms are common, *Retrovirology* 13, 7 (2016).
25. R. Tibshirani, Regression Shrinkage and Selection via the Lasso, *J. R. Stat. Soc.* 58, 267–288 (1996).



26. T. U. Consortium, UniProt: a worldwide hub of protein knowledge, *Nucleic Acids Res.* 47, D506–D515 (2019).
27. F. Loayza-Puch, K. Rooijers, L. C. M. Buil, J. Zijlstra, J. F. Oude Vrielink, R. Lopes, A. P. Ugalde, P. van Breugel, I. Hofland, J. Wesseling, O. van Tellingen, A. Bex, R. Agami, Tumour-specific proline vulnerability uncovered by differential ribosome codon reading, *Nature* 530, 490–494 (2016).
28. K. E. Varley, J. Gertz, B. S. Roberts, N. S. Davis, K. M. Bowling, M. K. Kirby, A. S. Nesmith, P. G. Oliver, W. E. Grizzle, A. Forero, D. J. Buchsbaum, A. F. LoBuglio, R. M. Myers, Recurrent read-through fusion transcripts in breast cancer, *Breast Cancer Res. Treat.* 146, 287–297 (2014).
29. S. Bobisse, R. Genolet, A. Roberti, J. L. Tanyi, J. Racle, B. J. Stevenson, C. Iseli, A. Michel, M.-A. L. Bitoux, P. Guillaume, J. Schmidt, V. Bianchi, D. Dangaj, C. Fenwick, L. Derré, I. Xenarios, O. Michielin, P. Romero, D. S. Monos, V. Zoete, D. Gfeller, L. E. Kandalaft, G. Coukos, A. Harari, Sensitive and frequent identification of high avidity neo-epitope specific CD8<sup>+</sup> T cells in immunotherapy-naive ovarian cancer, *Nat. Commun.* 9, 1–10 (2018).
30. M. S. Rooney, S. A. Shukla, C. J. Wu, G. Getz, N. Hacohen, Molecular and genetic properties of tumors associated with local immune cytolytic activity, *Cell* 160, 48–61 (2015).
31. K. B. Chiappinelli, P. L. Strissel, A. Desrichard, H. Li, C. Henke, B. Akman, A. Hein, N. S. Rote, L. M. Cope, A. Snyder, V. Makarov, S. Budhu, S. Buhu, D. J. Slamon, J. D. Wolchok, D. M. Pardoll, M. W. Beckmann, C. A. Zahnow, T. Merghoub, T. Mergoub, T. A. Chan, S. B. Baylin, R. Strick, Inhibiting DNA Methylation Causes an Interferon Response in Cancer via dsRNA Including Endogenous Retroviruses, *Cell* 162, 974–986 (2015).
32. D. Roulois, H. Loo Yau, R. Singhanian, Y. Wang, A. Danesh, S. Y. Shen, H. Han, G. Liang, P. A. Jones, T. J. Pugh, C. O'Brien, D. D. De Carvalho, DNA-Demethylating Agents Target Colorectal Cancer Cells by Inducing Viral Mimicry by Endogenous Transcripts, *Cell* 162, 961–973 (2015).
33. C. Alanio, F. Lemaitre, H. K. W. Law, M. Hasan, M. L. Albert, Enumeration of human antigen-specific naive CD8<sup>+</sup> T cells reveals conserved precursor frequencies, *Blood* 115, 3718–3725 (2010).
34. M. R. Corces, J. D. Buenrostro, B. Wu, P. G. Greenside, S. M. Chan, J. L. Koenig, M. P. Snyder, J. K. Pritchard, A. Kundaje, W. J. Greenleaf, R. Majeti, H. Y. Chang, Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution, *Nat. Genet.* 48, 1193–1203 (2016).
35. J.-D. Larouche, A. Trofimov, L. Hesnard, G. Ehx, Q. Zhao, K. Vincent, C. Durette, P. Gendron, J.-P. Laverdure, É. Bonneil, C. Côté, S. Lemieux, P. Thibault, C. Perreault, Widespread and tissue-specific expression of endogenous retroelements in human somatic tissues, *Genome Med.* 12, 1–16 (2020).
36. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2, *Nat. Methods* 9, 357–359 (2012).

37. A. Dobin, C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, T. R. Gingeras, STAR: ultrafast universal RNA-seq aligner, *Bioinformatics* 29, 15–21 (2013).
38. H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, 1000 Genome Project Data Processing Subgroup, The Sequence Alignment/Map format and SAMtools, *Bioinformatics* 25, 2078–2079 (2009).
39. R. Patro, G. Duggal, M. I. Love, R. A. Irizarry, C. Kingsford, Salmon: fast and bias-aware quantification of transcript expression using dual-phase inference, *Nat. Methods* 14, 417–419 (2017).
40. S. Hänzelmann, R. Castelo, J. Guinney, GSEA: gene set variation analysis for microarray and RNA-Seq data, *BMC Bioinformatics* 14, 7 (2013).
41. V. Thorsson, D. L. Gibbs, S. D. Brown, D. Wolf, D. S. Bortone, T.-H. Ou Yang, E. Portapardo, G. F. Gao, C. L. Plaisier, J. A. Eddy, E. Ziv, A. C. Culhane, E. O. Paull, I. K. A. Sivakumar, A. J. Gentles, R. Malhotra, F. Farshidfar, A. Colaprico, J. S. Parker, L. E. Mose, N. S. Vo, J. Liu, Y. Liu, J. Rader, V. Dhankani, S. M. Reynolds, R. Bowlby, A. Califano, A. D. Cherniack, D. Anastassiou, D. Bedognetti, Y. Mokrab, A. M. Newman, A. Rao, K. Chen, A. Krasnitz, H. Hu, T. M. Malta, H. Noushmehr, C. S. Pedamallu, S. Bullman, A. I. Ojesina, A. Lamb, W. Zhou, H. Shen, T. K. Choueiri, J. N. Weinstein, J. Guinney, J. Saltz, R. A. Holt, C. S. Rabkin, Cancer Genome Atlas Research Network, A. J. Lazar, J. S. Serody, E. G. Demicco, M. L. Disis, B. G. Vincent, I. Shmulevich, The Immune Landscape of Cancer, *Immunity* 48, 812–830.e14 (2018).
42. J. H. Friedman, T. Hastie, R. Tibshirani, Regularization Paths for Generalized Linear Models via Coordinate Descent, *J. Stat. Softw.* 33, 1–22 (2010).
43. M. Sill, T. Hielscher, N. Becker, M. Zucknick, c060: Extended Inference with Lasso and Elastic-Net Regularized Cox and Generalized Linear Models, *J. Stat. Softw.* 62, 1–22 (2014).
44. F. Madeira, Y. M. Park, J. Lee, N. Buso, T. Gur, N. Madhusoodanan, P. Basutkar, A. R. N. Tivey, S. C. Potter, R. D. Finn, R. Lopez, The EMBL-EBI search and sequence analysis tools APIs in 2019, *Nucleic Acids Res.* 47, W636–W641 (2019).
45. T. J. O'Donnell, A. Rubinsteyn, M. Bonsack, A. B. Riemer, U. Laserson, J. Hammerbacher, MHCflurry: Open-Source Class I MHC Binding Affinity Prediction, *Cell Syst.* 7, 129–132.e4 (2018).
46. Cancer Genome Atlas Network, Comprehensive molecular portraits of human breast tumours, *Nature* 490, 61–70 (2012).
47. K. Krug, E. J. Jaehnig, S. Satpathy, L. Blumenberg, A. Karpova, M. Anurag, G. Miles, P. Mertins, Y. Geffen, L. C. Tang, D. I. Heiman, S. Cao, Y. E. Maruvka, J. T. Lei, C. Huang, R. B. Kothadia, A. Colaprico, C. Birger, J. Wang, Y. Dou, B. Wen, Z. Shi, Y. Liao, M. Wiznerowicz, M. A. Wyczalkowski, X. S. Chen, J. J. Kennedy, A. G. Paulovich, M. Thiagarajan, C. R. Kinsinger, T. Hiltke, E. S. Boja, M. Mesri, A. I. Robles, H. Rodriguez, T. F. Westbrook, L. Ding, G. Getz, K. R. Clauser, D. Fenyö, K. V. Ruggles, B. Zhang, D. R. Mani, S. A. Carr, M. J. Ellis, M. A. Gillette, S. C. Avanessian, S. Cai, D. Chan, X. Chen, N. J. Edwards, A. N. Hoofnagle, M. H. Kane, K. A. Ketchum, E. Kuhn, D. A. Levine, S. Li, D. C. Liebler, T. Liu, J. Luo, S. Madhavan, C. Maher, J.

- E. McDermott, P. B. McGarvey, M. Oberti, A. Pandey, S. H. Payne, D. F. Ransohoff, R. C. Rivers, K. D. Rodland, P. Rudnick, M. E. Sanders, K. M. Shaw, I.-M. Shih, R. J. C. Slebos, R. D. Smith, M. Snyder, S. E. Stein, D. L. Tabb, R. R. Thangudu, S. Thomas, Y. Wang, F. M. White, J. R. Whiteaker, G. A. Whiteley, H. Zhang, Z. Zhang, Y. Zhao, H. Zhu, L. J. Zimmerman, Proteogenomic Landscape of Breast Cancer Tumorigenesis and Targeted Therapy, *Cell* 183, 1436-1456.e31 (2020).
48. R. Adusumilli, P. Mallick, Data Conversion with ProteoWizard msConvert, *Methods Mol. Biol.* 1550, 339–368 (2017).
49. B. Wen, X. Wang, B. Zhang, PepQuery enables fast, accurate, and convenient proteomic validation of novel genomic alterations, *Genome Res.* 29, 485–493 (2019).
50. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2, *Genome Biol.* 15, 550 (2014).
51. A. Zhu, J. G. Ibrahim, M. I. Love, Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences, *Bioinformatics* 35, 2084–2092 (2019).
52. A. Gros, E. Tran, M. R. Parkhurst, S. Ilyas, A. Pasetto, E. M. Groh, P. F. Robbins, R. Yossef, A. Garcia-Garijo, C. A. Fajardo, T. D. Prickett, L. Jia, J. J. Gartner, S. Ray, L. Ngo, J. R. Wunderlich, J. C. Yang, S. A. Rosenberg, Recognition of human gastrointestinal cancer neoantigens by circulating PD-1+ lymphocytes, *J. Clin. Invest.* 129, 4992–5004 (2019).
53. K. K. Jensen, V. Rantos, E. C. Jappe, T. H. Olsen, M. C. Jespersen, V. Jurtz, L. E. Jessen, E. Lanzarotti, S. Mahajan, B. Peters, M. Nielsen, P. Marcatili, TCRpMHCmodels: Structural modelling of TCR-pMHC class I complexes, *Sci. Rep.* 9, 14530 (2019).
54. K. R. Abhinandan, A. C. R. Martin, Analysis and improvements to Kabat and structurally correct numbering of antibody variable domains, *Mol. Immunol.* 45, 3832–3839 (2008).
55. M. S. Klausen, M. V. Anderson, M. C. Jespersen, M. Nielsen, P. Marcatili, LYRA, a webserver for lymphocyte receptor structural modeling, *Nucleic Acids Res.* 43, W349-355 (2015).
56. A. Fiser, A. Sali, Modeller: generation and refinement of homology-based protein structure models, *Methods Enzymol.* 374, 461–491 (2003).
57. A. Sali, T. L. Blundell, Comparative protein modelling by satisfaction of spatial restraints, *J. Mol. Biol.* 234, 779–815 (1993).
58. G. Launay, M. Ohue, J. Prieto Santero, Y. Matsuzaki, C. Hilpert, N. Uchikoga, T. Hayashi, J. Martin, Evaluation of CONSRANK-Like Scoring Functions for Rescoring Ensembles of Protein-Protein Docking Poses, *Front. Mol. Biosci.* 7, 559005 (2020).
59. R. Oliva, A. Vangone, L. Cavallo, Ranking multiple docking solutions based on the conservation of inter-residue contacts, *Proteins* 81, 1571–1584 (2013).
60. J. Hennecke, D. C. Wiley, T cell receptor-MHC interactions up close, *Cell* 104, 1–4 (2001).
61. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, UCSF Chimera--a visualization system for exploratory research and analysis, *J. Comput. Chem.* 25, 1605–1612 (2004).

62. A. Vangone, A. M. Bonvin, Contacts-based prediction of binding affinity in protein-protein complexes, *eLife* 4, e07454 (2015).
63. E. Driehuis, K. Kretzschmar, H. Clevers, Establishment of patient-derived cancer organoids for drug-screening applications, *Nat. Protoc.* 15, 3380–3409 (2020).

**Acknowledgments:** We would like to thank Benjamin Vincent, Christof Smith and Jason Garness for sharing data and fruitful exchanges. We also thank Gabriel Ichim (Centre de Recherche en Cancérologie de Lyon) for his help with the setup of the Nanolive experiments. We thank Brigitte Manship for manuscript editing.

**Funding:** This work was financially supported by the PRT-K grant for research in translational medicine for oncology, the Ligue Nationale contre le Cancer, the LabEx DEVweCAN and LYriCAN.

**Author contributions:**

**P. Bonaventura & V. Alcazer:** Conceptualization, data curation, formal analysis, funding acquisition, methodology, investigation, resources, software, validation, visualization, writing-original draft. **V. Mutez, N. Chuvin, Y. Estornes, E. Michel, R. Boulos:** Methodology, resources, investigation, validation, visualization, writing – review & editing. **A. Viari:** Resources, project administration, **J. Valladeau-Guilemond:** Resources, writing – review & editing. **C. Caux:** Methodology, resources, writing – review & editing. **S. Depil:** Conceptualization, funding acquisition, methodology, resources, validation, project administration, writing – review & editing

**Competing interests:**

SD is founder of ErVaccine Technologies. VM, NC, EM, RB and YE are employees of ErVaccine Technologies. PB and VA are consultant for Ervaccine Technologies. The other authors have no conflict of interest to declare.

**Data and materials availability:** All data used in this article are available upon request to the corresponding author.

<i>herv_id</i>	baseMean	Log2 FC	FC	padj	P1	P2	P3	P4	P5	P6
<i>herv_2953</i>	1.42	3	7.99	6.14E-17						x
<i>herv_4833</i>	507.59	2.4	5.28	2.84E-128	x	x				x
<i>herv_3232</i>	144.11	2.01	4.02	2.33E-29						x

<i>herv_4873</i>	32.5	1.56	2.94	4.04E-25			x	x	x	x
<i>herv_6069</i>	677.18	1.31	2.48	3.27E-19	x	x	x	x	x	x
<i>herv_2025</i>	8.18	1.21	2.32	1.05E-11	x	x	x	x	x	x
<i>herv_6079</i>	35.73	1.16	2.23	2.28E-30		x	x	x	x	x
<i>herv_2704</i>	28.63	1.1	2.14	3.63E-16	x	x	x	x	x	x
<i>herv_1741</i>	24.46	0.93	1.9	1.70E-14			x	x		x
<i>herv_3192</i>	15.98	0.87	1.82	5.01E-22	x	x		x		x
<i>herv_3288</i>	24.52	0.83	1.78	1.92E-18						x
<i>herv_2794</i>	8.4	0.77	1.7	1.49E-13					x	
<i>herv_2288</i>	32.98	0.71	1.64	1.79E-08						x
<i>herv_2582</i>	8.63	0.69	1.61	9.33E-09	x	x	x	x	x	x
<i>herv_4679</i>	157.18	0.69	1.61	9.88E-29						x
<i>herv_2476</i>	415.1	0.64	1.56	1.64E-16						x
<i>herv_4695</i>	30.94	0.6	1.51	4.43E-11	x	x	x	x	x	x
<i>herv_3652</i>	74.36	0.6	1.51	1.61E-13	x	x	x		x	x

**Table 1. Peptide-containing HERVs overexpressed in the basal breast cancer subtype.**

Previously identified peptides 1 to 6 have been located among HERVs overexpressed in the basal breast cancer subtype in TCGA patients. Log<sub>2</sub> fold changes have been calculated and shrunk using DESEQ2.

## SUPPLEMENTAL DATA

### Materials and Methods

Fig. S1. Differential HERV expression in breast cancer subtypes (Varley's database) and peptide mass spectra

Fig. S2. TCGA pancancer flow-chart and methylation analysis

Fig. S3. TCGA basal breast-cancer cyt-HERVs and peptide analysis

Fig. S4 HERV-derived epitopes CD8<sup>+</sup> T cell responses and gating strategies

Fig. S5. Dextramer staining after feeder expansion and analysis of 3D structural models

Fig. S6 Flow cytometry assays on T cells co-cultured with tumor cells

Fig. S7 HERV expression in TNBC patients and organoids cell surface markers

Table S1. Blast results of the 6 HERV peptides aligned against the human proteome.

Table S2: Number of samples from each cancer type included in the study

Supplementary videos 1, 2 and 3: 3D microscopy imaging using Nanolive technology : P1-specific T cells co-cultured with MDA-MB-231 (video 1) or HMECs (video 2) and dextramer-negative non-specific T cells co-cultured with MDA-MB-231 (video 3).



## **SUPPLEMENTARY METHODS**

### **Methylation analysis**

Pre-processed TCGA pancancer Beta-value from Illumina 450K methylation data were accessed from the GDC portal from the PanCanAtlas description page. Riboseq raw fastq files were accessed from the GEO portal under the accession number GSE69923 (27). Samples under glutamine starvation conditions were omitted from the analysis.

For each cyt-HERV, a local methylation status was calculated according to the mean beta-value of the 10 nearest surrounding probes. Spearman's correlations were then calculated independently for each cyt-HERV. HERVs with a Spearman's  $R < -0.3$  and an adjusted  $p < 0.05$  were considered to be associated with a local demethylation status, and HERVs with a Spearman's  $R > 0.3$  and an adjusted  $p < 0.05$  with a local methylation status.

### **Binding affinity analysis**

The 6 selected and control peptides were synthesized to perform a high throughput quantitative binding assay to MHC allele HLA-A2. ProImmune's Class I REVEAL® Rapid Epitope Discovery System detection is based on the presence or absence of the native conformation of the MHC-peptide complex. Each peptide is given a score relative to the positive control peptide, which is a known T cell epitope. The score is reported as a percentage of the signal generated by the test peptide versus the positive control peptide. REVEAL® on- and off-rate assays give a quantitative picture of the binding properties of individual peptides. On and off-rates for peptides that have passed the MHC-binding assay are measured at six points over 48 hours (on-rate) and 24 hours (off-rate). Results are presented as half-life values. Peptides are given a kinetic score based on the results of both rates; a higher score indicates a better candidate T cell epitope. Finally the 'R' score is defined for each peptide. This incorporates the kinetic score with the MHC-peptide binding assay score; the higher the R score, the better the epitope.

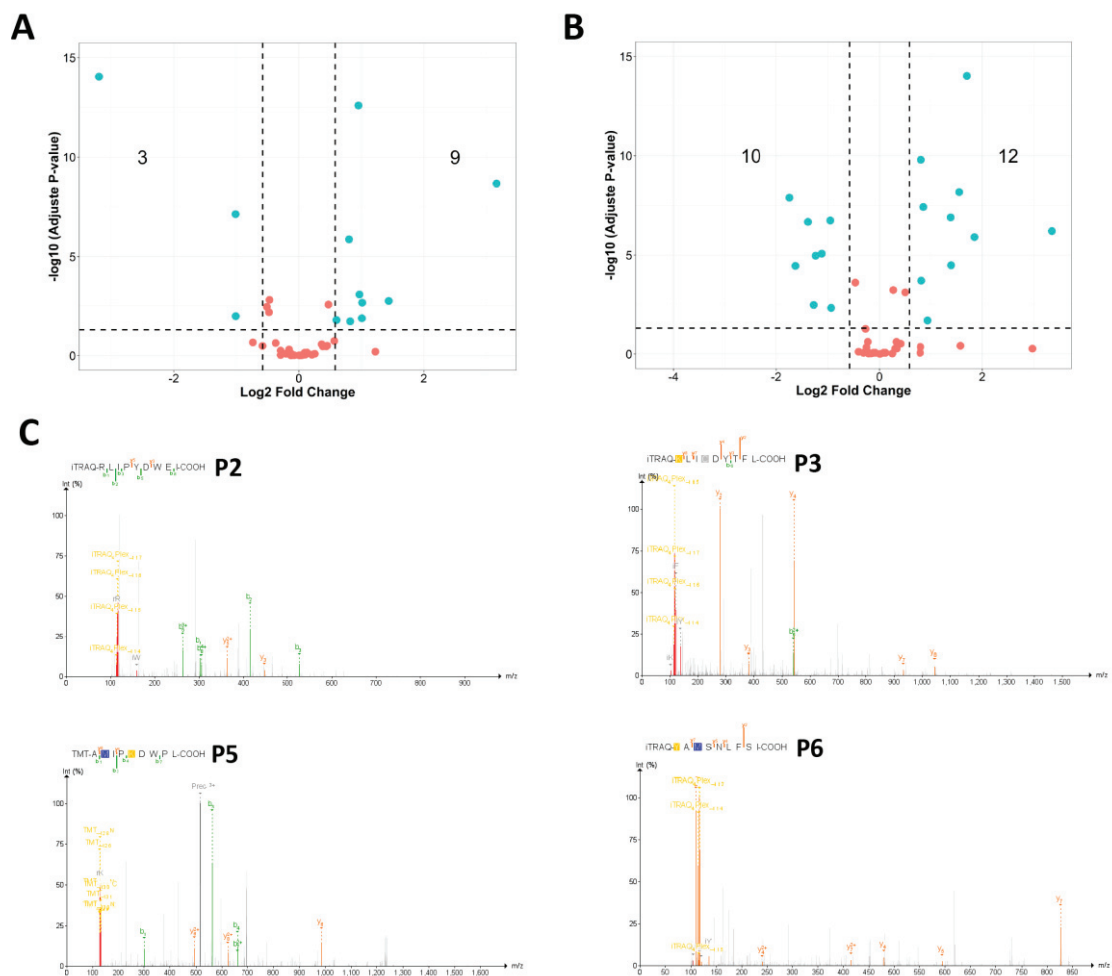
### **Dendritic cell-based priming assay**

Monocytes were isolated from PBMCs by positive selection of CD14<sup>+</sup> cells (Myltenyi, GE, EU). The negative fraction was considered as peripheral blood lymphocytes (PBLs). Cells were frozen in FBS 10% DMSO and kept at -80°C. Monocyte-derived dendritic cells (MoDCs) were generated from 6-day cultures of CD14<sup>+</sup> monocytes in complete RPMI medium supplemented with 10 % FCS and recombinant human GM-CSF (100 ng/mL) and IL-4 (50 ng/mL). MoDCs were pulsed with peptides overnight (10µg/mL) with 10 ng/ml of LPS and washed before co-

culture with isolated PBLs for 6 days (MoDCs:T-cells ratio 1:10) in 96 round wells plate. After 6 days T cells were counted and restimulated with autologous MoDCs for 6 more days.

SUPPLEMENTARY FIGURE

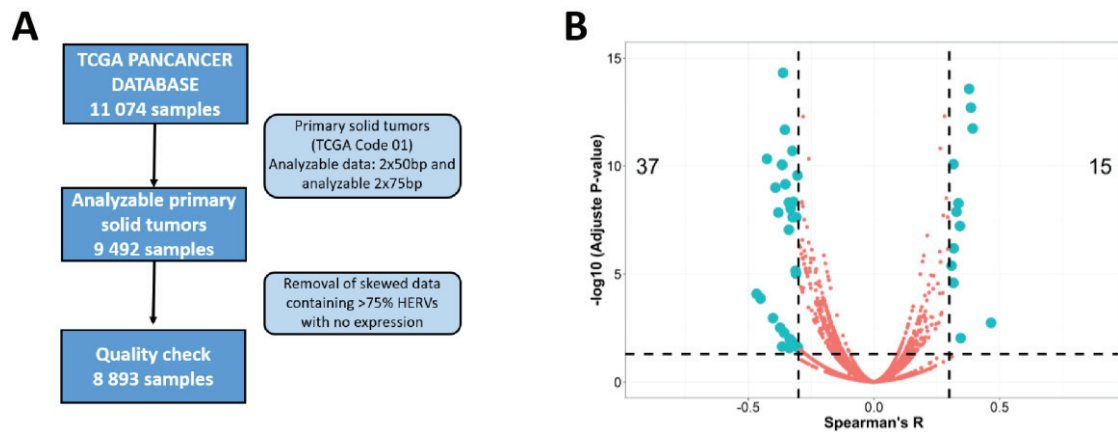
Supplementary figure 1



Supp Fig. 1. Differential HERVs expression in breast cancer subtypes (Varley’s database) and peptide mass spectra

- (A) Volcano plot of HERV expression in Estrogen Receptor positive (ER+) breast cancer samples versus pooled peritumoral and normal breast tissue (Varley’s dataset).
- (B) Volcano plot of HERV expression in triple negative breast cancer (TNBC) samples versus pooled peritumoral and normal breast tissue (Varley’s dataset).
- (C) MS/MS detection of P2, P3, P5 and P6 epitopes in samples from the CPTAC or TCGA breast cancer prospective dataset. MS/MS spectrum is identified by Pepquery analysis (Peptide Spectrum Match, p values 0.0010, 0.0010, 0.0010 and 0.0031, respectively)

## Supplementary figure 2



Supp Fig. 2. TCGA pancancer flow-chart and methylation analysis

- (A)** Flow-chart of included samples for the pancancer study. Only primary solid tumors samples were included. Studies with insufficient data quality were excluded, leading to the final analysis of 8,893 primary tumor samples from 29 different cancer types.
- (B)** Correlation between *cyt*-HERVs expression and local methylation state in TCGA pancancer dataset. For each HERV, the mean beta-coefficient of the 10 nearest methylation probes was calculated. Spearman's correlation between each individual HERV in each cancer type and its median surrounding beta-coefficient is shown.

*Cyt*-HERVs: HERVs associated with cytotoxic response, TCGA: The Cancer Genome Atlas



- (C)** Mean expression of the 54 cyt-HERVs from TCGA basal subtypes is represented in cell line samples from Varley's independent dataset.
- (D)** Mean expression of the 54 cyt-HERVs from TCGA basal subtypes in TNBC cell line samples from the CCLE database.
- (E)** Number of different HERVs containing the top 25 peptides. Affinity is colored on a red-blue scale.

*Cyt-HERVs: HERVs associated with cytotoxic response, ORFs: Open Reading Frames, TCGA: The Cancer Genome Atlas, TNBC: Triple-Negative Breast Cancer*

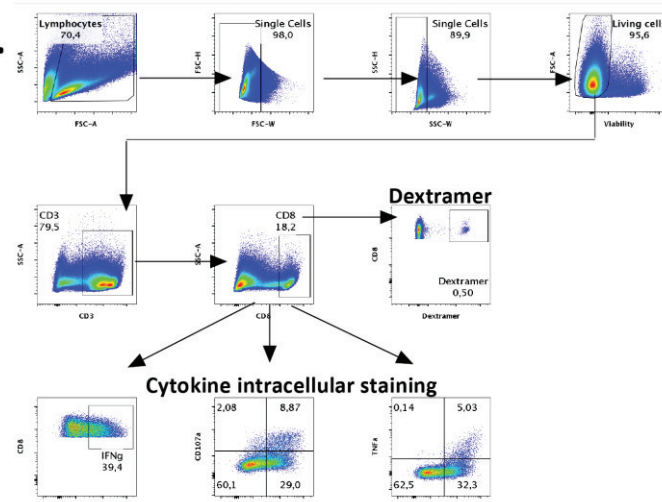


## Supplementary figure 4

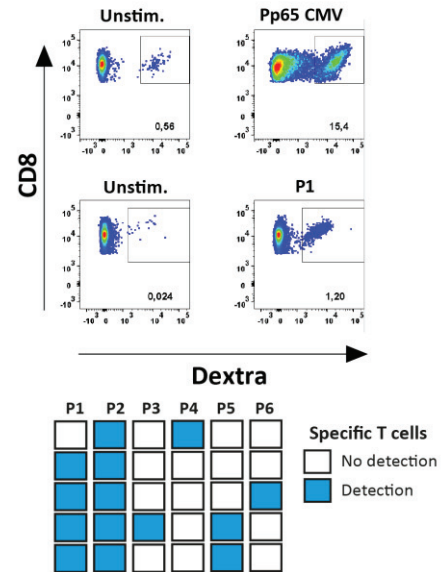
A

Peptide ID	Sequence	REVEAL® score	On-rate T1/2 (h)	Off-rate T1/2 (h)	Kinetic score	R-score
1	FLQFKTWWI	88.70	8.35	12.22	1.46	1.30
2	RLIPYDWEI	45.11	5.78	20.49	3.55	1.60
3	KLIDCYTFL	80.30	11.39	40.61	3.57	2.86
4	YLSFKILL	58.96	13.18	20.76	1.58	0.93
5	AMIPKDWPL	77.53	7.59	35.08	4.62	3.58
6	YAMSNLFSI	55.28	5.42	46.95	8.67	4.79
Positive control	-	100.00	11.83 +/- 3.00	85.58 +/- 34.42	7.21 +/- 0.48	7.24

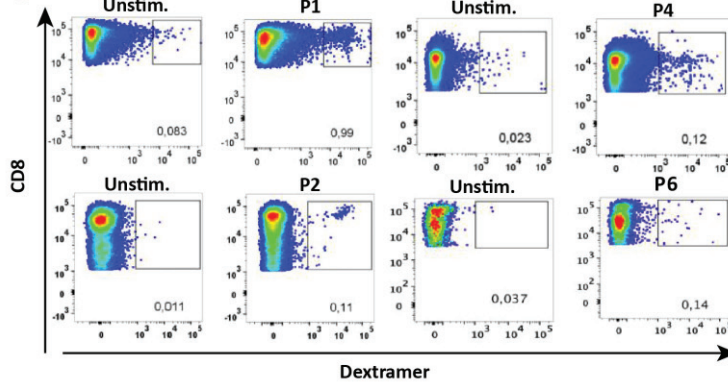
B



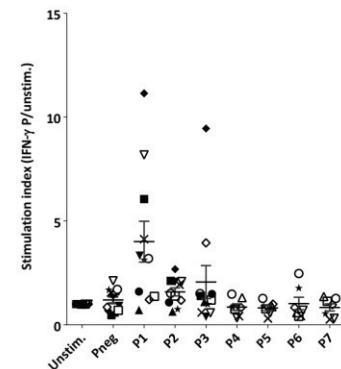
D



C



E



### Supp Fig. 4: HERV-derived epitopes CD8<sup>+</sup> T cell responses and gating strategies

(A) Summary table of the binding, rate data and kinetic scores for the 6 peptides. Reveal scores  $\geq 45\%$  indicate good binders (see supplementary methods for the definitions).

(B) Gating strategy for the analysis of flow cytometry data for dextramer staining protocol and polychromatic intracellular cytokine (IFN- $\gamma$  and TNF- $\alpha$ ) and CD107a CD8<sup>+</sup> T cells.

Gates on lymphocytes morphology, single cells in SSC and FSC, CD3<sup>+</sup> T cells, CD8<sup>+</sup> T

cells to identify double CD8<sup>+</sup>dextramer<sup>+</sup> cells. The same gating strategy was used to identify IFN- $\gamma$ <sup>+</sup> TNF- $\alpha$ <sup>+</sup> CD107a<sup>+</sup> CD8<sup>+</sup> T cells.

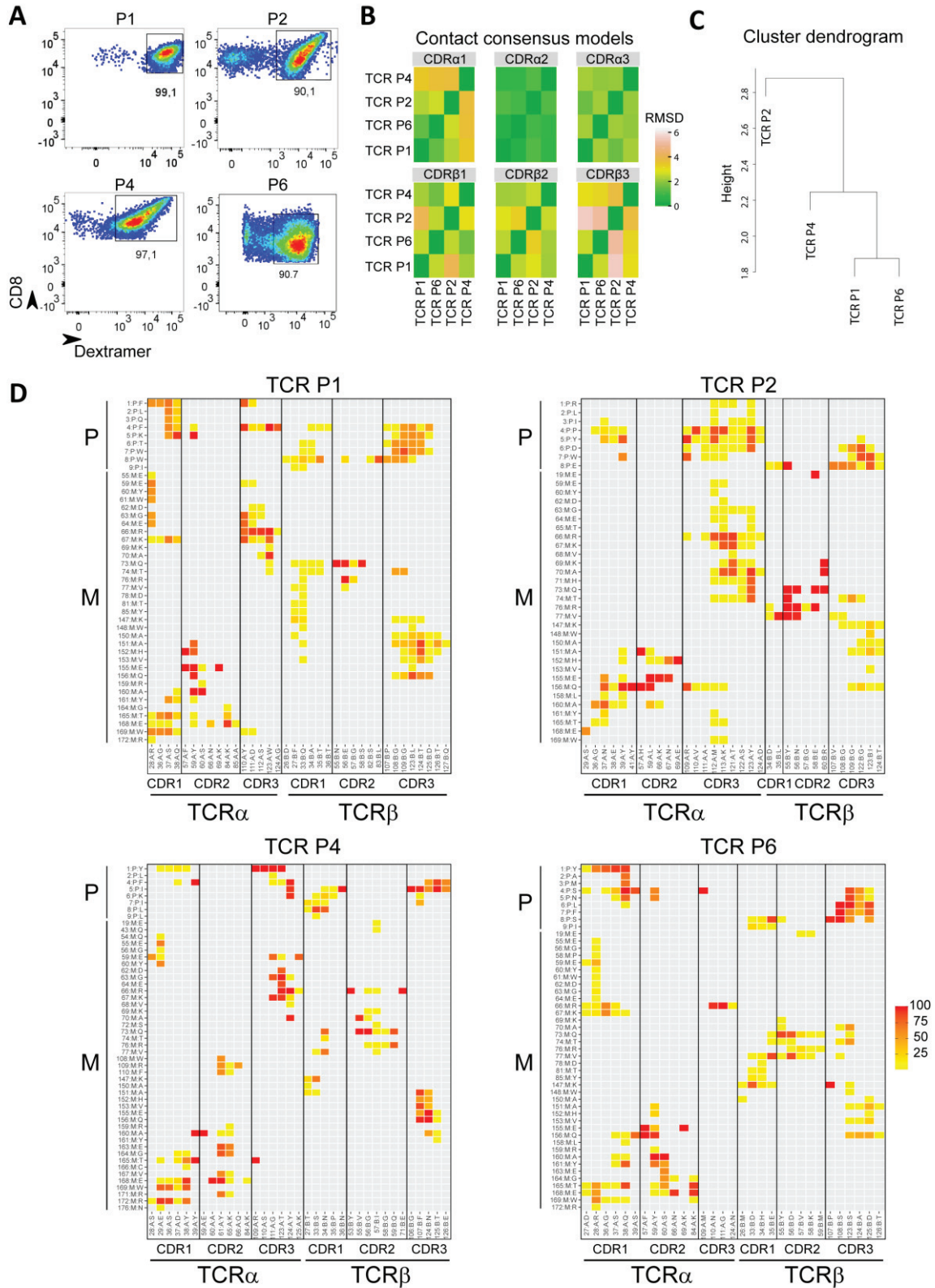
**(C)** Representative panels of dextramer staining in unstimulated (left panel for each peptide) and peptide-stimulated (right panel for each peptide) CD8<sup>+</sup> T cells for different peptides (P1, P2, P4, P6) after stimulation with peptides and TLR ligands using healthy donors PBMCs

**(D)** Representative panels of dextramer staining in unstimulated (left panels) and peptide-stimulated (right panels) CD8<sup>+</sup> T cells using monocyte-derived dendritic cells. Down summary table of the results obtained in 5 HLA-A2-positive healthy donors: blue squares indicate detection of epitope-specific CD8<sup>+</sup> T cells; white squares indicate a lack of detection

**(E)** Summary (individual values and median) of IFN- $\gamma$  stimulation index in 11 healthy donors PBMCs after stimulation with peptides and TLR ligands. IFN- $\gamma$  stimulation index is defined as the ratio between specific IFN- $\gamma$ <sup>+</sup>CD8<sup>+</sup> T cells (peptide stimulated) and unspecific IFN- $\gamma$ <sup>+</sup>CD8<sup>+</sup> T cells (non-stimulated with the peptide).

