



HAL
open science

Impact des anomalies moléculaires dans l'histoire naturelle de la leucémie lymphoïde chronique

Jasmine Chauzeix

► **To cite this version:**

Jasmine Chauzeix. Impact des anomalies moléculaires dans l'histoire naturelle de la leucémie lymphoïde chronique. Médecine humaine et pathologie. Université de Limoges, 2018. Français. NNT : 2018LIMO0051 . tel-02405662

HAL Id: tel-02405662

<https://theses.hal.science/tel-02405662>

Submitted on 11 Dec 2019

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.

Université de Limoges

ED 615 - Sciences Biologiques et Santé (SBS)

CRIBL – UMR CNRS 7276, INSERM 1262, axe mécanismes moléculaires de la lymphomagénèse

Thèse pour obtenir le grade de
Docteur de l'Université de Limoges
Biologie, médecine et santé

Présentée et soutenue par
Jasmine CHAUZEIX

Le 11 décembre 2018

Impact des anomalies moléculaires dans l'histoire naturelle de la leucémie lymphoïde chronique

Thèse dirigée par Dr Nathalie GACHARD et Pr Jean FEUILLARD

JURY :

Président du jury

M. Michel Cogné, PU-PH, UMR CNRS 7276, UMR INSERM 1262, Faculté de Médecine de Limoges

Rapporteurs

Mme Mary Callanan, PU-PH, UMR INSERM U1209, Faculté de Médecine de Dijon

Mme Fanny Baran-Marszak, PU-PH, UMR INSERM U978, Faculté de Médecine Paris 13

Examineur

M. Philippe Ruminy, UMR INSERM U918, Centre Henri Becquerel, Rouen

Invité

M. David Rizzo, MCU-PH, UMR CNRS 7276, UMR INSERM 1262, Faculté de Médecine de Limoges

A ma fille Flora, qui m'a accompagnée pendant toute l'écriture de cette thèse



Remerciements

Je tiens à remercier tout d'abord Madame le Professeur Fanny Baran-Marszak et Madame le Professeur Mary Callanan d'avoir accepté d'être les rapportrices de mon travail de thèse. Je remercie également le Docteur Philippe Ruminy d'avoir accepté d'être membre de ce jury.

Je remercie Monsieur le Professeur Michel Cogné d'avoir accepté d'être le président de ce jury. Je le remercie également pour les échanges scientifiques réalisés au sein du laboratoire m'ayant permis d'améliorer ma compréhension des mécanismes de réarrangements des gènes des immunoglobulines et d'alimenter ma réflexion sur les remaniements atypiques étudiés dans ce travail.

Je remercie Madame le Docteur Nathalie Gachard et Monsieur le Professeur Jean Feuillard, mes co-directeurs de thèse, pour avoir supervisé ce travail de manière active et pour leur réactivité. Merci pour vos relectures attentives de ce mémoire et vos conseils avisés.

Je remercie également le Docteur David Rizzo d'avoir accepté d'être membre de ce jury mais aussi pour son aide précieuse au cours de ce travail de thèse. Merci pour tes relectures de ce manuscrit, tes conseils et ta gentillesse.

Je tiens à remercier Mesdames Mélanie Deveza, Catherine Soria et Anne Thibault pour leur aide technique en cytogénétique. Sans vous, ce travail n'aurait pas pu avancer de cette façon. Merci également pour votre sympathie. Je remercie également Mme Marie-Pierre Laforêt pour ses conseils techniques en biologie moléculaire.

Je remercie Mme Anne-Sophie Lia et M. Paco Derouault pour leur collaboration dans ce travail. Merci pour ces échanges fructueux et pour votre sympathie.

Je remercie Mme Elodie Marcellaud, M. Liam Crowther et Mme Lucie Donaty pour leur contribution à ce travail au cours de leur stage effectué au laboratoire.

Je remercie également l'ensemble des membres du service d'hématologie biologique du CHU de Limoges, biologistes et techniciens, pour leur sympathie et ces moments de convivialité et d'échange. Je remercie particulièrement mes collègues biologistes qui m'ont permis de me libérer du temps pour élaborer ce travail.

Je remercie l'ensemble de l'équipe de l'UMR CNRS7276/INSERM1262 pour leur aide technique et les échanges scientifiques m'ayant permis d'avancer dans ce travail.

Je remercie mes amis d'école, de fac ou autres qui sont toujours là pour moi et qui comptent énormément pour moi, même si le temps peut parfois nous manquer.

Je remercie ma famille et ma belle-famille pour leur soutien, notamment ces derniers mois. Je remercie mon frère pour ses conseils. Je remercie tout particulièrement mes parents qui m'ont toujours soutenue et m'ont encore montré ces derniers mois leur implication à mes côtés. Sans vous, je n'aurais jamais pu terminer ce travail dans les temps. Merci infiniment pour tout ce que vous faites pour moi.

Je remercie Sébastien pour son soutien ces dernières années. Merci de me changer les idées et de m'encourager au quotidien. Enfin, merci à notre petite Flora. Tu es une petite fille formidable, mon petit rayon de soleil.

Sommaire

Sommaire	3
Introduction.....	24
Chapitre 1 : Oncogénèse des lymphomes B.....	26
A- Physiologie de la maturation B normale.....	26
1- Introduction	26
2- Structure des immunoglobulines.....	27
3- Structure des gènes des immunoglobulines.....	29
4- Le récepteur de la cellule B ou BCR	42
5- Lymphopoïèse B normale	45
B- Origine cellulaire des lymphomes B.....	55
C- Voies d'oncogénèse dans les lymphomes B.....	56
1- Dérégulation oncogénique via translocation	57
2- Mutations géniques	69
3- Autres mécanismes d'activation de proto-oncogènes.....	70
4- Inactivation de gènes suppresseurs de tumeur	70
5- Transformation par les virus.....	71
6- Activation du BCR	72
7- Inactivation du BCR.....	72
8- Activation par l'antigène.....	73
9- Rôle du microenvironnement.....	74
10- Rôle des ARN longs non codants	75



11- Modèles de lymphomagénèse.....	77
Chapitre 2 : la leucémie lymphoïde chronique	79
A- Définition.....	79
B- Epidémiologie.....	79
C- Etiologies	80
1- Facteurs de risque génétiques.....	80
2- Facteurs de risque environnementaux.....	81
3- Facteurs liés au mode de vie	82
4- Antécédents médicaux	82
D- Physiopathologie.....	82
1- Origine cellulaire.....	82
2- Sélection par l'antigène	83
3- Altérations génétiques.....	83
4- Microenvironnement.....	89
E- Clinique.....	94
1- Présentation clinique classique	94
2- Les variants	95
3- Les complications.....	96
F- Diagnostic	99
1- Numération-Formule sanguine.....	99
2- Cytométrie en flux	100
3- Autres.....	100
G- Facteurs pronostiques	101



1- Cytogénétique.....	101
2- Génétique de la partie variable des chaînes lourdes des immunoglobulines (IGHV)	108
3- Mutations récurrentes hors <i>IGHV</i>	111
4- Marqueurs sériques.....	119
5- Marqueurs immunophénotypiques.....	122
6- Temps de doublement de la lymphocytose	124
7- Classifications pronostiques	124
H- Traitement	127
1- Critères de mise en place thérapeutique	127
2- Evaluation de la réponse au traitement	129
3- Mesure de la maladie résiduelle.....	130
4- Les différentes options thérapeutiques disponibles	130
5- Choix du traitement en fonction des patients.....	137
Chapitre 3 : Problématiques et objectifs	139
Chapitre 4 : Matériels et méthodes	140
A- Patients et échantillons.....	140
1- Etude de translocations impliquant le gène des immunoglobulines	140
2- LLC et électrophorèse des protéines sériques.....	142
3- Evaluation du logiciel Cov'Cop dans la LLC	142
B- Techniques conventionnelles.....	144
C- Techniques spécialisées	146
1- FISH à façon.....	146
2- LM-PCR et séquençage	148



3- CGH-array.....	150
4- Séquençage haut débit (NGS).....	151
Chapitre 5 : Résultats	158
A- Etude de translocations atypiques dans les lymphomes B.....	158
1- Patient 1.....	158
2- Patient 2.....	177
3- Patient 3.....	183
B- Article 1 : Normal serum protein electrophoresis and mutated IGHV genes detect very slowly evolving chronic lymphocytic leukemia patients, accepté le 29 Mars 2018 dans Cancer medicine	193
1- Introduction	193
2- Article.....	194
3- Discussion	202
C- Article 2 : « Updates to « COV’COP » to easily detect CNVs in inherited and cancer diseases », manuscrit soumis dans Bioinformatics le 11/06/2018 et refusé le 18/06/2018, en attente de resoumission dans un autre journal	203
1- Introduction	203
2- Article.....	204
3- Discussion	208
D- Evaluation du logiciel CovCop dans la LLC	219
1- Introduction	219
2- Résultats.....	219
3- Courte discussion.....	227
Chapitre 7 : Discussion générale et perspectives.....	229



A- Remaniements atypiques impliquant les gènes des immunoglobulines dans la LLC/lymphome lymphocytaire.....	229
1- Mécanisme de survenue des remaniements impliquant les gènes des immunoglobulines	229
2- Mécanisme de dérégulation oncogénique	232
3- Modèles de lymphomagénèse.....	233
4- Perspectives	234
B- Impact pronostique de l'électrophorèse des protéines sériques dans la LLC	235
C- Utilisation du séquençage haut débit pour la détection d'anomalies de nombre	236
Conclusion	238
Références bibliographiques.....	239
Sommaire	289
Résumé.....	238