*IFN: Interferon, PBMCs: Peripheral Blood Mononuclear Cells, TLR: Toll-Like Receptor, TNF: Tumor Necrosis Factor*

# Supplementary figure 5



Supp Fig. 5: dexramer staining after feeder expansion and analysis of 3D structural models

**(A)** Representative plots of dextramer staining of specific CD8<sup>+</sup> T cells after feeder expansion.

Dextramer staining was between 90 and 99% for TCR sequencing.

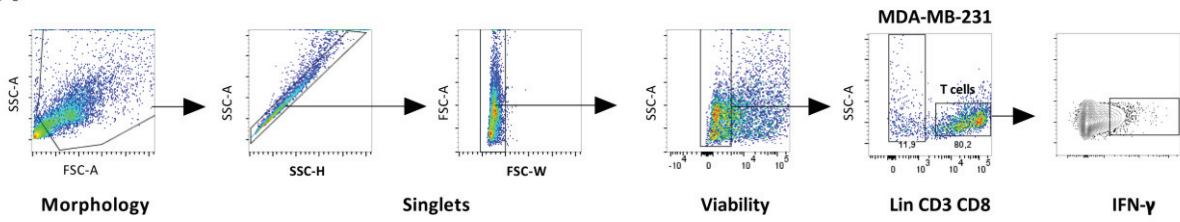
**(B)** RMSD matrix between CDR loop conformations on the refined representative models. The RMSD is averaged over the 6 CDR regions with equal weights.

**(C):** Clustering tree obtained from the RMSD matrix in B. High structural similarity is observed between TCR P1 and TCR P6 CDR loops, which share the same TCR alpha sequences. On the contrary, TCR P4 and TCR P2 CDR loops are structurally more distant.

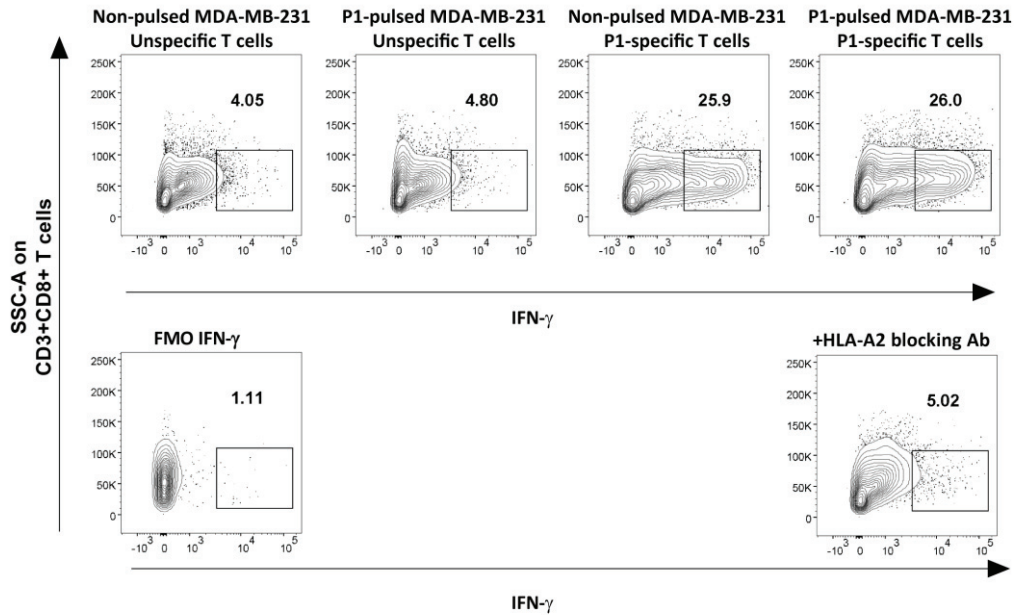
**(D)** Frequencies of inter-residue contacts obtained from the subsets of 25\*4 models generated during the refinement protocol, that were used to identify the representative models.

## Supplementary figure 6

**A**



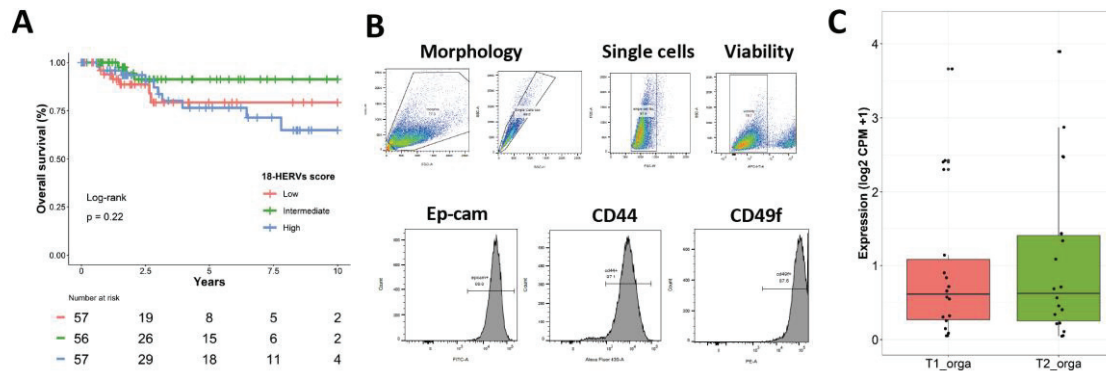
**B**



### Supp Fig. 6. Flow cytometry assays on T cells co-cultured with tumor cells

- (A) Gating strategy and representative plots for FACS analysis of IFN- $\gamma$  produced by T cell co-cultured with MDA-MB-231
- (B) Representative plots of IFN- $\gamma$  analysis of CD8<sup>+</sup> specific T cells co-cultured with MDA-MB-231 in the different culture conditions.

## Supplementary figure 7



### Supp Fig. 7. HERV expression in TNBC patients and organoids cell surface markers

**(A)** Overall survival according to 18-HERVs score in TCGA TNBC patients (all HLA subtypes, n=170). Patients were divided in three groups: blue line, high expression (n=57); green line, intermediate expression (n=56); red line, low expression (n=57)

**(B)** Flow cytometry analysis of organoid-constituting cells including Ep-cam, CD44 and CD49f epithelial markers

**(C)** Boxplot of mean expression of the 18 Cyt-HERVs expression in organoids RNAseq at early (T1) or late (T2) culture passage. Mean expression is represented in log<sub>2</sub> CPM +1.

*HLA: Human Leukocyte Antigen, TCGA: The Cancer Genome Atlas, TNBC: Triple-Negative Breast Cancer*



## ***SUPPLEMENTARY TABLES***

**Table S1#:** Blast results of the 6 HERVs peptides aligned against the human proteome.

**Table S2#:** Number of samples from each cancer type included in the study

LAML, ESCA and STAD have been completely removed from the study due to skewed distribution. For normal tissue comparison, only studies with at least 10 peritumoral samples have been included.

*LAML: Acute Myeloid Leukemia, ESCA: Esophageal Carcinoma, STAD: Stomach Adenocarcinoma*





## **ARTICLE 2 - HERVs characterize the different subpopulations of normal and leukemic cells and represent a source of shared epitopes for cancer immunotherapy**

*Vincent Alcazer*<sup>1,2\*</sup>†, *Paola Bonaventura*<sup>2\*</sup>, *Laurie Tonon*<sup>3</sup>, *Emilie Michel*<sup>4</sup>, *Virginie Mutez*<sup>4</sup>, *Alain Viari*<sup>3</sup>, *Klaus H. Metzeler*<sup>5,6</sup>, *Wolfgang Hiddemann*<sup>5</sup>, *Aarif M. N. Batch*<sup>7,8</sup>, *Tobias Herold*<sup>6</sup>, *Christophe Caux*<sup>2</sup>, *Stéphane Depil*<sup>2,4,9,10</sup>†

1 Department of Hematology, Hospices Civils de Lyon, Lyon Sud Hospital, Pierre-Bénite, France.

2 Cancer Research Center of Lyon, INSERM U1052 and CNRS UMR5286, Lyon, France.

3 Synergie Lyon Cancer Foundation, Gilles Thomas Bioinformatics Center, Centre Léon Bérard, Lyon, France.

4 Ervaccine Technologies, Centre Leon Bérard, Lyon, France.

5 Laboratory for Leukemia Diagnostics, Department of Medicine III, University Hospital, LMU Munich, Munich, Germany

6 Dept. of Hematology and Cell Therapy, University of Leipzig, Leipzig, Germany

7 Institute of Medical Data Processing, Biometrics and Epidemiology (IBE), Faculty of Medicine, LMU Munich, Munich, Germany

8 Data Integration for Future Medicine (DiFuture, [www.difuture.de](http://www.difuture.de)), LMU Munich, Munich, Germany

9 Department of Clinical Research, Centre Léon Bérard, Lyon, France.

10 University Claude Bernard Lyon 1, Lyon, France.

\* These authors equally contributed to this work.

† **Junior corresponding author:** Dr. Vincent Alcazer, Department of Hematology, Hospices Civils de Lyon, Lyon Sud Hospital, 165 chemin du Grand Revoyet, 69310 Pierre-Bénite, France. [Vincent.alcazer@chu-lyon.fr](mailto:Vincent.alcazer@chu-lyon.fr)

† **Senior corresponding author:** Prof. Stéphane DEPIL, Centre Leon Bérard, 69008 Lyon, France. [Stephane.depil@lyon.unicancer.fr](mailto:Stephane.depil@lyon.unicancer.fr)

==== Article actuellement en revue dans *The Journal of Clinical Investigation* ====

## Résumé

Bien qu'intéressante pour rechercher des épitopes, l'approche *hervQuant* ne capte qu'une partie du transcriptome HERVs. Dans ce deuxième travail, nous avons voulu évaluer l'intérêt d'une approche plus globale de quantification des HERVs permettant à la fois de tirer profit de l'information fournie par l'ensemble du rétrotranscriptome HERVs tout en conservant la possibilité de retrouver des cibles thérapeutiques.

Nous avons quantifié l'expression des HERVs dans une base de données RNAseq de cellules triées issues de moelles osseuses normales et leucémiques en utilisant l'approche Telescope, permettant une quantification précise de l'ensemble du rétrotranscriptome HERV regroupé en 14 968 loci. D'une façon surprenante, le *clustering* hiérarchique non supervisé basé uniquement sur l'expression des HERVs nous a permis de classer correctement les différentes sous-populations cellulaires médullaires normales. Ce *clustering* est encore améliorée lorsque nous utilisons les zones de chromatine ouverte centrées sur les HERVs fournies par les données ATAC-seq. L'ajout des échantillons de cellules leucémiques montre un regroupement de ces dernières avec leurs cellules normales d'origine, permettant ainsi de définir des profils cellulaires différents au sein même des populations tumorales. Pour démontrer le potentiel d'utilisation des HERVs en tant que marqueur d'identité cellulaire, nous avons ensuite établi une signature de cellules souches leucémiques (LSC pour *Leukemic Stem Cell*) établie sur 25 HERVs fortement corrélé au score validé LSC17 dans 4 jeux de données indépendants. Cette signature permet de classer correctement les populations de LSC avec une AUC de 0,80, *versus* 0,76 pour la signature LSC17. Dans un second temps, nous avons quantifié l'expression des HERVs dans 788 échantillons de moelles osseuses de patients présentant une LAM au diagnostic issus de 4 cohortes indépendantes. La classification non supervisée basée uniquement sur les HERVs les plus variables et conservés entre les cohortes nous a permis de définir 9 clusters, indépendamment de l'étude ou du groupe de risque ELN2017. Ces 9 clusters présentent un pronostic différent, indépendamment des facteurs de risques connus en analyse multivariée, avec un fond biologique également varié comme montré par l'analyse des voies de signalisation liées au cancer.

Enfin, nous avons utilisé une base de contrôle de plus de 1 000 échantillons provenant de 42 tissus sains différents pour définir des HERVs exprimés de façon spécifique dans la LAM. Nous avons ainsi identifié 125 HERVs spécifiques de la LAM à partir desquels nous avons pu prioriser 104 épitopes LT CD8<sup>+</sup> sélectionnés sur des critères d'affinité pour le HLA-A\*02 et d'expression forte et partagée par différents HERVs. Parmi ces épitopes, nous avons pu retrouver 8 peptides pour lesquels nous avons des données d'immunogénicité chez le donneur sain. A partir d'échantillons de moelle osseuse de patients présentant une LAM au diagnostic, nous avons mis en expansion les lymphocytes infiltrant la moelle pendant 14 jours et recherché la présence de LT CD8<sup>+</sup> spécifiques dirigés contre ces 8 peptides via l'utilisation de dextramères. D'une façon intéressante, des réponses spécifiques dirigées contre P1 ont été retrouvées chez les 9 des 10 patients testés.

Ce travail montre que le rétrotranscriptome HERV peut être utilisé pour caractériser les populations cellulaires, ce qui peut être d'une importance notable pour la recherche de biomarqueurs ou l'établissement de signatures. Par ailleurs, des réponses LT CD8<sup>+</sup> dirigées contre des peptides exprimés par des HERVs spécifiques de la LAM peuvent être mise en évidence dans les moelles osseuses de patient au diagnostic, suggérant leur intérêt en tant que cibles thérapeutiques.

# HERVs characterize normal and leukemia stem cells and represent a source of shared epitopes for cancer immunotherapy

## Short title: HERVs as biomarkers and source of epitopes in AML

*Vincent Alcazer<sup>1,2\*</sup>+, Paola Bonaventura<sup>2\*</sup>, Laurie Tonon<sup>3</sup>, Emilie Michel<sup>4</sup>, Virginie Mutez<sup>4</sup>, Nicolas Chavin<sup>4</sup>, Yann Estornes<sup>4</sup>, Alain Viari<sup>3</sup>, Klaus H. Metzeler<sup>5,6</sup>, Wolfgang Hiddemann<sup>5</sup>, Aarif M. N. Batch<sup>7,8</sup>, Tobias Herold<sup>5</sup>, Christophe Caux<sup>2</sup>, Stéphane Depil<sup>2,4,9,10</sup>+*

### Affiliations:

1 Department of Hematology, Hospices Civils de Lyon, Lyon Sud Hospital, Pierre-Bénite, France.

2 Cancer Research Center of Lyon, INSERM U1052 and CNRS UMR5286, Lyon, France.

3 Synergie Lyon Cancer Foundation, Gilles Thomas Bioinformatics Center, Centre Léon Bérard, Lyon, France.

4 Ervaccine Technologies, Centre Leon Bérard, Lyon, France.

5 Laboratory for Leukemia Diagnostics, Department of Medicine III, University Hospital, LMU Munich, Munich, Germany

6 Dept. of Hematology and Cell Therapy, University of Leipzig, Leipzig, Germany

7 Institute of Medical Data Processing, Biometrics and Epidemiology (IBE), Faculty of Medicine, LMU Munich, Munich, Germany

8 Data Integration for Future Medicine (DiFuture, [www.difuture.de](http://www.difuture.de)), LMU Munich, Munich, Germany

9 Department of Clinical Research, Centre Léon Bérard, Lyon, France.

10 University Claude Bernard Lyon 1, Lyon, France.

**\* These authors equally contributed to this work.**

**+ Corresponding authors**

**Junior corresponding author:** Vincent Alcazer, Department of Hematology, Hospices Civils de Lyon, Lyon Sud Hospital, Pierre-Bénite, France. [Vincent.alcazer@chu-lyon.fr](mailto:Vincent.alcazer@chu-lyon.fr)

**Senior corresponding author:** Prof. Stéphane DEPIL, Centre Leon Bérard, Lyon, France. [Stephane.depil@lyon.unicancer.fr](mailto:Stephane.depil@lyon.unicancer.fr)



**Keywords:** Human Endogenous Retroviruses, Acute Myeloid Leukemia, Leukemic Stem Cells, Epitopes, Cancer immunotherapy

**Summary word count:** 150

**Main text word count:** 3,960

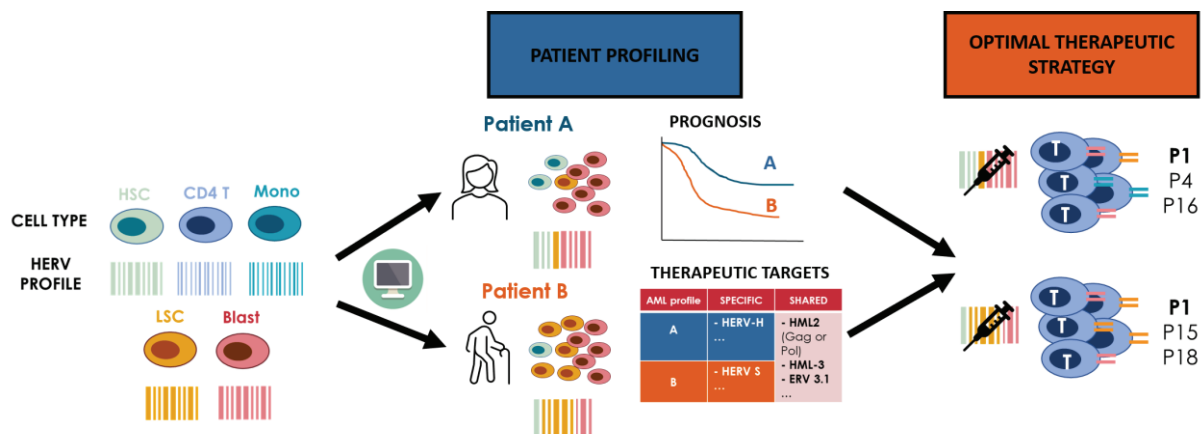
**Figures:** 5

**References:** 48

## BRIEF SUMMARY

- The HERV retrotranscriptome can be used to characterize leukemia stem cells or distinct AML subtypes
- HERVs represent a reservoir of non-conventional T cell epitopes in AML

## GRAPHICAL ABSTRACT



## ABSTRACT

Human endogenous retroviruses (HERVs) represent 8% of the human genome. The expression of HERVs and their immune impact have not been extensively studied in Acute Myeloid Leukemia (AML). In this study, we used a complete database of 14,968 HERV functional units to provide a thorough analysis of HERV expression in normal and AML bone marrow cells. We show that the HERV retrotranscriptome accurately characterizes normal and leukemic cell subpopulations, including leukemia stem cells, in line with different epigenetic profiles. We then show that HERV expression delineates AML subtypes with different prognoses. We finally propose a method to select and prioritize CD8<sup>+</sup> T cell epitopes derived from AML-specific HERVs and we show that lymphocytes infiltrating patient marrow at diagnosis contain naturally occurring CD8<sup>+</sup> T cells against these HERV epitopes. These results suggest that HERVs may represent promising biomarkers and an important reservoir of alternative tumor-specific T cell epitopes relevant for cancer immunotherapy.

## INTRODUCTION

Human endogenous retroviruses (HERVs) represent 8% of the human genome (1). These sequences are remnants of ancestral germline infections by exogenous retroviruses (2). The original sequence of a HERV is that of an exogenous retrovirus, with two promoter long-terminal repeat (LTR) sequences surrounding the virus open-reading frames (ORFs): gag, pro, pol and env (3). However, after millions of years of evolution, these ORFs have been deeply altered, and there is currently no description of any autonomous fully infectious HERV (4).

The long-standing belief is that HERVs are repressed by epigenetic mechanisms and are thus not expressed, or only poorly, in normal tissues (5). However, recent studies have shown that HERV expression can be detected in a vast range of normal tissues (6). Different pathological conditions can lead to aberrant HERV expression, as it has now been largely described in auto-immune diseases (7–10) and in cancers (11–14), where HERVs have been the subject of many studies over the last years. Indeed, it was reported that HERVs could participate in oncogenesis by inducing chromosomal instability, promoting aberrant gene expression with their LTR or by impacting the immune system with their RNA and protein products (15). HERVs could thus play a prominent role in cancer immunity, increasing tumor immunogenicity by promoting (i) an innate immune response triggered by the viral defense pathway induced by their nucleic acid intermediates, and (ii) an adaptive immune response by forming a pool of tumor-associated antigens (16).

Acute Myeloid Leukemia (AML) is a heterogeneous disease characterized by the clonal expansion of myeloid progenitor and stem cells (17). While some AML subtypes are characterized by recurrent genetic translocations or mutations associated with particular prognoses, most AMLs present a normal or complex karyotype, and identifying key factors that predict treatment resistance in these patients represents a major challenge (17, 18). Aside from disease stratification, AML also belongs to malignancies with the lowest mutational burdens (19), and finding tumor-specific antigens for immunotherapeutic approaches remains very difficult as the frequency of mutations creating neoantigens is expected to be low (20). In this context, HERV-derived antigens could represent a unique source of alternative tumor-specific antigens (i.e. antigens that do not originate from single nucleotide variation in a coding region (21)) that could be exploited for the development of new immunotherapies.

To date, little is known about the expression of HERVs in AML and its relevance as either a biomarker or a therapeutic target. Evidence of HERV-K /HML-2 expression in AML cells was shown as early as 1993 and confirmed in the early 2000s (22, 14). Few studies then focused on

HERVs in AML until the late 2010s, with the demonstration that azacytidine (Aza) activates the transcription of different HERVs, potentially contributing to its clinical effects (23). The exact role of HERVs in Aza therapy is however a matter of debate, with recent evidence arguing in favor of a HERV-independent therapeutic effect (24). More recently, a link was established between HERVs and the expression of surrounding genes in AML, suggesting a regulatory role of these retroelements (25). However, few data exist on HERV expression and their immune impact in AML, with studies relying on non-exhaustive quantification methods such as polymerase chain reaction (PCR) and focusing only on a few HERV loci (14) or globally quantifying HERVs at the family level together with transposable elements (26).

In this study, we thoroughly assessed HERV expression in AML and normal blood and bone marrow cells. Using a recent method to exhaustively quantify the HERV retrotranscriptome in next-generation sequencing data, we show that the latter can accurately define normal and leukemic cell populations, including leukemia stem cells (LSCs) that can be characterized by a 47-HERV signature. We also show that leukemic cells present a distinct epigenetic profile compared to their normal cell counterparts, with a significant correlation between the expression of HERVs located in open chromatin regions and surrounding cancer-associated genes. We then show that the HERV retrotranscriptome can be used to discriminate AML profiles from bulk RNA-seq data, distinguishing known but also new AML subtypes of different prognoses. Finally, we show that HERVs specifically expressed in AML cells represent a reservoir of T cell epitopes able to elicit specific immune responses in AML patients.

## **RESULTS**

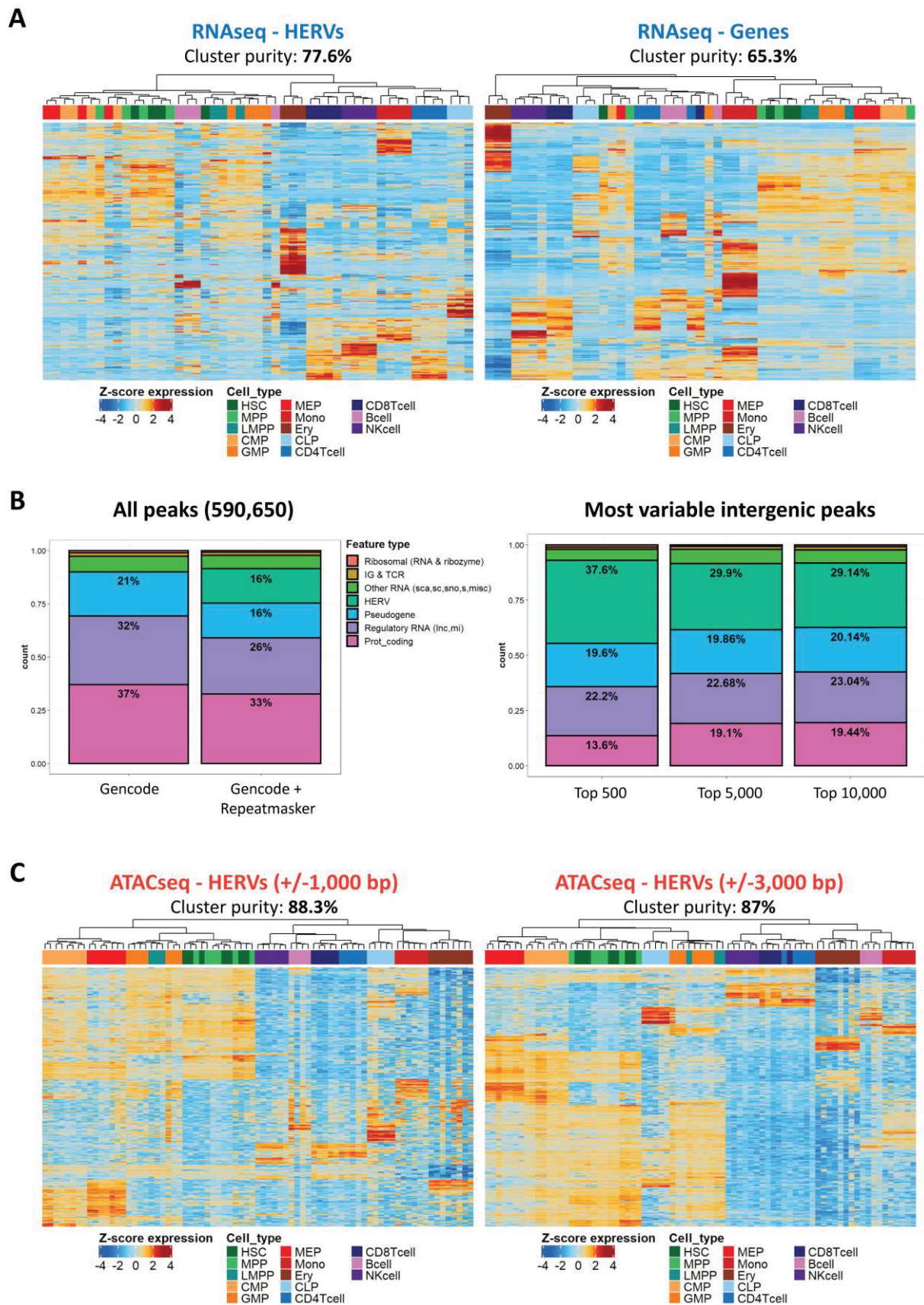
### **The HERV retrotranscriptome accurately defines normal hematopoietic cell populations**

As a first step, we examined HERV expression in different normal hematopoietic cell populations, assuming that distinct HERV profiles may characterize the main cell types. Using a custom pipeline based on Telescope (27), we quantified the expression of 14,968 HERV loci in public RNA-seq data from sorted bone-marrow and peripheral blood cell populations from 9 healthy donors (n = 49 samples, Figure S1A) (28). Unsupervised hierarchical clustering based on the top 20% most variable HERVs showed a robust classification of normal hematopoietic cell types with a cluster purity of 77.6% and a corrected Rand Index of 0.61 (Figure 1A). The same approach based on genes reached a cluster purity of 65.3% with a corrected Rand Index of 0.47 (Figure 1A).

We then sought to improve the clustering using ATAC-seq data through the analysis of peaks from open chromatin regions. Using the HOMER package, we applied a classic human genome annotation from gencode (v33) to annotate the set of 590,650 significant non-overlapping peaks from open chromatin regions previously defined in bone marrow and peripheral blood cells sorted from healthy donors (n = 80 samples) (28). As previously described, unsupervised hierarchical clustering based on promoter elements (peaks between -100bp and 1,000 bp away from a transcription start-site (TSS)) and intergenic elements (peaks more than 1,000 bp away from any other feature) significantly improved cluster classification, with a purity reaching 81.8% (Figure S1B). We then re-annotated these significant peaks with a custom reference consisting of the same Gencode annotation concatenated with the previously used 14,968 HERV loci from Repeatmasker. This new annotation showed that 16% of the significant peaks correspond to HERV regions (Figure 1B). One important previously reported finding is that classification based exclusively on intergenic elements (the so-called “distal regulatory elements”) is sufficient to classify normal hematopoietic cell populations (28). Enhanced annotation of these distal regulatory elements revealed an enrichment in HERVs, with up to 37.6% of the top 500 variable intergenic peaks corresponding to a HERV region (Figure 1B). A plot of the total aggregated count from these regions showed a gaussian distribution surrounding HERV TSS, confirming the good quality of the ATAC-seq signal (Figure S1C). Clustering of samples based on active HERV regions (AHR, defined by peaks surrounding HERVs regions +/- 1000 or 3000 bp) further improved the clustering, reaching 88.3% cluster purity (Figure 1C).

Altogether these results show that the HERV retrotranscriptome can be used to characterize normal immature and mature hematopoietic cell populations. The improved clustering obtained with AHR defined on ATAC-seq data suggests that this retrotranscriptomic signature reflect epigenetic footprints associated with cell differentiation.





**Figure 1**



## **Figure 1: Definition of normal hematopoietic cell populations based on the HERV retrotranscriptome**

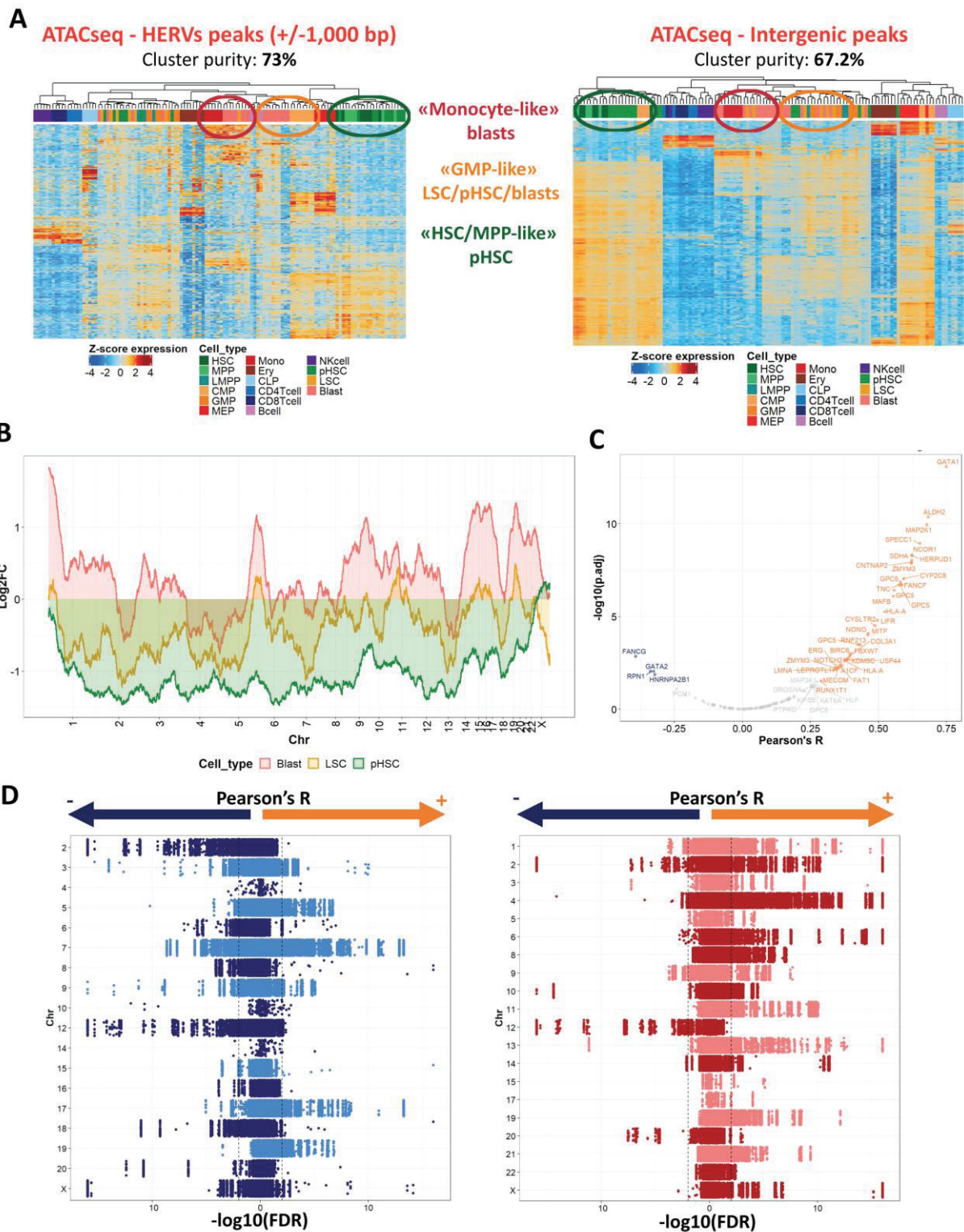
- A. Unsupervised hierarchical clustering of normal hematopoietic cell populations based on HERV (left) or gene (right) expression in RNA-seq. Clustering was performed with the ward.D2 method based on the maximum distance.
- B. Annotation of all significant ATAC-seq peaks (left) and top variable intergenic peaks (right) with a custom annotation of Gencode v33 with HERV references from Repeatmasker.
- C. Unsupervised hierarchical clustering of normal hematopoietic cell populations based on ATAC-seq peaks from HERV regions (+/- 1,000 (left) or 3,000 (right) bp from HERVs' TSS). Clustering was performed with the ward.D2 method based on the maximum distance.

*CLP: Common Lymphoid Progenitor, CMP: Common Myeloid Progenitor, Ery: Erythrocyte, GMP: Granulocyte-Macrophage Progenitor, HSC: Hematopoietic Stem cell, LMPP: lymphoid-primed multipotent progenitor, MEP: Megakaryocyte-Erythroid Progenitor, MPP: Multipotent Progenitor*

### **Acute myeloid leukemia cells show distinct HERV profiles close to their normal cell of origin**

We next evaluated how the HERV retrotranscriptome may help to distinguish AML cells. We performed the same clustering approach, integrating the 32 RNA-seq and 45 ATAC-seq bone marrow samples from 12 AML patients at diagnosis to our previous samples (Figure S1A) (28). Unsupervised clustering based on the top 20% variable AHR (+/- 1,000 bp from a HERV TSS) in ATAC-seq resulted in a good classification of normal and AML cells, with a slight increase in cluster purity compared to the top 20% most variable intergenic peaks (Figure 2A). Clustering based on HERV expression in RNA-seq yielded comparable results (Figure S2A). Interestingly, leukemic blast cells (blasts) clustered with either monocytes or granulocyte-monocyte progenitor (GMP) cells, LSCs with either GMP or lymphoid-primed multipotent progenitor (LMPP) cells and pre-leukemic hematopoietic stem cells (pHSCs) with either GMP or HSC/multipotent progenitor (MPP) cells, suggesting a clustering with their cell of origin as previously described (28). Cluster purity based on the original categories of cells did not consider these similarities and was thus a poor indicator of clustering performance in this case. Differential ATAC-count analysis centered on extended AHR (+/- 20,000 bp from a HERV TSS) revealed distinct profiles between AML LSCs, blasts and pHSCs compared to their normal counterparts, with globally a

more open chromatin in blasts and more closed in LSCs and pHSCs (Figure 2B). To further characterize the role of HERVs in these AHR, we computed correlations between RNA expression of each HERV present in an AHR and its respective surrounding genes located at +/- 50,000 bp. Strikingly, we mostly found positive correlations between HERV expression and their surrounding genes (Figure S2B). Annotation of the genes with a pre-established list of cancer-associated genes from the Cancer Gene Census database (29) revealed several genes positively correlated with HERVs expressed in AHR (Figure 2C). Of note, the highest correlation was found for GATA1 with ERVLB4\_Xp11.23b (Pearson's R: 0.74, adjusted p-value: 8.11e-14) (Table S1). Using TCGA LAML RNA-seq data, we then explored the association between each HERV located in an AHR and gene copy number variation (CNV) on the same cytoband. This highlighted several HERVs correlated both positively and negatively with deletions, and mostly positively with amplifications on the same cytoband (Figure 2D). These results demonstrate that HERV expression profiles differ according to the AML cell type, and suggest that HERVs are associated with gene regulation.