Table des abréviations

ABC : Activated B-Cell

ADN : Acide DesoxyriboNucléique

AID : Activation-Induced Cytidine Deaminase

AP : Site Apurique/Apyrimidique

APE1 : AP Endonuclease 1

API2 : Apoptosis Inhibitor 2

APRIL : a proliferation-inducing ligand

ARID1A : AT-Rich Interaction Domain 1A

ARN : Acide RiboNucléique

ATM : Ataxia Telangiectasia Mutated

B3GNTL1 : UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase like 1

BAC : Bacterial Artificial Chromosome

BAFF : B-cell Activating Factor

bam : Binary Alignment Map

BCAP : B-Cell Adaptor for phosphoinositide 3-kinase

BCL-2 : B-Cell Lymphoma 2

BCL-6 : B-Cell Lymphoma 6

BCR : B-Cell Receptor

BER : Base Excision Repair

BIRC3 : baculoviral IAP repeat containing 3

BL : Lymphome de Burkitt



BLIMP-1 : B-Lymphocyte-Induced maturation Protein 1

BLK : BLK proto-oncogene, Src family tyrosine kinase

BLNK : B-cell Linker Protein

BOB1 : B-cell specific Octamer Protein 1

BTK : Bruton Tyrosine Kinase

C : Constant

CARD11 : Caspase recruitment domain-containing protein 11

Cas9 : CRISPR-associated protein 9

CASC19 : Cancer Susceptibility 19

CCAT1 : Colon Cancer Associated Transcript 1

CDR : Complementary Determining Region

CGH : Comparative Genomic Hybridization

CHD2 : Chromodomain Helicase DNA binding protein 2

ChIP-Seq : Chromatin ImmunoPrecipitation Sequencing

CIITA : Class II major histocompatibility complex TransActivator

CIN85 : Cbl interacting protein of 85 kD

CLP : Progéniteur Lymphoïde Commun

CMH : Complexe Majeur d'Histocompatibilité

CMV : CytomégaloVirus

CRISPR : Clustered Regularly Interspaced Short Palindromic Repeats

CSH : Cellule Souche Hématopoïétique

CSK : C-terminal src tyrosine Kinase

CSR : Class Switch Recombination



D : Diversité

DDX3X : DEAD-box helicase 3 X-linked

DLBCL : Lymphome B diffus à grandes cellules (ou Diffuse Large B-Cell Lymphoma)

DLEU1 et 2 : Deleted in Lymphocytic leukemia 1 et 2

E2A : Immunoglobulin enhancer binding factors E12/E47

EBF1 : Early B-cell Factor 1

EBV : Epstein-Barr Virus

ELISA : Enzyme-Linked ImmunoSorbent Assay

EPS : Electrophorèse des Protéines Sériques

ERK : Extracellular signal-Regulated Kinase

EZH2 : Enhancer of Zest Homolog 2

FADD : Fas-Associated protein with Death Domain

FCR : Fludarabine-Cyclophosphamide-Rituximab

FDC : Follicular Dendritic Cells (ou Cellules Folliculaires Dendritiques)

FISH : Fluorescence In Situ Hybridization

FL : Lymphome folliculaire

FLIP : FADD-Like interleukine-1- β -converting enzyme-Inhibitory Protein

Flt-3 : fms-related tyrosine kinase 3

FOXO3 : Forkhead box O3

FR : Framework Region

FYN : FYN proto-oncogene, Src family tyrosine kinase

GAB1/2 : GRB2-associated-binding protein 1 et 2

GAS5 : Growth Arrest Specific 5



GC : Centre germinatif

G-CSF : Granulocyte-Colony Stimulating Factor

Gfi1 : Growth factor independent 1 transcription repressor

GRB2 : Growth factor Receptor Bound protein 2

GWAS : Genome Wide Association Studies ou études d'association génétique pangénomique

H3K4me3 : histone 3 lysine 4 triméthylée

HACE1 : HECT domain and Ankyrin repeat Containing E3 ubiquitin ligase 1

HECT domain : Homologous to the E6 associated protein carboxyl terminus domain

HHV8 : Human Herpes Virus 8

HLA : Human Leukocyte Antigen

HMGB-1 ou 2 : High Mobility Group Box 1 ou 2

HTLV1 : Human T-Lymphotropic Virus type 1

IAP : Inhibitor of Apoptosis Protein

IGH : chaîne lourde des immunoglobulines (immunoglobulin heavy chain complex)

IGHV : locus de la chaîne lourde variable des immunoglobulines

IGK : chaîne légère Kappa des immunoglobulines

IGL : chaîne légère Lambda des immunoglobulines

IKB α : nuclear factor of kappa light polypeptide gene enhancer in B-cell inhibitor alpha

IKK : I κ B Kinase

IL7 : Interleukine 7

IRF4 : Interferon-Regulatory Factor 4

ITAM : Immunoreceptor Tyrosine-based Activation Motif

ITIM : Immunoreceptor Tyrosine-based Inhibitory Motif



iwCLL : international workshop on Chronic Lymphocytic Leukemia

J : Jonction

Kde : Kappa deleting element

KLHL6 : Kelch-like family member 6

LAL-B : Leucémie Aiguë Lymphoblastique B

LDH : Lactate DésHydrogénase

LEF1-AS1 = Lymphoid Enhancer binding Factor 1-Antisens RNA1

lincRNA : long intergenic non coding RNA

LLC : Leucémie lymphoïde chronique

LMP1 : Latent Membrane Protein 1

LMP2A : Latent Membrane Protein 2A

LM-PCR : Ligation Mediated-Polymerase Chain Reaction

LMPP : Lymphoid-primed Multipotent Progenitor

lncRNAs : ARNs longs non codants

LNH : Lymphomes Non Hodgkiniens

Lymphocytes Th : Lymphocytes T helper

LYN : Lyn proto-oncogene, Src family tyrosine kinase

MALAT1 : Metastasis-Associated in Lung Adenocarcinoma Transcript 1

MALT : Mucosa Associated Lymphoid Tissue

MALT1 : Mucosa-associated lymphoid tissue lymphoma translocation gene 1

MAPK : Mitogen-Activated Protein Kinases

MBC : Cellules B mémoires (ou Memory B Cells)

MBL : Lymphocytose B Monoclonale (ou Monoclonal B-cell Lymphocytosis)



Mbr : Major breakpoint region

MCL : Lymphome du manteau

MDSC : Myeloid-Derived Suppressor Cells

METRNL : Meteorin-like, glial cell differentiation regulator

miARN : MicroARN

MITF : Microphthalmia-associated Transcription Factor

MM : Myélome multiple

MMR : Mismatch Repair

MSH2 ou 6 : MutS Homolog 2 ou 6

mTOR : mammalian Target Of Rapamycin

MYB-AS1 : MYB AntiSens RNA1

MYD88 : Myeloid Differentiation primary response 88

NBC : Cellules B naïves (ou Naïve B-Cells)

NEAT1 : Nuclear Enriched Abundant Transcript 1

NFAT : Nuclear Factor of activated T-cells

NFκB : Nuclear Factor Kappa B

NFKBIE : NFκB Inhibitor Epsilon

NGS : Next Generation Sequencing

NHEJ : Non Homologous End Joining

N-nucléotides : Non-templated nucleotides

OBF1 : Octamer-binding transcription Factor (OCT)-Binding Factor 1

OCT2 : Organic Cation Transporter 2

OMS : Organisation Mondiale de la Santé



ORF : Open Reading Frame

OS : survie globale (Overall Survival)

PCAT1 : Prostate Cancer Associated Transcript 1

PCL : Plasma Cell Leukemia

PCR : Polymerase Chain Reaction

PD-1 : Programmed cell Death 1

PDL : Programmed Death Ligand

PI3K : Phosphoinositide 3 kinase

PIM1 : Pim proto-oncogene

PIP2 : phosphatidylinositide biphosphate

PIP3 : phosphatidylinositide triphosphate

PKC β : Protéine Kinase C Béta

PLC γ : Phospholipase C gamma

P-nucléotide : Palindromique-nucléotide

POT1 : Protection Of Telomeres 1

PRC2 : Polycomb Repressive Complex 2

PRDM1 : PR Domain zinc finger protein 1

PTEN : Phosphatase and Tensin homolog

PU.1 : Purine box factor 1

PVT1 : Plasmocytoma Variant translocation gene 1

QMPSF : Quantitative Multiplex PCR of Short Fluorescent Fragments

RAG : Recombination Activating Gene

RAGE : Receptor for Advanced Glycation end-product



ROC : Receiving Operator Characteristic

ROR1 : Receptor tyrosine kinase like Orphan Receptor 1

RSS : Recombination Signal Sequence

RT-MLPA : Reverse Transcription- Multiplex Ligation-dependent Probe Amplification

RT-PCR : Reverse Transcription-Polymerase Chain Reaction

SCF : Stem Cell Factor

SETD2 : SET Domain containing 2

SF3B1 : Splicing Factor 3B subunit 1

SFK : SRC Family Kinases

SHIP1/2 : SH2-containing Inositol 5' Polyphosphatase 1 ou 2

SHM : Somatic Hypermutation

SMAD1-AS1 = Smad family member 1 AntiSens RNA1

SNP : Single Nucleotide Polymorphism

SPE : Electrophorèse des Protéines Sériques

SRC : Sarcoma kinase gene

SYK : Spleen Tyrosine Kinase

T-nucléotides : templated nucleotides

TARC : Thymus and Activation-Regulated Chemokine

TCF3 : Transcription Factor 3

TCL1 : T-Cell Leukemia/Lymphoma

TdT : Terminal deoxynucleotide Transferase

TFS : Survie sans traitement (Time to First Treatment ou Treatment Free Survival)

TI : T-Indépendant



TIR8 : Toll Interleukine-1 Receptor 8

TLR : Toll-Like Receptor

TNF : Tumor Necrosis Factor

TRAF1 et 2 : Tumor necrosis factor Receptor-Associated Factors 1 et 2

UNG : Uracil-DNA-Glycosylase

UPD : UniParental Disomy

V : Variable

VAF : Variant Allelic Frequency

vcf : Variant Call Format

VHC : Virus de l'hépatite C

VIH : Virus de l'Immunodéficience Humaine

XBP1 : X-Box binding Protein 1

XPO1 : Exportine 1 gene

ZAP70 : ζ -Associated Protein



Liste des figures

Figure 1 : Schéma d'une immunoglobuline avec en rouge les parties variables et en bleu les parties constantes	27
Figure 2 : Schéma détaillé d'une immunoglobuline, adapté selon <i>Schroeder et al.</i> ³ représentant les chaînes lourdes et légères, ainsi que les différents domaines des immunoglobulines et les interactions avec ponts disulfure.	28
Figure 3 : Schéma d'un RSS adapté selon <i>Schatz et al.</i> ⁵	30
Figure 4 : Mécanisme de recombinaison VDJ entre segments RSS, adapté selon <i>Schatz et al.</i> ⁵	31
Figure 5 : Mécanisme de recombinaison VDJ : action du complexe RAG au niveau de l'ADN selon <i>Schatz et al.</i> ⁵	32
Figure 6 : Schéma de recombinaison VDJ au niveau du locus IGH avec la première étape de réarrangement entre les segments D _H et J _H , puis la deuxième étape au niveau des régions V _H -D _H J _H sous l'action du complexe RAG (pour Recombination Activating Gene) au niveau des RSS (pour Recombination Signal Sequence).	33
Figure 7 : Schéma de commutation de classe isotypique (ou CSR pour Class Switch Recombination) sous l'action d'AID (pour Activation-Induced Cytidine Deaminase), générant le segment d'immunoglobuline réarrangé et un cercle d'excision correspondant à la région coupée.	35
Figure 8 : Mécanisme de désamination de cytidine via AID dans les processus de SHM et CSR d'après <i>Hwang et al.</i> ²¹	36
Figure 9 : Représentation du chromosome 14 avec localisation du locus IGH, adapté d'après <i>Lefranc et al.</i> ²²	37
Figure 10 : Schéma simplifié du locus IGH en 14q32, adapté d'après <i>Schroeder et al.</i> , <i>Van Dongen et al.</i> , et <i>Pinaud et al.</i> ^{3,23,24}	37
Figure 11 : Segment IGH réarrangé et diversité jonctionnelle, adapté d'après <i>Schroeder et al.</i> ⁵	38
Figure 12 : Représentation du chromosome 2 avec localisation du locus IGK, adapté d'après <i>Lefranc et al.</i> ²⁶	39
Figure 13 : Schéma simplifié du locus IGK en 2p11.2, adapté selon <i>Schroeder et al.</i> et <i>Van Dongen et al.</i> ^{3,23}	40



Figure 14 : Représentation du chromosome 22 avec localisation du locus IGL, adapté d'après Lefranc <i>et al.</i> ²⁸	41
Figure 15 : Schéma simplifié du locus IGL en 22q11.2, adapté selon Schroeder <i>et al.</i> et Van Dongen <i>et al.</i> ^{3,23}	41
Figure 16 : Représentation du récepteur de la cellule B (BCR), adapté selon Immunologie, Janeway, De Boeck, 2009 ⁸	42
Figure 17 : Représentation schématique de la voie de signalisation du BCR adapté d'après Seda <i>et al.</i> et Young et Staudt ^{30,31}	43
Figure 18 : Maturation lymphocytaire B précoce adapté selon Meffre <i>et al.</i> et Le système immunitaire, De Boeck ^{33,10}	46
Figure 19 : Expression des marqueurs immunophénotypiques au cours des différents stades de développement de la cellule B, adapté d'après Armitage <i>et al.</i> , Ye et Graf, et LeBien et Fates ^{2,34,35}	46
Figure 20 : Réarrangement VJ sur la chaîne légère Kappa adapté selon Le Système Immunitaire, De Boeck ¹⁰ . Deux recombinaisons VJ successives sur le même allèle <i>IGK</i>	49
Figure 21 : Ordre de réarrangements des gènes des immunoglobulines au cours du développement lymphocytaire B précoce, adapté d'après Le système immunitaire, De Boeck ¹⁰	50
Figure 22 : Différenciation lymphocytaire B au niveau des centres germinatifs ganglionnaires, adapté d'après Küppers ⁴¹	52
Figure 23 : Origine cellulaire des lymphomes B adapté d'après Küppers ⁴¹	56
Figure 24 : Mécanismes de dérégulation génique via translocation dans les lymphomes B, adapté d'après Armitage <i>et al.</i> ²	58
Figure 25 : Modèle de translocation via transposition médiée par RAG, faisant intervenir des extrémités franches, adapté d'après Roth et Craig ⁵³	62
Figure 26 : Modèle de translocation par transposition via insertion médiée par RAG, adapté d'après Küppers et Dalla-Favera ⁴⁸	63
Figure 27 : Mécanisme de la translocation <i>IGH/BCL2</i> au cours du processus de recombinaison VDJ (exemple de cassure dans les régions JH pour <i>IGH</i> et au niveau de la Mbr pour <i>BCL2</i>)... 64	
Figure 28 : Mécanisme de la translocation <i>IGH-MYC</i> faisant intervenir la recombinaison de classe, adapté d'après Blum <i>et al.</i> ⁵⁷	65
	18



Figure 29 : IncRNAs exprimés au cours de la différenciation B et dans les hémopathies lymphoïdes B, adapté d'après Nobili <i>et al.</i> ¹¹⁰	76
Figure 30 : Les 3 modèles de lymphomagénèse, adapté d'après Weigert et Weinstock ⁷⁶	78
Figure 31 : Image de frottis sanguin coloré au MGG dans un cas de LLC.	99
Figure 32 : Schéma de la sonde IGH Kreatech adapté selon les indications du fournisseur.	144
Figure 33 : Schéma de la sonde IGH LSI Vysis, adapté selon les indications du fournisseur.	145
Figure 34 : Schéma de la sonde IGL Dako, adapté selon les indications du fournisseur.	145
Figure 35 : Schéma de la sonde MYC Dako, adapté selon les indications du fournisseur.	146
Figure 36 : Schéma de la sonde MYC Metasystem, adapté selon les indications du fournisseur.	146
Figure 37 : Principe de la technique de Nick translation	147
Figure 38 : Schéma de la localisation des BAC et fosmidés utilisés au niveau d'IGH.	148
Figure 39 : Principe de la LM-PCR	149
Figure 40 : Caryotype du patient n°1.	158
Figure 41 : Image de FISH IGH (sonde Kreatech), concernant le patient n°1.	159
Figure 42 : A- Représentation schématique à partir du logiciel IGV du locus IGH (en vert) et des fosmidés ciblant les régions constantes d'IGH (en bleu). B- Image de FISH métaphasique avec ces fosmidés marqués en Aqua.....	162
Figure 43 : Image de FISH métaphasique avec une sonde télomérique du bras long du chromosome 17.	163
Figure 44 : Représentation schématique à partir du logiciel IGV du locus <i>IGH</i> (en vert) et du fragment d'excision selon le remaniement VDJ mis en évidence (en orange), ainsi que les 3 BACs utilisés pour vérifier la délocalisation du fragment d'excision (en bleu).....	164
Figure 45 : Image de FISH métaphasique avec les BACs RP11379F22 et RP11-47P23 marqués respectivement en vert et rouge.	165
Figure 46 : Image de M-FISH réalisée chez le patient n°1.....	166
Figure 47 : Idéogrammes des chromosomes 13, 14 et 17 (multiFISH).	166



Figure 48 : Profil en CGH-array du patient n°1.	167
Figure 49 : Visualisation à partir du logiciel IGV des séquences obtenues par séquençage du génome entier au niveau du chromosome 14, position 105864259.....	168
Figure 50 : Visualisation à partir du logiciel IGV des séquences obtenues par séquençage du génome entier au niveau du chromosome 14, position 105894509.....	169
Figure 51 : Visualisation à partir du logiciel IGV des séquences obtenues par séquençage du génome entier au niveau du chromosome 17, centrée sur les positions 82971126 à 82971132.	170
Figure 52 : Résultat du séquençage par technique Sanger du premier point de cassure impliquant IGHJ4 et la position 82971132 sur le chromosome 17.....	171
Figure 53 : Résultat du séquençage par Sanger du deuxième point de cassure impliquant IGHD4-17 et la position 82971126 sur le chromosome 17.....	171
Figure 54 : Schémas représentant le réarrangement identifié chez le patient 1.	172
Figure 55 : Schéma récapitulatif des différents résultats obtenus concernant le remaniement entre les chromosomes 14 et 17 chez le patient n°1.....	173
Figure 56 : Schéma de l’hypothèse en deux étapes pour le remaniement observé chez le patient n°1.....	175
Figure 57 : Caryotype de la patiente n°2.....	177
Figure 58 : Image de FISH avec la sonde IGL Dako concernant la patiente numéro 2.	178
Figure 59 : Visualisation à partir du logiciel IGV des séquences obtenues par séquençage du génome entier au niveau du chromosome 17, position 83175424 concernant la deuxième patiente.	179
Figure 60 : Visualisation à partir du logiciel IGV des séquences obtenues par séquençage du génome entier au niveau du chromosome 22, position 22850521 concernant la deuxième patiente.	179
Figure 61 : Séquence sens obtenue par séquençage Sanger confirmant la jonction du der(22)t(22 ;17).....	180
Figure 62 : Schéma représentant la translocation entre les bras longs des chromosomes 17 et 22 mise en évidence chez la patiente n°2.....	181
Figure 63 : Caryotype du patient n°3.	183



Figure 64 : Image de FISH métaphasique avec la sonde IGH Kreatech.....	184
Figure 65 : Image de FISH métaphasique avec la sonde IGL Dako.....	184
Figure 66 : Image de FISH métaphasique avec les fosmides G248P85777C9, G248P8983A5 et G248P82502G1 marqués tous les trois en Aqua, ainsi que la sonde IGH Vysis.....	186
Figure 67 : Image de FISH métaphasique avec les fosmides G248P85777C9, G248P8983A5 marqués respectivement en rouge et vert.	187
Figure 68 : Visualisation à partir du logiciel IGV des séquences obtenues par séquençage du génome entier au niveau du chromosome 14, position 105647500 concernant la patiente n°3.	188
Figure 69 : Visualisation à partir du logiciel IGV des séquences obtenues par séquençage du génome entier au niveau du chromosome 8, position 127196845.....	188
Figure 70 : Séquence sens obtenue par séquençage Sanger pour le der(8)t(8;14).....	189
Figure 71 : Image de FISH métaphasique avec la sonde MYC Metasystem pour la patiente n°3.	190
Figure 72 : Schéma représentant la translocation entre les bras long des chromosomes 8 et 14, ainsi que le dérivé 8 obtenu pour la patiente n°3.....	191
Figure 73 : Capture d'écran du logiciel Cov'Cop2 montrant un exemple d'analyse d'une puce comportant 5 patients.....	209
Figure 74 : Visualisation graphique du logiciel Cov'Cop2. Exemple de détection d'amplicons mal amplifiés.	210
Figure 75 : Exemple de visualisation graphique fournie par le logiciel Cov'Cop2 avec intégration d'une courbe Loess en vert.	210
Figure 76 : Visualisation des graphiques obtenus avec le logiciel Cov'Cop2 en appliquant l'option « CUSUM chart ».....	211
Figure 77 : Représentation graphique avec le logiciel Cov'Cop2 d'une duplication du gène AARS avec l'option de visualisation de la valeur moyenne des ratios normalisés par gène.	212
Figure 78 : Visualisation des graphiques obtenus avec Cov'Cop avec en abscisse les régions chromosomiques et en ordonnée les ratios normalisés obtenus.	213
Figure 79 : Visualisation graphique à partir du logiciel Cov'Cop2 de la duplication du gène AARS.	214