**Figure 2**

Figure 2: AML cells show distinct epigenetic profiles compared to their normal counterpart

A. Unsupervised hierarchical clustering of normal and leukemic hematopoietic cell populations based on HERV-centered (left) or global intergenic (right) open regions in

ATAC-seq. Clustering was performed with the ward.D2 method based on the maximum distance.

- B. Differential ATAC-count analysis between AML and normal BM cells. Log2FC between each AML subpopulation (blast, LSC, pHSC) and normal BM cells is shown on the y axis, and chromosomal location on the x axis. Only values with an FDR < 0.05 were considered.
- C. Correlation between RNA expression of HERVs located in active chromatin regions and the surrounding cancer-associated genes (+/- 50,000 bp). Significant Pearson's R are represented in orange (positive) or blue (negative). P-values were corrected using the FDR method.
- D. Correlation between RNA expression of HERVs located in active chromatin regions and CNV: deletions (left) and amplifications (right). Individual chromosomes are represented on the y axis,  $-\log_{10}(\text{FDR})$  on the x axis. Pearson's R were independently calculated for each HERV and all the alterations present on the same cytoband.

*AML: Acute Myeloid Leukemia, BM: Bone Marrow, CLP: Common Lymphoid Progenitor, CMP: Common Myeloid Progenitor, CNV: Copy Number Variation, Ery: Erythrocyte, FDR: False Discovery Rate, GMP: Granulocyte-Macrophage Progenitor, HSC: Hematopoietic Stem cell, LMPP: lymphoid-primed multipotent progenitor, LSC: Leukemic Stem Cell, MEP: Megakaryocyte-Erythroid Progenitor, MPP: Multipotent Progenitor, pHSC: pre-leukemic Hematopoietic Stem Cell*

## **HERV expression defines subtypes of AML with distinct cancer hallmarks and outcomes**

Having established that the HERV retrotranscriptome characterizes particular cell types, we then wondered whether HERV expression could define distinct AML profiles in bulk RNA-seq data. We explored HERV expression in 4 independent RNA-seq datasets (TCGA, AMLCG, LEUCEGENE and BEAT), retaining only bone marrow samples from AML patients at diagnosis (n = 788) (Table S2). For each dataset, we selected the top 2,000 most variable HERVs based on the scaled DESEQ2 VST normalized count (Figure S3A). We merged the 4 datasets, keeping only the intersect between each top 2,000 candidate HERVs, resulting in 961 variable HERVs conserved across the 4 datasets. Unsupervised hierarchical clustering guided by the average silhouette (Figure S3B) and Bayesian Information Criterion (BIC) evolution defined 9 clusters that were not dependent on the study (Figure 3A and Figure S3C). These 9 clusters were associated with significant differences in overall survival among intensively treated patients (Figure 3B), independently of established prognostic factors such as age, ELN2017 and white

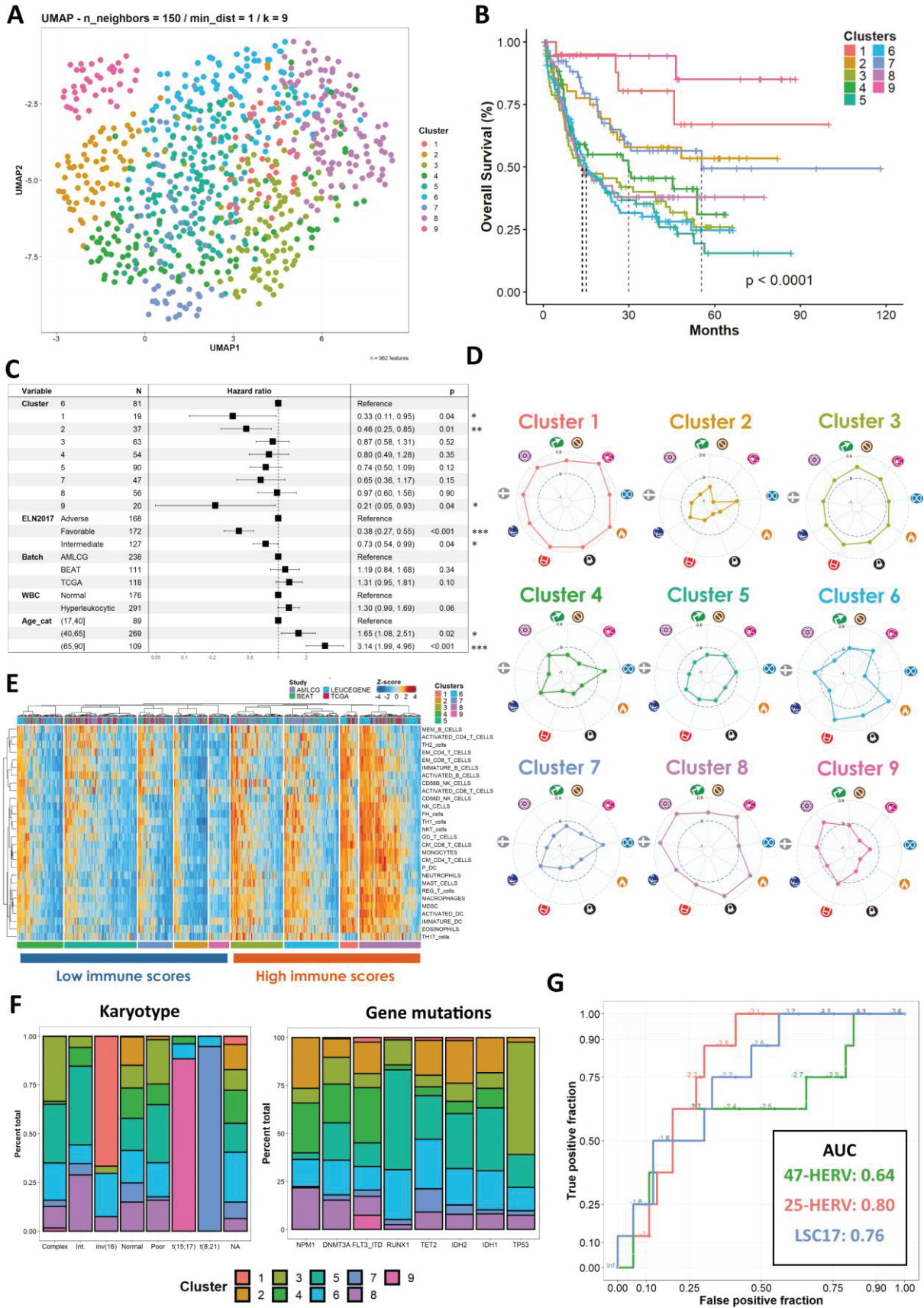
blood count, integrated in a multivariate Cox model (Figure 3C). These clusters also presented distinct cancer hallmark profiles (Figure 3D), as assessed by single sample gene-set variation analysis (GSVA) based on cancer hallmark signatures (Table S3). GSVA of immune signatures revealed no clear immune subtype profiles, distinguishing clusters with globally low or high immune scores (Figure 3E).

We then assessed whether these clusters were associated with known recurrent translocations or gene mutations. We found a clear enrichment in *inv(16)*, *t(8;21)* and *t(15;17)* in clusters 1, 7 and 9, respectively (Figure 3F). Other karyotypes (such as complex, poor or intermediate abnormalities) showed no particular association with any cluster, underlining the heterogeneous composition of these groups. Regarding gene mutations, we observed a tendency for an enrichment in *RUNX1* mutation in cluster 5 and in *TP53* mutation in cluster 3 (Figure 3F).

Focusing on HERVs discriminating each cluster (i.e. HERVs overexpressed in a given cluster compared to all other clusters), we found that clusters 8 and 9 had the highest number of different discriminating HERVs, whereas only few HERVs discriminated clusters 3 and 5 (Figure S3D). HERVs from the *HERV-H*, *ERV-L*, *MER4*, *HERV-L* and *HERV-K* families were the most frequent HERVs to discriminate clusters (Figure S3E). When reported to the total number of HERVs per family, *HERV-S*, *HERV-E*, *HERV-P*, *ERV1* and *HARLEQUIN* families were the most frequently represented (Figure S3F).

To further demonstrate the value of HERV expression as a promising biomarker, we next established a LSC signature based on HERV expression. We calculated correlations between each individual HERV and the previously published LSC17 score (30) in the 4 independent datasets. 47 HERVs with a significant correlation with the LSC17 score (adjusted *p*-value < 0.05) in at least 2 independent datasets were used to establish a new LSC score based on HERV expression only (see methods). These 47 HERVs were further refined to select the minimal number of HERVs allowing a good LSC discrimination. The resulting LSC signature based on 25 different HERVs showed a very good discrimination in the independent validation set of sorted AML cells, separating LSCs from all other cells with an area under the curve (AUC) of 0.80 versus 0.64 for the non-refined signature and 0.76 for the LSC17 score (Figure 3G). Both the full 47-HERV and the 25-HERV LSC signatures showed a significant impact on overall survival, mirroring the LSC17 signature's performances (Figure S3G). Altogether these results show that HERVs represent an important source of genetic information that can be used to define different AML subtypes as well as cell-specific signatures, as highlighted herein for LSCs.







### **Figure 3: The HERV retrotranscriptome defines AML subtypes with distinct prognosis and cancer hallmarks**

- A. UMAP representation of the 788 AML patients. Clusters defined by the unsupervised hierarchical clustering approach are shown.
- B. Overall survival of intensively treated patients according to the 9 clusters in the whole cohort.
- C. Multivariate Cox analysis of overall survival of intensively treated patients. Known risk factor (Age, ELN2017 and WBC), study (batch) and clusters are integrated in the multivariate model.
- D. Cancer hallmark profiles of each cluster. Each cancer hallmark is represented by its symbol as defined in (267). For each cluster, hallmark scores are calculated by ssGSVA. Mean scores are represented on a radar plot.
- E. Heatmap of immune signature enrichment. Patients were grouped into clusters before performing unsupervised hierarchical clustering. Z-scores of ssGSVA signature enrichment are represented.
- F. Bar chart of genomic alteration (left) and mutation (right) distribution according to clusters. For each category, the total count and the percentage is represented.
- G. ROC-curve of LSC classification according to the 47-HERV (green), 25-HERV (red) and 18-gene (blue) LSC signature.

*LSC: Leukemic Stem Cell, ROC: Receiver Operating Characteristic Curve, ssGSVA: Single Sample Genes-set Variation Analysis, WBC: White Blood Count*

### **AML-specific HERVs contain several open-reading frames representing a source of shared tumor antigens**

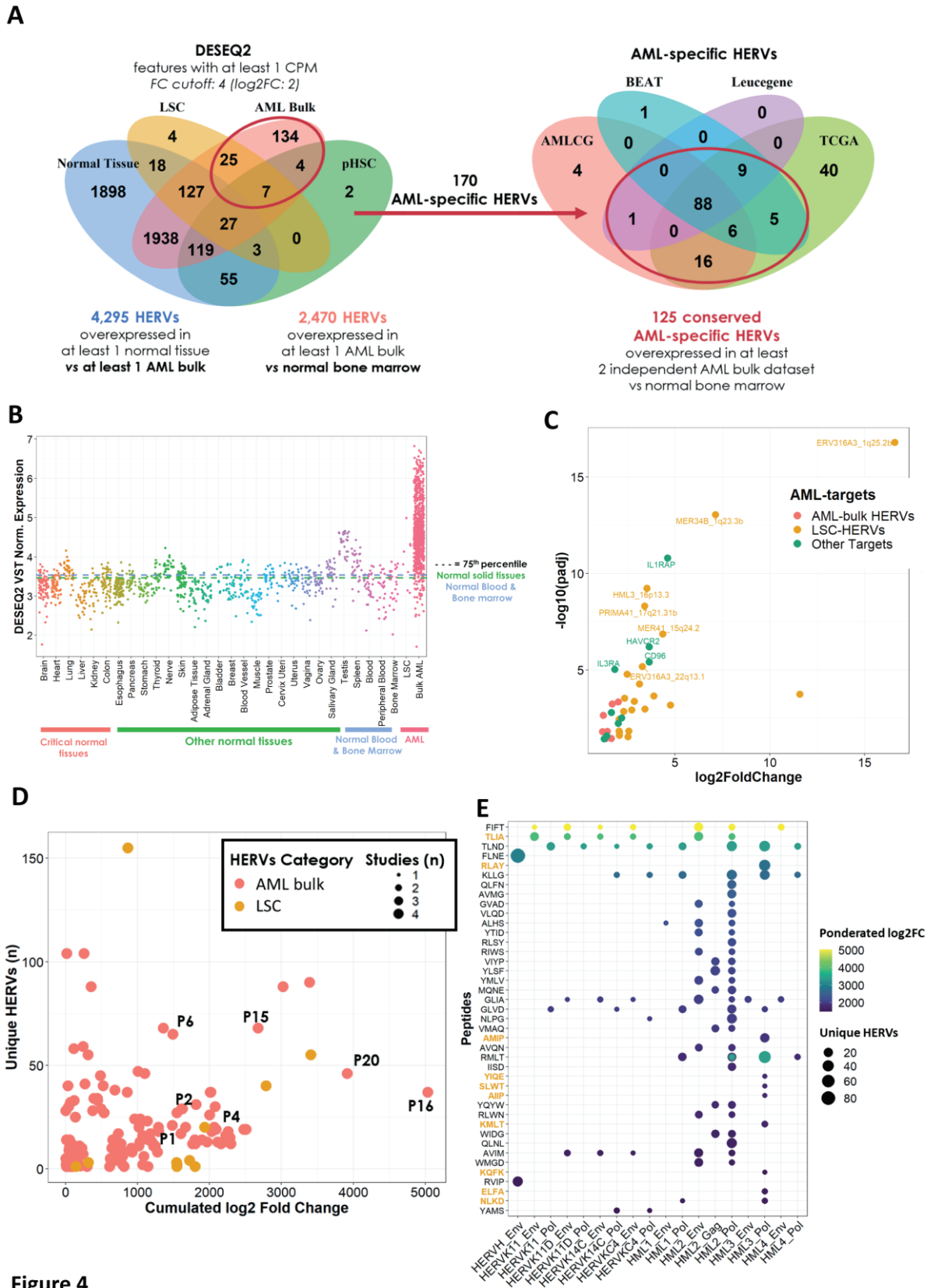
Considering the specific expression of some HERVs in AML, we evaluated if they may represent a source of shared tumor antigens for immunotherapeutic approaches. To define leukemia-specific targets, we used a dataset of more than 1,000 normal samples from 42 different sites extracted from the Genotype-Tissue Expression (GTEx) database (Figure S4A) in addition to the normal bone marrow and peripheral blood samples used above. We performed differential HERV expression analysis between AML and normal tissues, considering transcripts with a shrunken fold change  $> 4$  and a base mean of at least 1 normalized count per million as overexpressed. Initially using normal bone marrow cells as a reference, we found 2,470 HERVs overexpressed in at least one of the 4 bulk AML datasets. To define AML-specific HERVs, we

setup an additional filter to remove HERVs that could be expressed at higher levels in any of the 42 normal tissues than in AML bone marrows (see methods). Among the 2,470 HERV candidates, 170 were never overexpressed in any normal tissue and were further considered to be AML-specific (Figure 4A). We then cross-validated the results between each of our 4 datasets to retain only HERVs that were overexpressed in at least 2 independent datasets, finally selecting a set of 125 conserved AML-specific HERVs (Figure 4A). The mean AML expression of these 125 AML-specific HERVs highly exceeded in most of cases the 75th percentile expression in normal tissues, with the exception of testis (Figure 4B). When focusing on the top overexpressed HERVs, we found candidates with a particularly high differential expression, with log<sub>2</sub>FC up to 17.47 and 16.58 for ERV316A3\_1q25.2b in bulk AML cells and LSCs, respectively (Figure S4B, Figure 4C). Several HERVs had an expression level in LSCs compared to normal bone marrow cells at least as high as other current LSC therapeutic targets (Figure 4C).

To screen for potential HERV-derived epitopes, we next defined all the putative ORFs by translating these 125 AML-specific HERVs into the 6 possible frames. To reduce the number of false positives, ORFs of at least 10 amino-acids were then aligned against a manually annotated database of 71 known existing HERV proteins derived from curated Uniprot references. ORFs with at least a 75% identity with an existing HERV protein and an E-value < 0.01 were selected for epitope screening. Using these criteria, we found that only 33/125 AML-specific HERVs possessed at least 1 ORF (Figure S4C). Two independent recent tools trained on mass spectrometry data were then used for epitope prediction: netMHCpan 4.1 (31) and MHC flurry 2.0 (32). HLA-A\*02 epitope screening revealed 262 unique peptides predicted as strong HLA-A\*02 binders with both independent tools.

To set up a list of peptides to prioritize, we further ranked these results. Considering that epitope immunogenicity is linked to its tumor abundance (33), we established an additive model taking into account (i) the mean expression of HERVs containing the peptide sequence, (ii) the fold change compared to normal bone marrow and (iii) the number of different HERVs containing the peptide sequence. We relocated each peptide in all the different HERVs containing its sequence and added the fold changes pondered by the base mean expression of each individual HERV. This additional step filtered-out 158 peptides that were either shared with HERVs present in normal tissue (pondered negatively in our additive model) or overexpressed but with a very low base expression. Among the 104 remaining peptides (Figure 4D), the top expressed candidates were mainly from Gag and Pol proteins of HERV-K/HML-2 family (Figure 4E and Figure S4D).

Overall, these results show that HERVs represent a reservoir of potential CD8+ T cell epitopes overexpressed in AML that may represent original therapeutic targets.



**Figure 4**  
Figure 4: HERVs as a source of shared epitopes in AML

- A. Venn diagram representing the results of the differential HERV expression analysis. HERV expression between AML samples and normal tissues (blue) or AML samples and normal bone-marrow samples (red) was first calculated, leading to the identification of 170 AML-specific HERVs that are overexpressed in at least 1 AML dataset and never overexpressed in any normal tissue (left). Among these 170 AML-specific HERVs, 125 were found in at least 2 independent datasets and retained for further analyses (right).
- B. Scatter plot showing the mean expression of the 125 AML-specific HERVs across normal and AML tissues. Dotted lines represent the 75<sup>th</sup> percentile expression of normal solid and hematopoietic tissues.
- C. Scatter plot of individual HERV fold change between AML bulk datasets and normal bone marrow. Features (*i.e. genes or HERVs*) overexpressed in both LSC and AML bulk datasets are shown in orange, and features expressed only in bulk are shown in red. Other currently investigated LSC targets are represented in green. Only features with a FC  $\geq 4$  and a FDR  $\leq 0.05$  are represented.
- D. Scatter plot of individual peptides in the additive model. For each peptide, cumulative log<sub>2</sub>FC between AML cells and normal BM (x axis) and number of unique HERVs containing the peptide sequence (y axis) is represented. Cumulative log<sub>2</sub> FC were pondered by the base mean expression of the corresponding HERV.
- E. Dot plot showing the family and protein type of the most expressed peptides from the additive model. Peptides are ranked by cumulated fold change. Peptides overexpressed in LSC are represented in orange.

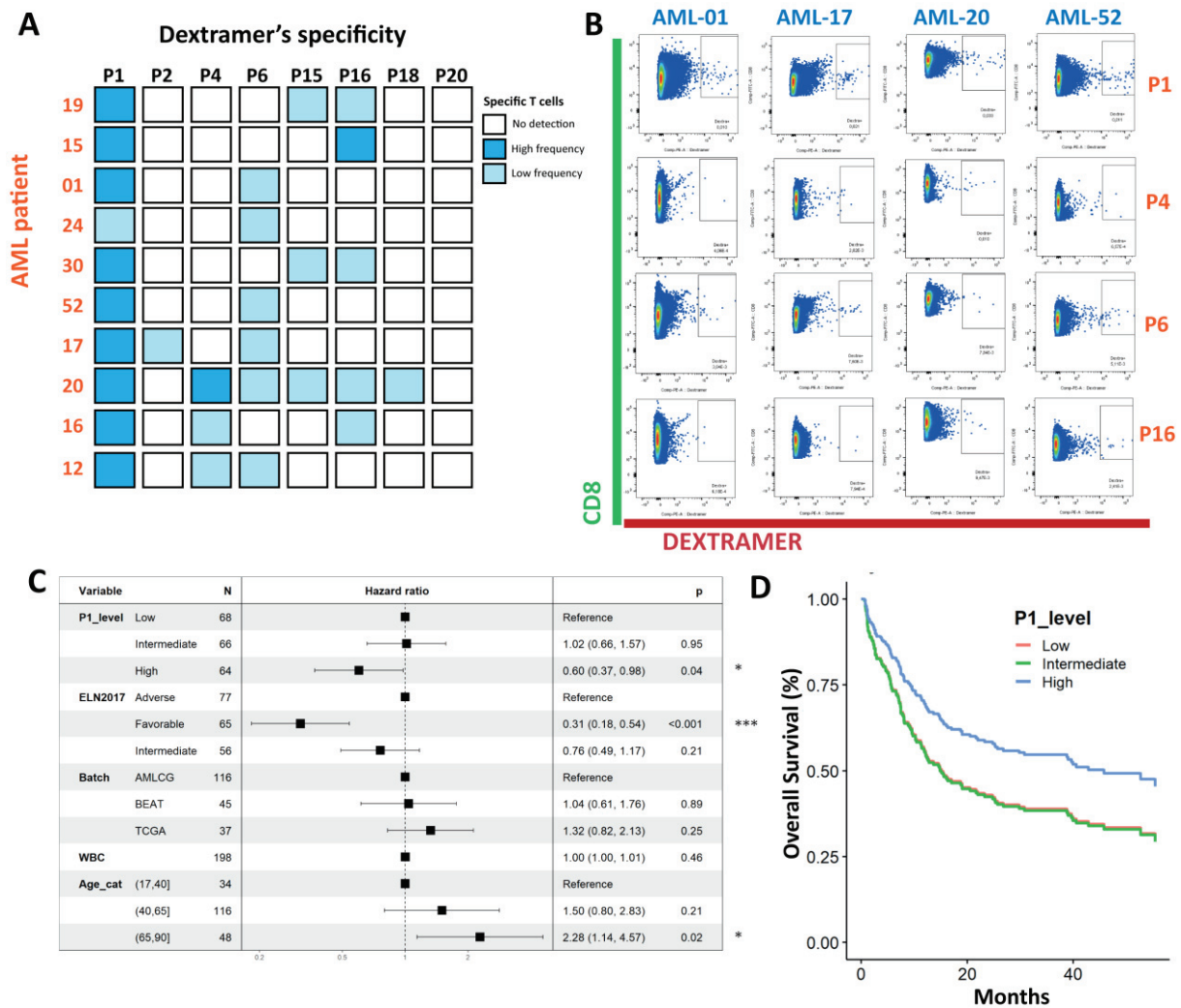
*FC: Fold Change, FDR: False Discovery Rate, LSC: Leukemic Stem Cell*

### **T cells specific for HERV-derived epitopes are naturally present in AML patients**

We then examined whether the predicted HERV-derived HLA-A\*02 epitopes could elicit an immune response in AML patients. We selected 8 different HLA-A\*02 epitopes among the top candidates in the additive model, focusing on peptides already identified as immunogenic in our lab (P1, P2, P4 and P6) or peptides preselected in a preliminary study in solid tumors but for which the immunogenicity has so far not been confirmed (P15 P16 P18 P20) (Bonaventura et al, submitted). Bone marrow-infiltrating lymphocytes (MILs) were expanded from bone marrow mononuclear cells (BMMCs) of HLA-A\*02 AML patients at diagnosis. High-dose IL-2 stimulation induced expansion of CD45<sup>+</sup> cells from 14.4 % to 74.3% (mean values) among living

cells in 14 days (Figure S5A). Using HLA-A\*02-peptide dextramers, we screened for the presence of specific CD8<sup>+</sup> T cells against the selected epitopes. Strikingly, P1-specific CD8<sup>+</sup> T cells were found in all patients (n=10) at high frequencies (Figure 5A and 5B). Specific CD8<sup>+</sup> T cell responses against P4, P6, P15 and P16 were also detected in some patients at lower frequencies (Figure 5A and 5B). To further explore the potential impact of P1 expression in AML cells, we assessed the prognostic value of the expression of HERVs containing its sequence among HLA-A\*02 patients. Multivariate Cox analysis revealed that patients with a high expression of P1-containing HERVs in leukemic cells had a better overall survival than patients with a low or intermediate expression, regardless of age, ELN2017, batch and white blood count (Figure 5C).

Overall, these results show that HERV-derived epitopes elicit a CD8<sup>+</sup> T cell response that can be detected in the bone marrow of AML patients. Furthermore, a higher expression of HERVs containing immunogenic epitopes may be associated with a better clinical outcome.



**Figure 5**  
 Figure 5: Specific CD8<sup>+</sup> T cells against HERV-derived epitopes are found in bone marrow-infiltrating lymphocytes (MILs) of AML patients at diagnosis

- Summary table of specific CD8<sup>+</sup> T cell responses found among MILs in patients. Low frequencies are represented in light blue, high frequencies in dark blue.
- Representative dextramer results of 4 AML patients. Bone marrow CD8<sup>+</sup> T cell populations are gated among single living cells after 14 days of expansion. CD8 staining is represented on the y axis, dextramer staining is represented on the x axis.
- Multivariate Cox analysis of overall survival in intensively HLA-A\*02 treated patients. Known risk factor (age, ELN2017 and WBC), study (batch) and P1 expression level (represented by the mean expression of HERVs containing the peptide sequence) are integrated in the multivariate model.
- Adjusted overall survival curves according to P1 expression level in intensively treated HLA-A\*02 patients. Predicted survival is based on the previously established cox model.

MILs: Marrow Infiltrating Lymphocytes, P1: Peptide 1



## DISCUSSION

Representing 8% of the human genome, HERVs are a great source of genetic information. However, as the results of their analysis are largely dependent on the chosen analytical strategy, this signal can often be missed or misinterpreted. For instance, using a partial database limited to HERVs corresponding to known Uniprot references (34) can occult a major part of the HERV retrotranscriptome, and restricting HERV signal to uniquely mapped reads strongly reduces the signal from conserved, recently integrated families prone to multimap (35). Consistently, we used a custom-designed pipeline based on Telescope (27), a recently developed tool allowing the accurate quantification of a complete base of 14,968 HERV loci. One of the main advantages of Telescope resides in its reassignment algorithm to correctly reassign a high number of multimaps and thus keep the highest possible signal from recently integrated and conserved HERV families. This approach provided a thorough analysis of HERV expression in normal and AML cell populations, providing a robust classification of specific cell-types only based on HERV transcriptomes.

Chromatin accessibility represents a major factor of cell differentiation that may capture cell identity more accurately than gene expression in blood cells (28). Furthermore, HERVs may actively shape the chromatin architecture (36). In line with these observations, we show here that a very efficient clustering of normal hematopoietic cell populations is obtained with the use of AHR located in distal regulatory elements. In the same way, AHR profiles allowed not only the discrimination of leukemia cells from normal blood cell populations, but also between leukemia cells clustering with their cell-of-origin. Thus, pHSCs, LSCs and leukemic blast cells displayed a distinct epigenetic profile with distinct AHR compared to their normal counterparts. Using the unique information provided by the HERV retrotranscriptome, we built a signature based on 47 HERVs that allowed a robust classification of LSCs among normal and leukemic bone-marrow cells. Misclassified cells were mostly represented by blast cells. A LSC-HERV signature could represent an original tool to either evaluate AML prognosis (i.e. as a surrogate marker of the remaining LSCs) or minimal residual disease during follow-up.

A positive correlation was found between HERV expression from AHR and their surrounding genes, including cancer-associated genes. This may reflect the presence of these elements in actively transcribed DNA regions. Alternatively, HERVs may exert a regulatory role in gene transcription in these regions, as already described (37). In this context, it is tempting to speculate that some of these HERVs may act as enhancers for AML oncogenes, as reported recently by Deniz et al. (25). Interestingly, LSCs and pHSCs showed fewer AHR compared to blasts. This

result is consistent with a previous report demonstrating that transposable elements are silenced in LSCs (26). This may be due to the fact that silencing of HERVs might promote LSC immune escape, since HERVs have been described to induce both innate and adaptive immune responses (16). However, as we show herein, some HERVs containing conserved ORFs are still overexpressed in LSCs compared to normal cells, suggesting that it may be possible to therapeutically boost an immune response against HERV epitopes in LSCs.

Relying only on HERV expression, we managed to retrieve particular AML subtypes but also to define new subtypes within normal and complex karyotype groups. Importantly, the prognostic differences found across these AML subtypes was independent of the currently used ELN2017 score, based on the karyotype and mutated genes. These results suggest that HERVs could be used to improve risk stratification and treatment resistance prediction in patients with no genetic or molecular abnormalities associated with well-defined prognosis and resistance profiles. Association of HERVs to gene expression could improve gene expression-based predictors such as PS29MRC (18). Unfortunately, the lack of thoroughly annotated data (with at least the exact stem cell transplantation date and type when performed) precluded the establishment of a specific predictive score, which would require a dedicated study.

Finding specific tumor antigens in AML remains a major challenge for immunotherapeutic approaches. Because AML belongs to malignancies with the lowest mutational burdens (19), the frequency of mutations creating neoantigens is expected to be low. In this context, HERV-derived antigens represent a unique source of non-conventional epitopes (21) that could be exploited for the development of new immunotherapies. Our bioinformatics-based approach allowed us to identify several CD8<sup>+</sup> T cell epitopes from AML-specific HERVs, i.e. HERVs expressed at high levels in AML cells and never expressed in normal tissue. The presence of CD8<sup>+</sup> T cells specific for HERV epitopes in the bone marrow of AML patients at diagnosis confirmed that these epitopes are naturally processed and are immunogenic. Our results substantiate a recent report from Saini et al. showing that HERVs represent a reservoir of CD8<sup>+</sup> T cell epitopes in myeloid malignancies (38). In this latter study, a high-throughput epitope screening method was set up based on barcode dextramers. It would be interesting to combine both approaches to obtain an optimized epitope detection pipeline associated with a more systematic functional screening method. Of note, the P1 epitope (FLQFKTWWI), which seemed to be particularly immunogenic in our study and in another in the context of solid tumors (39) was screened by Saini et al. but not detected with the barcoded dextramers. This discrepancy could be explained by the use of peripheral blood samples or a lower sensitivity in the absence of

prior expansion of bone marrow T cells (40), especially in AML, a disease characterized by a massive leukemic cell infiltration in the bone marrow.

In conclusion, our study unveils the HERV retrotranscriptome as a powerful tool to characterize and classify specific cell populations and to provide new tumor-specific epitopes for the development of immunotherapeutic strategies. This approach, demonstrated here in AML, could easily be extended to any disease to provide a better risk-stratification, establish relevant predictive signatures for therapeutic responses or identify original therapeutic targets.

## **METHODS**

### **Raw RNA-seq data**

Raw RNA-seq data files were accessed from the NCBI Gene Expression Omnibus (GEO) portal, under the accession numbers GSE74246 for the sorted hematopoietic normal and AML cells from Corces et al. (28), GSE49642, GSE52656, GSE62190, GSE66917, GSE67039 and GSE106272 for the LEUCEGENE datasets. TCGA LAML (41) and BEAT-AML (42) data were accessed from the NCI Genomic Data Commons (GDC) data portal (<https://portal.gdc.cancer.gov/>). Raw data for the AMLCG cohort (18) were directly provided by the AMLCG group. HLA genotyping of TCGA, BEAT and AMLCG cohorts was assessed using archasHLA v0.2.0 (43).

### **HERV and gene expression quantification**

HERV expression was quantified using a custom pipeline derived from Telescope (27). Briefly, RNA-seq reads were aligned to a custom transcriptome using bowtie2 v2.2.1 (44) with custom parameters to retain multimaps (-k 100 --very-sensitive-local --score-min "L,0,1.6"). The custom transcriptome consisted in the hg38 reference transcriptome with 14,968 HERV transcriptional units compiled from RepeatMasker annotations (27). SAM outputs were converted to BAM files using SAMtools v1.4 (45). HERV and gene expression were then calculated using Telescope (27) and HTSeq 0.12.3 (46), respectively. Raw counts were concatenated and normalized independently for each dataset using DESEQ2 v1.28.0 with variance stabilizing transformation (VST) (47).

### **ATACseq data**

Significant peaks called from ATACseq data analysis were retrieved from the original paper (28). Briefly, peaks were called using MACS2 and filtered using a custom blacklist. A final set of

590,650 significant peaks were defined among a list of non-overlapping maximally significant 500 bp peaks ranked by their summit significance value. These significant peaks were re-annotated using HOMER with the command “annotatePeaks.pl” and two different references: Gencode v33 only and Gencode v33 with the previously used HERV annotation. Regions containing significant peaks around +/- 1,000 or 3,000 bp of a HERV TSS were considered to be active HERV regions.

### **Unsupervised hierarchical clustering**

For the sorted cells, DESEQ2 VST normalized expression data were directly used for unsupervised hierarchical clustering. Cluster purity was used as an external validation criterion and was calculated by first creating a confusion matrix between assigned cluster number and annotated cell type before adding the maximum values from each row (i.e. assigned cluster) and dividing by the total number of samples.

A benchmark of distances (euclidean, maximum and pearson) and methods (ward.D2, single, complete, average and centroid) was performed to identify the optimal method leading to the best cluster purity using a pre-defined number of clusters according to the original annotation.

For bulk datasets, DESEQ2 VST normalized expression data were independently calculated and further center-scaled for each dataset to correct the potential batch effect. Unsupervised hierarchical clustering was then performed using the average silhouette width and the Bayesian Information Criterion as internal validation markers.

### **Differential ATAC-count analysis**

For differential ATAC-count analysis, raw ATAC-seq counts were retrieved from the original paper (28). Differential expression analysis between each AML population (LSC, pHSC and Blasts) and their normal counterparts (HSC, GMP, LMPP and monocytes) was performed using DESEQ2, with cell type as a covariate. Differentially expressed regions surrounding a HERV TSS (+/- 20,000 bp) and with a FDR < 5% were retained for the final plot. The rolling mean of 1,000 sequential regions, ordered by chromosome location, was then represented.

### **HERVs, genes and copy number variation correlations**

HERVs located in extended AHR (peaks +/- 20,000 bp of a HERV TSS) were selected for correlation analysis. For each HERV, a list of surrounding genes located at +/- 50,000 bp of their TSS was established. Pearson's correlations were calculated between the RNA expression of each HERV and each of its surrounding gene, independently. P-values were corrected with the FDR

method. Genes were then annotated using a published list of cancer-related genes from the Cancer Gene Census (29). The same list of HERVs was then used to perform correlations with CNV from the same cytoband. TCGA LAML CNV data were retrieved from the NCI GDC portal and used as is to calculate Pearson's correlations with HERVs from the same cytoband.

### **Survival analysis**

Intensively treated patients with available survival data were retained for the survival analysis. Patients receiving hypomethylating agent therapies or supportive care were discarded from this analysis. Overall survival curves were estimated with the Kaplan-Meier method. Only known prognostic factors (Age, ELN2017, white blood count), batch and clusters were integrated in the final multivariate cox model.

For the second survival analysis (shown in figure 5), intensively treated HLA-A\*02 patients were selected and stratified according to the P1 expression level, cut in terciles. The same covariates were included in a multivariate cox model of overall survival. Adjusted survival curves were then drawn based on this model. Survival analyses were performed using the R survival and survminer packages. Cox model proportional-hazards assumptions were assessed using Schoenfeld's test and inspection of residual plots.