Figure 80 : Visualisation graphique à partir du logiciel Cov'Cop2 du chromosome 13 chez 3 patients.....	215
Figure 81 : Visualisation graphique avec le logiciel Cov'Cop2 de la simulation de la délétion d'un gène (visualisé par un trait rouge-orangé sous les graphiques), avec différentes proportions de cellules contenant la délétion.....	216
Figure 82 : Visualisation graphique avec le logiciel Cov'Cop2 de la délétion du gène 38 (soit <i>ATM</i>) chez différents patients présentant un infiltration clonale variant entre 0 et 100 % (estimation par caryotype et/ou FISH).....	217
Figure 83 : Visualisation d'une trisomie 12 par le logiciel Cov'Cop2 et l'analyse des VAF. ...	220
Figure 84 : Tables de contingence entre l'approche par NGS (Cov'Cop2 et/ou VAF) et techniques conventionnelles.	221
Figure 85 : Graphiques représentant par anomalie cytogénétique les faux négatifs et vrais positifs par technique NGS par rapport à la technique conventionnelle (cytogénétique +/- QMPSF) selon le pourcentage de cellules porteuses de l'anomalie dans l'échantillon.	222
Figure 86 : Visualisation graphique des VAF des SNP détectés par NGS : cas d'un faux positif pour la del(11q).....	223
Figure 87 : Visualisation des graphiques des résultats avec le logiciel CovCop2 et les VAF : exemple d'un cas de disomie uniparentale touchant <i>TP53</i>	224
Figure 88 : Classement des patients selon le score de Rossi à partir de la cytogénétique (en bleu) ou du NGS seul (en rouge) pour la caractérisation de la del(13q), la trisomie 12, la del(11q) et la del(17p).....	225
Figure 89 : Courbes de survie (TFS) selon le score de Rossi avec classement pour les anomalies cytogénétiques tenant compte des techniques conventionnelles ou du NGS seul.	226
Figure 90 : Visualisation avec le logiciel Cov'Cop2 d'un cas ininterprétable avec grande dispersion des ratios normalisés.....	227



Liste des tableaux

Tableau 1 : Calcul du score de Matutes-Moreau	100
Tableau 2 : Caractéristiques principales des principaux subset dans la LLC, adapté d'après Stamatopoulos <i>et al.</i> ⁴¹⁹	110
Tableau 3 : Classification de Rai selon Rai <i>et al.</i> ⁵⁴⁶	125
Tableau 4 : Classification de Binet selon Binet <i>et al.</i> ⁵⁴⁸	125
Tableau 5 : Critères et calcul du CLL-IPI selon The International CLL Working Group ⁴⁹⁷	126
Tableau 6 : Catégories de risque selon le score d'après le CLL-IPI ⁴⁹⁷	127
Tableau 7 : Caractéristiques des patients de l'étude sur l'évaluation du logiciel Cov'Cop dans la LLC.....	143
Tableau 8 : Enzymes de restriction utilisées pour la technique de LM-PCR et séquence reconnue correspondante (sens 5' → 3'). Le tiret dans la séquence indique le site de restriction.	149
Tableau 9 : Liste des cibles incluses dans le panel NGS lymphoïde.....	154



Introduction

La leucémie lymphoïde chronique (LLC) est une hémopathie maligne B. Il s'agit du lymphome avec phase leucémique le plus fréquent chez l'adulte dans les pays occidentaux. Elle est caractérisée par une grande hétérogénéité dans son évolution naturelle, certains patients nécessitant une chimiothérapie rapidement après le diagnostic et d'autres n'ayant jamais besoin de traitement spécifique et ayant une survie similaire à celle de la population générale. Il a donc très vite apparu nécessaire de déterminer des marqueurs pronostiques, afin d'essayer de détecter les patients à risque d'évoluer et nécessitant un suivi rapproché. Les classifications pronostiques permettent également de stratifier les patients en groupes les plus homogènes possibles dans le cadre de protocoles d'études cliniques.

Cette variabilité dans l'histoire naturelle de la LLC est sous-tendue par une grande hétérogénéité génétique. Il a été retrouvé dès les années 1990 des anomalies chromosomiques récurrentes qui sont généralement recherchées au diagnostic ou avant toute mise en place thérapeutique. Il s'agit des délétions du bras court du chromosome 17, du bras long du chromosome 11 ou du chromosome 13, et de la trisomie 12. Il existe également des remaniements tels que les translocations, impliquant dans la majorité des cas un élément régulateur du gène des immunoglobulines et aboutissant à la dérégulation d'un oncogène. Certaines de ces translocations sont bien connues, avec un partenaire identifié, tel que les oncogènes *BCL2*, *BCL6* ou *MYC*. D'autres impliquent un partenaire inconnu. Dans ce travail, nous nous sommes intéressés à ce type de translocations pour lesquelles le gène partenaire est inconnu, et constitue un probable nouvel oncogène. Nous avons également essayé de comprendre le mécanisme de survenue de ces translocations. Enfin, nous avons tenté de mettre en évidence le mode de dérégulation intervenant dans chaque cas. La découverte de nouveaux oncogènes ou éléments dérégulés est intéressante sur le plan de la physiopathologie, et peut également permettre de découvrir de nouvelles cibles thérapeutiques.

Par ailleurs, avec l'avènement des techniques moléculaires à haut débit, un grand nombre de mutations a été mis en évidence dans la LLC. Leur impact pronostique et leur place dans la décision thérapeutique est actuellement étudié. Certaines mutations ont un impact pronostique bien établi (*SF3B1*, *NOTCH1* et *BIRC3* par exemple), tandis que d'autres sont encore en cours d'évaluation (*MYD88*). Cependant, ces mutations ne sont actuellement pas prises en compte pour la décision thérapeutique. En pratique, seul le statut muté/déléte de *TP53* modifie le type de chimiothérapie à instaurer mais ne constitue pas à lui seul un critère de traitement. A côté de ces mutations génétiques, il existe des marqueurs immunophénotypiques ou sériques. Parmi ceux-ci, l'électrophorèse des protéines sériques permet la détection de pic monoclonal ou une diminution globale des gammaglobulines (ou hypogammaglobulinémie). Cet examen est généralement prescrit chez les patients atteints de LLC au diagnostic et au cours du suivi, essentiellement pour la détection d'une hypogammaglobulinémie qui, si elle est associée à des infections récidivantes, peut nécessiter une prise en charge thérapeutique spécifique. Il existe maintenant un grand nombre de



marqueurs pronostiques, la plupart étant des marqueurs péjoratifs d'évolution. Dans ce travail, nous nous sommes intéressés à l'impact pronostique de deux marqueurs de bon pronostic : l'électrophorèse des protéines sérique normale et le statut IGHV muté dans l'ère du séquençage haut débit (ou NGS pour Next Generation Sequencing).

De plus, nous avons collaboré avec l'EA6308 dans l'amélioration de « Cov'Cop », un logiciel de détection des variations des nombres de copies. Celui-ci est adapté aux résultats obtenus par NGS basé sur la réalisation de librairies par amplicons. Il a initialement été développé pour une interprétation dans les pathologies héréditaires. Nous avons donc contribué à son amélioration et son application dans la détection d'anomalies du nombre de copies dans un cadre somatique. Enfin, nous avons évalué ses performances dans la détection des principales anomalies cytogénétiques récurrentes de la LLC.

Dans ce manuscrit de thèse, nous décrivons les mécanismes d'oncogénèse dans les lymphomes B. Un chapitre sera consacré à la LLC, avec une revue des données actuelles en particulier concernant la physiopathologie et les marqueurs pronostiques. Nous consacrerons un chapitre aux méthodes utilisées dans ce travail. Nous exposerons ensuite les résultats et montrerons qu'il existe des translocations originales impliquant les gènes des immunoglobulines dans la LLC et pour lesquelles nous pourrions discuter le mécanisme de survenue et le potentiel rôle oncogénique. Les articles de la thèse portant sur l'impact pronostique de l'électrophorèse des protéines sériques dans la LLC ainsi que sur l'amélioration du logiciel Cov'Cop. Nous présenterons également les résultats de l'évaluation des performances d'une approche par séquençage haut débit pour la détection des variations des nombres de copies dans la LLC. Enfin une discussion générale des résultats à la lumière de la littérature et une conclusion seront proposées.



Chapitre 1 : Oncogénèse des lymphomes B

Les lymphomes B correspondent à l'expansion d'un clone lymphocytaire B bloqué à un certain stade de maturation, à l'exception de la maladie de Waldenström pour laquelle il existe un continuum de cellules clonales comprenant des lymphocytes B, des lymphoplasmocytes et des plasmocytes. La classification OMS (pour Organisation Mondiale de la Santé), qui fait référence actuellement, tend à proposer la contrepartie normale pour une hémopathie donnée, c'est-à-dire le stade de maturation auquel la cellule est supposée correspondre¹. Le clone malin dériverait alors d'un lymphocyte B bloqué à un stade de maturation donné. Dans ce chapitre, nous allons donc revoir la physiologie normale du lymphocyte B avec son processus de maturation et les modifications opérées aux différents stades, puis nous expliquerons les principales voies d'oncogénèse identifiées dans les lymphomes B, en insistant plus particulièrement sur le mécanisme de translocation impliquant les gènes des immunoglobulines.

A- Physiologie de la maturation B normale

1- Introduction

Les lymphocytes B sont impliqués dans la réponse immunitaire adaptative². Leur rôle principal est de reconnaître les antigènes étrangers comme les agents infectieux et de produire, une fois activés, une grande quantité d'anticorps spécifiques pour permettre l'élimination de ces antigènes. Les anticorps correspondent aux immunoglobulines. Celles-ci sont constituées de chaînes lourdes et légères, dont les parties variables sont codées par différents segments appelés V (pour variable), D (pour diversité) et J (pour jonction), le segment D n'étant présent que pour les chaînes lourdes (Figure 1). Ces segments doivent subir des réarrangements génomiques pour permettre la formation de régions variables fonctionnelles². Cette structure particulière permet la production d'une importante diversité d'anticorps différents, phénomène amplifié par le processus d'hypermutation somatique au cours du développement du lymphocyte B.



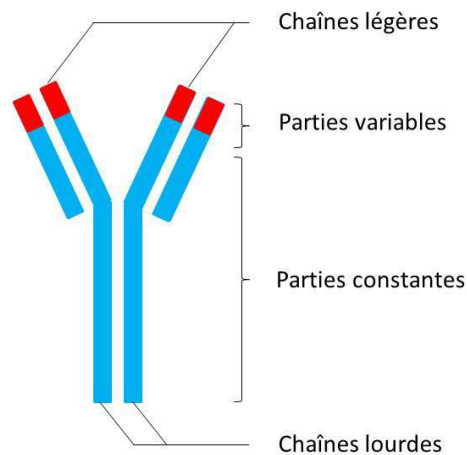


Figure 1 : Schéma d'une immunoglobuline avec en rouge les parties variables et en bleu les parties constantes

Les immunoglobulines peuvent être produites par le lymphocyte B et exprimées à la surface de celui-ci ou être libérées pour aller neutraliser des antigènes à distance³.

2- Structure des immunoglobulines

Les immunoglobulines sont constituées de deux chaînes lourdes identiques et deux chaînes légères identiques (Kappa ou Lambda) (Figure 1)^{3,4}.

Chaque chaîne comprend un domaine amino-terminal (-NH₂) côté variable et un domaine carboxylique (-COOH) côté constant. La structure de chacun de ces domaines est constituée de 2 feuillets β repliés en épingle grâce à un pont disulfure entre 2 résidus cystéine³. Tandis que les chaînes légères sont constituées d'un seul domaine constant, les chaînes lourdes peuvent quant à elles en contenir trois ou quatre. Lorsque les chaînes lourdes contiennent trois domaines constants (CH), les premier (CH1) et deuxième (CH2) sont séparés par un espaceur appelé charnière (« hinge ») (Figure 2).



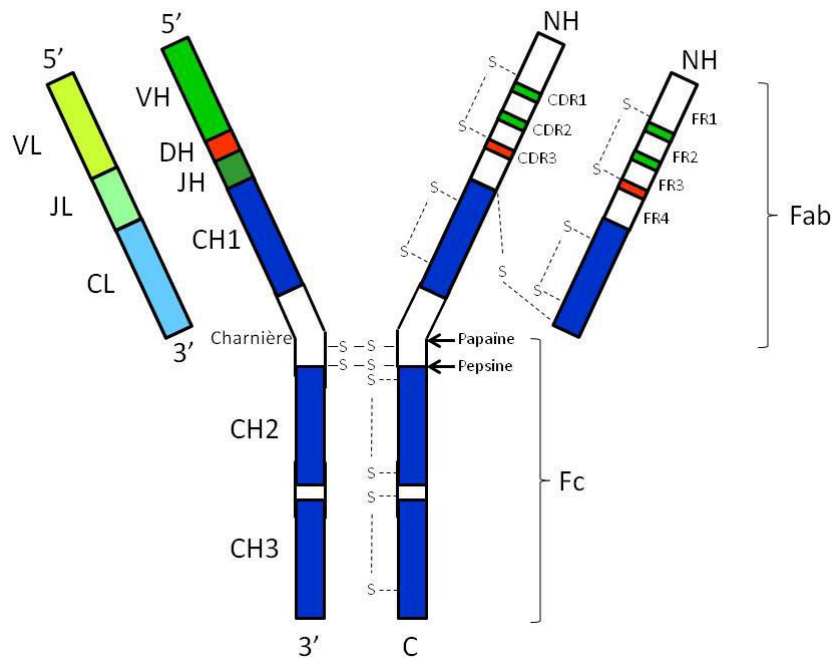


Figure 2 : Schéma détaillé d'une immunoglobuline, adapté selon Schroeder *et al.*³ représentant les chaînes lourdes et légères, ainsi que les différents domaines des immunoglobulines et les interactions avec ponts disulfure.

Les flèches noires correspondent au site de digestion par les enzymes Papaine ou Pepsine. (C = région carboxy-terminale ; NH = région amino-terminale).

Chaque chaîne légère est reliée à une chaîne lourde par un pont disulfure. Il existe également des liaisons non covalentes entre les régions variables lourdes et légères⁴. Les chaînes lourdes sont reliées entre elles par des ponts disulfures situés à peu près à leur moitié, et par des liaisons non covalentes au niveau des parties constantes C-terminales.

Chez l'Homme, environ 60 % des lymphocytes B expriment une chaîne légère de type Kappa et dans 40 % des cas de type Lambda².

On distingue le fragment liant l'antigène ou Fab (pour Antigen binding Fragment) et le fragment constant Fc (pour fragment cristallisable, en référence aux expériences réalisées après digestion enzymatique). Le fragment Fab permet ainsi la reconnaissance de l'antigène et est constitué d'une chaîne légère entière avec le domaine variable et le domaine constant, ainsi que d'une partie de la chaîne lourde comprenant le domaine variable et le domaine CH1. Les autres domaines CH de la chaîne lourde correspondent au fragment Fc. Deux enzymes permettent de couper entre les fragments Fab et Fc : la papaïne permet la libération de chaque Fab d'une part et du Fc entier d'autre part, tandis que la pepsine entraîne la libération du dimère (Fab)₂ (c'est-à-dire des 2 Fab liés par un pont disulfure) et du fragment Fc entier.



On désigne par paratope la portion de l'immunoglobuline qui interagit avec l'antigène. A l'inverse, l'épitope, est la partie de l'antigène qui est liée à l'immunoglobuline³. Les idiotypes sont des déterminants individuels contenus dans les parties variables de l'immunoglobuline. Les allotypes correspondent à des déterminants communs à différents membres d'une même espèce (par exemple chez l'Homme) et pouvant différer d'autres sujets, à cause de polymorphismes génétiques de différents allèles³. Les isotypes sont des déterminants communs et correspondent à une portion de la région constante des immunoglobulines ; cette région permet de séparer les immunoglobulines en plusieurs classes.

3- Structure des gènes des immunoglobulines

Les chaînes lourdes et légères des immunoglobulines sont codées par plusieurs gènes séparés. Le locus IGH (pour Immunoglobulin Heavy chain) est situé sur le bras long du chromosome 14 et code la chaîne lourde des immunoglobulines. Le locus *IGK* (pour Immunoglobuline Kappa) est situé sur le bras court du chromosome 2 et code la chaîne légère Kappa des immunoglobulines. Le locus *IGL* (pour Immunoglobuline Lambda) est situé sur le bras long du chromosome 22 et code la chaîne légère Lambda des immunoglobulines. Chaque locus code les parties variables et constantes par des éléments indépendants. La séquence primaire de la région variable des immunoglobulines est séparée en 3 domaines hypervariables appelés les CDR (pour Complementary Determining Region) situés entre quatre régions de séquence stable appelés FR (pour Framework Region) (Figure 2)³.

a- Réarrangement physiologique du gène des immunoglobulines

i- Recombinaison V(D)J

Chaque segment V est classiquement constitué d'un promoteur, d'un exon principal, d'un intron puis d'un autre exon qui code les trois premières FR ainsi que les CDR1, CDR2 entiers, la région amino-terminale du CDR3, et un signal de recombinaison appelé RSS (pour Recombination Signal Sequence)³.

Chaque segment J commence par son propre signal de recombinaison puis code la région carboxy-terminale du CDR3 et la FR4³.

La formation du domaine V est liée aux RSS qui sont à proximité des segments de gènes à réarranger.



Chaque RSS contient une séquence hautement conservée de 7 nucléotides ou heptamère (CACAGTG) séparé par un espaceur de 12 ou 23 nucléotides d'un nonamère (constitué de 9 nucléotides) un peu moins bien conservé (ACAAAAACC) (Figure 3). Ces espaceurs permettent de placer du même côté les heptamères et nonamères, après un ou deux tours d'hélice d'ADN, permettant ainsi une localisation spatiale favorable au réarrangement³. Chez l'Homme, au niveau du locus *IGH* par exemple, les gènes *VH* et *JH* ont des RSS avec espaceur de 23 nucléotides, tandis que les gènes *DH* ont des RSS avec espaceur de 12 nucléotides². Cela permet d'éviter le réarrangement entre 2 gènes *VH* ou directement entre *VH* et *JH*.

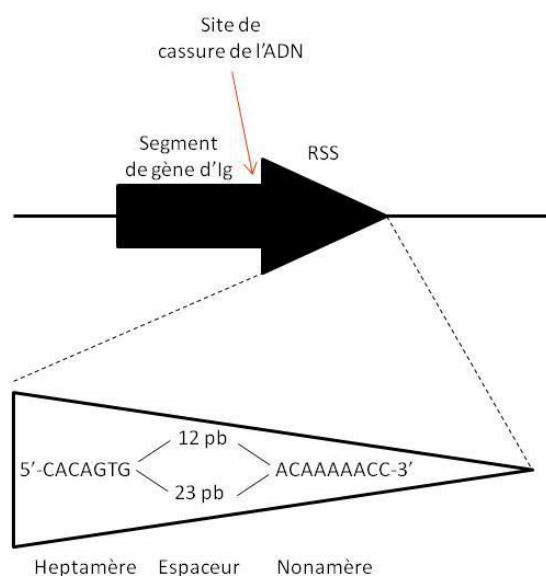


Figure 3 : Schéma d'un RSS adapté selon Schatz *et al.*⁵.

Le RSS est situé juste après le segment d'immunoglobuline et contient 2 éléments hautement conservés : l'heptamère et le nonamère séparés par un espaceur de 12 ou 23 pb.

L'initiation de la recombinaison V(D)J nécessite l'activation des gènes *RAG-1* et *RAG-2* (pour Recombination Activating Gene). Ces 2 gènes sont quasiment exclusivement exprimés dans les lymphocytes B en développement^{6,7}. Les enzymes *RAG-1* et *RAG-2* vont entraîner des cassures double-brin de l'ADN (« Double Strand Break ») à la fin du segment de gène à réarranger. Ces cassures double-brin vont ensuite être réparées par le processus de jonction d'extrémités non homologues (ou NHEJ pour Non Homologous End Joining). Le processus NHEJ permet la recombinaison précise des RSS et moins précise à la fin des régions codantes, pouvant ainsi insérer des nucléotides à la fin des régions codantes³.

Classiquement, le processus V(D)J commence par la reconnaissance par *RAG-1* d'un espaceur de 12 nucléotides, puis *RAG-2* va s'associer à *RAG-1* et l'heptamère pour former un complexe. Ensuite, un second complexe *RAG-1/RAG-2* va reconnaître un espaceur de 23 nucléotides et



s'associer au second RSS. Les 2 complexes RAG-1/RAG-2 vont pouvoir interagir. Ce processus est facilité par les protéines HMGB-1 et HMGB-2 (pour High Mobility Group Box 1 ou 2), protéines liées à la chromatine et qui permettent le repliement de l'ADN (Figure 4).