### **Cancer hallmark and immune signatures GSVA**

For each cancer hallmark (48), a unique gene signature was established (Table S3) based on The Molecular Signatures Database (MSigDb) Hallmark Gene Set Collection (49). When not available in MSigDb, hallmark signatures were established from Gene Ontology (GO) signatures, as previously described (50). The signature for the immune evasion hallmark was retrieved from Hubert et al. (51). Individual enrichment scores were calculated from each patient by single sample gene-set variation analysis (ssGSVA) (52), and scaled by study. The mean score for each cluster was then calculated and shown in a radar plot.

Immune signatures were obtained from Thorsson et al. (53) and calculated by ssGSVA for each sample. Unsupervised hierarchical clustering was then performed on study-scaled ssGSVA scores in each cluster.

### **HERV-LSC signature**

To establish the HERV-LSC signature, correlations between the expression of each unique HERV and the validated LSC17 score (30) were computed independently in the 4 bulk RNA-seq datasets. 47 HERVs showing a significant correlation (FDR-adjusted p-value < 0.05) with the

LSC17 score in at least 2 datasets and with concordant results (i.e. correlated in the same direction in each dataset) were retained to build signatures. Signatures were calculated as the mean expression of all HERVs pondered (multiplied) by the sum of their Pearson's correlation coefficients. To select a minimal number of HERVs, the resulting 47-HERVs were ranked according to the absolute sum of their Pearson's correlation coefficients, and signatures were iteratively built by removing the last candidate (i.e. the one with the lowest summed correlation). Signatures' performances were then evaluated in the independent testing set of sorted cells from Corces et al. with a classification approach for LSC alone or LSC and Blasts. ROC curves and AUC were drawn and calculated with the plotROC R package. The signature showing the best performances for LSC classification with the lowest number of HERVs was eventually selected, leading to the final 25-HERVs signature.

### **Differential HERV expression analysis**

Differential expression analysis was performed using DESEQ2 (47). Raw counts of HERVs and genes from all normal and AML datasets were merged and integrated into the same DESEQ object, using study (i.e. batch) as a covariate in the design formula. Differential expression analysis was performed for all the 4 independent bulk AML datasets and the sorted LSC and pHSC populations against each of the 42 normal tissues. Fold change were shrunk with the apeglm method (54). Features with a fold change superior to 4 ( $\log_2FC > 2$ ) and a base mean of at least 1 normalized count per million were considered to be overexpressed.

### **Open-reading frame detection and peptide selection**

ORFs were defined using EMBOSS's sixpack v6.6.0.0 (55) to translate HERV sequences into the 6 possible frames. ORFs of at least 10 amino-acids were then aligned to a reference of known HERV proteins (Gag, Pro, Pol, Env, Rec and Np9 from different HERV families). This reference was established from Uniprot (56), referencing all existing manually reviewed HERV protein sequences, resulting in 71 references (research equation: keyword:"endogenous retrovirus" AND reviewed:yes AND organism:"Homo sapiens (Human) [9606]", last accession: 12/11/2020). Blast was used for the alignment, with optimal parameters for retroviruses (Word size = 3, composition-based statistics, no "low-complexity-region" filter). ORFs aligning with at least 75% identity with a known HERV protein and an E-value  $< 0.01$  were considered for further analysis.

Retained ORFs were then screened for predicted HLA-A\*02 strong binders using MHCflurry v2.0 (32) and netMHCpan 4.1 (31). Peptides with an affinity rank  $\leq 0.5$ th percentile for both



tools were considered to be strong-binders. The human proteome was downloaded on Uniprot (ID: UP000005640) to validate the absence of match before peptide synthesis and in vitro validation.

### **Additive model**

Differential expression at the peptide level between normal bone marrow and AML cells was estimated by relocating each peptide in all the potential HERVs with predicted ORFs coding for one of the 262 peptides. Each peptide could thus have one or several differential expression data according to the number of HERVs they were found in. This total number of HERVs was taken into account by considering the total sum of fold change of each HERV containing the peptide of interest, leading to an estimation of the cumulated fold change per peptide. To avoid prioritization of targets with artificially high fold-changes due to very low expression levels, each fold change was pondered by the base mean of its respective HERV, leading to the final additive model.

### **Biological samples**

Bone marrow samples were collected from AML patients at diagnosis at the Centre Hospitalier Lyon Sud in Lyon, France. Sample collection was approved from the institutional review board and ethics committee (20.01.31.72653 - 21/20\_3) and after obtaining the written informed consent of patients, in accordance with the Declaration of Helsinki. BMMCs were obtained by Ficoll density gradient centrifugation (Eurobio, FR, EU) and immediately cryoconserved in fetal bovine serum (FBS) with 10% dimethyl sulfoxide (DMSO).

### **MILs growth**

Bone marrow mononuclear cells (BMMCs) were rapidly thawed at 37°C and grown in RPMI medium (Gibco, FR, EU) supplemented with 8% human AB-serum (Etablissement Français du Sang, FR, EU) and high doses (6,000 UI/mL) of IL-2 (PROLEUKIN aldesleukine, Novartis Pharma, CH, EU) after a 2-hour resting period. Plates were then incubated for 14 days, with medium replacement when needed.

### **Dextramer and flow cytometry analysis**

After 14 days of growth, cells were washed in 2 mL washing buffer (PBS + 2% FBS + 2mM EDTA (Sigma Alderich, MI, US)) and stained for 10 minutes with dextramers (Immudex ApS, DK, EU) at room temperature prior to viability and surface marker staining. Washing was performed twice to avoid non-specific dextramer staining. Peptide 18, for which we did not find

any specific response in any patient, was further considered as an irrelevant peptide and a negative control.

## **AUTHOR CONTRIBUTIONS**

**V. Alcazer:** Conceptualization, data curation, formal analysis, funding acquisition, methodology, investigation, resources, software, validation, visualization, writing-original draft. **P. Bonaventura, Michel, V. Mutez :** Methodology, resources, investigation, validation, visualization, writing – review & editing. **L. Tonon, A. M. N. Batcha:** Software, investigation, writing – review & editing. **A. Viari:** Resources, project administration, **G. Sauvageau, K. H. Metzeler, W. Hiddemann T. Herold, N. Chuvin, Y. Estornes:** Resources, writing – review & editing. **C. Caux:** Methodology, resources, writing – review & editing. **S. Depil:** Conceptualization, funding acquisition, methodology, resources, validation, project administration, writing – review & editing

## **ACKNOWLEDGMENT**

The results/or some of the results obtained in this publication are based on data generated by the Leucegene group essentially located at IRIC in Montreal, Canada and supported by Genome Canada and Genome Québec. This study thus relied on human AML specimens provided by the BCLQ, Montreal, QC.

This work was supported by the INSERM “poste d’accueil” research grant (to V. Alcazer), the Agence de biomédecine and Fondation ARC pour la Recherche sur le Cancer and Ligue Contre le cancer.

This work is supported by a grant of the Wilhelm-Sander-Stiftung (no. 2013.086.2) to TH and by the BMBF grant 01ZZ1804B (DIFUTURE) to A.M.N.B.

The authors would like to thank François Mallet and Nicolas Dulphy for their feedback and help throughout the study, and Brigitte Manship for her help with English editing.

## **CONFLICT OF INTERESTS**

SD is founder of ErVaccine Technologies. PB and VA are consultant for ErVaccine Technologies. EM and VM are employees of ErVaccine Technologies. The other authors have no conflict of interests to disclose.

## REFERENCES

1. Consortium IHGS. Initial sequencing and analysis of the human genome. *Nature* 2001;409(6822):35057062.
2. Johnson WE. Origins and evolutionary consequences of ancient endogenous retroviruses. *Nat. Rev. Microbiol.* 2019;17(6):355–370.
3. Vargiu L et al. Classification and characterization of human endogenous retroviruses; mosaic forms are common. *Retrovirology* 2016;13(1):7.
4. Kassiotis G, Stoye JP. Immune responses to endogenous retroelements: taking the bad with the good. *Nat. Rev. Immunol.* 2016;16(4):207–219.
5. Alcazer V, Bonaventura P, Depil S. Human Endogenous Retroviruses (HERVs): Shaping the Innate Immune Response in Cancers. *Cancers* 2020;12(3). doi:10.3390/cancers12030610
6. Larouche J-D et al. Widespread and tissue-specific expression of endogenous retroelements in human somatic tissues. *Genome Med.* 2020;12(1):1–16.
7. Hervé CA, Lugli EB, Brand A, Griffiths DJ, Venables PJW. Autoantibodies to human endogenous retrovirus-K are frequently detected in health and disease and react with multiple epitopes. *Clin. Exp. Immunol.* 2002;128(1):75–82.
8. Mameli G et al. Brains and peripheral blood mononuclear cells of multiple sclerosis (MS) patients hyperexpress MS-associated retrovirus/HERV-W endogenous retrovirus, but not Human herpesvirus 6. *J. Gen. Virol.* 2007;88(1):264–274.
9. Brudek T et al. B cells and monocytes from patients with active multiple sclerosis exhibit increased surface expression of both HERV-H Env and HERV-W Env, accompanied by increased seroreactivity. *Retrovirology* 2009;6(1):104.
10. Li W et al. Human endogenous retrovirus-K contributes to motor neuron disease. *Sci. Transl. Med.* 2015;7(307):307ra153-307ra153.
11. Ma W et al. Human Endogenous Retroviruses-K (HML-2) Expression Is Correlated with Prognosis and Progress of Hepatocellular Carcinoma. *BioMed Res. Int.* 2016;2016:1–9.
12. Strissel PL et al. Reactivation of codogenic endogenous retroviral (ERV) envelope genes in human endometrial carcinoma and prestages: Emergence of new molecular targets. *Oncotarget* 2012;3(10):1204–1219.
13. Johanning GL et al. Expression of human endogenous retrovirus-K is strongly associated with the basal-like breast cancer phenotype [Internet]. *Sci. Rep.* 2017;7. doi:10.1038/srep41960
14. Depil S, Roche C, Dussart P, Prin L. Expression of a human endogenous retrovirus, HERV-K, in the blood cells of leukemia patients. *Leukemia* 2002;16(2):254–259.
15. Burns KH. Transposable elements in cancer. *Nat. Rev. Cancer* 2017;17(7):415–424.
16. Attermann AS, Bjerregaard A-M, Saini SK, Grønbæk K, Hadrup SR. Human endogenous retroviruses and their implication for immunotherapeutics of cancer. *Ann. Oncol.* 2018;29(11):2183–2191.
17. De Kouchkovsky I, Abdul-Hay M. ‘Acute myeloid leukemia: a comprehensive review and 2016 update’. *Blood Cancer J.* 2016;6(7):e441–e441.
18. Herold T et al. A 29-gene and cytogenetic score for the prediction of resistance to induction treatment in acute myeloid leukemia. *Haematologica* 2018;103(3):456–465.

19. Alexandrov LB et al. Signatures of mutational processes in human cancer. *Nature* 2013;500(7463):415–421.
20. Schumacher TN, Schreiber RD. Neoantigens in cancer immunotherapy. *Science* 2015;348(6230):69–74.
21. Smith CC et al. Alternative tumour-specific antigens. *Nat. Rev. Cancer* 2019;19(8):465–478.
22. Brodsky I et al. Expression of HERV-K proviruses in human leukocytes. *Blood* 1993;81(9):2369–2374.
23. Tobiasson M et al. Comprehensive mapping of the effects of azacitidine on DNA methylation, repressive/permissive histone marks and gene expression in primary cells from patients with MDS and MDS-related disease. *Oncotarget* 2017;8(17):28812–28825.
24. Kazachenka A et al. Epigenetic therapy of myelodysplastic syndromes connects to cellular differentiation independently of endogenous retroelement derepression. *Genome Med.* 2019;11(1):86.
25. Deniz Ö et al. Endogenous retroviruses are a source of enhancers with oncogenic potential in acute myeloid leukaemia. *Nat. Commun.* 2020;11(1):3506.
26. Colombo AR et al. Suppression of Transposable Elements in Leukemic Stem Cells [Internet]. *Sci. Rep.* 2017;7. doi:10.1038/s41598-017-07356-9
27. Bendall ML et al. Telescope: Characterization of the retrotranscriptome by accurate estimation of transposable element expression. *PLoS Comput. Biol.* 2019;15(9):e1006453.
28. Corces MR et al. Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution. *Nat. Genet.* 2016;48(10):1193–1203.
29. Sondka Z et al. The COSMIC Cancer Gene Census: describing genetic dysfunction across all human cancers. *Nat. Rev. Cancer* 2018;18(11):696–705.
30. Ng SWK et al. A 17-gene stemness score for rapid determination of risk in acute leukaemia. *Nature* 2016;540(7633):433–437.
31. Reynisson B, Alvarez B, Paul S, Peters B, Nielsen M. NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. *Nucleic Acids Res.* 2020;48(W1):W449–W454.
32. O'Donnell TJ, Rubinsteyn A, Laserson U. MHCflurry 2.0: Improved Pan-Allele Prediction of MHC Class I-Presented Peptides by Incorporating Antigen Processing. *Cell Syst.* 2020;11(4):418–419.
33. Wells DK et al. Key Parameters of Tumor Epitope Immunogenicity Revealed Through a Consortium Approach Improve Neoantigen Prediction. *Cell* 2020;183(3):818-834.e13.
34. Mayer J, Blomberg J, Seal RL. A revised nomenclature for transcribed human endogenous retroviral loci. *Mob. DNA* 2011;2(1):7.
35. Iñiguez LP et al. Transcriptomic analysis of human endogenous retroviruses in systemic lupus erythematosus. *Proc. Natl. Acad. Sci.* 2019;116(43):21350–21351.
36. Zhang Y et al. Transcriptionally active HERV-H retrotransposons demarcate topologically associating domains in human pluripotent stem cells. *Nat. Genet.* 2019;51(9):1380–1388.

37. Feschotte C. Transposable elements and the evolution of regulatory networks. *Nat. Rev. Genet.* 2008;9(5):397–405.
38. Saini SK et al. Human endogenous retroviruses form a reservoir of T cell targets in hematological cancers. *Nat. Commun.* 2020;11(1):5660.
39. Rakoff-Nahoum S et al. Detection of T lymphocytes specific for human endogenous retrovirus K (HERV-K) in patients with seminoma. *AIDS Res. Hum. Retroviruses* 2006;22(1):52–56.
40. D’Ippolito E, Wagner KI, Busch DH. Needle in a Haystack: The Naïve Repertoire as a Source of T Cell Receptors for Adoptive Therapy with Engineered T Cells. *Int. J. Mol. Sci.* 2020;21(21):8324.
41. The Cancer Genome Atlas Research Network. Genomic and Epigenomic Landscapes of Adult De Novo Acute Myeloid Leukemia. *N. Engl. J. Med.* 2013;368(22):2059–2074.
42. Tyner JW et al. Functional genomic landscape of acute myeloid leukaemia. *Nature* 2018;562(7728):526.
43. Orenbuch R et al. arcasHLA: high-resolution HLA typing from RNAseq. *Bioinforma. Oxf. Engl.* 2020;36(1):33–40.
44. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 2012;9(4):357–359.
45. Li H et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 2009;25(16):2078–2079.
46. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* 2015;31(2):166–169.
47. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15(12):550.
48. Hanahan D, Weinberg RA. Hallmarks of Cancer: The Next Generation. *Cell* 2011;144(5):646–674.
49. Liberzon A et al. The Molecular Signatures Database Hallmark Gene Set Collection. *Cell Syst.* 2015;1(6):417–425.
50. Loeffler-Wirth H et al. A modular transcriptome map of mature B cell lymphomas. *Genome Med.* 2019;11(1):27.
51. Hubert M et al. IFN-III is selectively produced by cDC1 and predicts good clinical outcome in breast cancer [Internet]. *Sci. Immunol.* 2020;5(46). doi:10.1126/sciimmunol.aav3942
52. Hänzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-Seq data. *BMC Bioinformatics* 2013;14(1):7.
53. Thorsson V et al. The Immune Landscape of Cancer. *Immunity* 2018;48(4):812–830.e14.
54. Zhu A, Ibrahim JG, Love MI. Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences. *Bioinformatics* 2019;35(12):2084–2092.
55. Madeira F et al. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* 2019;47(W1):W636–W641.
56. Consortium TU. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* 2019;47(D1):D506–D515.

## SUPPLEMENTAL DATA

### SUPPLEMENTAL TABLES

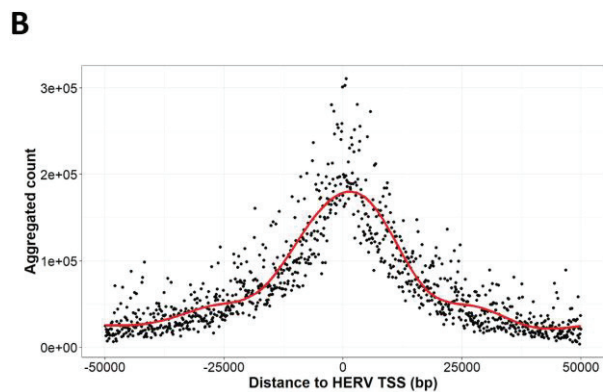
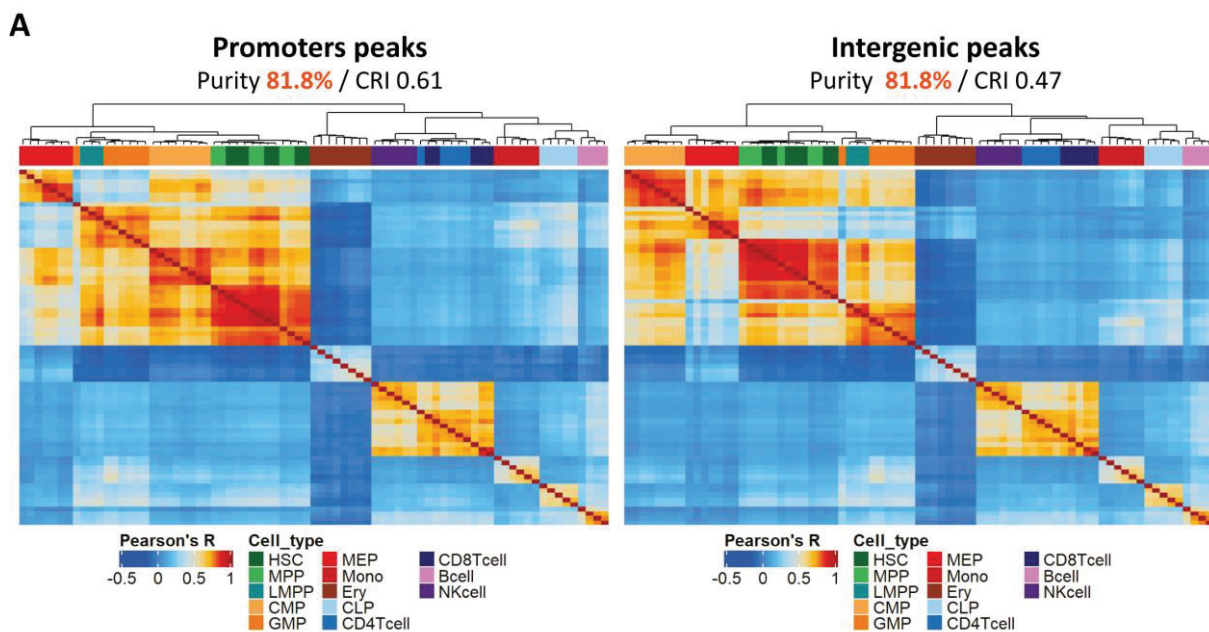
Supplemental table 1: Correlations between HERVs in active chromatin regions and surrounding cancer-related genes.

Supplemental table 2: List of included patients from the 4 independent studies

Supplemental table 3: Cancer hallmark signatures used in the gene-set variation analysis

Supplemental table 4: List of 47 HERVs used in the LSC signature

### SUPPLEMENTAL FIGURES



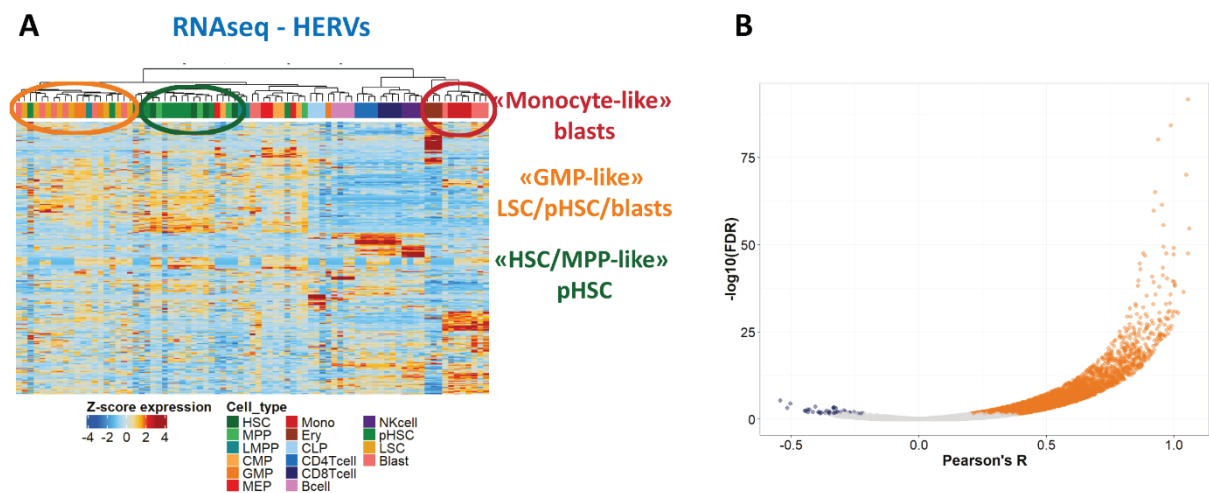


## Supplemental figure 1

D. Unsupervised hierarchical clustering of normal hematopoietic cell populations based on promoters TSS (left) or intergenic peaks (right) found in ATAC-seq.

E. Aggregated count surrounding HERV TSS in ATAC-seq.

*CLP: Common Lymphoid Progenitor, CMP: Common Myeloid Progenitor, CRI: Corrected Rand Index, Ery: Erythrocyte, FDR: False Discovery Rate, GMP: Granulocyte-Macrophage Progenitor, HSC: Hematopoietic Stem cell, LMPP: lymphoid-primed multipotent progenitor, MEP: Megakaryocyte-Erythroid Progenitor, MPP: Multipotent Progenitor, TSS: Transcription Start Site*

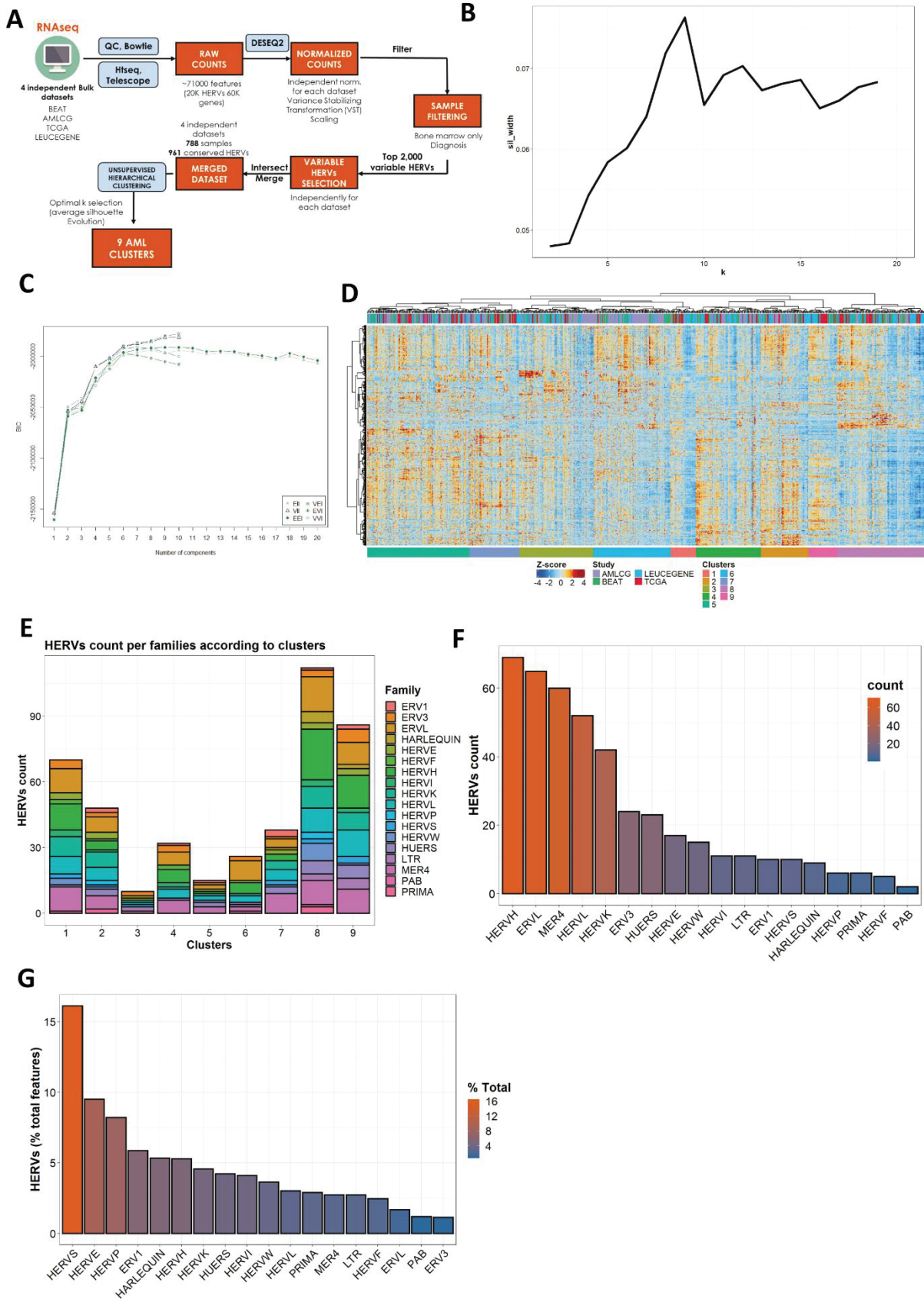


## Supplemental figure 2

E. Unsupervised hierarchical clustering of normal and leukemic hematopoietic cell populations based on HERV expression in RNA-seq.

F. Correlation between the RNA expression of HERVs located in active chromatin regions and the surrounding genes (+/- 50,000 bp). Significant Pearson's R are represented in orange (positive) or blue (negative).

*CLP: Common Lymphoid Progenitor, CMP: Common Myeloid Progenitor, CNV: Copy Number Variation, Ery: Erythrocyte, FDR: False Discovery Rate, GMP: Granulocyte-Macrophage Progenitor, HSC: Hematopoietic Stem cell, LMPP: lymphoid-primed multipotent progenitor, LSC: Leukemic Stem Cell, MEP: Megakaryocyte-Erythroid Progenitor, MPP: Multipotent Progenitor, pHSC: pre-leukemic Hematopoietic Stem Cell*

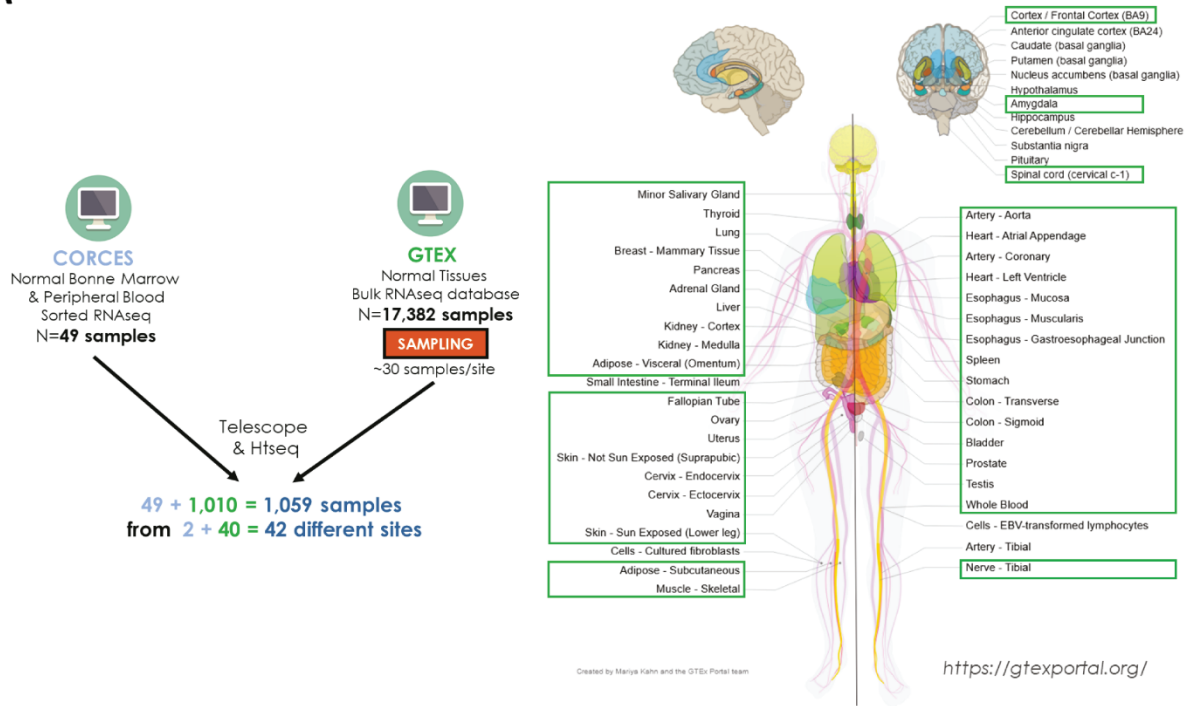


### Supplemental figure 3

- A. Global Workflow. HERV and Gene expression were independently calculated and DESEQ2 VST normalized in the 4 datasets. HERV expression was then center-scaled for each dataset to get rid of the batch effect. The top 2,000 variable HERVs were then selected for each dataset, keeping only bone marrow AML samples at diagnosis (n=788). The conserved variable HERVs were then selected, resulting in the final analysis of 961 HERVs. Unsupervised hierarchical clustering was then performed, using the silhouette evolution and BIC to select the optimal number of clusters.
- B. Average silhouette evolution according to the number of clusters k.
- C. BIC evolution according to the number of clusters k.
- D. Heatmap showing a visual representation of the 9 clusters from the unsupervised hierarchical clustering performed on the 961 conserved variable HERVs across the 4 datasets.
- E. Total number of discriminating HERVs (count) per clusters. HERVs overexpressed in a given cluster compared to all the others are represented.
- F. Most represented families among cluster-discriminating HERVs (total count).
- G. Most represented families among cluster-discriminating HERVs (percent of total family members).