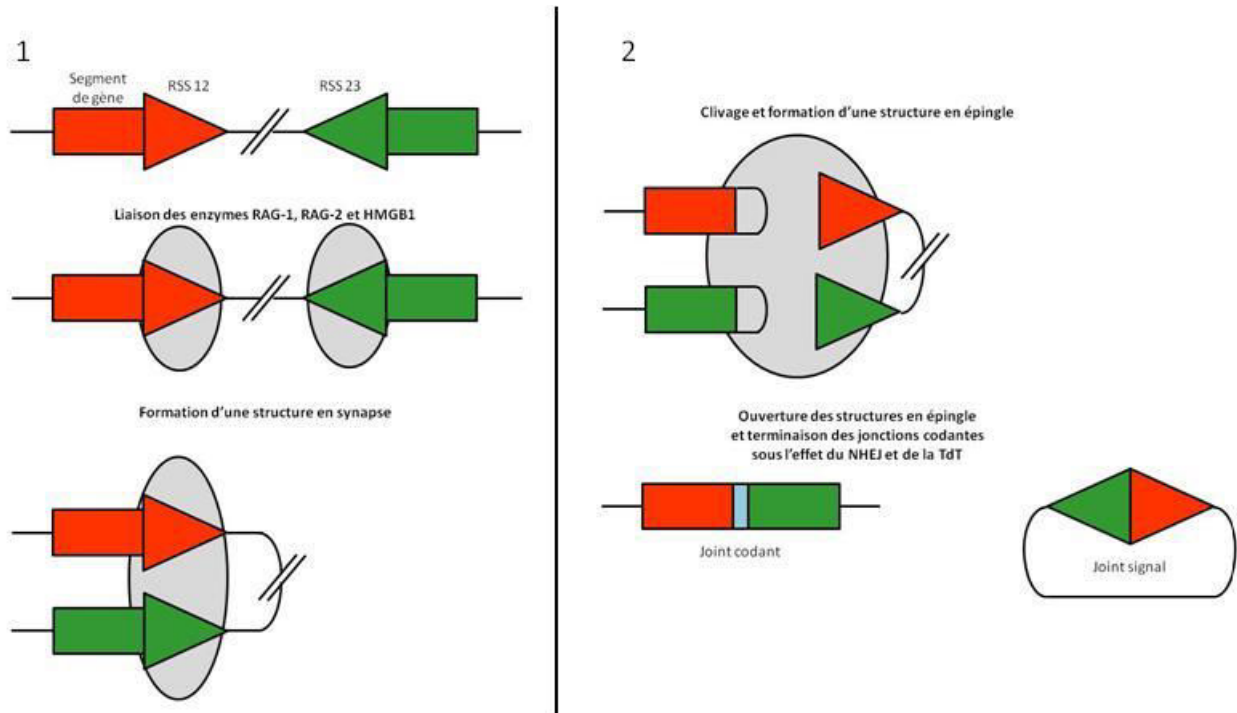


Figure 4 : Mécanisme de recombinaison VDJ entre segments RSS, adapté selon Schatz *et al.*⁵

Le double complexe RAG-1/RAG-2 formé va ensuite couper l'ADN au niveau de l'heptamère. L'extrémité 3'OH va alors se lier à l'extrémité 5'-Phosphate et créer une structure en épingle (Figure 5). La cassure à la fin des RSS permet la formation de signaux précis. Cependant, la formation de boucle en épingle après les régions codantes va nécessiter une nouvelle cassure de l'ADN, généralement au niveau des 4 à 5 derniers nucléotides de la structure en épingle. Ce phénomène peut générer quelques modifications supplémentaires. Des P-nucléotides (pour Palindromique-nucléotides) peuvent être insérés. Il s'agit de séquences palindromiques insérées aux extrémités des segments géniques⁸. Cela correspond à un ou deux nucléotides contigus au joint codant et de séquence complémentaire inverse de ce segment⁹. De plus, la TdT (pour Terminal deoxynucleotide Transferase) peut entraîner l'addition au hasard de quelques nucléotides appelés N-nucléotides (pour Non-templated nucleotides)³.



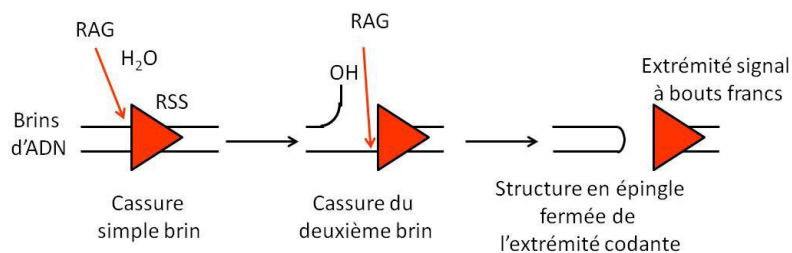


Figure 5 : Mécanisme de recombinaison VDJ : action du complexe RAG au niveau de l'ADN selon Schatz *et al.*⁵.

Les cassures double-brin ainsi générées vont alors être réparées par le processus NHEJ qui implique notamment les protéines Ku70, Ku80, XRCC4 et des protéines kinases dépendantes de l'ADN.

A cause des modifications introduites à la fin des segments de gènes réarrangés, le cadre de lecture correct, défini par le codon de départ dans le peptide leader du gène V, peut-être retenu ou perdu au cours du réarrangement V(D)J. Dans les réarrangements dits « in-frame », le segment J et le segment C suivant sont lus selon un cadre ouvert de lecture correct et peuvent donc être traduits en protéines. On parle de réarrangement productif¹⁰. Cependant, si le cadre de lecture correct est perdu, le segment C n'est pas lu et le réarrangement est dit « out of frame ». Cela génère fréquemment des codons stop prématurés dans la séquence du segment C². Dans ce cas, le réarrangement est dit non productif¹⁰. A chaque réarrangement, la probabilité d'avoir un réarrangement productif est de 1 sur 3¹⁰.

Ce processus ordonné permet le réarrangement des gènes des immunoglobulines avec production de chaînes légères et lourdes, déterminant le devenir de la cellule. Suivant le locus d'immunoglobuline, plusieurs réarrangements successifs peuvent avoir lieu (Figure 6 et Figure 20). Les spécificités de chaque locus sont décrites plus loin.



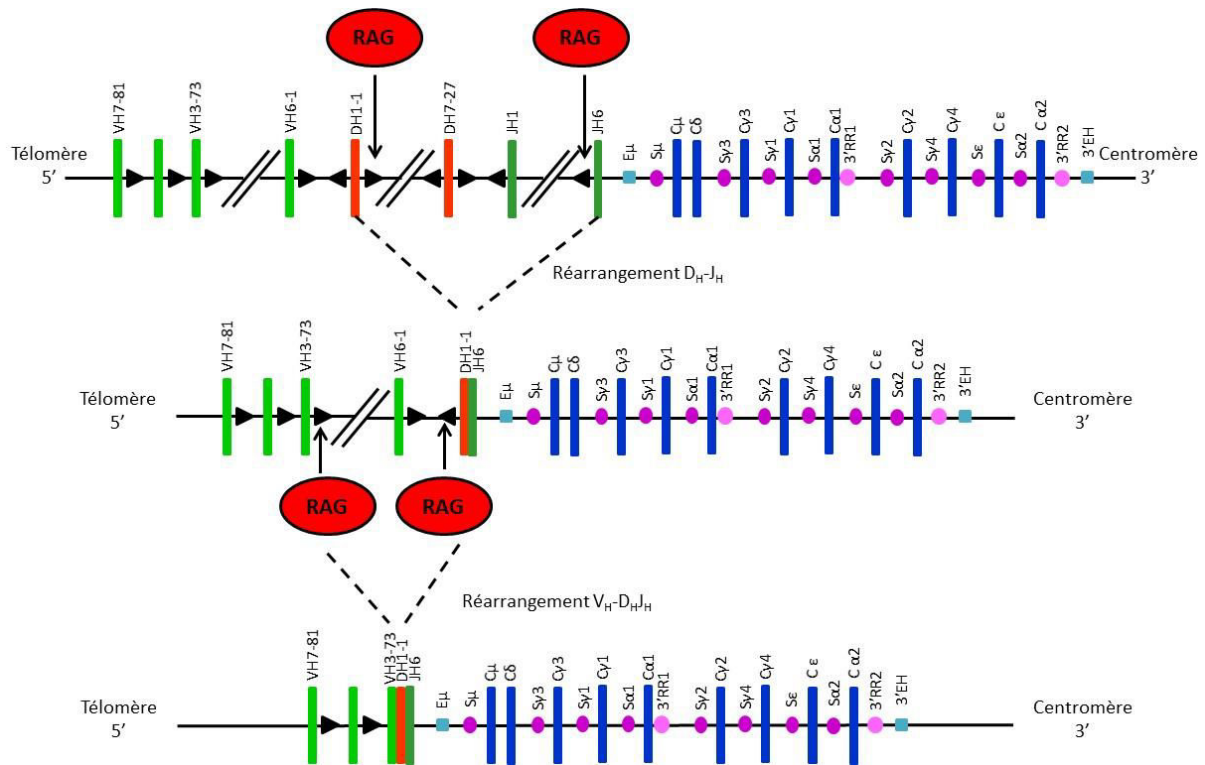


Figure 6 : Schéma de recombinaison VDJ au niveau du locus IGH avec la première étape de réarrangement entre les segments D_H et J_H, puis la deuxième étape au niveau des régions V_H-D_HJ_H sous l'action du complexe RAG (pour Recombination Activating Gene) au niveau des RSS (pour Recombination Signal Sequence).

ii- Hypermutation somatique (ou SHM pour Somatic HyperMutation)

Ce processus se produit uniquement après contact avec l'antigène. Il consiste à introduire des mutations ponctuelles au niveau de la région variable des immunoglobulines à une fréquence élevée (10^{-3} à 10^{-4} par paire de base et par division cellulaire). La région constante des gènes des immunoglobulines n'est pas touchée par ce processus. Les mutations observées sont principalement des modifications d'un seul nucléotide, le plus souvent de type transition^a, mais des délétions ou duplications de taille variable peuvent être rencontrées^{2,11}. L'apparition de délétion et duplication serait due à la génération de cassures double brin au cours de l'hypermutation somatique^{11,12}. Cela permet d'augmenter la diversité des immunoglobulines.

^a Mutation ponctuelle entraînant la substitution d'une base purique par une autre base purique (A → G ou G → A) ou d'une base pyrimidique par une autre base pyrimidique (T → C ou C → T)



Ainsi, avec l'aide des lymphocytes T, les lymphocytes B situés au niveau du centre germinatif des ganglions vont subir ce processus d'hypermutation somatique. Chez l'Homme, la première partie des SHM va concerner des hotspots mutationnels au niveau de motifs RGYW (Purine/Guanine/Pyrimidine/Adénosine), et la seconde partie va consister en l'introduction d'erreurs au niveau de l'ADN, pouvant induire des mésappariements entre brins muté et sauvage. Ainsi, l'exposition à l'antigène va entraîner des mutations permettant une maturation lymphocytaire en augmentant l'affinité pour l'antigène³.