*BIC : Bayesian information criterion, QC: Quality Check*

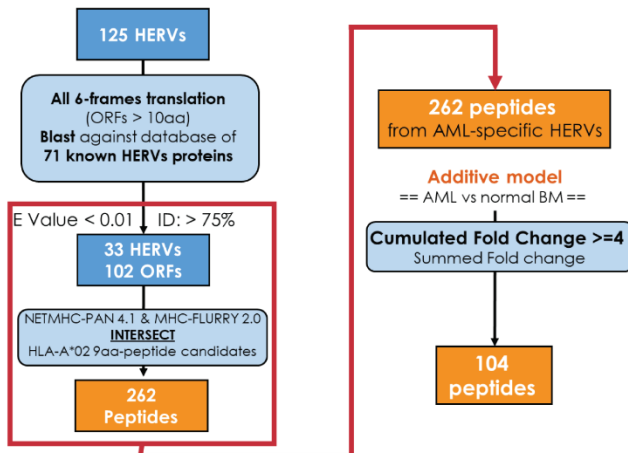
**A**



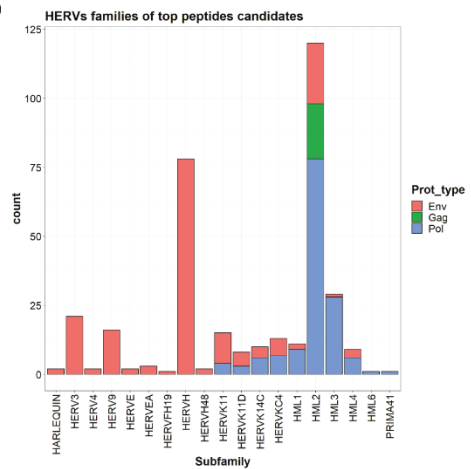
**B**

Top 20 overexpressed features in AML bulk						Top 20 overexpressed features in LSC					
Feature_id	BaseMean	Log2_FC	FC	p.value	p.adj	Feature_id	BaseMean	Log2_FC	FC	p.value	p.adj
ERV316A3_1q25.2b	4.33	17.47	181549.07	1.65E-59	3.86E-58	ERV316A3_1q25.2b	4.33	16.58	97966.44	7.43E-20	1.62E-17
ERV316A3_8q11.21a	2.22	12.58	6122.9	5.46E-12	1.76E-11	ERV316A3_8q11.21a	2.22	11.58	3061.45	1.30E-05	1.82E-04
MER34B_1q23.3b	404.91	10.66	1618	1.24E-70	4.53E-69	MER34B_1q23.3b	404.91	7.13	140.07	7.43E-16	9.13E-14
HUERSP3_8q11.1	58.12	10.24	1209.34	5.32E-23	4.67E-22	HERVH_12q15b	5.35	4.75	26.91	5.90E-05	6.77E-04
HML3_16p13.3	288.51	6.25	76.11	4.14E-81	1.90E-79	LITRAP	1039.51	4.59	24.08	1.94E-13	1.59E-11
ERVLB4_8q11.21b	3.42	5.61	48.84	4.60E-04	0.001	MER41_15q24.2	3.1	4.35	20.39	4.14E-09	1.36E-07
ERV316A3_22q13.1	32.22	5.28	38.85	9.90E-22	4.87E-21	MER41_7q34b	294.26	3.88	14.72	1.66E-05	2.25E-04
ERVLB4_17q25.1	31.21	4.9	29.86	5.30E-22	7.60E-21	HAVCR2 (TIM3)	353.21	3.64	12.47	2.34E-08	6.50E-07
HERVL18_11q12.2	12.86	4.68	25.63	3.43E-13	1.24E-12	CD96	944.54	3.63	12.38	1.72E-07	3.88E-06
ERVLE_8q12.3f	72.8	4.53	23.1	6.92E-09	2.41E-08	HML3_16p13.3	288.51	3.51	11.39	9.91E-12	6.00E-10
ERVLB4_22q11.21a	5.73	4.37	20.68	2.27E-09	6.47E-09	ERVLE_14q11.2d	4.44	3.41	10.63	1.01E-04	0.0010
HERVL11q12.2	336.09	4.26	19.16	5.88E-22	4.25E-21	PRIMA41_17q21.31b	102.65	3.4	10.56	1.02E-10	4.91E-09
LTR25_12p11.22a	6.12	4.18	18.13	1.59E-10	6.17E-10	ERV316A3_22q13.1	32.22	3.28	9.71	3.15E-07	6.65E-06
HERV9_3p24.3	15.99	4.07	16.8	1.97E-08	8.79E-08	HERVH_12p11.22a	90.21	3.11	8.63	3.24E-06	5.37E-05
ERVLE_11q13.3b	5.79	4.05	16.56	2.40E-07	6.01E-07	ERV316A3_11q13.1b	13.98	2.84	7.16	3.52E-05	4.31E-04
LTR23_9q22.33b	38.15	3.97	15.67	7.70E-24	6.11E-23	MER61_6q24.1b	19.13	2.71	6.54	1.14E-04	0.0012
HERVW_3p21.31	6.43	3.91	15.03	2.79E-10	1.34E-09	HERVIP10FH_14q11.2	5.66	2.56	5.9	0.0022	0.0146
HML3_4q31.3	420.96	3.68	12.82	2.12E-21	1.76E-20	ERVLB4_8q11.21b	3.42	2.51	5.7	0.0053	0.0297
HERVH_12p11.22a	90.21	3.68	12.82	6.77E-10	1.96E-09	HERVW_1p33a	10.92	2.47	5.54	8.89E-07	1.68E-05
HERVH_12q15b	5.35	3.66	12.64	0.0114	0.0213	HML3_4q31.3	420.96	2.33	5.03	2.25E-05	2.93E-04

**C**



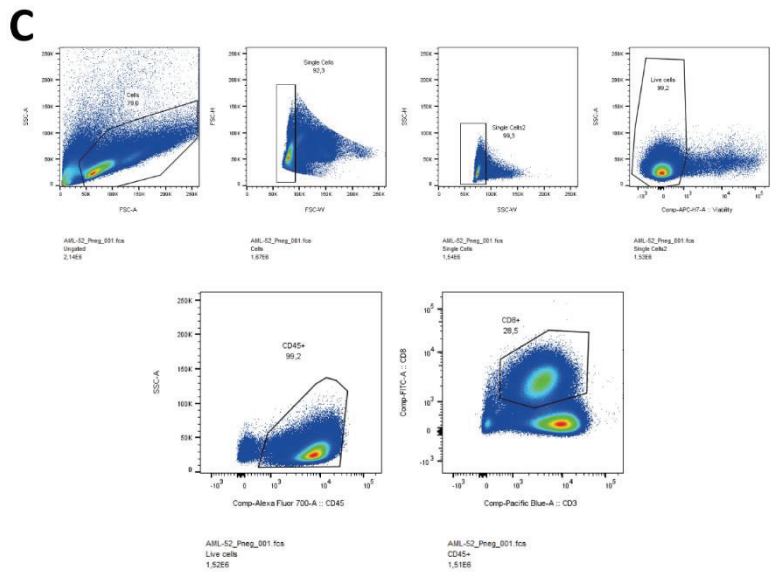
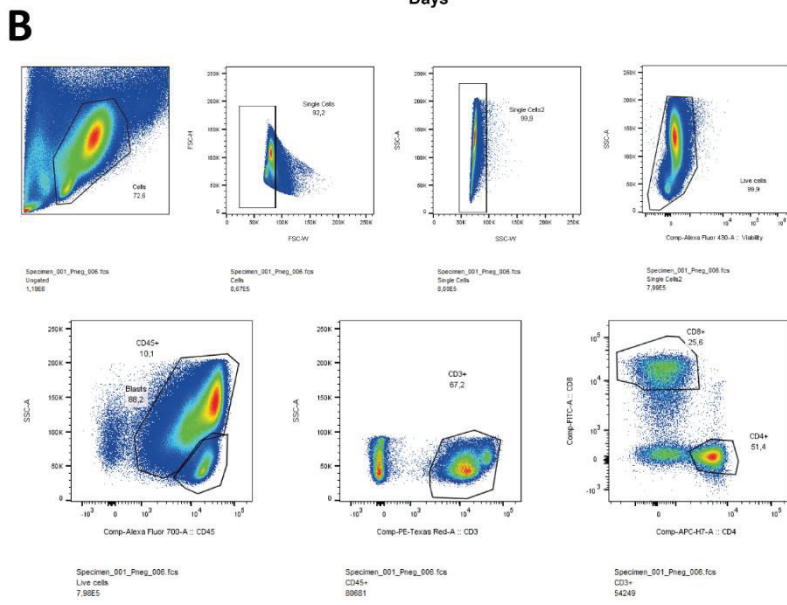
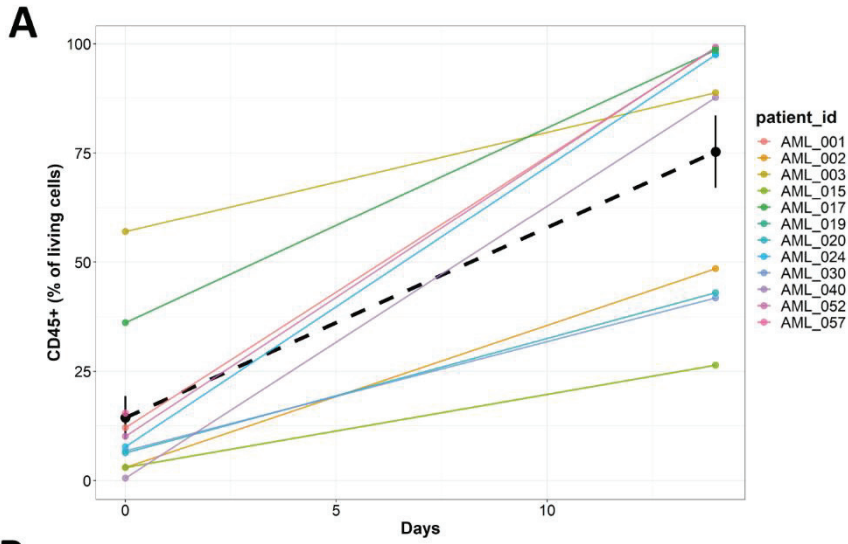
**D**



#### Supplemental figure 4

- A. Global Workflow and graphical representation of the sampled normal tissues.
- B. Table of the top 20 overexpressed features in the bulk datasets (left) or in the LSC (right) compared to normal bone marrow.
- C. Workflow representing the different steps to identify CD8+ T cells epitopes from HERVs sequences. HERV sequences are translated in the 6 frames and then aligned to a database of 71 known HERV proteins (Gag, Pol and Env). ORFs with at least 75% identity with a known HERV protein and a E-value < 0.01 are then retained for epitope prediction. Two different tools, netMHC-pan 4.2 and MHC-Flurry 2.0 are used to predict the top 9amino-acids HLA-A\*02 peptide candidates. To further prioritize these candidates, an additive model is proposed to take into account each peptide' expression level.
- D. Number of HERVs per families containing the top peptides candidates.

*LSC : Leukemic Stem Cell, ORF: Open Reading Frame*





## Supplemental Figure 5

A. Percentage of CD45+ cells among living cells before (day 0) and after 14 days of expansion in high dose IL2.

B. Gating strategy of MILs at day 0

C. Gating strategy of TILs at day 14

*MILs : Bone Marrow Infiltrating Lymphocytes*



## DISCUSSION

---



Nos travaux viennent étayer le rationnel de l'utilisation des HERVs en immunoncologie. Dans le premier article, nous proposons pour la première fois une approche bioinformatique systématique, applicable à grande échelle et reproductible pour l'identification d'épitopes lymphocytaires T CD8<sup>+</sup> au sein des HERVs les plus exprimés dans une tumeur donnée. Nous proposons une liste d'épitopes partagés et fortement exprimés dans plusieurs cancers, et validons ensuite cette approche en montrant que les épitopes sélectionnés permettent d'induire une réponse LT CD8<sup>+</sup> de forte avidité et cytotoxique contre des lignées cellulaires et organoïdes de cancer du sein triple négatif. Dans le deuxième article, nous allons plus loin en montrant que chaque population cellulaire possède sa propre signature rétrotranscriptomique, ouvrant ainsi la voie à l'utilisation des HERVs pour la mise au point de signatures cellulaires plus précises pouvant favoriser l'identification de profils transcriptomique prédictifs d'une issue clinique d'intérêt. Nous discuterons ici ces résultats en nous concentrant d'abord sur l'aspect de ciblage thérapeutique de ces épitopes partagés avant de nous concentrer sur l'utilisation des HERVs en tant que biomarqueurs. Enfin, nous proposerons de raffiner le modèle actuel de l'impact physiologique et physiopathologique des HERVs en fonction de leur expression, que nous tenterons de résumer en un seul schéma.

## **I. Cible thérapeutique**

Comme nous l'avons abordé dans l'introduction, un bon épitope candidat pour une approche immunothérapeutique est un épitope à la fois exprimé en quantité suffisante à la surface de la cellule tumorale, non (ou faiblement) présent dans le tissu sain et possédant une forte immunogénicité. Bien que ne reposant que sur des données de séquençage ARN, notre stratégie d'identification d'épitopes LT CD8<sup>+</sup> présente plusieurs originalités qui en font sa force, l'évaluation biologique que nous développons ensuite venant confirmer leur validité.

### **A. Spécificité et expression par le tissu sain**

L'expression potentielle de peptides issus des HERVs par le tissu normal est probablement la limite majeure à leur utilisation en tant que cibles thérapeutiques. Si certaines équipes comme le groupe de Claude Perreault proposent des filtres très restrictifs (39,82), d'autres auteurs ne proposent qu'un contrôle partiel par RT-PCR (268) ou immuno-histochimie (196) dans certains tissus sains, voir même aucun contrôle comme dans l'étude récente utilisant des multimères à code-barres ADN pour identifier des épitopes T CD8<sup>+</sup> dans différentes hémopathies myéloïdes (199).

Afin de minimiser ce risque, nous avons développé un filtre stringent basé sur le comparatif de l'expression des HERVs entre la tumeur et son tissu péri-tumoral. L'utilisation de la base pancancer nous aura par ailleurs permis de filtrer de façon systématique tout HERV davantage exprimé dans un tissu péri-tumoral par rapport à sa tumeur correspondante. Cela signifie par exemple que si un HERV est retrouvé surexprimé dans le tissu péri-tumoral de thyroïde par rapport à son cancer correspondant, il est exclu de l'ensemble de l'analyse et ne pourra pas être retenu comme épitope dans le cancer du sein. Bien que permettant de filtrer de nombreux HERVs, l'utilisation du tissu péri-tumoral en tant que contrôle est discutable, celui-ci pouvant présenter différentes altérations (inflammation, infiltration immunitaire, hypervascularisation, présence éventuelle de cellules pré-néoplasiques) susceptibles de modifier son profil HERV. A cet effet, nous avons contrôlé l'expression de nos HERVs cibles dans une

base indépendante de séquençage ARN comprenant des échantillons de cancers du sein, de tissu péri-tumoral mais également du vrai tissu sain issu de réduction mammaire. Ce comparatif nous a permis de confirmer d'une part la surexpression de nos cibles dans le tissu cancéreux, mais également de voir que l'expression dans le tissu sain est encore plus basse que dans le tissu péri-tumoral (**Article 1** - Figure 3D). L'utilisation de tissu péri-tumoral en contrôle semble donc plutôt pénaliser l'analyse.

Dans le deuxième article basé sur Telescope, nous avons pu utiliser du tissu normal comme contrôle, nous permettant de mettre en place une base de contrôle de tissus sains constituée de 1 059 échantillons provenant de 42 tissus sains différents (**Article 2**, Figure S4A). Cette base nous a permis de filtrer tout HERV plus exprimé dans n'importe quel tissu sain par rapport aux échantillons de moelle osseuse de patients atteints de LAM. Cette approche, plus restrictive, est particulièrement intéressante dans les cas où une cible pourrait être surexprimée par rapport à son tissu sain correspondant, mais sous-exprimée par rapport à un tissu sain à distance (exemple : surexpression d'un HERV dans la moelle osseuse de LAM par rapport à la moelle osseuse saine, mais sous-expression par rapport au tissu hépatique sain). De plus, le focus réalisé sur une seule pathologie en utilisant 4 bases de données indépendantes nous permet de mieux contrôler les faux positifs. Ainsi nous avons pu voir dans cet article qu'une quarantaine de HERVs surexprimés dans la base du TCGA n'étaient retrouvés dans aucune autre base, correspondant ainsi probablement à un biais technique ou effet *batch*.

## **B. Sélection des épitopes**

### **1) De la séquence nucléotidique au peptide**

Une fois notre liste de HERVs établie, les meilleurs épitopes parmi ces candidats restent à définir. Cela passe habituellement par la traduction des séquences nucléotidiques en protéines potentielles. En l'absence de *gold standard*, de nombreux auteurs proposent de réaliser une traduction systématique des séquences nucléotidiques dans les 6 cadres ouverts de lecture afin de considérer tous les ORFs potentiels (269,270). La taille minimale d'une protéine étant de 10 acides aminés (271), le nombre de séquences générées peut être très important, avec par exemple une base de ~3,2 Gb obtenue lors de la traduction systématique de l'ensemble du génome humain, soit 70 fois plus que le protéome connu de référence de la base Ensemble (~45 mb) (272). Certains auteurs proposent d'utiliser un seuil arbitraire plus restrictif de 70 ou 100 acides aminés par exemple, permettant de réduire le nombre d'ORFs générés (265). Dans notre approche, nous filtrons les ORFs de nos séquences HERVs d'intérêt en les alignant sur des protéines connues de HERVs. Dans le premier article, nous cherchions à identifier des peptides très conservés issus de HERV-K : nous nous sommes donc restreints aux protéines Gag et Pol connues de HERV-K, permettant ainsi de passer de 76 934 ORFs définis pour nos 192 Cyt-HERVs à 27 ORFs uniques ayant une identité >90% avec les protéines de référence. Pour le deuxième article, nous avons généralisé l'approche à l'ensemble des protéines publiées de HERVs (71 références validées dans Uniprot) sur la base plus exhaustive de Telescope. Sur plus de 3 millions d'ORFs reconstruits sur l'ensemble de la base (14 968 HERVs), seuls 1 615 possèdent une identité > 90% avec les protéines connues de HERVs. Cette approche nous permet ainsi de réduire drastiquement le nombre d'ORFs testés, tout en pouvant jouer sur le paramètre d'identité pour obtenir des protéines plus ou moins dégénérées (au risque d'augmenter le nombre de faux positifs). Toujours dans la LAM, nous pouvons ainsi définir plusieurs stratégies en partant de nos



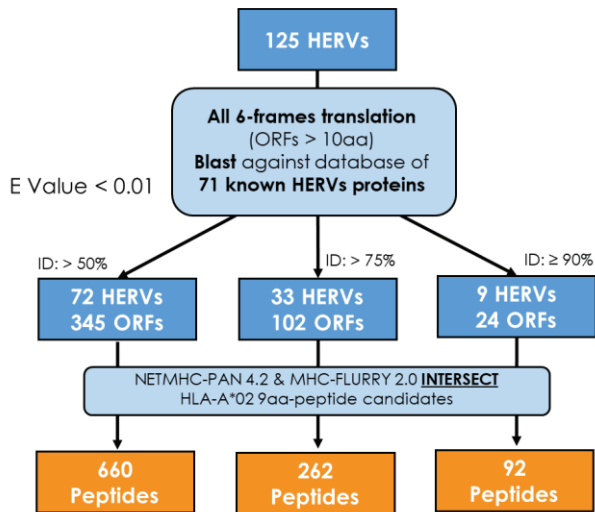
125 HERVs exprimés de façon spécifique dans la LAM, menant à la validation finale de 660 peptides pour les 345 ORFs avec au moins 50% d'identité ou 92 peptides pour les 24 ORFs avec au moins 90% d'identité (Figure 7).

Une approche d'emblée plus restreinte aurait pu se discuter, en se limitant à une base d'ORFs pré-identifiés comme celle de Retrotector ou gEVE par exemple (100,184), voir à une base très restreinte de protéines connues comme avec l'approche « Mayer ». Bien que permettant de limiter davantage les faux positifs, ces approches sont cependant limitées et excluent toute découverte d'épitope présent dans des ORFs encore non annotés.

## **2) Approche prédictive**

Les épitopes sont ensuite définis à partir de ces ORFs via des outils permettant de prédire les séquences ayant le plus de chance d'être présentées à la surface de la cellule par le CMH de classe I, selon des critères d'affinité pour le CMH, de chance d'être clivés à un niveau donné ainsi que de transport par la protéine TAP (84,85). Ces outils sont régulièrement critiqués du fait des performances limitées obtenues dans le contexte des antigènes issus de mutation, avec moins de 1% des épitopes menant effectivement à une réponse immunogène (273–275). Par rapport à ces études, notre approche diffère en de nombreux points :

- Dans les cibles choisies tout d'abord, le fait d'utiliser des séquences d'origine virales que sont les HERVs permettant de maximiser des critères tels que la dissimilarité avec le soi, augmentant fortement la chance d'avoir des cibles immunogènes (276). Des données de spectrométrie de masse comparatives suggèrent par ailleurs que les HERVs auraient une capacité supérieure à produire des MAPs par rapport aux autres rétroéléments (138). Si cette différence peut facilement se concevoir pour les SINE (qui sont dépourvus de séquence codante pour des protéines), la différence avec les LINE est moins évidente à expliquer.
- Le niveau d'expression ensuite, directement associé à la présentation par le CMH (277). Contrairement aux études sur les antigènes issus de mutations, nous prenons ici en compte le niveau d'expression du HERVs. Dans le deuxième article, nous proposons également un modèle additif permettant de prioriser les épitopes ayant un fort niveau d'expression. Cela paraît d'autant plus important que la densité d'épitope est également associée à l'immunogénicité (278). Il n'existe cependant pas de seuil clairement établi pour guider les choix. Dans l'étude de Ehx G. *et al.* que nous détaillerons un peu plus loin, les auteurs montrent en se basant sur des données couplées de séquençage ARN et spectrométrie de masse qu'une séquence exprimée à moins de 8,55 *reads* par centaine de millions de reads mappés (RPHM) a une probabilité < 5% de générer un MAP (82). D'autres études proposent des seuils arbitraire de 1 TPM, permettant également de ne pas artificiellement augmenter les *fold change* des études comparatives sur des séquences très faiblement exprimées (265). Nous nous sommes alignés sur cette approche, en choisissant de ne pas considérer les séquences ayant moins de 1 CPM d'expression moyenne



**Figure 7 : Sélection des ORFs à partir des séquences HERVs**

Les HERVs présélectionnés sur des critères d'expression forte dans le tissu tumoral et faible/absente dans les tissus sains sont traduits dans les 6 sens de lecture. L'ensemble des ORFs de plus de 10 acides aminés est ensuite aligné face à une base de données de 71 protéines HERVs connues issue d'Uniprot.

ID : Identité, ORFs : Open Reading Frames (Cadre Ouvert de Lecture)

mutations, les auteurs se sont aperçus que, bien que les cibles vaccinales aient été initialement prédites sur des critères d'affinité pour le CMH de classe I, la réponse LT CD4<sup>+</sup> était finalement prédominante chez les patients, avec une réponse LT CD4<sup>+</sup> polyfonctionnelle détectée pour 60% des épitopes prédits vs 20% seulement pour les LT CD8<sup>+</sup> (282). Ces résultats rejoignent d'autres données chez la souris soulignant l'importance de la réponse LT CD4<sup>+</sup> dans la réponse immunitaire dirigée contre les antigènes tumoraux issus de mutations (283). Bien que les outils de prédiction épitopiques pour le CMH de classe II soient moins performants du fait de la nature même de la relation CMH-peptide (les peptides étant plus long que ceux présentés par le CMH de classe I d'une part, et les positions d'ancrages plus variées, multipliant ainsi le nombre de possibilités), ceux-ci progressent avec des performances meilleures dans les développements récents (279,284,285). Ces différentes données nous ont poussé à faire le choix de nous concentrer sur des motifs épitopiques présentés par le CMH de classe I pour la preuve de concept dans un premier temps. Nous prévoyons pour la suite d'intégrer ces motifs dans des peptides plus longs dont la sélection sera guidée par les outils récents de prédiction pour le CMH de classe II, afin de valider l'immunogénicité de ces longs peptides en monitorant les réponses LT CD4<sup>+</sup> et CD8<sup>+</sup>.

### 3) Approche protéomique (spectrométrie de masse)

La spectrométrie de masse est actuellement la méthode de référence pour démontrer la présence d'un épitope effectivement présenté à la surface de la cellule par le CMH de classe I. L'approche classique dite « *shotgun* » couple l'isolation des peptides d'un échantillon donné par chromatographie liquide au séquençage de leur spectre de masse en tandem par un spectromètre de masse (272). Un des points critiques de cette approche est que l'identification des peptides est réalisée par comparaison des spectres de l'échantillon à une base de spectres possibles établie à partir d'un protéome de référence. En plus de restreindre fortement les peptides potentiels,

- Les outils utilisés enfin, les nouvelles versions des principaux outils que nous utilisons tels que netMHCpan 4.1 et MHC Flurry 2.0 étant entraînées sur des données de spectrométrie de masse avec de très bonnes performances (279,280).

Notre approche exploite donc pleinement les connaissances et outils disponibles pour optimiser la chance de détecter un épitope présenté par le CMH de classe I. Il n'en demeure pas moins plusieurs limitations, en premier lieu desquelles l'absence d'étude des épitopes présentés par le CMH de classe II. Le rôle des LT CD4<sup>+</sup> dans l'immunité antitumorale est pourtant majeur, que ce soit par les populations LT *helper* mais également avec les populations LT CD4<sup>+</sup> cytotoxiques, de plus en plus étudiées (17,281). Dans un des essais cliniques fondateurs de vaccination thérapeutique utilisant des pools de peptides longs définis à partir d'antigènes issus de

L'utilisation du protéome canonique en base de référence est exposée à un taux important de faux positifs, la recherche de cibles non exprimées dans l'échantillon augmentant la taille de la base de données, directement corrélée au nombre de peptide identifiés. Ces limites ont ouvert le champ de la protéogénomique, permise par les gains en puissance de calcul et les avancées bio-informatiques, restreignant désormais la recherche de spectres protéiques au sein d'une base ciblée sur le profil d'expression génomique de l'échantillon (83,272). En plus de valider la présence d'un transcrit fortement exprimé au niveau protéique, cette approche permet également de découvrir de nouvelles protéines et ainsi alimenter les bases de données telles qu'Uniprot (183).

La plupart des équipes référentes dans le domaine dont celles de Claude Perreault et de Michal Bassani Sternberg s'intéressent désormais aux portions non codantes du génome, source de la majorité des MAPs (39,83,277). Comme nous l'avons vu, les MAPs dérivent préférentiellement de produits ribosomiaux défectueux (DRIPs) (21,23). Les HERVs, en particulier ceux des familles les plus anciennement intégrées, étant dégénérés, nous pourrions nous attendre à ce qu'ils soient enclins à produire des peptides instables de ce type.

L'étude récente de Ehx G. *et al.* dans la LAM vient même intégrer des paramètres d'expression ARN et des filtres de tissu sain (cellules souches hématopoïétiques et mTECs) pour augmenter la probabilité d'identification de MAPs dits d'intérêt, spécifiques de la tumeur et fortement exprimés par celle-ci (82). D'une façon intéressante, aucun antigène muté n'était présent dans les 58 MAPs identifiés comme étant les plus prometteurs (exprimés à moins de 8,55 RPHM dans les tissus sains et surexprimés dans les moelles osseuses leucémiques par rapport aux moelles osseuses normales par un facteur 5 minimum), dont 86% étaient dérivés de portions non codantes du génome et 48% secondaires à une rétention intronique, phénomène particulièrement important dans la LAM (286). Ces MAPs présentent également l'avantage d'être partagés, les patients présentant au moins deux mutations parmi les gènes NPM1, FLT3 ou DNMT3A étant parmi ceux exprimant le plus de MAPs différents, en lien avec des profils aberrants de méthylation.

Malgré la performance de ces approches de spectrométrie de masse et leur importante spécificité, celles-ci peuvent manquer de sensibilité notamment pour les épitopes présents en faible densité à la surface cellulaire (272). Une autre méthode de spectrométrie de masse dite ciblée permet de marquer les peptides d'intérêt afin d'orienter l'acquisition des spectres de masse et quantifier de façon plus précise la présence d'un ou plusieurs peptides cibles (287). Ce type d'approche bien que limitée à certaines cibles permet d'atteindre une sensibilité maximale pour la détection d'épitopes dont on connaît déjà la séquence. Une étude récente publiée dans Science a utilisé cette approche pour quantifier un antigène issu d'une mutation de TP53 (p53<sup>R175H</sup>) sur différentes lignées cellulaires, allant jusqu'à la détection de 1,5 copies de l'antigène mutant présenté par le HLA-A\*02 :01 à la surface de la lignée TYK-nu (carcinome ovarien), contre plus de 700 pour leur condition contrôle (lignée COS-7 transfectée avec les gènes HLA-A :02 :01 et TP53<sup>mut</sup>) (288). Bien que la densité d'un tel épitope pourrait amener à questionner son intérêt en tant que cible tumorale, la mise au point par les auteurs d'un anticorps de forte affinité ciblant le complexe HLA-peptide et permettant la lyse efficace de ces mêmes lignées malgré une densité très faible d'antigène présenté ouvre le champ de nombreuses possibilités thérapeutiques pour des anticorps bispécifiques ou des lymphocytes CAR-T.

## C. Immunogénicité

Qu'ils aient été sélectionnés par une approche prédictive ou validés par spectrométrie de masse, l'immunogénicité des épitopes doit ensuite être validée. Là encore des outils existent pour prédire l'immunogénicité d'un épitope : celui de la base de donnée des épitopes immuns (IEDB) par exemple propose un score d'immunogénicité à partir de caractéristiques identifiées comme associées à une forte immunogénicité dans une base d'épitopes, telles que la structure des acides aminés et leur position (des acides aminés possédant de larges chaînes aromatiques, d'autant plus si ils sont en position 4 à 6, augmentant de façon majeure l'immunogénicité) (289). Ces outils sont cependant souvent restreints à des bases d'épitopes microbiens, et nos différents essais avec des épitopes dont nous connaissions l'immunogénicité potentielle ne se sont pas révélés concluants.

### 1) Annotation de l'immunogénicité

L'exhaustivité apportée par une quantification plus systématique des HERVs est contrebalancée par un risque majoré d'identification de faux positifs. Sur la seule base de 66 références HERVs de Mayer, l'utilisation d'outils de prédiction épitopique trouve plus de 200 épitopes LT CD8<sup>+</sup> de forte affinité pour le HLA-A\*02 parmi l'ensemble des ORFs potentiels. En utilisant la base *hervQuant* dans le cancer du sein, plus de 35 000 ORFs peuvent être définis à partir des ~300 HERVs surexprimés dans l'ensemble des sous-types de cancer du sein, ce qui amène à plusieurs centaines de peptides prédits de forte affinité pour le seul HLA-A\*02.

En permettant de détecter plus de 1 000 épitopes différents dans un seul et même échantillon, la technique des dextramères identifiés par code-barre ADN prend tout son intérêt dans ce type de configuration (198). Dans une étude récente déjà citée utilisant cette technologie dans différentes hémopathies myéloïdes (199), Saini S. *et al.* parviennent ainsi à détecter des réponses LT CD8<sup>+</sup> dirigées contre 29 épitopes parmi les 1 169 épitopes potentiels prédits à partir de la base de Mayer pour 4 HLA différents. Certains épitopes déjà décrits comme immunogènes ne sont cependant pas détectés par leur approche, en premier lieu desquels notre premier peptide (P1) pour lequel nous détectons une réponse LT CD8<sup>+</sup> spécifique dans quasiment l'ensemble des patients atteints de LAM testés (n=10, **Article 2**, Figure 5B). Plusieurs facteurs pourraient participer au manque de sensibilité de cette approche, notamment l'absence de stimulation des LT spécifiques, la recherche réalisée sur des PBMCs plutôt que sur des TILs ou encore les potentielles interactions entre les différents dextramères regroupés dans un seul et même échantillon.

Notre méthode de validation épitopique gagne en sensibilité ce qu'elle perd en débit. En vérifiant chaque épitope à l'aide de plusieurs tests (contrôle du *binding* HLA, test d'immunogénicité chez le donneur sain, recherche de LT spécifiques chez le patient après expansion puis test de cytotoxicité à l'aide d'un clone), cette approche ne permet pas de valider un nombre important de cibles. C'est précisément ce besoin de sélectionner un nombre restreint de cibles à valider qui nous a mené à développer des modèles visant à classer et prioriser les candidats à partir d'une liste de HERVs surexprimés dans le cancer.

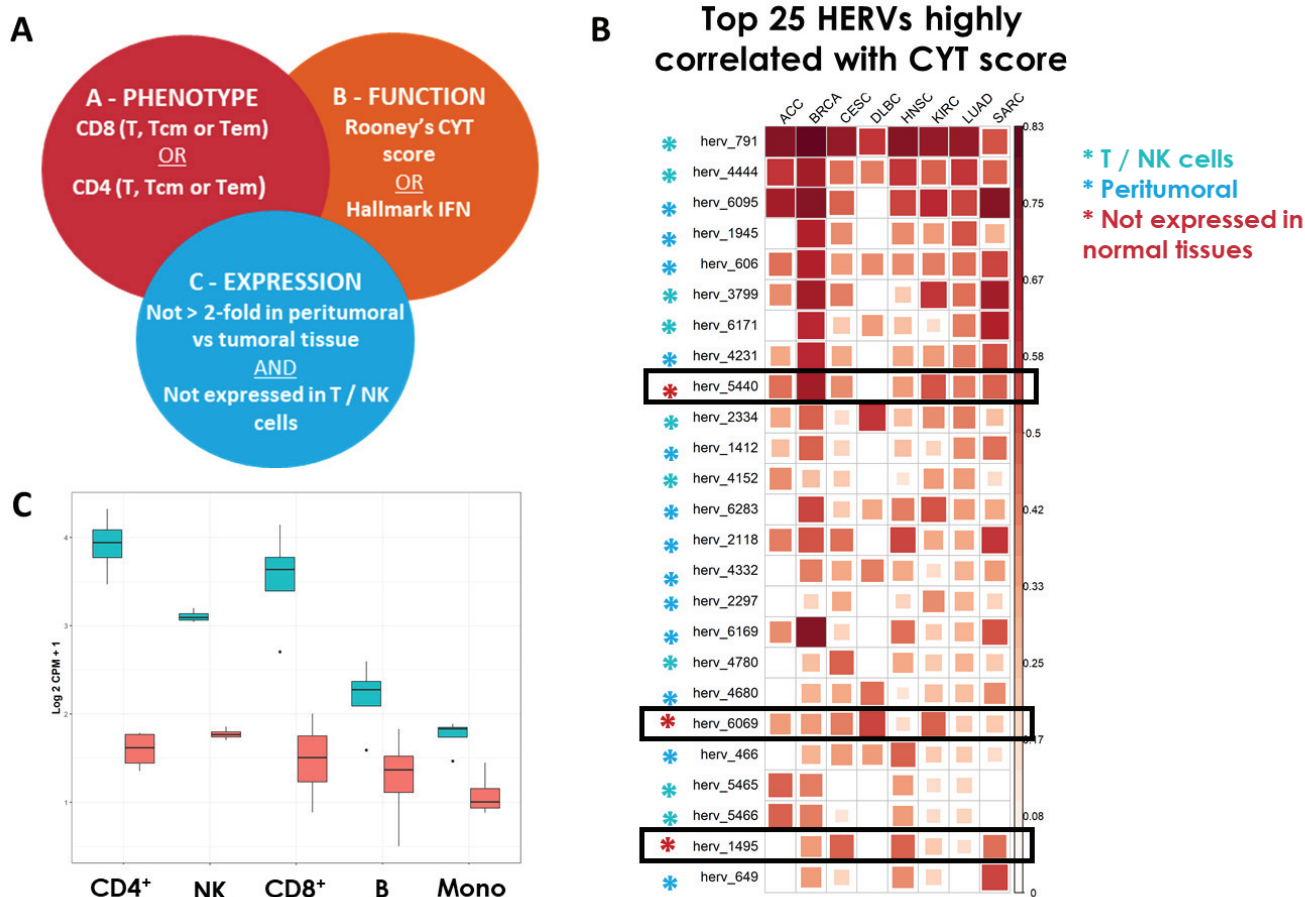
Afin d'enrichir notre liste en HERVs possédant des épitopes, le premier objectif était ainsi d'annoter les candidats qui sont associés à une réponse cytotoxique LT CD8<sup>+</sup>. L'idée de cette annotation n'est pas nouvelle en soit : en 2015, Rooney M. *et al.* avaient déjà décrit une corrélation positive entre l'expression de 7 HERVs de la famille HERV-K et un score d'activité

cytolytique (le CYT score, correspondant à la moyenne géométrique de l'expression de Granzyme A et Perforine) dans les cancers du sein et du rein (261). Cette analyse était cependant très succincte, la liste de HERVs n'étant d'une part pas exhaustive (base de Mayer), et les coefficients de corrélation de Spearman ne dépassant pas au mieux 0.4. Toujours à partir de la base de Mayer, Panda A. *et al.* sont allés un peu plus loin en définissant deux critères permettant de retenir un HERV comme potentiellement immunogènes, via la corrélation de son expression avec (i) au moins une signature d'activation immunitaire (soit l'Immune Score de l'outil ESTIMATE (290) soit l'expression de CD8A) et (ii) au moins un gène d'une voie de checkpoint immunitaire (soit PD1 ou un de ses ligands, CTLA4 ou un de ses ligands, ou un des gènes de la voie BTLA/HVEM (291)). Bien qu'intéressante, cette association repose là encore sur de simples corrélations avec des seuils par ailleurs très larges puisque toute corrélation dont le coefficient de Spearman était supérieur à 0 avec une p-value < 0.05 était retenue comme significative.

En s'inspirant de ces résultats, nous avons défini de nouveaux critères pour pouvoir annoter les HERVs associés à une réponse LT<sup>+</sup> CD8<sup>+</sup> cytotoxique de façon plus précise. Pour être retenu, chaque candidat doit être associé à une signature phénotypique T CD4<sup>+</sup> ou CD8<sup>+</sup> (Figure 8A, critère A) ET une signature de fonctionnalité (CYT-score ou signature IFN-gamma hallmark, Figure 8A, critère B) sans être exprimé dans aucuns des tissus sain péri-tumoraux analysés ou par les lymphocytes T/NK (Figure 8A, critère C). Ce double filtre phénotype/fonctionnalité nous permet de gagner en spécificité lors du choix des candidats, ne retenant que les HERVs associés à une signature T CD8<sup>+</sup> fonctionnelle dans au moins un cancer. Le filtre C est quant à lui indispensable pour capter des HERVs exprimés par la cellule tumorale et non pas par le tissu sain ou même par les cellules immunitaires elles-mêmes. En effet, si nous prenons les 25 HERVs les plus fortement corrélés au CYT-score selon l'approche de Rooney (Figure 8B), nous nous apercevons que la grande majorité est en fait soit exprimé par le tissu sain péri-tumoral, soit par les lymphocytes T et NK eux-mêmes (Figure 8C). Seuls 3 HERVs échappent au filtre C (annotés par une étoile rouge sur la figure 3B, boîtes à moustache rouges sur la figure 3C), constituant ainsi d'excellents candidats.

Parallèlement à ces critères plus restrictifs, notre calcul de l'association est également différent. Plutôt que de proposer de simples corrélations, nous avons préféré apprécier la force associative via un modèle de régression pénalisé de type L1 (lasso). Ce type d'approche est particulièrement adapté aux données dites de haute dimension où il existe de très nombreux prédicteurs (ici 3 173 HERVs), avec une colinéarité potentiellement très forte (c'est-à-dire une corrélation entre plusieurs prédicteurs) (292). La pénalisation du modèle par la technique du lasso permet par ailleurs de réaliser une sélection des variables, en n'attribuant un coefficient qu'aux variables les plus significatives et fortement associées à la variable étudiée. Pour rendre plus stables nos prédictions, nous avons enfin opté pour une approche inspirée de l'apprentissage automatique, où chaque modèle est établi indépendamment pour chaque sous-type de cancer via une validation croisée répétée 10 fois.





**Figure 8 : Annotation de l'immunogénicité potentielle des HERVs**

A : Critères d'annotation permettant de classer un HERVs comme potentiellement immunogènes. Les HERVs surexprimés dans un cancer donné doivent être associé à au moins une signature phénotypique (A) et une signature de fonctionnalité (B), sans être surexprimé par les lymphocytes T et NK ni dans aucun tissu sain (C).

B : Top 25 des HERVs les plus fortement corrélés au Cyt-Score dans 8 cancers du TCGA. Les HERVs exprimés par les LT et NK sont annotés d'une étoile verte, ceux exprimés par le tissu péri-tumoral d'une étoile bleue. Les HERVs que nous retenons comme Cyt-HERVs sont annotés en rouge et encadrés.

C : Comparatif d'expression dans les cellules sanguines périphériques des HERVs annotés comme Cyt-HERVs (rouge) comparés aux autres uniquement corrélés au Cyt-score (bleu).

Cette approche nous permet de tirer profit de la grande quantité de données fournies par la base pancancer du TCGA en fournissant ainsi des forces associatives stables robustes. Par rapport à une analyse dans un cancer unique, l'annotation d'immunogénicité sur une base pancancer nous permet également de récupérer certains HERVs potentiellement immunogènes mais non associés à une réponse cytotoxique dans certaines tumeurs du fait d'un microenvironnement immunosuppresseur ou non favorable (tumeurs froides, en particulier tumeurs à exclusion immunitaire).

## 2) Validation de l'immunogénicité

Notre approche de validation de l'immunogénicité se fait en 2 temps. Dans la première partie, nous réalisons des tests classiques de *priming* et de fonctionnalité à partir de lymphocytes de donneurs sains afin de vérifier l'immunogénicité du peptide. Pour les peptides ayant passé la première étape, nous établissons ensuite des clones de LT CD8<sup>+</sup> toujours à partir de PBMCs de donneurs sains afin de tester l'avidité et la cytotoxicité contre des lignées cellulaires d'intérêt



exprimant la cible. Notre approche n'incluant pour le moment pas d'analyse protéomique par spectrométrie de masse, nous utilisons cette étape pour valider indirectement la présence effective de l'épitope à la surface de la cellule cible.

Si cette validation a l'avantage de fournir de nombreuses données de fonctionnalité et confirmer l'intérêt de la cible par la démonstration de l'étape finale espérée (la destruction de la cellule tumorale par le LT spécifique), elle présente plusieurs limites. L'établissement des clones lui-même tout d'abord, qui est un processus long et laborieux, exposé à un risque non négligeable d'échec à différentes étapes (mortalité cellulaire pendant les phases de tri cellulaire, expansions répétées exposant à l'épuisement du clone, contamination par un clone non spécifique prenant l'ascendant sur le clone voulu...).

La spécificité du TCR ensuite qui, du fait de la cross-réactivité théorique, ne garantit pas la présentation spécifique de la cible à la surface de la cellule. Ce phénomène de cross-réactivité n'est pas facile à quantifier : si certaines études extrapolant à partir de données expérimentales rapportent jusqu'à 1 millions de peptides différents pouvant être reconnus par un même TCR (293), d'autres peinent à corroborer ces fréquences élevées (294). Par exemple dans un article étudiant plus de 30 000 interactions CMH-peptide-TCR pour 15 clones de LT CD8<sup>+</sup> humains et murins, seule 1 réaction croisée a pu être détectée (295). Il n'en reste pas moins que cette cross-réactivité peut être un problème majeur si elle n'est pas évaluée, comme l'ont montré les toxicités sévères et décès suites à l'utilisation de TAAAs en tant que cibles thérapeutiques (voir Introduction – partie II-B-1). A partir des 20 acides aminés existant, un nombre très important de peptides peut être généré dans les tailles correspondant aux épitopes LT CD8<sup>+</sup>, de l'ordre de 10<sup>10</sup> pour des peptides de 8 acides aminés à 10<sup>18</sup> pour des peptides de 14 acides aminés, sans compter les éventuelles modifications post-traductionnelles, ce qui rend la vérification systématique de l'ensemble des possibilités difficile même si l'on considère que seuls 1 à 3% seront processés et effectivement présentés par le CMH (296,297). De multiples approches *in vitro* ont été développées pour apprécier la cross-réactivité d'un TCR donné, telles que la constitution d'une librairie de combinaisons peptidiques regroupées par pools (298), l'encodage directe d'une séquence CMH-peptide dans une librairie de levures interrogée par des tétramères (299) ou encore l'utilisation de multimères identifiés par code-barres ADN (297). Parallèlement, la mise au point d'outils *in silico* prédictifs de la réactivité des TCR avance progressivement, bien que les résultats ne soient pas encore parfaits (300,301). Pour tenter de pallier ces différentes limites, nous avons développé une approche de modélisation du TCR afin de tenter de prédire les réactions croisées potentielles. Notre groupe est également en train de développer la transduction de TCR au sein de LT CD8<sup>+</sup> afin de pouvoir tester directement la cross-réactivité en s'affranchissant des problématiques de culture et expansion des clones.

La présence effective de l'antigène enfin, qui peut être discutée même en cas de lyse du fait de cette cross-réactivité théorique. Pour ce point particulier, notre groupe est actuellement en train de mettre un place un modèle d'extinction de la cible par utilisation de petits ARN en épingle à cheveux (shRNA pour *short hairpin RNA*) afin de démontrer la spécificité de cette cytotoxicité. L'identification finale de nos cibles par spectrométrie de masse ciblée est également en cours d'évaluation.

### 3) Tolérance immunitaire

L'expression de cibles thérapeutiques potentielles par les cellules thymiques peut questionner sur leur potentielle tolérance centrale et donc moindre immunogénicité. L'expression de HERVs par les cellules cTECs responsable de la sélection positive ne pose pas de soucis théorique du fait du rôle des antigènes présentés par les cTECs dans la stimulation chronique des TCR : en effet, la réexposition périphérique à des peptides utilisés pour la sélection positive procure un signal de stimulation basal faible aux LT permettant d'une part d'éviter leur mort en l'absence de signaux activateurs et d'autre part de répondre rapidement à un stimulus antigénique étranger (47,205). Cette expression mériterait ainsi d'explorer le rôle des HERVs en tant qu'antigènes « publics » permettant de maintenir l'homéostasie des LT.

D'un autre côté, l'expression par les mTECs, responsables de la sélection négative, pourrait être théoriquement plus embarrassante. Nos résultats montrent l'expression de différents HERVs par les cellules mTECs, venant ainsi corroborer une autre étude récente (138). Ces données sont à prendre en compte lors de la définition de cibles thérapeutiques. Ainsi dans le premier article, nous avons contrôlé l'expression des HERVs contenant nos peptides dans les cellules mTECs, montrant leur expression faible au même niveau que le tissu sain (**Article 1**, Figure 3D).

Cette question de la tolérance est primordiale pour apporter des réponses relatives à l'expression de HERVs possédant nos cibles dans le tissu sain. Plus que l'expression basale, le niveau d'expression et potentiellement de surexpression ainsi que les mécanismes de régulation post-transcriptionnels pourraient être des éléments clés. En effet, si l'on prend les membres de la famille HERV-K par exemple, il n'est pas rare de détecter une expression à faible niveau dans le tissu sain, celle-ci ne s'accompagnant par définition d'aucune réaction auto-immune. Dès lors, la présence de peptides immunogènes parmi ces séquences, associée à la possibilité que la réponse induite chez les donneurs sains soit en partie liée à une expansion de lymphocytes mémoires amène à poser deux hypothèses principales : soit l'expression basale est trop faible pour entraîner une réponse spécifique, soit celle-ci est suffisante mais associée à des mécanismes de tolérance périphérique protecteurs. Les différentes propriétés oncogéniques associées à l'expression de protéines notamment de HERV-K pouvant impacter l'homéostasie du tissu sain en cas d'expression significative plaideraient plutôt en faveur de la première hypothèse, bien que tout reste encore à démontrer dans le domaine. Enfin, comme nous l'avons vu, la densité épitopique ainsi que la quantité de molécules de CMH de classe I présente sont des facteurs décisifs pour la réponse immune (278,288). Une expression ARN détectée dans un contexte de faible densité de CMH de classe I, comme dans le cerveau, la rétine ou le muscle par exemple (302), pourrait ainsi n'avoir qu'un impact très limité.

### D. Applications potentielles

Les applications thérapeutiques potentielles de l'utilisation de ces épitopes partagés pour des approches d'immunothérapie ont déjà été discutées à plusieurs reprises dans l'introduction, les différentes revues et les articles. Rappelons que l'utilisation potentielle de ces cibles dans des approches de type vaccin thérapeutique ou TCR conçus (*engineered*) devra s'intégrer dans une stratégie plus globale pour être pleinement efficace : chimiothérapie/radiothérapie de débulking, traitement hypométhylant favorisant l'expression des cibles et la réponse inflammatoire, puis thérapie ciblée contre les HERVs afin de cibler les cellules résiduelles et prévenir les rechutes,

éventuellement associé à un traitement par ICPI. Comme discuté dans les revues 2 et 3, l'approche vaccinale semble être idéale dans des situations de faible masses tumorales (traitement adjuvant, prévention des rechutes après une première ligne...) tandis que l'infusion de lymphocytes à TCR conçus pourrait être plus adaptée aux situations de forte masse tumorale, avec la possibilité de fournir directement des effecteurs en nombre conséquent. En parallèle des cibles partagées, l'identification de cibles strictement spécifiques de la cellule tumorale, dont l'absence dans le tissu sain pourrait idéalement être contrôlé par spectrométrie de masse, est également une perspective prometteuse pour minimiser les risques de tolérance et de réaction croisée. La définition de cibles de surfaces via les protéines d'enveloppe par exemple pourrait également permette de coupler une approche anticorps ou lymphocytes CAR-T.

Un des champs encore peu explorés pour le moment est le rôle des HERVs dans la senescence. Plusieurs articles rapportent une augmentation globale de la transcription d'ETs pendant la senescence, participant au phénotype sécrétoire associé à la senescence (SASP pour *Senescence-associated secretory phenotype*) (303,304). Les rétrotransposons, et en particulier les LINE-1, seraient en grande partie responsable du phénotype inflammatoire associé à la senescence tardive, activé par la reconnaissance de leur ADNc par la voie cGAS-STING (304). L'accumulation d'ADNc est d'ailleurs favorisée par la répression de la DNase TREX1 (305). L'utilisation de lamivudine, un inhibiteur de RT, chez des souris âgées de 20 à 26 mois permet par ailleurs de prévenir le développement de différents signes de vieillissement tels que l'atrophie musculaire ou l'inflammation chronique (304). Il semblerait que les patterns d'hypométhylation en lien avec le vieillissement soient associés à une augmentation préférentielle de l'expression de certains ETs, en particulier des familles Alu et HERV-K (306). Contrairement à la déméthylation observée dans le cancer survenant plutôt dans des zones portant des marques de méthylation des histones associées à un état de répression chromatique (H3K9me3), cette déméthylation associée à la sénescence survient préférentiellement dans des régions portant des marques associées à l'activation transcriptionnelle (H3K4me1) (307).

Concernant les HERVs en particulier, peu d'études existent à ce jour. Dans un article en pre-print publié sur bioRxiv, une équipe Chinoise montre à l'aide de différents modèles que des altérations épigénétiques sont responsables de la dérégulation de HERVs principalement de la famille HERV-K au cours de la sénescence et également du vieillissement des tissus (308). L'expression de HERVs dans ce contexte conduit à une activation des voies de l'immunité innée par l'ARN double brin rétroviral, comme cela avait déjà été rapporté par d'autres équipes (309), mais également de l'ADN rétrotranscrit par la RT et reconnu par la voie cGAS-STING. D'autre part, cette expression s'accompagne de la production de particules rétrovirales capables de conférer un phénotype sénescence à des cellules plus jeunes, l'utilisation d'anticorps bloquant ces particules annulant cet effet.

Nos travaux viennent apporter une approche complémentaire à ces données. La dérégulation des HERVs par la cellule sénescence pourrait, en plus de l'activation de la réponse IFN de type I, entraîner l'expression d'épitopes à la surface de la cellule permettant sa clairance par l'immunité adaptative. Cette hypothèse reste bien entendu à démontrer dans une étude spécifique.

## II. Biomarqueur & quantification

Dans notre deuxième article, nous montrons pour la première fois qu'une population cellulaire peut être finement caractérisée par son rétrotranscriptome HERV, ces données pouvant être plus informatives que le transcriptome canonique en provenance des gènes dans certaines situations. Ces résultats ouvrent la voie à l'utilisation des HERVs pour de nombreuses applications, permettant ainsi d'apprécier l'hétérogénéité d'un échantillon tumoral par du séquençage ARN *bulk*.

### A. Caractérisation des populations cellulaires

La possibilité de pouvoir caractériser une population cellulaire par son rétrotranscriptome est probablement le point majeur de notre deuxième article. En nous basant sur ces données, nous montrons qu'une signature basée sur l'expression de 25 HERVs permet de classer correctement les cellules souches leucémiques (AUC 0,80 versus 0,76 pour le score LSC17 (310)) dans une base de données indépendantes de populations cellulaires triées (**Article 2**, Figure 3). Si cette signature tire sa force de la variété des cohortes d'entraînement utilisées et de sa validité dans un jeu de donnée totalement indépendant, elle n'en reste pas moins perfectible. Sa mise au point par exemple, établie par corrélation avec le score LSC17 dans 4 jeux de données bulk, mériterait d'être affinée sur des données de population de LSC triées voire de *single-cell* RNAseq. La population de cellules souches leucémiques dans le jeu de validation, définies par le phénotype Lin<sup>-</sup>CD34<sup>+</sup>CD38<sup>-</sup>CD99<sup>+</sup>TIM3<sup>+</sup>, ne représente par ailleurs qu'une portion des cellules souches leucémiques (311). Des ajustements restent à prévoir donc, et plus qu'une alternative à la signature LSC17 déjà établie (310), l'utilisation des HERVs pourrait être une source d'information complémentaire en vue d'établir une signature compilant gènes et HERVs pouvant refléter au mieux la biologie et l'identité du pool de LSC d'un patient donné. Une telle signature pourrait prendre tout son intérêt pour l'évaluation pronostique des patients ou le suivi de la maladie résiduelle, notamment chez les patients dépourvus de marqueurs spécifiques de maladie (caryotypes normaux par exemple, sans mutations particulières retrouvée en NGS).

La caractérisation de l'environnement immunitaire des LAM au diagnostic en lien avec le rétrotranscriptome HERV était un des objectifs de notre analyse de donnée. Malgré l'utilisation de différentes signatures, notre analyse finale n'a pu que souligner l'existence de deux profils généraux de LAM (**Article 2**, Figure 3E) : un profil plutôt infiltré (*high immune scores*) et un profil plutôt non-infiltré (*low immune scores*). Ces données rejoignent des résultats récents montrant que le pronostic des LAM peut être stratifié selon leur infiltration immunitaire (*immune-infiltrated* versus *immune-depleted*), ce profil étant particulièrement prédictif pour les LAM du groupe favorable ou mauvais selon la classification pronostique de l'ELN (312). Dans les deux cas, la prise en compte du pourcentage de blaste, interférant avec les signatures immunitaires, devrait probablement permettre de sensibiliser l'analyse pour faire ressortir des sous-groupes plus fins d'infiltration par des populations immunitaires spécifiques.

Un des autres points soulignés par nos résultats est l'existence d'un rétrotranscriptome HERV pouvant également caractériser les tissus normaux. Bien que le modèle actuel considère l'expression de HERVs faible voire nulle à l'état basal, plusieurs études commencent à s'intéresser au rétrotranscriptome HERVs des tissus sains et également aux peptides issus de ces derniers. L'étude déjà mentionnée de Larouche J-D. *et al.* ayant analysé l'expression d'ETs dans 32

différents tissus sains issus notamment de la base GTEX, retrouve une expression globale particulièrement importante dans les mTECs ainsi que dans deux tissus n'exprimant pas le CMH de classe I : le testicule et une lignée de cellules souches embryonnaires (138). La quantification des ETs par famille n'a cependant pas conduit à la définition d'une signature spécifique propre à chaque tissu sain dans cette étude.

## B. Signature prédictive de réponse thérapeutique

Ces résultats nous poussent à envisager de proposer l'utilisation des HERVs en tant que source d'information génomique importante pour des approches de prédiction. Cette idée n'est pas nouvelle en soit, de nombreuses études ayant déjà proposé d'utiliser les HERVs notamment de la famille HERV-K en tant que biomarqueurs prédictifs de réponse à un traitement ou bien encore de rechute (313). La plupart de ces études ne proposent cependant pas d'approche systématisée et clairement reproductible, se basant souvent sur des quantifications globales de familles par RT-PCR, ou ne contrôlant pas les facteurs confondants potentiels tels que le sous-type histologique et moléculaire du cancer étudié.

L'équipe à l'origine du pipeline *hervQuant* a par exemple identifié un HERV (HERV\_4700 dans l'annotation *Retrotector*) avec une forte expression dans le cancer du rein à cellules claires par rapport au rein normal, ainsi que des preuves de traduction dans des bases de données de Riboseq (séquençage d'ARN en cours de traduction par le ribosome) (257). Les auteurs montrent dans un second temps que l'expression avant tout traitement des séquences *gag*, *pol* et *env* de ce HERV évaluée par RT-PCR dans une base indépendante de patients traités par anti-PD1 est plus élevée chez les patients répondeurs par rapport aux patients non répondeurs, suggérant un possible marqueur prédictif de réponse thérapeutique.

Dans un autre article, Solovyov A. *et al.* ont stratifié par une classification non supervisée basée sur l'expression des HERVs pré-traitement un groupe de patients présentant un carcinome urothélial et traités par anti-PD-L1 (260). Leurs résultats montrent que le cluster présentant une forte expression de HERVs répond mieux au traitement par anti-PD-L1, la même approche basée sur les gènes d'une signature de défense virale n'étant quant à elle pas prédictive d'une quelconque réponse.

Plus récemment, une équipe anglaise a proposé un score prédictif de la rechute dans le cancer de l'ovaire via la sélection de HERVs associés à la survie sans progression par un modèle de Cox pénalisé de type L1 (Lasso) (314). Bien que l'approche soit intéressante, les résultats finaux restent limités et sont non significatifs dans un modèle de Cox multivarié prenant en compte notamment l'âge, le stade et le grade de la tumeur dans la cohorte indépendante. Plusieurs facteurs ont pu limiter l'analyse réalisée : (i) la quantification des HERVs tout d'abord, se basant sur une référence plutôt complète basée sur HERVd mais ne proposant pas de gestion des *multimaps* (seuls les premiers alignements sont conservés) ; (ii) la normalisation ensuite en Fragments par kilobase par millions de *reads* mappés (FPKM), pas forcément adaptée aux comparaisons inter-échantillons avec un filtre arbitraire de 10 FPKM minimum dans 2 échantillons pour chaque HERV ; (iii) le choix des HERVs enfin, seuls 23 HERVs sur les 32 de leur score étant exprimés dans la cohorte de validation. Une quantification plus fine basée sur Telescope par exemple, associée à une meilleure gestion du *batch effect* pour éliminer certains HERVs pouvant être exprimés dans une cohorte mais non retrouvés dans d'autres permettrait sans doute d'améliorer ce score.



### C. Approches *single-cell*

La prochaine étape dans le développement bioinformatique sera d'aller explorer le rétrotranscriptome HERV au sein de données de séquençage ARN sur cellule unique (*single-cell RNAseq*). Au cours de ces travaux, nous avons essayé différentes approches basées sur *hervQuant* ou *Telescope* pour apprécier l'expression des HERVs sur des jeux de données *single-cell*. Nous avons notamment quantifié le rétrotranscriptome HERV à partir du séquençage ARN de cellules uniques de LAM sur les données publiées par Van Galen *et al.* (16 patients, 5 donneurs sains, plus de 38 000 cellules sur 40 échantillons totaux au diagnostic et après traitement) (315). Bien que nous ayons pu capter l'expression de certains HERVs fortement exprimés (tels que certains loci ERV3-1), le manque de sensibilité observé sur ces données issues de la technologie *SeqWell* ne nous a pas permis de clustériser les sous-populations hématologiques en se basant uniquement sur le rétrotranscriptome HERV.

La quantification des HERVs sur données *single-cell* est pour l'instant exposée à plusieurs limites, la principale étant le manque de sensibilité ne permettant pas de capter l'expression de HERVs moyennement ou faiblement exprimés. Les méthodes les plus sensibles actuellement dites de bas-débit (*CEL-seq2* ou *Smart-seq2*) permettent ainsi de détecter en moyenne ~8 000 gènes différents par cellules, tandis que les méthodes dites de haut-débit tournent plutôt autour de 3 000 - 4 000 gènes différents par cellules (*10x Chromium*, *Dropseq* et *SeqWell*) (316), ce qui laisse peu de place pour les 14 000 transcrits HERVs que nous quantifions avec l'approche *Telescope*.

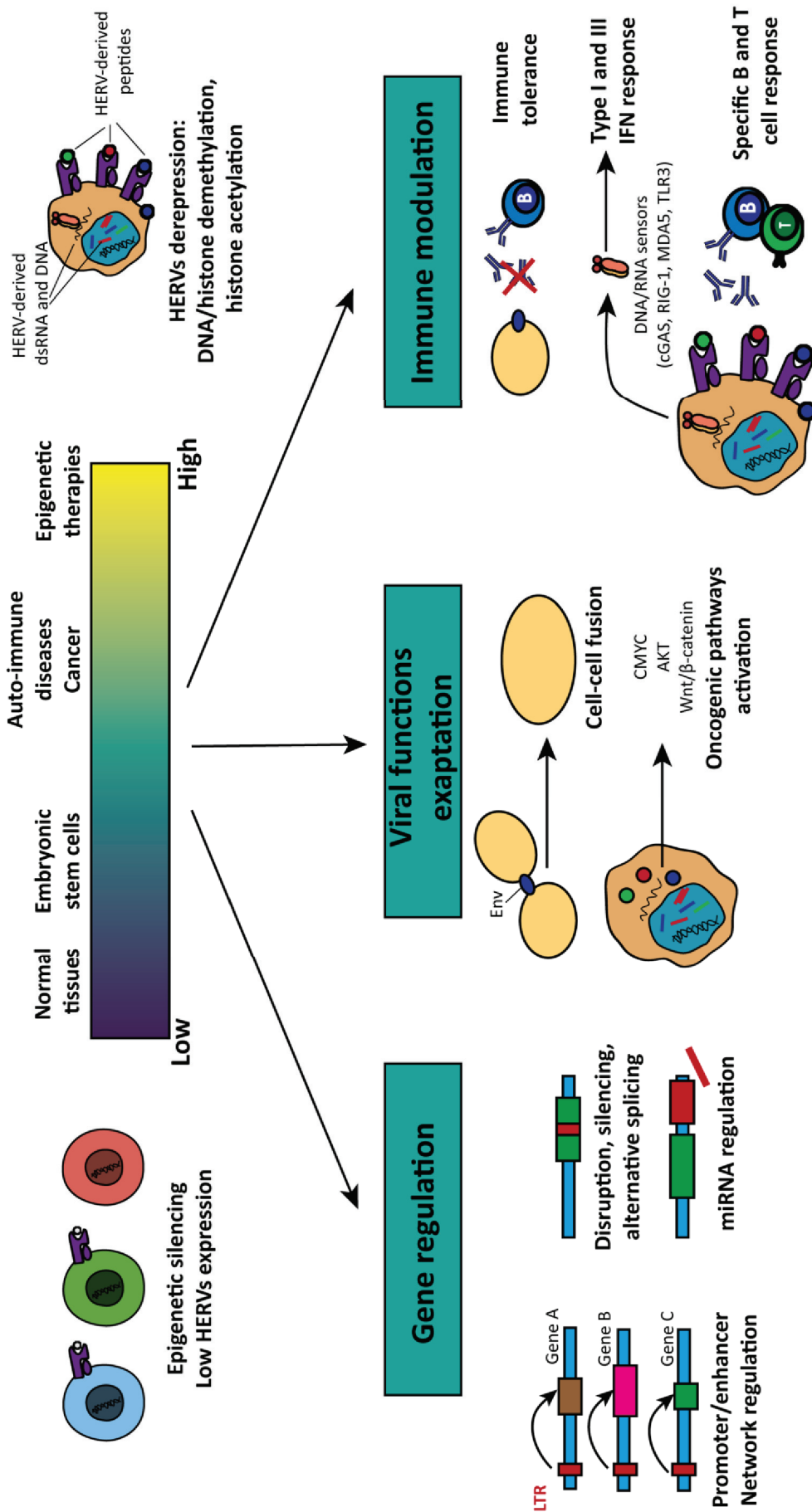
La deuxième limite vient de la méthode de quantification en elle-même, pour le moment non optimisée pour données *single-cell*. Plusieurs équipes commencent à se pencher sur le sujet, les différentes études menant là encore à la détection de quelques séquences plutôt qu'à l'appréciation précise du rétrotranscriptome dans son ensemble, souffrant de la limite technologique (317). Beaucoup reste à faire donc et la mise au point de méthodes de quantification des HERVs en *single-cell* sera probablement contemporaine de l'amélioration de la profondeur des techniques de séquençage.

### III. Proposition d'un modèle

L'ensemble de ces données ajoutées aux connaissances actuelles passées en revue au cours de cette thèse nous permet de proposer un raffinement du modèle actuel de l'impact physiologique et pathologique des HERVs (Figure 9). De la dichotomie initiale « absence d'expression dans les tissus sains – expression dans les tumeurs », nous pouvons voir que chaque tissu peut en fait exprimer des HERVs, avec une empreinte rétrotranscriptomique et chromatinique propre à chaque population cellulaire. Le niveau global d'expression est en revanche variable selon les situations, plutôt faible dans les tissus sains, augmenté dans certains tissus tels que le testicule ou les cellules souches embryonnaires, et bien plus élevé dans certaines conditions de dérégulation comme dans les maladies auto-immunes ou les cancers. Les thérapies épigénétiques telles que les DNMTi et les HDACi pourraient encore augmenter le niveau d'expression. Les HERVs peuvent exercer de multiples effets en condition physiologique ou pathologique : (i) régulation génique, via l'intégration de leur LTR en amont d'un gène agissant en tant que promoteur ou *enhancer*, la modification directe d'un gène par l'intégration au sein de sa séquence (codon stop prématuré, signal d'épissage alternatif...) ou encore la fourniture de sites



de liaison pour une régulation post-transcriptionnelle par des micro ARN ; (ii) exaptation de fonctions virales telles que la fusion cellulaire dans le syncytiotrophoblaste permise par une protéine d'enveloppe ou l'activation de voies de signalisations oncogéniques ; (iii) modulation immunitaire par stimulation d'une réponse IFN de type I et III, l'induction d'une réponse lymphocytaire T et B spécifique ou encore la mise en place d'une réponse immunosuppressive. Le lien exact entre niveau d'expression et activité des HERVs n'est pas clairement défini et pourrait varier selon les situations, une faible expression pouvant suffire à l'activité de régulation génique là où une plus forte expression est attendue pour la mise en place d'une réponse immunitaire. Ce schéma nous montre ainsi que la limite entre effet physiologique et pathologique peut parfois être très fine, la bonne interprétation de l'expression d'un HERV donné nécessitant une analyse dédiée adaptée à chaque situation.



**Figure 8 : Modèle récapitulatif de l'expression des HERVs ainsi que de leur impact physio-pathologique**

Le niveau d'expression des HERVs varie selon les situations : plutôt faible dans les tissus sains (répression épigénétique), augmenté dans certains tissus tels que le testicule ou les cellules souches embryonnaires et plutôt élevé dans certaines situations pathologiques telles que les maladies auto-immunes ou les cancers (dépression). Les HERVs peuvent avoir de multiples impacts physiologiques et pathologiques via leur activité de régulation génique ou l'induction d'une certaine instabilité génomique par rétrotransposition, via l'exaptation de certaines de leurs fonctions virales ou bien via l'impact immunologique de leurs intermédiaires nucléotidiques et protéiques. Il n'existe pas de corrélation directe entre niveau d'expression et activité d'un HERV, bien que celle-ci puisse exister dans certaines situations (comme pour l'induction de réponses lymphocytaires T ou B spécifiques par exemple).

## CONCLUSION

---



A l'ère de la démocratisation des techniques de séquençage haut-débit, l'exploration des portions dites « non codantes » du génome devient un axe de recherche majeur, indispensable à la bonne compréhension des différents éléments constituant notre patrimoine génétique. Représentant 8% du génome humain, les HERVs constituent une grande partie de ce patrimoine dont l'acquisition, fruit d'une intégration progressive dans le génome des mammifères au cours de millions d'années d'évolution, aura été à double tranchant : véritable moteur évolutif permettant l'exaptation de fonctions d'origine virale d'un côté, source d'instabilité génomique de l'autre. Cet effet à double tranchant se reflète en contexte tumoral, les HERVs fournissant à la fois un substrat à la cellule tumorale et de potentielles cibles thérapeutiques.

Ce travail vient renforcer le rationnel de ciblage des HERVs par des approches d'immunothérapie dans le cancer. En développant une méthode de sélection d'épitopes applicable à grande échelle puis en montrant l'immunogénicité d'une sélection d'épitopes partagés, nous fournissons tous les éléments démontrant l'intérêt de telles cibles d'origine virale. En parallèle, nos résultats montrent que les HERVs peuvent être utilisés comme marqueurs d'identité cellulaire. L'abondance des HERVs dans notre génome fournit en effet une importante quantité de matériel génétique variant selon le profil de la cellule. Si certains HERVs participent probablement à la régulation génique de façon active, d'autres ne pourraient représenter que de simples transcrits marquant une zone d'activité chromatinienne. Dans les deux cas, cette information peut être utilisée pour mieux préciser le profil cellulaire et pourrait se révéler d'une importance capitale pour la recherche de biomarqueurs ou les besoins de stratifications de certaines maladies.

Nos résultats ouvrent de nombreuses perspectives : côté ciblage thérapeutique, un travail est en cours afin de pouvoir proposer, pour chaque cancer, une liste d'épitopes LT CD8<sup>+</sup> prioritaires, combinant des cibles partagées et plus spécifiques et prenant en compte l'ensemble des molécules HLA. L'utilisation de TCR conçus pour optimiser la validation des cibles et mieux comprendre les relations CMH-peptide-TCR est également en cours, ainsi qu'une ouverture sur les épitopes de surface en vue de développer des approches basées sur l'utilisation d'anticorps thérapeutiques ou de lymphocytes CAR-T. Côté biomarqueur, le développement d'une signature robuste de cellules souches leucémiques basée sur plusieurs sources de données, l'adaptation des outils de quantification aux données *single-cell* et l'application de stratégies de stratification basées sur les HERVs à des problématiques de vie réelle font parties des projets en cours. De nombreuses choses restent à faire, et nul doute que les développements technologiques futurs viendront accélérer encore un champ de recherche en pleine croissance.





## REFERENCES

---

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* nov 2018;68(6):394-424.
2. Oiseth SJ, Aziz MS. Cancer immunotherapy: a brief review of the history, possibilities, and challenges ahead. *J Cancer Metastasis Treat.* 31 oct 2017;3(10):250.
3. Coley W. The treatment of malignant tumors by repeated inoculations of erysipelas : with a report of ten original cases. *Am J Med Sci.* 1893;105:487.
4. Dobosz P, Dzieciatkowski T. The Intriguing History of Cancer Immunotherapy. *Front Immunol.* 17 déc 2019;10:2965.
5. Haslam A, Prasad V. Estimation of the Percentage of US Patients With Cancer Who Are Eligible for and Respond to Checkpoint Inhibitor Immunotherapy Drugs. *JAMA Netw Open.* 3 mai 2019;2(5):e192535.
6. Lindahl T, Wood RD. Quality control by DNA repair. *Science.* 3 déc 1999;286(5446):1897-905.
7. Kim R, Emi M, Tanabe K. Cancer immunoediting from immune surveillance to immune escape. *Immunology.* mai 2007;121(1):1-14.
8. Schreiber RD, Old LJ, Smyth MJ. Cancer Immunoediting: Integrating Immunity's Roles in Cancer Suppression and Promotion. *Science.* 25 mars 2011;331(6024):1565-70.
9. Balkwill F, Mantovani A. Inflammation and cancer: back to Virchow? *Lancet.* 17 févr 2001;357(9255):539-45.
10. Matzinger P. Tolerance, Danger, and the Extended Family. *Annu Rev Immunol.* :55.
11. Pardoll D. Does the Immune System See Tumors as Foreign or Self? *Annu Rev Immunol.* 1 avr 2003;21(1):807-39.
12. Dighe AS, Richards E, Old LJ, Schreiber RD. Enhanced in vivo growth and resistance to rejection of tumor cells expressing dominant negative IFN $\gamma$  receptors. *Immunity.* 1 sept 1994;1(6):447-56.
13. Shankaran V, Ikeda H, Bruce AT, White JM, Swanson PE, Old LJ, et al. IFN $\gamma$  and lymphocytes prevent primary tumour development and shape tumour immunogenicity. *Nature.* avr 2001;410(6832):1107-11.
14. Dunn GP, Bruce AT, Ikeda H, Old LJ, Schreiber RD. Cancer immunoediting: from immuno- surveillance to tumor escape. *Nat Immunol.* 2002;3(11):8.
15. Phan TG, Croucher PI. The dormant cancer cell life cycle. *Nat Rev Cancer.* juill 2020;20(7):398-411.
16. Chen DS, Mellman I. Oncology Meets Immunology: The Cancer-Immunity Cycle. *Immunity.* juill 2013;39(1):1-10.

17. Takeuchi A, Saito T. CD4 CTL, a Cytotoxic Subset of CD4+ T Cells, Their Differentiation and Function. *Front Immunol* [Internet]. 23 févr 2017 [cité 24 mars 2021];8. Disponible sur: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5321676/>
18. Chen DS, Mellman I. Elements of cancer immunity and the cancer-immune set point. *Nature*. janv 2017;541(7637):321-30.
19. Bonaventura P, Shekarian T, Alcazer V, Valladeau-Guilemond J, Valsesia-Wittmann S, Amigorena S, et al. Cold Tumors: A Therapeutic Challenge for Immunotherapy. *Front Immunol*. 8 févr 2019;10:168.
20. Coulie PG, Van den Eynde BJ, van der Bruggen P, Boon T. Tumour antigens recognized by T lymphocytes: at the core of cancer immunotherapy. *Nat Rev Cancer*. févr 2014;14(2):135-46.
21. Laumont CM, Perreault C. Exploiting non-canonical translation to identify new targets for T cell-based cancer immunotherapy. *Cell Mol Life Sci*. 1 févr 2018;75(4):607-21.
22. Granados DP, Laumont CM, Thibault P, Perreault C. The nature of self for T cells—a systems-level perspective. *Curr Opin Immunol*. 1 juin 2015;34:1-8.
23. Yewdell JW, Antón LC, Bennink JR. Defective ribosomal products (DRiPs): a major source of antigenic peptides for MHC class I molecules? *J Immunol*. 1 sept 1996;157(5):1823-6.
24. Schubert U, Antón LC, Gibbs J, Norbury CC, Yewdell JW, Bennink JR. Rapid degradation of a large fraction of newly synthesized proteins by proteasomes. *Nature*. avr 2000;404(6779):770-4.
25. Gross L. Intradermal Immunization of C3H Mice against a Sarcoma That Originated in an Animal of the Same Line. *Cancer Res*. 1 mai 1943;3(5):326-33.
26. Old LJ, Boyse EA. Immunology of experimental tumors. *Annu Rev Med*. 1964;15:167-86.
27. Holden HT, Haskill JS, Kirchner H, Herberman RB. Two Functionally Distinct Anti-Tumor Effector Cells Isolated from Primary Murine Sarcoma Virus-Induced Tumors. *J Immunol*. 1 août 1976;117(2):440-6.
28. Plata F, Sordat B. Murine sarcoma virus (MSV)-induced tumors in mice. I. Distribution of MSV-immune cytolytic T lymphocytes in vivo. *Int J Cancer*. 15 févr 1977;19(2):205-11.
29. Brunner KT, MacDonald HR, Cerottini JC. Quantitation and clonal isolation of cytolytic T lymphocyte precursors selectively infiltrating murine sarcoma virus-induced tumors. *J Exp Med*. 1 août 1981;154(2):362-73.
30. Vose BM, Bonnard GD. Human tumour antigens defined by cytotoxicity and proliferative responses of cultured lymphoid cells. *Nature*. mars 1982;296(5855):359-61.
31. Townsend ARM, Rothbard J, Gotch FM, Bahadur G, Wraith D, McMichael AJ. The epitopes of influenza nucleoprotein recognized by cytotoxic T lymphocytes can be defined with short synthetic peptides. *Cell*. 28 mars 1986;44(6):959-68.