Le processus d'hypermutation somatique peut toucher d'autres gènes que les immunoglobulines. Ainsi, il a été montré en particulier que trois gènes pouvaient être touchés à une fréquence significative. Il s'agit des gènes *BCL6*, *CD95* (ou *FAS*) et *CD79b*^{2,13,14,15,16}.

iii- Recombinaison de classe ou Class switch recombination (CSR)

La recombinaison de classe ou commutation isotypique permet de modifier la classe d'immunoglobuline produite (IgM au départ pouvant devenir IgG, IgE, IgA, IgD). Cette recombinaison se produit au niveau du locus IGH. Il existe 9 gènes constants de la chaîne lourde (*CH*) situés en aval des régions VDJ³. Ces gènes sont constitués d'une série d'exons, chacun codant un domaine CH. Chaque gène constant peut coder 2 types différents de domaines terminaux, suivant un épissage alternatif, permettant une expression de l'immunoglobuline à la surface du lymphocyte B ou bien aboutissant à sa sécrétion sous forme soluble. Chaque gène *CH* (à l'exception de *CH δ*) est précédé d'une région répétée d'ADN appelée région Switch et d'un premier exon non traduit. La recombinaison de classe a lieu entre le Switch μ et l'une des sept autres régions Switch des gènes constants (Figure 7). Ce processus permet la juxtaposition de la même chaîne variable lourde VDJ avec une autre région de chaîne lourde constante d'immunoglobuline, entraînant la délétion de la région des gènes des chaînes constantes initialement exprimées (IgM-IgD) jusqu'au switch de la chaîne lourde impliquée. Ce processus est initié par la transcription de promoteurs situés en amont des gènes *CH*¹⁷.



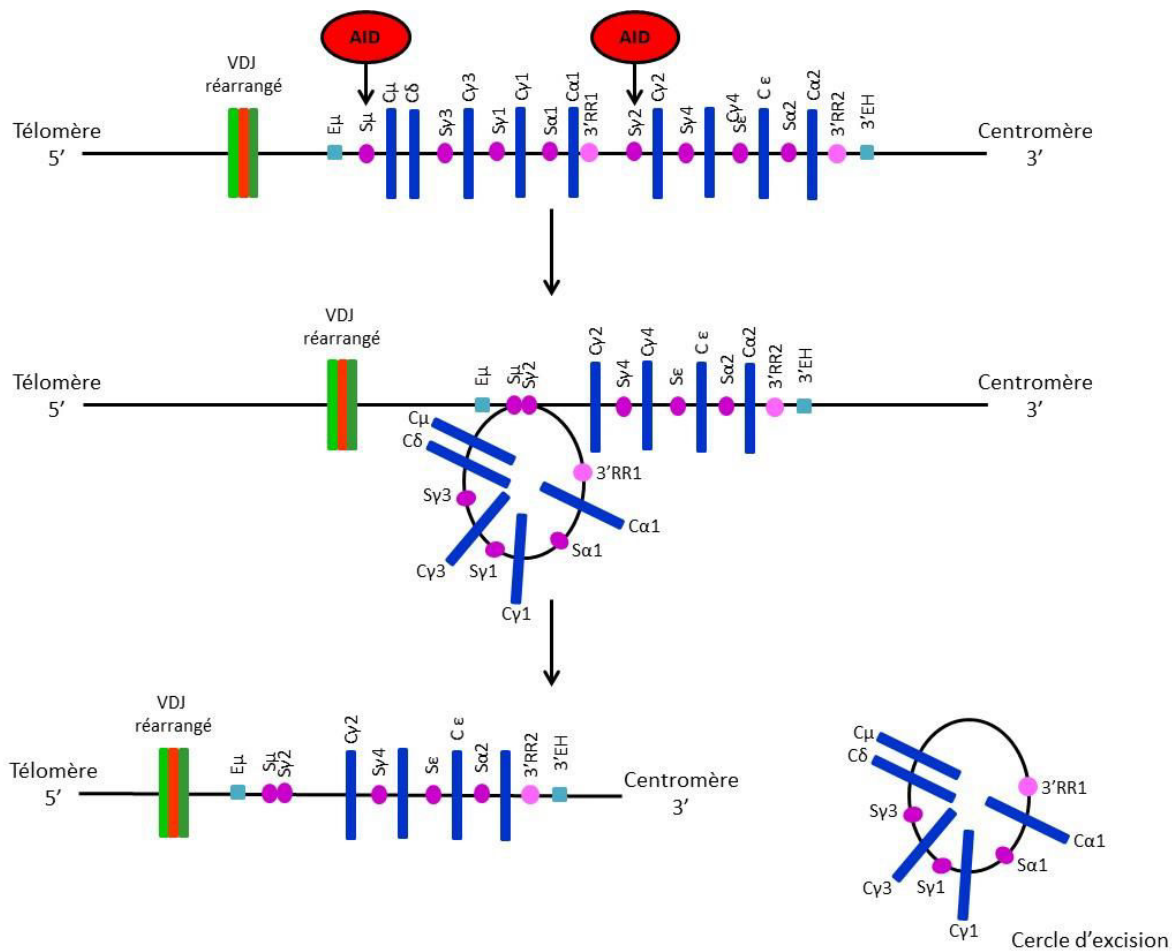


Figure 7 : Schéma de commutation de classe isotypique (ou CSR pour Class Switch Recombination) sous l'action d'AID (pour Activation-Induced Cytidine Deaminase), générant le segment d'immunoglobuline réarrangé et un cercle d'excision correspondant à la région coupée.

iv- Le rôle de l'enzyme AID (pour Activation-Induced Cytidine Deaminase)

L'enzyme AID joue un rôle clé dans les phases de recombinaison de classe et d'hypermutation somatique au niveau des centres germinatifs des ganglions. Le rôle clé d'AID dans ces deux processus a été mis en évidence d'une part par le fait qu'un déficit en AID entraîne le syndrome d'hyperIgM autosomique récessif caractérisé par un manque de recombinaison de classe isotypique et d'hypermutation somatique, et d'autre part par le fait qu'une expression ectopique d'AID est suffisante pour causer l'hypermutation somatique et la commutation de classe au niveau de structures transgéniques dans les cellules B^{18,19,20}. Il s'agit d'une enzyme ADN-simple brin, c'est-à-dire qu'elle ne va interagir qu'avec un seul brin d'ADN. Lors de la transcription d'un domaine V ou d'une région Switch d'un gène constant, le brin d'ADN sens



va être ouvert. AID peut alors venir agir et désaminer une Cytidine en Uracile, aboutissant ensuite à des mésappariements désoxyUridine/désoxyGuanosine. Le système de réparation par excision de base (ou « base excision repair ») via l'enzyme Uracyl-DNA glycosylase va enlever cette base. Enfin, les systèmes de réparation de l'ADN vont permettre d'aboutir soit à la recombinaison de classe, soit à l'hypermutation somatique. Les protéines MSH2 et MSH6 (pour MutS Homolog 2 ou 6) appartenant au système de réparation des mésappariements (ou « mismatch repair ») peuvent également intervenir au niveau des mésappariements désoxyUridine/désoxyGuanosine (Figure 8).

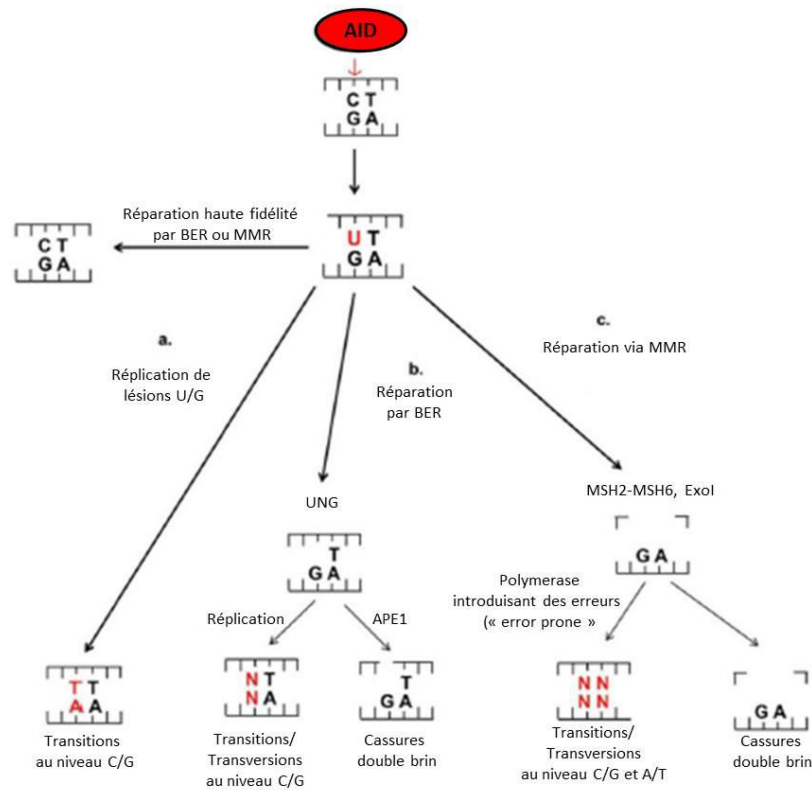


Figure 8 : Mécanisme de désamination de cytidine via AID dans les processus de SHM et CSR d'après Hwang *et al.*²¹.

AID désamine une Cytidine (C) en Uracile (U). La lésion U/G peut être réparée de façon fidèle via les systèmes de réparation d'excision de base (BER pour Base Excision Repair) ou de réparation de mésappariement (MMR pour Mismatch Repair). Les systèmes de réparation peuvent générer des modifications : a. La réplication au niveau U/G entraîne une transition. b. L'Uracil-DNA-Glycosylase (UNG) du système BER coupe la base U, créant un site apurinique/aprimidinique ou AP, et la réplication entraîne une transition ou une transversion au niveau C/G. N représente une base parmi A,T,G,C. AP endonucléase 1 (APE1) peut entraîner une cassure à ce niveau et générer une cassure double brin. c. MSH2-MSH6 du système MMR reconnaît le mésappariement U/G et Exo1 va couper ce site. Une polymérase introduisant des erreurs (ou « error prone polymerase ») va resynthétiser l'ADN au niveau de ce site en générant des mutations au niveau A/T également. Il peut aussi y avoir la génération de cassure double brin.



b- Locus IGH

Le locus *Immunoglobulin Heavy* ou *IGH* est situé sur le bras long du chromosome 14 en 14q32.33. Il s'étend sur environ 1250 kb²². Il s'agit du locus d'immunoglobuline le plus complexe. Il est antisens sur le chromosome 14 (c'est-à-dire que la partie 5' est située du côté télomérique et que la région 3' est localisée du côté centromérique (Figure 9). Il comprend environ 80 segments VH différents situés à proximité du télomère du chromosome 14³. Ces différents segments sont regroupés en 7 familles. A la suite de la région VH se trouvent les segments DH. Il en existe 27 chez l'Homme. On trouve ensuite les 6 segments JH (Figure 10). Chaque segment VH et JH est accolé à un RSS dit « deux tours », soit avec un espaceur de 23 pb, tandis que les segments DH sont entourés de part et d'autre par un RSS dit « un tour », soit avec un espaceur de 12 pb³. Cette structure permet d'éviter les mauvais réarrangements entre segments V et J directement.