32. Lurquin C, Van Pel A, Mariamé B, De Plaen E, Szikora JP, Janssens C, et al. Structure of the gene of tum- transplantation antigen P91A: the mutated exon encodes a peptide recognized with Ld by cytolytic T cells. *Cell*. 28 juill 1989;58(2):293-303.
33. Rammensee HG, Falk K, Rötzschke O. Peptides Naturally Presented by MHC Class I Molecules. *Annu Rev Immunol*. 1 avr 1993;11(1):213-44.
34. van der Bruggen P, Traversari C, Chomez P, Lurquin C, De Plaen E, Van den Eynde B, et al. A gene encoding an antigen recognized by cytolytic T lymphocytes on a human melanoma. *Science*. 13 déc 1991;254(5038):1643-7.
35. Knuth A, Danowski B, Oettgen HF, Old LJ. T-cell-mediated cytotoxicity against autologous malignant melanoma: analysis with interleukin 2-dependent T-cell cultures. *Proc Natl Acad Sci U S A*. juin 1984;81(11):3511-5.
36. Sahin U, Tureci O, Schmitt H, Cochlovius B, Johannes T, Schmits R, et al. Human neoplasms elicit multiple specific immune responses in the autologous host. *Proc Natl Acad Sci*. 5 déc 1995;92(25):11810-3.
37. Cox AL, Skipper J, Chen Y, Henderson RA, Darrow TL, Shabanowitz J, et al. Identification of a peptide recognized by five melanoma-specific human cytotoxic T cell lines. *Science*. 29 avr 1994;264(5159):716-9.
38. Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi A, et al. Landscape of transcription in human cells. *Nature*. 6 sept 2012;489(7414):101-8.
39. Laumont CM, Vincent K, Hesnard L, Audemard É, Bonneil É, Laverdure J-P, et al. Noncoding regions are the main source of targetable tumor-specific antigens. *Sci Transl Med [Internet]*. 5 déc 2018 [cité 16 janv 2020];10(470). Disponible sur: <https://stm.sciencemag.org/content/10/470/eaau5516>
40. Haen SP, Löffler MW, Rammensee H-G, Brossart P. Towards new horizons: characterization, classification and implications of the tumour antigenic repertoire. *Nat Rev Clin Oncol*. oct 2020;17(10):595-610.
41. Coulie PG, Brichard V, Van Pel A, Wölfel T, Schneider J, Traversari C, et al. A new gene coding for a differentiation antigen recognized by autologous cytolytic T lymphocytes on HLA-A2 melanomas. *J Exp Med*. 1 juill 1994;180(1):35-42.
42. Bakker AB, Schreurs MW, Tafazzul G, de Boer AJ, Kawakami Y, Adema GJ, et al. Identification of a novel peptide derived from the melanocyte-specific gp100 antigen as the dominant epitope recognized by an HLA-A2.1-restricted anti-melanoma CTL line. *Int J Cancer*. 4 juill 1995;62(1):97-102.
43. Inoue K, Ogawa H, Sonoda Y, Kimura T, Sakabe H, Oka Y, et al. Aberrant Overexpression of the Wilms Tumor Gene (WT1) in Human Leukemia. *Blood*. 15 févr 1997;89(4):1405-12.
44. Pittet MJ, Valmori D, Dunbar PR, Speiser DE, Liénard D, Lejeune F, et al. High Frequencies of Naive Melan-a/Mart-1-Specific Cd8+ T Cells in a Large Proportion of Human Histocompatibility Leukocyte Antigen (Hla)-A2 Individuals. *J Exp Med*. 6 sept 1999;190(5):705-16.

45. Alanio C, Lemaitre F, Law HKW, Hasan M, Albert ML. Enumeration of human antigen-specific naive CD8+ T cells reveals conserved precursor frequencies. *Blood*. 6 mai 2010;115(18):3718-25.
46. Goldrath AW, Bogatzki LY, Bevan MJ. Naive T cells transiently acquire a memory-like phenotype during homeostasis-driven proliferation. *J Exp Med*. 21 août 2000;192(4):557-64.
47. Zippelius A, Pittet MJ, Batard P, Rufer N, de Smedt M, Guillaume P, et al. Thymic Selection Generates a Large T Cell Pool Recognizing a Self-Peptide in Humans. *J Exp Med*. 18 févr 2002;195(4):485-94.
48. Stone JD, Harris DT, Kranz DM. TCR affinity for p/MHC formed by tumor antigens that are self-proteins: impact on efficacy and toxicity. *Curr Opin Immunol*. avr 2015;33:16-22.
49. Yarchoan M, Johnson BA, Lutz ER, Laheru DA, Jaffee EM. Targeting neoantigens to augment antitumour immunity. *Nat Rev Cancer*. 2017;17(4):209-22.
50. Tian S, Maile R, Collins EJ, Frelinger JA. CD8+ T Cell Activation Is Governed by TCR-Peptide/MHC Affinity, Not Dissociation Rate. *J Immunol*. 1 sept 2007;179(5):2952-60.
51. Morgan RA, Yang JC, Kitano M, Dudley ME, Laurencot CM, Rosenberg SA. Case Report of a Serious Adverse Event Following the Administration of T Cells Transduced With a Chimeric Antigen Receptor Recognizing ERBB2. *Mol Ther*. 1 avr 2010;18(4):843-51.
52. Morgan RA, Chinnasamy N, Abate-Daga D, Gros A, Robbins PF, Zheng Z, et al. Cancer regression and neurological toxicity following anti-MAGE-A3 TCR gene therapy. *J Immunother*. févr 2013;36(2):133-51.
53. Joseph CG, Darrah E, Shah AA, Skora AD, Casciola-Rosen LA, Wigley FM, et al. Association of the Autoimmune Disease Scleroderma with an Immunologic Response to Cancer. *Science*. 10 janv 2014;343(6167):152-7.
54. Simpson AJG, Caballero OL, Jungbluth A, Chen Y-T, Old LJ. Cancer/testis antigens, gametogenesis and cancer. *Nat Rev Cancer*. août 2005;5(8):615-25.
55. Boël P, Wildmann C, Sensi ML, Brasseur R, Renaud JC, Coulie P, et al. BAGE: a new gene encoding an antigen recognized on human melanomas by cytolytic T lymphocytes. *Immunity*. févr 1995;2(2):167-75.
56. De Backer O, Arden KC, Boretti M, Vantomme V, De Smet C, Czekay S, et al. Characterization of the GAGE genes that are expressed in various human cancers and in normal testis. *Cancer Res*. 1 juill 1999;59(13):3157-65.
57. Gibbs ZA, Whitehurst AW. Emerging Contributions of Cancer/Testis Antigens to Neoplastic Behaviors. *Trends Cancer*. oct 2018;4(10):701-12.
58. Scanlan MJ, Simpson AJG, Old LJ. The cancer/testis genes: review, standardization, and commentary. *Cancer Immun*. 23 janv 2004;4:1.
59. Weber J, Salgaller M, Samid D, Johnson B, Herlyn M, Lassam N, et al. Expression of the MAGE-1 tumor antigen is up-regulated by the demethylating agent 5-aza-2'-deoxycytidine. *Cancer Res*. 1 avr 1994;54(7):1766-71.

60. De Smet C, Lurquin C, Lethé B, Martelange V, Boon T. DNA methylation is the primary silencing mechanism for a set of germ line- and tumor-specific genes with a CpG-rich promoter. *Mol Cell Biol.* nov 1999;19(11):7327-35.
61. Fijak M, Meinhardt A. The testis in immune privilege. *Immunol Rev.* oct 2006;213:66-81.
62. Gubin MM, Zhang X, Schuster H, Caron E, Ward JP, Noguchi T, et al. Checkpoint blockade cancer immunotherapy targets tumour-specific mutant antigens. *Nature.* nov 2014;515(7528):577-81.
63. Matsushita H, Vesely MD, Koboldt DC, Rickert CG, Uppaluri R, Magrini VJ, et al. Cancer exome analysis reveals a T-cell-dependent mechanism of cancer immunoediting. *Nature.* févr 2012;482(7385):400-4.
64. Yadav M, Jhunjhunwala S, Phung QT, Lupardus P, Tanguay J, Bumbaca S, et al. Predicting immunogenic tumour mutations by combining mass spectrometry and exome sequencing. *Nature.* nov 2014;515(7528):572-6.
65. Roudko V, Bozkus CC, Orfanelli T, McClain CB, Carr C, O'Donnell T, et al. Shared Immunogenic Poly-Epitope Frameshift Mutations in Microsatellite Unstable Tumors. *Cell.* 10 déc 2020;183(6):1634-1649.e17.
66. Smith CC, Selitsky SR, Chai S, Armistead PM, Vincent BG, Serody JS. Alternative tumour-specific antigens. *Nat Rev Cancer.* août 2019;19(8):465-78.
67. Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature.* 18 juill 2012;487(7407):330-7.
68. Yarchoan M, Hopkins A, Jaffee EM. Tumor Mutational Burden and Response Rate to PD-1 Inhibition. *N Engl J Med.* 21 déc 2017;377(25):2500-1.
69. Keskin DB, Anandappa AJ, Sun J, Tirosh I, Mathewson ND, Li S, et al. Neoantigen vaccine generates intratumoral T cell responses in phase Ib glioblastoma trial. *Nature.* janv 2019;565(7738):234-9.
70. Hopkins AM, Rowland A, Kichenadasse G, Wiese MD, Gurney H, McKinnon RA, et al. Predicting response and toxicity to immune checkpoint inhibitors using routinely available blood and clinical markers. *Br J Cancer.* sept 2017;117(7):913-20.
71. Jayasinghe RG, Cao S, Gao Q, Wendl MC, Vo NS, Reynolds SM, et al. Systematic Analysis of Splice-Site-Creating Mutations in Cancer. *Cell Rep.* 3 avr 2018;23(1):270-281.e3.
72. Frankiw L, Baltimore D, Li G. Alternative mRNA splicing in cancer immunotherapy. *Nat Rev Immunol.* nov 2019;19(11):675-87.
73. Smart AC, Margolis CA, Pimentel H, He MX, Miao D, Adeegbe D, et al. Intron retention is a source of neoepitopes in cancer. *Nat Biotechnol.* nov 2018;36(11):1056-8.
74. Van den Eynde BJ, Gaugler B, Probst-Kepper M, Michaux L, Devuyst O, Lorge F, et al. A New Antigen Recognized by Cytolytic T Lymphocytes on a Human Kidney Tumor Results from Reverse Strand Transcription. *J Exp Med.* 20 déc 1999;190(12):1793-800.



75. Yip BH, Dolatshad H, Roy S, Pellagatti A, Boulwood J. Impact of Splicing Factor Mutations on Pre-mRNA Splicing in the Myelodysplastic Syndromes. *Curr Pharm Des.* 2016;22(16):2333-44.
76. Saulquin X, Scotet E, Trautmann L, Peyrat M-A, Halary F, Bonneville M, et al. +1 Frameshifting as a novel mechanism to generate a cryptic cytotoxic T lymphocyte epitope derived from human interleukin 10. *J Exp Med.* 4 févr 2002;195(3):353-8.
77. Malarkannan S, Horng T, Shih PP, Schwab S, Shastri N. Presentation of Out-of-Frame Peptide/MHC Class I Complexes by a Novel Translation Initiation Mechanism. *Immunity.* 1 juin 1999;10(6):681-90.
78. Bullock TN, Patterson AE, Franlin LL, Notidis E, Eisenlohr LC. Initiation codon scanthrough versus termination codon readthrough demonstrates strong potential for major histocompatibility complex class I-restricted cryptic epitope expression. *J Exp Med.* 6 oct 1997;186(7):1051-8.
79. Hanada K-I, Yewdell JW, Yang JC. Immune recognition of a human renal cancer antigen through post-translational protein splicing. *Nature.* 15 janv 2004;427(6971):252-6.
80. Liepe J, Marino F, Sidney J, Jeko A, Bunting DE, Sette A, et al. A large fraction of HLA class I ligands are proteasome-generated spliced peptides. *Science.* 21 oct 2016;354(6310):354-8.
81. Shukla GC, Singh J. Mutations of RNA splicing factors in hematological malignancies. *Cancer Lett.* 28 nov 2017;409:1-8.
82. Ehx G, Larouche J-D, Durette C, Laverdure J-P, Hesnard L, Vincent K, et al. Atypical acute myeloid leukemia-specific transcripts generate shared and immunogenic MHC class-I-associated epitopes. *Immunity [Internet].* 18 mars 2021 [cité 25 mars 2021];0(0). Disponible sur: [https://www.cell.com/immunity/abstract/S1074-7613\(21\)00087-X](https://www.cell.com/immunity/abstract/S1074-7613(21)00087-X)
83. Chong C, Müller M, Pak H, Harnett D, Huber F, Grun D, et al. Integrated proteogenomic deep sequencing and analytics accurately identify non-canonical peptides in tumor immunopeptidomes. *Nat Commun.* 10 mars 2020;11(1):1293.
84. Alcazer V, Bonaventura P, Tonon L, Wittmann S, Caux C, Depil S. Neoepitopes-based vaccines: challenges and perspectives. *Eur J Cancer.* févr 2019;108:55-60.
85. Depil S, Bonaventura P, Alcazer V, Tonon L. Nouvelles approches vaccinales en cancérologie. *Bull Cancer (Paris).* déc 2018;105:S113-20.
86. Melero I, Gaudernack G, Gerritsen W, Huber C, Parmiani G, Scholl S, et al. Therapeutic vaccines for cancer: an overview of clinical trials. *Nat Rev Clin Oncol.* sept 2014;11(9):509-24.
87. Alcazer V, Delenda C, Poirot L, Depil S. Développement des CAR-T dans les tumeurs solides. *Bull Cancer (Paris).* déc 2018;105:S178-87.
88. Alcazer V, Depil S. Les CAR-T cells en héματο-oncologie. *Onko+.* 29 sept 2019;11(89).
89. Aftab BT, Sasu B, Krishnamurthy J, Gschweng E, Alcazer V, Depil S. Toward “off-the-shelf” allogeneic CAR T cells. *Adv Cell Gene Ther.* 2020;3(3):e86.

90. Consortium IHGS. Initial sequencing and analysis of the human genome. *Nature*. 15 févr 2001;409(6822):350-57062.
91. Cordaux R, Batzer MA. The impact of retrotransposons on human genome evolution. *Nat Rev Genet*. oct 2009;10(10):691-703.
92. Craig N, Craigie R, Gellert M, Lambowitz A. Mobile DNA II. Mobile DNA II. American Society of Microbiology; 2002.
93. Huang CRL, Burns KH, Boeke JD. Active Transposition in Genomes. *Annu Rev Genet*. 2012;46:651-75.
94. Gibbs RA, Rogers J, Katze MG, Bumgarner R, Weinstock GM, Mardis ER, et al. Evolutionary and Biomedical Insights from the Rhesus Macaque Genome. *Science*. 13 avr 2007;316(5822):222-34.
95. Deininger PL, Batzer MA. Alu Repeats and Human Disease. *Mol Genet Metab*. 1 juill 1999;67(3):183-93.
96. Chen J-M, Stenson PD, Cooper DN, Férec C. A systematic analysis of LINE-1 endonuclease-dependent retrotranspositional events causing human genetic disease. *Hum Genet*. sept 2005;117(5):411-27.
97. Kazazian HH, Wong C, Youssoufian H, Scott AF, Phillips DG, Antonarakis SE. Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man. *Nature*. mars 1988;332(6160):164-6.
98. Hancks DC, Kazazian HH. SVA retrotransposons: Evolution and genetic instability. *Semin Cancer Biol*. août 2010;20(4):234-45.
99. Xiong Y, Eickbush TH. Origin and evolution of retroelements based upon their reverse transcriptase sequences. *EMBO J*. oct 1990;9(10):3353-62.
100. Vargiu L, Rodriguez-Tomé P, Sperber GO, Cadeddu M, Grandi N, Blikstad V, et al. Classification and characterization of human endogenous retroviruses; mosaic forms are common. *Retrovirology*. 22 janv 2016;13(1):7.
101. Jern P, Coffin JM. Effects of retroviruses on host genome function. *Annu Rev Genet*. 2008;42:709-32.
102. Kassiotis G, Stoye JP. Immune responses to endogenous retroelements: taking the bad with the good. *Nat Rev Immunol*. avr 2016;16(4):207-19.
103. Boller K, Schönfeld K, Lischer S, Fischer N, Hoffmann A, Kurth R, et al. Human endogenous retrovirus HERV-K113 is capable of producing intact viral particles. *J Gen Virol*. 2008;89(2):567-72.
104. Magiorkinis G, Gifford RJ, Katzourakis A, Ranter JD, Belshaw R. Env-less endogenous retroviruses are genomic superspreaders. *Proc Natl Acad Sci*. 5 août 2012;109(19):7385-90.

105. Escalera-Zamudio M, Greenwood AD. On the classification and evolution of endogenous retrovirus: human endogenous retroviruses may not be « human » after all. *APMIS Acta Pathol Microbiol Immunol Scand.* févr 2016;124(1-2):44-51.
106. Zhang Y, Maksakova IA, Gagnier L, Lagemaat LN van de, Mager DL. Genome-Wide Assessments Reveal Extremely High Levels of Polymorphism of Two Active Families of Mouse Endogenous Retroviral Elements. *PLOS Genet.* 29 févr 2008;4(2):e1000007.
107. Wildschutte JH, Williams ZH, Montesion M, Subramanian RP, Kidd JM, Coffin JM. Discovery of unfixed endogenous retrovirus insertions in diverse human populations. *Proc Natl Acad Sci U S A.* 19 avr 2016;113(16):E2326-2334.
108. Li W, Lin L, Malhotra R, Yang L, Acharya R, Poss M. A computational framework to assess genome-wide distribution of polymorphic human endogenous retrovirus-K In human populations. *PLOS Comput Biol.* 28 mars 2019;15(3):e1006564.
109. Stoye JP, Coffin JM. The four classes of endogenous murine leukemia virus: structural relationships and potential for recombination. *J Virol.* 1 sept 1987;61(9):2659-69.
110. Gifford RJ, Blomberg J, Coffin JM, Fan H, Heidmann T, Mayer J, et al. Nomenclature for endogenous retrovirus (ERV) loci. *Retrovirology.* 28 août 2018;15(1):59.
111. Lyn D, Deaven LL, Istock NL, Smulson M. The Polymorphic ADP-Ribosyltransferase (NAD<sup>+</sup>) Pseudogene 1 in Humans Interrupts an Endogenous pol-like Element on 13q34. *Genomics.* 1 nov 1993;18(2):206-11.
112. Werner T, Brack-Werner R, Leib-Mösch C, Backhaus H, Erfle V, Hehlmann R. S71 is a phylogenetically distinct human endogenous retroviral element with structural and sequence homology to simian sarcoma virus (SSV). *Virology.* janv 1990;174(1):225-38.
113. Seifarth W, Skladny H, Krieg-Schneider F, Reichert A, Hehlmann R, Leib-Mösch C. Retrovirus-like particles released from the human breast cancer cell line T47-D display type B- and C-related endogenous retroviral sequences. *J Virol.* 1 oct 1995;69(10):6408-16.
114. Tristem M. Identification and Characterization of Novel Human Endogenous Retrovirus Families by Phylogenetic Screening of the Human Genome Mapping Project Database. *J Virol.* 15 avr 2000;74(8):3715-30.
115. Mayer J, Blomberg J, Seal RL. A revised nomenclature for transcribed human endogenous retroviral loci. *Mob DNA.* 2011;2(1):7.
116. Maksakova IA, Mager DL, Reiss D. Keeping active endogenous retroviral-like elements in check: the epigenetic perspective. *Cell Mol Life Sci.* nov 2008;65(21):3329-47.
117. Greenberg MVC, Bourc'his D. The diverse roles of DNA methylation in mammalian development and disease. *Nat Rev Mol Cell Biol.* oct 2019;20(10):590-607.
118. Li E, Zhang Y. DNA Methylation in Mammals. *Cold Spring Harb Perspect Biol.* 5 janv 2014;6(5):a019133.

119. Min B, Park JS, Jeong YS, Jeon K, Kang Y-K. Dnmt1 binds and represses genomic retroelements via DNA methylation in mouse early embryos. *Nucleic Acids Res.* 4 sept 2020;48(15):8431-44.
120. Szpakowski S, Sun X, Lage JM, Dyer A, Rubinstein J, Kowalski D, et al. Loss of epigenetic silencing in tumors preferentially affects primate-specific retroelements. *Gene.* 15 déc 2009;448(2):151-67.
121. Ohtani H, Liu M, Zhou W, Liang G, Jones PA. Switching roles for DNA and histone methylation depend on evolutionary ages of human endogenous retroviruses. *Genome Res.* août 2018;28(8):1147-57.
122. Hurst TP, Magiorkinis G. Epigenetic Control of Human Endogenous Retrovirus Expression: Focus on Regulation of Long-Terminal Repeats (LTRs). *Viruses.* 31 mai 2017;9(6).
123. Fukuda K, Shinkai Y. SETDB1-Mediated Silencing of Retroelements. *Viruses.* 30 mai 2020;12(6).
124. Rowe HM, Jakobsson J, Mesnard D, Rougemont J, Reynard S, Aktas T, et al. KAP1 controls endogenous retroviruses in embryonic stem cells. *Nature.* 14 janv 2010;463(7278):237-40.
125. Collins PL, Kyle KE, Egawa T, Shinkai Y, Oltz EM. The histone methyltransferase SETDB1 represses endogenous and exogenous retroviruses in B lymphocytes. *Proc Natl Acad Sci.* 7 juill 2015;112(27):8367-72.
126. Takikita S, Muro R, Takai T, Otsubo T, Kawamura YI, Dohi T, et al. A Histone Methyltransferase ESET Is Critical for T Cell Development. *J Immunol.* 15 sept 2016;197(6):2269-79.
127. Kato M, Takemoto K, Shinkai Y. A somatic role for the histone methyltransferase Setdb1 in endogenous retrovirus silencing. *Nat Commun.* déc 2018;9(1):1683.
128. Bruno M, Mahgoub M, Macfarlan TS. The Arms Race Between KRAB-Zinc Finger Proteins and Endogenous Retroelements and Its Impact on Mammals. *Annu Rev Genet.* 2019;53(1):393-416.
129. Turelli P, Castro-Diaz N, Marzetta F, Kapopoulou A, Raclot C, Duc J, et al. Interplay of TRIM28 and DNA methylation in controlling human endogenous retroelements. *Genome Res.* août 2014;24(8):1260-70.
130. Imbeault M, Helleboid P-Y, Trono D. KRAB zinc-finger proteins contribute to the evolution of gene regulatory networks. *Nature.* mars 2017;543(7646):550-4.
131. Ishak CA, Marshall AE, Passos DT, White CR, Kim SJ, Cecchini MJ, et al. An RB-EZH2 Complex Mediates Silencing of Repetitive DNA Sequences. *Mol Cell.* 15 déc 2016;64(6):1074-87.
132. Sheng W, LaFleur MW, Nguyen TH, Chen S, Chakravarthy A, Conway JR, et al. LSD1 Ablation Stimulates Anti-tumor Immunity and Enables Checkpoint Blockade. *Cell.* 26 2018;174(3):549-563.e19.

133. Hurst T, Pace M, Katzourakis A, Phillips R, Klenerman P, Frater J, et al. Human endogenous retrovirus (HERV) expression is not induced by treatment with the histone deacetylase (HDAC) inhibitors in cellular models of HIV-1 latency. *Retrovirology* [Internet]. 6 févr 2016 [cité 18 déc 2017];13. Disponible sur: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4744380/>
134. Sato T, Cesaroni M, Chung W, Panjarian S, Tran A, Madzo J, et al. Transcriptional Selectivity of Epigenetic Therapy in Cancer. *Cancer Res.* 15 janv 2017;77(2):470-81.
135. Foroushani AK, Chim B, Wong M, Rastegar A, Smith PT, Wang S, et al. Posttranscriptional regulation of human endogenous retroviruses by RNA-binding motif protein 4, RBM4. *Proc Natl Acad Sci.* 20 oct 2020;117(42):26520-30.
136. Ku Y, Park J-H, Cho R, Lee Y, Park H-M, Kim M, et al. Noncanonical immune response to the inhibition of DNA methylation by Stau1 via stabilization of endogenous retrovirus RNAs. *Proc Natl Acad Sci.* 30 mars 2021;118(13).
137. Chelmicki T, Roger E, Teissandier A, Dura M, Bonneville L, Rucli S, et al. m6A RNA methylation regulates the fate of endogenous retroviruses. *Nature.* mars 2021;591(7849):312-6.
138. Larouche J-D, Trofimov A, Hesnard L, Ehx G, Zhao Q, Vincent K, et al. Widespread and tissue-specific expression of endogenous retroelements in human somatic tissues. *Genome Med.* déc 2020;12(1):1-16.
139. Deniz Ö, Frost JM, Branco MR. Regulation of transposable elements by DNA modifications. *Nat Rev Genet.* juill 2019;20(7):417-31.
140. Medstrand P, Lagemaat LN van de, Dunn CA, Landry J-R, Svenback D, Mager DL. Impact of transposable elements on the evolution of mammalian gene regulation. *Cytogenet Genome Res.* 2005;110(1-4):342-52.
141. Ploegh HL. Viral strategies of immune evasion. *Science.* 10 avr 1998;280(5361):248-53.
142. Brattås PL, Jönsson ME, Fasching L, Wahlestedt JN, Shahsavani M, Falk R, et al. TRIM28 Controls a Gene Regulatory Network Based on Endogenous Retroviruses in Human Neural Progenitor Cells. *Cell Rep.* 3 janv 2017;18(1):1-11.
143. Fort A, Hashimoto K, Yamada D, Salimullah M, Keya CA, Saxena A, et al. Deep transcriptome profiling of mammalian stem cells supports a regulatory role for retrotransposons in pluripotency maintenance. *Nat Genet.* juin 2014;46(6):558-66.
144. Adoue V, Binet B, Malbec A, Fourquet J, Romagnoli P, van Meerwijk JPM, et al. The Histone Methyltransferase SETDB1 Controls T Helper Cell Lineage Integrity by Repressing Endogenous Retroviruses. *Immunity.* 19 mars 2019;50(3):629-644.e8.
145. Zhang Y, Li T, Preissl S, Amaral ML, Grinstein JD, Farah EN, et al. Transcriptionally active HERV-H retrotransposons demarcate topologically associating domains in human pluripotent stem cells. *Nat Genet.* sept 2019;51(9):1380-8.

146. Grow EJ, Flynn RA, Chavez SL, Bayless NL, Wossidlo M, Wesche DJ, et al. Intrinsic retroviral reactivation in human preimplantation embryos and pluripotent cells. *Nature*. juin 2015;522(7555):221-5.
147. Feschotte C. Transposable elements and the evolution of regulatory networks. *Nat Rev Genet*. mai 2008;9(5):397-405.
148. Jordan IK, Rogozin IB, Glazko GV, Koonin EV. Origin of a substantial fraction of human regulatory sequences from transposable elements. *Trends Genet*. 1 févr 2003;19(2):68-72.
149. Bejerano G, Lowe CB, Ahituv N, King B, Siepel A, Salama SR, et al. A distal enhancer and an ultraconserved exon are derived from a novel retroposon. *Nature*. 4 mai 2006;441(7089):87-90.
150. Chuong EB, Rumi MAK, Soares MJ, Baker JC. Endogenous retroviruses function as species-specific enhancer elements in the placenta. *Nat Genet*. mars 2013;45(3):325-9.
151. Lynch VJ. A copy-and-paste gene regulatory network. *Science*. 4 mars 2016;351(6277):1029-30.
152. Chuong EB, Elde NC, Feschotte C. Regulatory evolution of innate immunity through co-option of endogenous retroviruses. *Science*. 4 mars 2016;351(6277):1083-7.
153. Wang T, Zeng J, Lowe CB, Sellers RG, Salama SR, Yang M, et al. Species-specific endogenous retroviruses shape the transcriptional network of the human tumor suppressor protein p53. *Proc Natl Acad Sci U S A*. 20 nov 2007;104(47):18613-8.
154. Mi S, Lee X, Li X, Veldman GM, Finnerty H, Racie L, et al. Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. *Nature*. févr 2000;403(6771):785-9.
155. Blaise S, de Parseval N, Bénit L, Heidmann T. Genomewide screening for fusogenic human endogenous retrovirus envelopes identifies syncytin 2, a gene conserved on primate evolution. *Proc Natl Acad Sci U S A*. 28 oct 2003;100(22):13013-8.
156. Lokossou AG, Toudic C, Barbeau B. Implication of Human Endogenous Retrovirus Envelope Proteins in Placental Functions. *Viruses*. nov 2014;6(11):4609-27.
157. Vargas A, Toufaily C, LeBellego F, Rassart É, Lafond J, Barbeau B. Reduced expression of both syncytin 1 and syncytin 2 correlates with severity of preeclampsia. *Reprod Sci*. nov 2011;18(11):1085-91.
158. Moore PS, Chang Y. Why do viruses cause cancer? Highlights of the first century of human tumour virology. *Nat Rev Cancer*. déc 2010;10(12):878-89.
159. Kassiotis G. Endogenous Retroviruses and the Development of Cancer. *J Immunol*. 15 févr 2014;192(4):1343-9.
160. Maksakova IA, Romanish MT, Gagnier L, Dunn CA, Lagemaat LN van de, Mager DL. Retroviral Elements and Their Hosts: Insertional Mutagenesis in the Mouse Germ Line. *PLOS Genet*. 27 janv 2006;2(1):e2.



161. Stocking C, Kozak CA. Murine endogenous retroviruses. *Cell Mol Life Sci CMLS*. nov 2008;65(21):3383-98.
162. Fan H, Johnson C. Insertional oncogenesis by non-acute retroviruses: implications for gene therapy. *Viruses*. avr 2011;3(4):398-422.
163. Barbieri D, Elvira-Matelot E, Pelinski Y, Genève L, de Laval B, Yogarajah G, et al. Thrombopoietin protects hematopoietic stem cells from retrotransposon-mediated damage by promoting an antiviral response. *J Exp Med*. 3 avr 2018;215(5):1463-80.
164. Ishida Y, Zhao K, Greenwood AD, Roca AL. Proliferation of Endogenous Retroviruses in the Early Stages of a Host Germ Line Invasion. *Mol Biol Evol*. janv 2015;32(1):109-20.
165. Sarker N, Fabijan J, Owen H, Seddon J, Simmons G, Speight N, et al. Koala retrovirus viral load and disease burden in distinct northern and southern koala populations. *Sci Rep*. 14 janv 2020;10(1):263.
166. McEwen GK, Alquezar-Planas DE, Dayaram A, Gillett A, Tarlinton R, Mongan N, et al. Retroviral integrations contribute to elevated host cancer rates during germline invasion. *Nat Commun*. 26 févr 2021;12(1):1316.
167. Voisset C, Weiss RA, Griffiths DJ. Human RNA « rumor » viruses: the search for novel human retroviruses in chronic disease. *Microbiol Mol Biol Rev*. mars 2008;72(1):157-96, table of contents.
168. Lee E, Iskow R, Yang L, Gokcumen O, Haseley P, Luquette LJ, et al. Landscape of Somatic Retrotransposition in Human Cancers. *Science*. 24 août 2012;337(6097):967-71.
169. Lee YN, Bieniasz PD. Reconstitution of an Infectious Human Endogenous Retrovirus. *PLOS Pathog*. 26 janv 2007;3(1):e10.
170. Contreras-Galindo R, Kaplan MH, Contreras-Galindo AC, Gonzalez-Hernandez MJ, Ferlenghi I, Giusti F, et al. Characterization of Human Endogenous Retroviral Elements in the Blood of HIV-1-Infected Individuals. *J Virol*. 1 janv 2012;86(1):262-76.
171. Young GR, Eksmond U, Salcedo R, Alexopoulou L, Stoye JP, Kassiotis G. Resurrection of endogenous retroviruses in antibody-deficient mice. *Nature*. 29 nov 2012;491(7426):774-8.
172. Lamprecht B, Walter K, Kreher S, Kumar R, Hummel M, Lenze D, et al. Derepression of an endogenous long terminal repeat activates the CSF1R proto-oncogene in human lymphoma. *Nat Med*. mai 2010;16(5):571-9.
173. Ehlers A, Oker E, Bentink S, Lenze D, Stein H, Hummel M. Histone acetylation and DNA demethylation of B cells result in a Hodgkin-like phenotype. *Leukemia*. avr 2008;22(4):835-41.
174. Jang HS, Shah NM, Du AY, Dailey ZZ, Pehrsson EC, Godoy PM, et al. Transposable elements drive widespread expression of oncogenes in human cancers. *Nat Genet*. avr 2019;51(4):611-7.

175. Deniz Ö, Ahmed M, Todd CD, Rio-Machin A, Dawson MA, Branco MR. Endogenous retroviruses are a source of enhancers with oncogenic potential in acute myeloid leukaemia. *Nat Commun.* 14 2020;11(1):3506.
176. Villesen P, Aagaard L, Wiuf C, Pedersen FS. Identification of endogenous retroviral reading frames in the human genome. *Retrovirology.* 11 oct 2004;1:32.
177. Lemaître C, Tsang J, Bireau C, Heidmann T, Dewannieux M. A human endogenous retrovirus-derived gene that can contribute to oncogenesis by activating the ERK pathway and inducing migration and invasion. *PLoS Pathog.* juin 2017;13(6):e1006451.
178. Zhou F, Li M, Wei Y, Lin K, Lu Y, Shen J, et al. Activation of HERV-K Env protein is essential for tumorigenesis and metastasis of breast cancer cells. *Oncotarget.* 20 déc 2016;7(51):84093-117.
179. Serafino A, Balestrieri E, Pierimarchi P, Matteucci C, Moroni G, Oricchio E, et al. The activation of human endogenous retrovirus K (HERV-K) is implicated in melanoma cell malignant transformation. *Exp Cell Res.* 10 mars 2009;315(5):849-62.
180. Chen T, Meng Z, Gan Y, Wang X, Xu F, Gu Y, et al. The viral oncogene Np9 acts as a critical molecular switch for co-activating  $\beta$ -catenin, ERK, Akt and Notch1 and promoting the growth of human leukemia stem/progenitor cells. *Leukemia.* juill 2013;27(7):1469-78.
181. Alcazer V, Bonaventura P, Depil S. Human Endogenous Retroviruses (HERVs): Shaping the Innate Immune Response in Cancers. *Cancers.* mars 2020;12(3):610.
182. Lima-Junior DS, Krishnamurthy SR, Bouladoux N, Collins N, Han S-J, Chen EY, et al. Endogenous retroviruses promote homeostatic and inflammatory responses to the microbiota. *Cell* [Internet]. 23 juin 2021 [cité 7 juill 2021]; Disponible sur: <https://www.sciencedirect.com/science/article/pii/S0092867421006504>
183. Consortium TU. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* 8 janv 2019;47(D1):D506-15.
184. Nakagawa S, Takahashi MU. gEVE: a genome-based endogenous viral element database provides comprehensive viral protein-coding sequences in mammalian genomes. *Database* [Internet]. 1 janv 2016 [cité 7 juill 2021];2016(baw087). Disponible sur: <https://doi.org/10.1093/database/baw087>
185. Thibodeau J, Croteau G, Labrecque N, Sekaly RP. Les superantigènes bactériens : agents pathogènes ou outils thérapeutiques. *Médecine Sci.* 1993;9(5):605-13.
186. Sutkowski N, Conrad B, Thorley-Lawson DA, Huber BT. Epstein-Barr virus transactivates the human endogenous retrovirus HERV-K18 that encodes a superantigen. *Immunity.* oct 2001;15(4):579-89.
187. Barbulescu M, Turner G, Seaman MI, Deinard AS, Kidd KK, Lenz J. Many human endogenous retrovirus K (HERV-K) proviruses are unique to humans. *Curr Biol CB.* 26 août 1999;9(16):861-8.

188. Löwer R, Boller K, Hasenmaier B, Korbmacher C, Müller-Lantzsch N, Löwer J, et al. Identification of human endogenous retroviruses with complex mRNA expression and particle formation. *Proc Natl Acad Sci.* 15 mai 1993;90(10):4480-4.
189. Sauter M, Schommer S, Kremmer E, Remberger K, Dölken G, Lemm I, et al. Human endogenous retrovirus K10: expression of Gag protein and detection of antibodies in patients with seminomas. *J Virol.* janv 1995;69(1):414-21.
190. Towler EM, Gulnik SV, Bhat TN, Xie D, Gustschina E, Sumpter TR, et al. Functional characterization of the protease of human endogenous retrovirus, K10: can it complement HIV-1 protease? *Biochemistry.* 8 déc 1998;37(49):17137-44.
191. Muster T, Waltenberger A, Grassauer A, Hirschl S, Caucig P, Romirer I, et al. An Endogenous Retrovirus Derived from Human Melanoma Cells. *Cancer Res.* 15 déc 2003;63(24):8735-41.
192. Sauter M, Roemer K, Best B, Afting M, Schommer S, Seitz G, et al. Specificity of antibodies directed against Env protein of human endogenous retroviruses in patients with germ cell tumors. *Cancer Res.* 1 oct 1996;56(19):4362-5.
193. Schiavetti F, Thonnard J, Colau D, Boon T, Coulie PG. A Human Endogenous Retroviral Sequence Encoding an Antigen Recognized on Melanoma by Cytolytic T Lymphocytes. *Cancer Res.* 1 oct 2002;62(19):5510-6.
194. Kershaw MH, Hsu C, Mondesire W, Parker LL, Wang G, Overwijk WW, et al. Immunization against Endogenous Retroviral Tumor-associated Antigens. *Cancer Res.* 1 nov 2001;61(21):7920-4.
195. Sacha JB, Kim I-J, Chen L, Ullah JH, Goodwin DA, Simmons HA, et al. Vaccination with Cancer- and HIV Infection-Associated Endogenous Retrotransposable Elements Is Safe and Immunogenic. *J Immunol.* 1 août 2012;189(3):1467-79.
196. Krishnamurthy J, Rabinovich BA, Mi T, Switzer KC, Olivares S, Maiti SN, et al. Genetic Engineering of T Cells to Target HERV-K, an Ancient Retrovirus on Melanoma. *Clin Cancer Res.* 15 juill 2015;21(14):3241-51.
197. Zhou F, Krishnamurthy J, Wei Y, Li M, Hunt K, Johanning GL, et al. Chimeric antigen receptor T cells targeting HERV-K inhibit breast cancer and its metastasis through downregulation of Ras. *Oncoimmunology.* nov 2015;4(11):e1047582.
198. Bentzen AK, Marquard AM, Lyngaa R, Saini SK, Ramskov S, Donia M, et al. Large-scale detection of antigen-specific T cells using peptide-MHC-I multimers labeled with DNA barcodes. *Nat Biotechnol.* oct 2016;34(10):1037-45.
199. Saini SK, Ørskov AD, Bjerregaard A-M, Unnikrishnan A, Holmberg-Thydén S, Borch A, et al. Human endogenous retroviruses form a reservoir of T cell targets in hematological cancers. *Nat Commun.* 9 nov 2020;11(1):5660.
200. Hogquist KA, Baldwin TA, Jameson SC. Central tolerance: learning self-control in the thymus. *Nat Rev Immunol.* oct 2005;5(10):772-82.

201. Klein L, Kyewski B, Allen PM, Hogquist KA. Positive and negative selection of the T cell repertoire: what thymocytes see (and don't see). *Nat Rev Immunol*. juin 2014;14(6):377-91.
202. Yu W, Jiang N, Ebert PJR, Kidd BA, Müller S, Lund PJ, et al. Clonal Deletion Prunes but Does Not Eliminate Self-Specific  $\alpha\beta$  CD8<sup>+</sup> T Lymphocytes. *Immunity*. 19 mai 2015;42(5):929-41.
203. Inglesfield S, Cosway EJ, Jenkinson WE, Anderson G. Rethinking Thymic Tolerance: Lessons from Mice. *Trends Immunol*. 1 avr 2019;40(4):279-91.
204. Ebert PJR, Jiang S, Xie J, Li Q-J, Davis MM. An endogenous positively selecting peptide enhances mature T cell responses and becomes an autoantigen in the absence of microRNA miR-181a. *Nat Immunol*. nov 2009;10(11):1162-9.
205. Stefanová I, Dorfman JR, Germain RN. Self-recognition promotes the foreign antigen sensitivity of naive T lymphocytes. *Nature*. 28 nov 2002;420(6914):429-34.
206. Young GR, Ploquin MJ-Y, Eksmond U, Wadwa M, Stoye JP, Kassiotis G. Negative Selection by an Endogenous Retrovirus Promotes a Higher-Avidity CD4<sup>+</sup> T Cell Response to Retroviral Infection. *PLOS Pathog*. 10 mai 2012;8(5):e1002709.
207. Meylan F, De Smedt M, Leclercq G, Plum J, Leupin O, Marguerat S, et al. Negative thymocyte selection to HERV-K18 superantigens in humans. *Blood*. 1 juin 2005;105(11):4377-82.
208. Denner J. Immunosuppressive properties of retroviruses. *Eur J Immunol*. 1 janv 2016;46(1):253-5.
209. Morozov VA, Thi VLD, Denner J. The Transmembrane Protein of the Human Endogenous Retrovirus - K (HERV-K) Modulates Cytokine Release and Gene Expression. *PLOS ONE*. 7 août 2013;8(8):e70399.
210. Barcova M, Kacani L, Speth C, Dierich MP. gp41 envelope protein of human immunodeficiency virus induces interleukin (IL)-10 in monocytes, but not in B, T, or NK cells, leading to reduced IL-2 and interferon-gamma production. *J Infect Dis*. avr 1998;177(4):905-13.
211. Mangeney M, Renard M, Schlecht-Louf G, Bouallaga I, Heidmann O, Letzelter C, et al. Placental syncytins: Genetic disjunction between the fusogenic and immunosuppressive activity of retroviral envelope proteins. *Proc Natl Acad Sci*. 18 déc 2007;104(51):20534-9.
212. Tolosa JM, Schjenken JE, Clifton VL, Vargas A, Barbeau B, Lowry P, et al. The endogenous retroviral envelope protein syncytin-1 inhibits LPS/PHA-stimulated cytokine responses in human blood and is sorted into placental exosomes. *Placenta*. 1 nov 2012;33(11):933-41.
213. Holder BS, Tower CL, Forbes K, Mulla MJ, Aplin JD, Abrahams VM. Immune cell activation by trophoblast-derived microvesicles is mediated by syncytin 1. *Immunology*. juin 2012;136(2):184-91.

214. Hummel J, Kämmerer U, Müller N, Avota E, Schneider-Schaulies S. Human endogenous retrovirus envelope proteins target dendritic cells to suppress T-cell activation. *Eur J Immunol.* juin 2015;45(6):1748-59.
215. Renaudineau Y, Hillion S, Saraux A, Mageed RA, Youinou P. An alternative exon 1 of the CD5 gene regulates CD5 expression in human B lymphocytes. *Blood.* 15 oct 2005;106(8):2781-9.
216. Chikuma S, Yamanaka S, Nakagawa S, Ueda MT, Hayabuchi H, Tokifuji Y, et al. TRIM28 Expression on Dendritic Cells Prevents Excessive T Cell Priming by Silencing Endogenous Retrovirus. *J Immunol* [Internet]. 22 févr 2021 [cité 16 mars 2021]; Disponible sur: <https://www.jimmunol.org/content/early/2021/02/19/jimmunol.2001003>
217. Chung HY, Kim DH, Lee EK, Chung KW, Chung S, Lee B, et al. Redefining Chronic Inflammation in Aging and Age-Related Diseases: Proposal of the Senoinflammation Concept. *Aging Dis.* 1 avr 2019;10(2):367-82.
218. Doxiadis GGM, Groot N de, Bontrop RE. Impact of Endogenous Intronic Retroviruses on Major Histocompatibility Complex Class II Diversity and Stability. *J Virol.* 1 juill 2008;82(13):6667-77.
219. Kulski JK, Gaudieri S, Bellgard M, Balmer L, Giles K, Inoko H, et al. The evolution of MHC diversity by segmental duplication and transposition of retroelements. *J Mol Evol.* déc 1997;45(6):599-609.
220. Zhao Z, Zhang Z, Li J, Dong Q, Xiong J, Li Y, et al. Sustained TNF- $\alpha$  stimulation leads to transcriptional memory that greatly enhances signal sensitivity and robustness. *Shi X, Struhl K, Shi X, éditeurs. eLife.* 6 nov 2020;9:e61965.
221. Chiappinelli KB, Strissel PL, Desrichard A, Li H, Henke C, Akman B, et al. Inhibiting DNA Methylation Causes an Interferon Response in Cancer via dsRNA Including Endogenous Retroviruses. *Cell.* 27 août 2015;162(5):974-86.
222. Roulois D, Loo Yau H, Singhania R, Wang Y, Danesh A, Shen SY, et al. DNA-Demethylating Agents Target Colorectal Cancer Cells by Inducing Viral Mimicry by Endogenous Transcripts. *Cell.* 27 août 2015;162(5):961-73.
223. Ohtani H, Ørskov AD, Helbo AS, Gillberg L, Liu M, Zhou W, et al. Activation of a Subset of Evolutionarily Young Transposable Elements and Innate Immunity Are Linked to Clinical Responses to 5-Azacytidine. *Cancer Res.* 15 juin 2020;80(12):2441-50.
224. Brocks D, Schmidt CR, Daskalakis M, Jang HS, Shah NM, Li D, et al. DNMT and HDAC inhibitors induce cryptic transcription start sites encoded in long terminal repeats. *Nat Genet.* juill 2017;49(7):1052.
225. Schumacher TN, Schreiber RD. Neoantigens in cancer immunotherapy. *Science.* 3 avr 2015;348(6230):69-74.
226. Ghoneim HE, Fan Y, Moustaki A, Abdelsamed HA, Dash P, Dogra P, et al. De Novo Epigenetic Programs Inhibit PD-1 Blockade-Mediated T Cell Rejuvenation. *Cell.* 29 juin 2017;170(1):142-157.e19.

227. Bolis M, Paroni G, Fratelli M, Vallerga A, Guarrera L, Zanetti A, et al. All-Trans Retinoic Acid Stimulates Viral Mimicry, Interferon Responses and Antigen Presentation in Breast-Cancer Cells. *Cancers*. mai 2020;12(5):1169.
228. Mikhalkevich N, O'Carroll IP, Tkavc R, Lund K, Sukumar G, Dalgard CL, et al. Response of human macrophages to gamma radiation is mediated via expression of endogenous retroviruses. *PLoS Pathog*. févr 2021;17(2):e1009305.
229. Zhou X, Singh M, Santos GS, Guerlavais V, Carvajal LA, Aivado M, et al. Pharmacological activation of p53 triggers viral mimicry response thereby abolishing tumor immune evasion and promoting anti-tumor immunity. *Cancer Discov* [Internet]. 1 janv 2021 [cité 22 juill 2021]; Disponible sur: <http://cancerdiscovery.aacrjournals.org/content/early/2021/07/01/2159-8290.CD-20-1741>
230. Lanciano S, Cristofari G. Measuring and interpreting transposable element expression. *Nat Rev Genet*. déc 2020;21(12):721-36.
231. Sudmant PH, Rausch T, Gardner EJ, Handsaker RE, Abyzov A, Huddleston J, et al. An integrated map of structural variation in 2,504 human genomes. *Nature*. oct 2015;526(7571):75-81.
232. Goerner-Potvin P, Bourque G. Computational tools to unmask transposable elements. *Nat Rev Genet*. nov 2018;19(11):688-704.
233. Jurka J, Kapitonov VV, Pavlicek A, Klonowski P, Kohany O, Walichiewicz J. Repbase Update, a database of eukaryotic repetitive elements. *Cytogenet Genome Res*. 2005;110(1-4):462-7.
234. Hubley R, Finn RD, Clements J, Eddy SR, Jones TA, Bao W, et al. The Dfam database of repetitive DNA families. *Nucleic Acids Res*. 4 janv 2016;44(D1):D81-9.
235. Smit AFA, Hubley R, Greenberg P. RepeatMasker [Internet]. RepeatMasker Open-4.0. 2013 [cité 25 févr 2021]. Disponible sur: <http://www.repeatmasker.org>
236. Lerat E. Identifying repeats and transposable elements in sequenced genomes: how to find your way through the dense forest of programs. *Heredity*. juin 2010;104(6):520-33.
237. Hoen DR, Hickey G, Bourque G, Casacuberta J, Cordaux R, Feschotte C, et al. A call for benchmarking transposable element annotation methods. *Mob DNA*. 4 août 2015;6(1):13.
238. Pačes J, Pavlíček A, Pačes V. HERVD: database of human endogenous retroviruses. *Nucleic Acids Res*. 1 janv 2002;30(1):205-6.
239. Koning APJ de, Gu W, Castoe TA, Batzer MA, Pollock DD. Repetitive Elements May Comprise Over Two-Thirds of the Human Genome. *PLOS Genet*. 1 déc 2011;7(12):e1002384.
240. Caballero J, Smit AFA, Hood L, Glusman G. Realistic artificial DNA sequences as negative controls for computational genomics. *Nucleic Acids Res*. juill 2014;42(12):e99.
241. Sperber GO, Airola T, Jern P, Blomberg J. Automated recognition of retroviral sequences in genomic data—RetroTector©. *Nucleic Acids Res*. août 2007;35(15):4964-76.



242. Guo R, Li Y-R, He S, Ou-Yang L, Sun Y, Zhu Z. RepLong: de novo repeat identification using long read sequencing data. *Bioinformatics*. 1 avr 2018;34(7):1099-107.
243. Wang J, Song L, Grover D, Azrak S, Batzer MA, Liang P. dbRIP: a highly integrated database of retrotransposon insertion polymorphisms in humans. *Hum Mutat*. avr 2006;27(4):323-9.
244. Mir AA, Philippe C, Cristofari G. euL1db: the European database of L1HS retrotransposon insertions in humans. *Nucleic Acids Res*. 28 janv 2015;43(Database issue):D43-7.
245. Büscher K, Trefzer U, Hofmann M, Sterry W, Kurth R, Denner J. Expression of human endogenous retrovirus K in melanomas and melanoma cell lines. *Cancer Res*. 15 mai 2005;65(10):4172-80.
246. de Parseval N, Lazar V, Casella J-F, Benit L, Heidmann T. Survey of Human Genes of Retroviral Origin: Identification and Transcriptome of the Genes with Coding Capacity for Complete Envelope Proteins. *J Virol*. oct 2003;77(19):10414-22.
247. Okahara G, Matsubara S, Oda T, Sugimoto J, Jinno Y, Kanaya F. Expression analyses of human endogenous retroviruses (HERVs): tissue-specific and developmental stage-dependent expression of HERVs. *Genomics*. 1 déc 2004;84(6):982-90.
248. Forsman A, Yun Z, Hu L, Uzhamackis D, Jern P, Blomberg J. Development of broadly targeted human endogenous gammaretroviral pol-based real time PCRs Quantitation of RNA expression in human tissues. *J Virol Methods*. oct 2005;129(1):16-30.
249. Deininger P, Morales ME, White TB, Baddoo M, Hedges DJ, Servant G, et al. A comprehensive approach to expression of L1 loci. *Nucleic Acids Res*. 17 mars 2017;45(5):e31.
250. Flockerzi A, Maydt J, Frank O, Ruggieri A, Maldener E, Seifarth W, et al. Expression pattern analysis of transcribed HERV sequences is complicated by ex vivo recombination. *Retrovirology*. 6 juin 2007;4(1):39.
251. Pichon J-P, Bonnaud B, Mallet F. Quantitative multiplex degenerate PCR for human endogenous retrovirus expression profiling. *Nat Protoc*. 2006;1(6):2831-8.
252. Becker J, Pérot P, Cheynet V, Oriol G, Mugnier N, Mommert M, et al. A comprehensive hybridization model allows whole HERV transcriptome profiling using high density microarray. *BMC Genomics*. 8 avr 2017;18(1):286.
253. Oja M, Peltonen J, Blomberg J, Kaski S. Methods for estimating human endogenous retrovirus activities from EST databases. *BMC Bioinformatics*. 3 mai 2007;8(2):S11.
254. Tokuyama M, Kong Y, Song E, Jayewickreme T, Kang I, Iwasaki A. ERVmap analysis reveals genome-wide transcription of human endogenous retroviruses. *Proc Natl Acad Sci*. 11 déc 2018;115(50):12565-72.
255. Sexton CE, Han MV. Paired-end mappability of transposable elements in the human genome. *Mob DNA*. 10 juill 2019;10(1):29.

256. Chung N, Jonaid GM, Quinton S, Ross A, Sexton CE, Alberto A, et al. Transcriptome analyses of tumor-adjacent somatic tissues reveal genes co-expressed with transposable elements. *Mob DNA*. 3 sept 2019;10(1):39.
257. Smith CC, Beckermann KE, Bortone DS, De Cubas AA, Bixby LM, Lee SJ, et al. Endogenous retroviral signatures predict immunotherapy response in clear cell renal cell carcinoma. *J Clin Invest*. 1 nov 2018;128(11):4804-20.
258. Bendall ML, de Mulder M, Iñiguez LP, Lecanda-Sánchez A, Pérez-Losada M, Ostrowski MA, et al. Telescope: Characterization of the retrotranscriptome by accurate estimation of transposable element expression. *PLoS Comput Biol*. 2019;15(9):e1006453.
259. Decker CJ, Steiner HR, Hoon-Hanks LL, Morrison JH, Haist KC, Stabell AC, et al. dsRNA-Seq: Identification of Viral Infection by Purifying and Sequencing dsRNA. *Viruses* [Internet]. 14 oct 2019 [cité 25 févr 2021];11(10). Disponible sur: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6832592/>
260. Solovyov A, Vabret N, Arora KS, Snyder A, Funt SA, Bajorin DF, et al. Global Cancer Transcriptome Quantifies Repeat Element Polarization between Immunotherapy Responsive and T Cell Suppressive Classes. *Cell Rep*. avr 2018;23(2):512-21.
261. Rooney MS, Shukla SA, Wu CJ, Getz G, Hacohen N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell*. 15 janv 2015;160(1-2):48-61.
262. Göke J, Lu X, Chan Y-S, Ng H-H, Ly L-H, Sachs F, et al. Dynamic Transcription of Distinct Classes of Endogenous Retroviral Elements Marks Specific Populations of Early Human Embryonic Cells. *Cell Stem Cell*. 5 févr 2015;16(2):135-41.
263. Subramanian RP, Wildschutte JH, Russo C, Coffin JM. Identification, characterization, and comparative genomic distribution of the HERV-K (HML-2) group of human endogenous retroviruses. *Retrovirology*. 8 nov 2011;8:90.
264. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*. 4 août 2011;12(1):323.
265. Attig J, Young GR, Hosie L, Perkins D, Encheva-Yokoya V, Stoye JP, et al. LTR retroelement expansion of the human cancer transcriptome and immunopeptidome revealed by de novo transcript assembly. *Genome Res*. 19 sept 2019;31:88-100.
266. Kazachenka A, Young GR, Attig J, Kordella C, Lamprianidou E, Zoulia E, et al. Epigenetic therapy of myelodysplastic syndromes connects to cellular differentiation independently of endogenous retroelement derepression. *Genome Med*. 23 déc 2019;11(1):86.
267. Hanahan D, Weinberg RA. Hallmarks of Cancer: The Next Generation. *Cell*. 4 mars 2011;144(5):646-74.
268. Wentzensen N, Coy JF, Knaebel H-P, Linnebacher M, Wilz B, Gebert J, et al. Expression of an endogenous retroviral sequence from the HERV-H group in gastrointestinal cancers. *Int J Cancer*. 2007;121(7):1417-23.

269. Fermin D, Allen BB, Blackwell TW, Menon R, Adamski M, Xu Y, et al. Novel gene and gene model detection using a whole genome open reading frame analysis in proteomics. *Genome Biol.* 28 avr 2006;7(4):R35.
270. Khatun J, Yu Y, Wrobel JA, Risk BA, Gunawardena HP, Secret A, et al. Whole human genome proteogenomic mapping for ENCODE cell line data: identifying protein-coding regions. *BMC Genomics.* 28 févr 2013;14(1):141.
271. Fan K, Wang W. What is the minimum number of letters required to fold a protein? *J Mol Biol.* 9 mai 2003;328(4):921-6.
272. Nesvizhskii AI. Proteogenomics: concepts, applications and computational strategies. *Nat Methods.* nov 2014;11(11):1114-25.
273. Newey A, Griffiths B, Michaux J, Pak HS, Stevenson BJ, Woolston A, et al. Immunopeptidomics of colorectal cancer organoids reveals a sparse HLA class I neoantigen landscape and no increase in neoantigens with interferon or MEK-inhibitor treatment. *J Immunother Cancer.* 18 nov 2019;7(1):309.
274. Löffler MW, Mohr C, Bichmann L, Freudenmann LK, Walzer M, Schroeder CM, et al. Multi-omics discovery of exome-derived neoantigens in hepatocellular carcinoma. *Genome Med.* 30 avr 2019;11(1):28.
275. Mattos-Arruda LD, Vazquez M, Finotello F, Lepore R, Porta E, Hundal J, et al. Neoantigen prediction and computational perspectives towards clinical benefit: recommendations from the ESMO Precision Medicine Working Group. *Ann Oncol.* 1 août 2020;31(8):978-90.
276. Richman LP, Vonderheide RH, Rech AJ. Neoantigen Dissimilarity to the Self-Proteome Predicts Immunogenicity and Response to Immune Checkpoint Blockade. *Cell Syst.* 23 oct 2019;9(4):375-382.e4.
277. Pearson H, Daouda T, Granados DP, Durette C, Bonneil E, Courcelles M, et al. MHC class I-associated peptides derive from selective regions of the human genome. *J Clin Invest.* 1 déc 2016;126(12):4690-701.
278. Cosma GL, Eisenlohr LC. Impact of epitope density on CD8+ T cell development and function. *Mol Immunol.* 1 sept 2019;113:120-5.
279. Reynisson B, Alvarez B, Paul S, Peters B, Nielsen M. NetMHCpan-4.1 and NetMHCIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. *Nucleic Acids Res.* 2 juill 2020;48(W1):W449-54.
280. O'Donnell TJ, Rubinsteyn A, Laserson U. MHCflurry 2.0: Improved Pan-Allele Prediction of MHC Class I-Presented Peptides by Incorporating Antigen Processing. *Cell Syst.* 21 oct 2020;11(4):418-9.
281. Tay RE, Richardson EK, Toh HC. Revisiting the role of CD4 + T cells in cancer immunotherapy—new insights into old paradigms. *Cancer Gene Ther.* févr 2021;28(1):5-17.

282. Ott PA, Hu Z, Keskin DB, Shukla SA, Sun J, Bozym DJ, et al. An immunogenic personal neoantigen vaccine for patients with melanoma. *Nature*. juill 2017;547(7662):217-21.
283. Kreiter S, Vormehr M, van de Roemer N, Diken M, Löwer M, Diekmann J, et al. Mutant MHC class II epitopes drive therapeutic immune responses to cancer. *Nature*. 30 avr 2015;520(7549):692-6.
284. Racle J, Michaux J, Rockinger GA, Arnaud M, Bobisse S, Chong C, et al. Robust prediction of HLA class II epitopes by deep motif deconvolution of immunopeptidomes. *Nat Biotechnol*. nov 2019;37(11):1283-6.
285. Chen B, Khodadoust MS, Olsson N, Wagar LE, Fast E, Liu CL, et al. Predicting HLA class II antigen presentation through integrated deep learning. *Nat Biotechnol*. nov 2019;37(11):1332-43.
286. Dvinge H, Bradley RK. Widespread intron retention diversifies most cancer transcriptomes. *Genome Med*. 15 mai 2015;7(1):45.
287. Vidova V, Spacil Z. A review on mass spectrometry-based quantitative proteomics: Targeted and data independent acquisition. *Anal Chim Acta*. 29 avr 2017;964:7-23.
288. Hsiue EH-C, Wright KM, Douglass J, Hwang MS, Mog BJ, Pearlman AH, et al. Targeting a neoantigen derived from a common TP53 mutation. *Science* [Internet]. 5 mars 2021 [cité 21 avr 2021];371(6533). Disponible sur: <https://science.sciencemag.org/content/371/6533/eabc8697>
289. Calis JJA, Maybeno M, Greenbaum JA, Weiskopf D, De Silva AD, Sette A, et al. Properties of MHC Class I Presented Peptides That Enhance Immunogenicity. Asquith B, éditeur. *PLoS Comput Biol*. 24 oct 2013;9(10):e1003266.
290. Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-Garcia W, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun*. 2013;4:2612.
291. Panda A, Cubas AA de, Stein M, Riedlinger G, Kra J, Mayer T, et al. Endogenous retrovirus expression is associated with response to immune checkpoint blockade in clear cell renal cell carcinoma. *JCI Insight* [Internet]. 23 août 2018 [cité 25 févr 2019];3(16). Disponible sur: <https://insight.jci.org/articles/view/121522>
292. Tibshirani R. Regression Shrinkage and Selection via the Lasso. *J R Stat Soc*. 1996;58(1):267-88.
293. Wooldridge L, Ekeruche-Makinde J, van den Berg HA, Skowera A, Miles JJ, Tan MP, et al. A single autoimmune T cell receptor recognizes more than a million different peptides. *J Biol Chem*. 6 janv 2012;287(2):1168-77.
294. Petrova G, Ferrante A, Gorski J. Cross-Reactivity of T Cells and Its Role in the Immune System. *Crit Rev Immunol*. 2012;32(4):349-72.
295. Ishizuka J, Grebe K, Shenderov E, Peters B, Chen Q, Peng Y, et al. Quantitating T cell cross-reactivity for unrelated peptide antigens. *J Immunol*. 1 oct 2009;183(7):4337-45.

296. Sewell AK. Why must T cells be cross-reactive? *Nat Rev Immunol.* sept 2012;12(9):669-77.
297. Bentzen AK, Hadrup SR. T-cell-receptor cross-recognition and strategies to select safe T-cell receptors for clinical translation. *Immuno-Oncol Technol.* 1 sept 2019;2:1-10.
298. Wooldridge L, Laugel B, Ekeruche J, Clement M, Berg HA van den, Price DA, et al. CD8 Controls T Cell Cross-Reactivity. *J Immunol.* 15 oct 2010;185(8):4625-32.
299. Gee MH, Han A, Lofgren SM, Beausang JF, Mendoza JL, Birnbaum ME, et al. Antigen Identification for Orphan T Cell Receptors Expressed on Tumor-Infiltrating Lymphocytes. *Cell.* 25 janv 2018;172(3):549-563.e16.
300. Fischer DS, Wu Y, Schubert B, Theis FJ. Predicting antigen specificity of single T cells based on TCR CDR3 regions. *Mol Syst Biol.* 1 août 2020;16(8):e9416.
301. Lee CH, Salio M, Napolitani G, Ogg G, Simmons A, Koohy H. Predicting Cross-Reactivity and Antigen Specificity of T Cell Receptors. *Front Immunol [Internet].* 22 oct 2020 [cité 20 avr 2021];11. Disponible sur: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7642207/>
302. Boegel S, Löwer M, Bukur T, Sorn P, Castle JC, Sahin U. HLA and proteasome expression body map. *BMC Med Genomics.* 27 mars 2018;11(1):36.
303. Colombo AR, Elias HK, Ramsingh G. Senescence induction universally activates transposable element expression. *Cell Cycle.* 2018;17(14):1846-57.
304. Cecco MD, Ito T, Petrashen AP, Elias AE, Skvir NJ, Criscione SW, et al. L1 drives IFN in senescent cells and promotes age-associated inflammation. *Nature.* févr 2019;566(7742):73-8.
305. Takahashi A, Loo TM, Okada R, Kamachi F, Watanabe Y, Wakita M, et al. Downregulation of cytoplasmic DNases is implicated in cytoplasmic DNA accumulation and SASP in senescent cells. *Nat Commun.* 28 mars 2018;9(1):1249.
306. Jintaridith P, Mutirangura A. Distinctive patterns of age-dependent hypomethylation in interspersed repetitive sequences. *Physiol Genomics.* 1 avr 2010;41(2):194-200.
307. Pérez RF, Tejedor JR, Bayón GF, Fernández AF, Fraga MF. Distinct chromatin signatures of DNA hypomethylation in aging and cancer. *Aging Cell.* juin 2018;17(3):e12744.
308. Liu X, Liu Z, Sun L, Ren J, Wu Z, Jiang X, et al. Resurrection of human endogenous retroviruses during aging reinforces senescence. *bioRxiv.* 22 févr 2021;2021.02.22.432260.
309. Thiagarajan D, Capone S, Connor K, Aggarwal PK, Li X, Wu K, et al. Immune Clearance of Senescent Cells Via Activation of Endogenous Retroviruses and the Viral Recognition Pathway. *Blood.* 7 déc 2017;130(Supplement 1):3576-3576.
310. Ng SWK, Mitchell A, Kennedy JA, Chen WC, McLeod J, Ibrahimova N, et al. A 17-gene stemness score for rapid determination of risk in acute leukaemia. *Nature.* 15 2016;540(7633):433-7.

311. Thomas D, Majeti R. Biology and relevance of human acute myeloid leukemia stem cells. *Blood*. 23 mars 2017;129(12):1577-85.
312. Vadakekolathu J, Minden MD, Hood T, Church SE, Reeder S, Altmann H, et al. Immune landscapes predict chemotherapy resistance and immunotherapy response in acute myeloid leukemia. *Sci Transl Med* [Internet]. 3 juin 2020 [cité 22 juill 2021];12(546). Disponible sur: <https://stm.sciencemag.org/content/12/546/eaaz0463>
313. Curty G, Marston JL, de Mulder Rougvie M, Leal FE, Nixon DF, Soares MA. Human Endogenous Retrovirus K in Cancer: A Potential Biomarker and Immunotherapeutic Target. *Viruses*. juill 2020;12(7):726.
314. Natoli M, Gallon J, Lu H, Amgheib A, Pinato DJ, Mauri FA, et al. Transcriptional analysis of multiple ovarian cancer cohorts reveals prognostic and immunomodulatory consequences of ERV expression. *J Immunother Cancer*. 1 janv 2021;9(1):e001519.
315. van Galen P, Hovestadt V, Wadsworth Li MH, Hughes TK, Griffin GK, Battaglia S, et al. Single-Cell RNA-Seq Reveals AML Hierarchies Relevant to Disease Progression and Immunity. *Cell*. 7 mars 2019;176(6):1265-1281.e24.
316. Ding J, Adiconis X, Simmons SK, Kowalczyk MS, Hession CC, Marjanovic ND, et al. Systematic comparison of single-cell and single-nucleus RNA-sequencing methods. *Nat Biotechnol*. juin 2020;38(6):737-46.
317. Shao W, Wang T. Transcript assembly improves expression quantification of transposable elements in single-cell RNA-seq data. *Genome Res*. 1 janv 2021;31(1):88-100.



## ANNEXES

---

## Brevets déposés

**AC BET 21L1108** - Transcriptomic signature based on HERVs expression to characterize leukemic stem cells.

**AC BET 21L1109** - Use of a transcriptomic signature based on HERVs expression to characterize new acute myeloid leukemia subtypes.



# RÉSUMÉ

---

Les rétrovirus endogènes humains (HERVs) représentent 8% de notre génome. Réprimés par des mécanismes épigénétiques dans les cellules normales, différentes situations peuvent mener à leur dérégulation comme dans les maladies auto-immunes ou le cancer. Les HERVs sont impliqués dans de nombreuses fonctions physiologiques ou pathologiques via leur activité de régulation génique, l'exaptation de fonctions virales ou encore la modulation immunitaire induite par leurs intermédiaires nucléotidiques ou protéiques. Du fait de ce profil d'expression particulier et de leur origine virale, les protéines dérivées des HERVs pourraient fournir de très bonnes cibles immunothérapeutiques en oncologie. Il n'existe cependant à ce jour pas de méthode pour identifier les cibles les plus pertinentes dans un cancer donné. Par ailleurs, le profil d'expression des HERVs et leur impact immunitaire n'a à ce jour jamais été étudié de façon exhaustive dans certaines pathologies comme la leucémie aiguë myéloïde (LAM).

Cette thèse fait le point sur les techniques d'identification d'épitopes lymphocytaires T CD8<sup>+</sup> dérivés de HERVs, ainsi que sur l'utilisation du rétrotranscriptome HERV via une quantification bioinformatique sensible pour caractériser des populations cellulaires. Après plusieurs revues sur les approches immunothérapeutiques actuelles basées sur les lymphocytes T en oncologie, nous situons les épitopes dérivés des HERVs comme alternative prometteuse aux néoépitopes classiques issus de mutations géniques. Dans le premier article, nous proposons une méthode permettant de définir les meilleurs épitopes T CD8<sup>+</sup> à partir de la base pancancer du TCGA. Basée sur une approche d'apprentissage automatique permettant de détecter les HERVs associés à une réponse cytotoxique au sein des HERVs surexprimés par un cancer donné et non exprimés par le tissu sain, notre méthode permet de réduire d'environ 90% le nombre d'épitopes candidats à tester. Nous montrons ensuite dans un second temps qu'une sélection d'épitopes partagés sélectionnés via cette méthode peut induire une réponse lymphocytaire T CD8<sup>+</sup> spécifique et fonctionnelle chez les donneurs sains, avec une activité cytotoxique contre des lignées cellulaires de cancer du sein triple-négatif *in vitro*.

Dans le second article, nous utilisons une méthode récente de quantification sensible de l'ensemble des séquences HERVs du génome humain pour montrer que le rétrotranscriptome HERV permet de caractériser des populations cellulaires normales et leucémiques. Nous montrons qu'une telle approche peut être utilisée pour établir une signature de cellules souches leucémiques basée sur l'expression de 25 HERVs. Nous utilisons ensuite cette quantification au sein de 4 cohortes indépendantes de patients présentant une LAM au diagnostic pour définir de nouveaux sous-types de LAM via une approche de classification non supervisée. Enfin, nous montrons que des épitopes T CD8<sup>+</sup> peuvent être définis parmi des HERVs exprimés de façon spécifique dans la LAM, et que des réponses T CD8<sup>+</sup> spécifiques de ces épitopes peuvent être identifiées parmi les lymphocytes infiltrant la moelle osseuse de patients au diagnostic.

Dans l'ensemble, ces résultats fournissent le rationnel et la méthodologie pour l'utilisation d'épitopes dérivés de HERVs comme cibles immunothérapeutiques ainsi que du rétrotranscriptome HERV pour la caractérisation de populations cellulaires et la mise au point de signatures transcriptomiques.