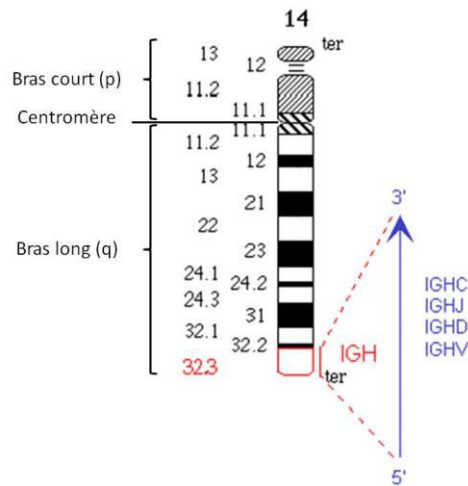


Figure 9 : Représentation du chromosome 14 avec localisation du locus IGH, adapté d'après Lefranc *et al.*²².

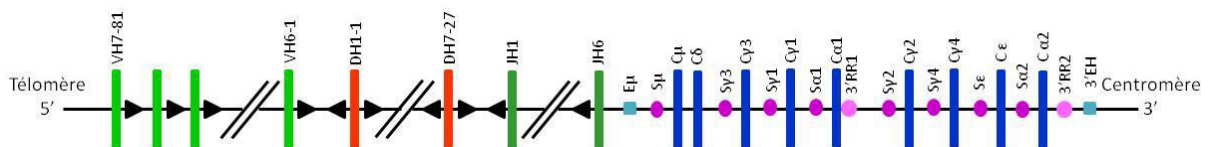


Figure 10 : Schéma simplifié du locus IGH en 14q32, adapté d'après Schroeder *et al.*, Van Dongen *et al.*, et Pinaud *et al.*^{3,23,24}.

Les flèches noires représentent les RSS. Les ronds violets indiquent les régions Switch. Les ronds roses représentent les 3'RR (pour 3' Regulatory Region). Enfin, les carrés bleus indiquent les régions enhancer.



La recombinaison VDJ au niveau du locus *IGH* débute par la recombinaison DH-JH, puis le segment VH est relié au segment DJ créé (Figure 6). Le segment V codera les portions FR1, FR2 et FR3, CDR1, CDR2 et la partie amino-terminale du CDR3. Le segment D codera le milieu du CDR3. Le segment J contient la partie carboxy-terminale du CDR3 et le FR4. La recombinaison possible entre la cinquantaine de segments V actifs, les 27 D et les 6 J permet la potentielle formation de 10^4 combinaisons VDJ différentes³. En plus de cela, il existe plusieurs mécanismes entraînant une diversité jonctionnelle au niveau VDJ. Tout d'abord, chaque segment DH peut se réarranger avec des phénomènes de délétion et d'inversion, et être épissé et traduit suivant 3 cadres de lecture. Ainsi, 6 segments différents pour chaque DH peuvent être produits. Ensuite, au moment du réarrangement, il y a création d'une cassure avec structure en épingle dont le clivage va aboutir à la formation de région palindromique appelée jonction P. De plus, les extrémités des segments réarrangés peuvent perdre un ou plusieurs nucléotides au cours du processus de recombinaison. Enfin, la TdT peut entraîner l'addition au hasard ou le remplacement de plusieurs nucléotides entre les segments V et D ou D et J générant ainsi la N-diversité (Figure 11). L'ensemble de ces processus va entraîner la formation d'une région VDJ réarrangée de structure et de taille variables, avec plus de 10^{10} combinaisons possibles. L'association par la suite avec les chaînes légères réarrangées entraîne 10^{16} combinaisons possibles avant sélection puis modification sous l'influence de l'exposition à l'antigène³.

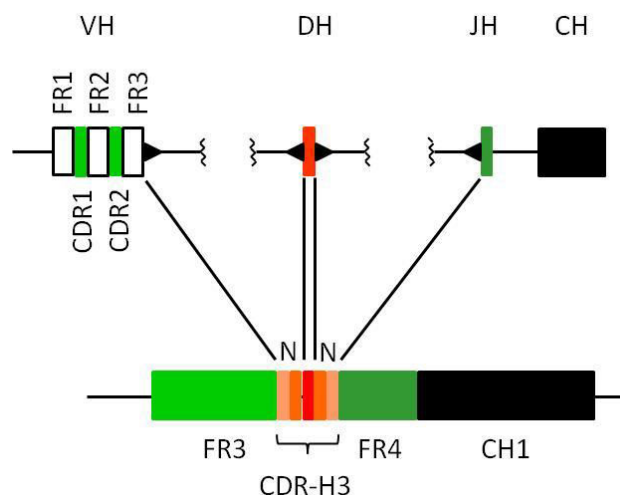


Figure 11 : Segment IGH réarrangé et diversité jonctionnelle, adapté d'après Schroeder *et al.*⁵. En orange, les régions d'introduction de nucléotides au cours du réarrangement VDJ (orange clair : insertion région codante VH ou JH et orange foncé : insertion région codante DH).

En aval des régions VDJ se situent les régions constantes CH, côté centromérique du chromosome 14 (Figure 10)²². Il existe 11 segments IGHC différents : il s'agit des gènes IGHM, IGHD, IGHG3, IGHG1, IGHEP1, IGHA1, IGHP, IGHG2, IGHG4, IGHE et IGHA2. Hormis IGHP et



IGHEP1 qui sont respectivement un ORF^a (pour Open Reading Frame) et un pseudogène^b, les 9 autres sont des gènes fonctionnels²².

A noter qu'il existe des orphons^c, c'est-à-dire des gènes *IGH* localisés en dehors du locus *IGH* en 14q32.33. Ils sont localisés en 15q11.2, 16p11.2, 9p24.2-p24.1 et 21p11.2^{22,25}.

c- Locus IGK

Le locus Immunoglobuline Kappa (IGK) est situé sur le bras court du chromosome 2 en 2p11.2 (Figure 12). Ce locus fait environ 1820 kb²⁶.

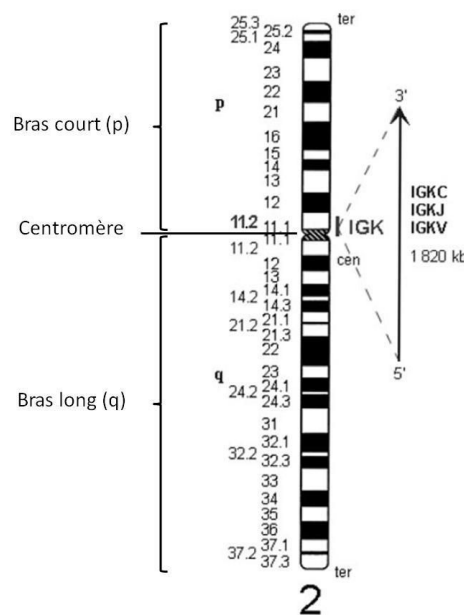


Figure 12 : Représentation du chromosome 2 avec localisation du locus IGK, adapté d'après Lefranc *et al.*²⁶.

^a ORF = Cadre de lecture ouvert c'est-à-dire succession de codons significants pour une séquence codante

^b Pseudogène = séquence non exprimée d'ADN présentant une grande homologie avec un gène actif, dont il dérive par duplication, mutation ou encore rétrotranscription

^c Orphon = Gène non fonctionnel correspondant à un pseudogène ou un ORF situé dans une région chromosomique différente du locus principal



La région variable de la chaîne légère Kappa sera codée par le réarrangement entre les segments V_k et J_k , tandis que la région constante sera codée par un seul exon C_k . Le locus IGK comprend 76 V_k et 5 J_k différents, situés en amont des régions constantes C_k ²⁶. Environ un tiers des segments V_k contiennent des mutations frameshift ou générant des codons stop, les rendant inutilisables pour la formation de protéines fonctionnelles³. Moins de 30 segments V_k différents ont été retrouvés dans des immunoglobulines fonctionnelles. Ils sont regroupés en 7 familles basées sur des similarités de séquence et de structure^{3,26}. Les segments V_k sont dispersés selon deux clusters dupliqués : le premier situé directement en amont des segments J_k présente une orientation identique à celle des segment V_k , tandis que le deuxième situé encore plus en amont présente une orientation inversée (segments D_k) (Figure 13)²³.

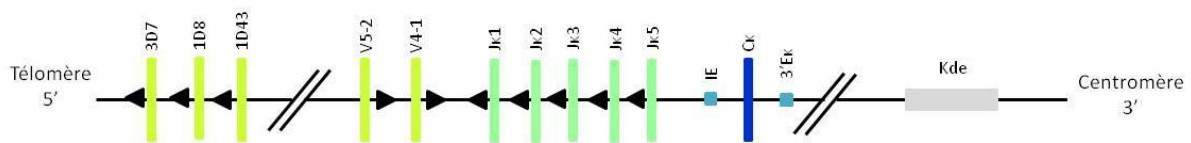


Figure 13 : Schéma simplifié du locus IGK en 2p11.2, adapté selon Schroeder et al. et Van Dongen et al.^{3,23}.

Les flèches noires indiquent les RSS. Les carrés bleus représentent les enhancers (avec IE pour enhancer intronique et 3'Ek pour Enhancer 3' Kappa). Le segment gris représente la région Kde (pour Kappa deleting element).

La recombinaison V_k - J_k entraîne la formation de 140 combinaisons possibles. Le segment V_k code les régions CDR1, CDR2, et la partie amino-terminale du CDR3, ainsi que les FR1, FR2 et FR3. Le segment J_k code la partie carboxy-terminale du CDR3 et le FR4. L'extrémité de chaque segment réarrangé peut perdre 1 à 5 nucléotides au cours du processus. De plus, la TdT peut introduire un ou plusieurs nucléotides pour remplacer l'un ou l'ensemble des nucléotides enlevés au cours du réarrangement, ou ajouter de façon aléatoire plusieurs nucléotides à la séquence originale. Ces processus permettent une diversité jonctionnelle 20 fois supérieure. Cette diversité se situe donc au niveau du CDR3.

Il existe également un segment appelé Kde (pour Kappa deleting element) situé approximativement 24 kb en aval de C_k , pouvant être impliqué dans des réarrangements V_k -Kde, ainsi qu'un RSS isolé au niveau de l'intron J_k - C_k (intron RSS-Kde)²³. Les réarrangements impliquant ces deux régions entraîneront une inactivation fonctionnelle de l'allèle κ .

Il existe des orphons *IGK* en dehors du locus IGK. Ils sont localisés en 1q21.1, 2p11.1, 2q11.2, 2q12-14, 9q12, 9q13, 10q11.21, 22q11, Yq11.21 et sur le chromosome 15^{26,27}.



d- Locus IGL

Le locus Immunoglobuline Lambda (ou IGL) est localisé sur le bras long du chromosome 22 en 22q11.2. Le locus fait environ 1050 kb²⁸. L'orientation du gène *IGL* est sens, c'est-à-dire que la région 5' est située côté centromérique et la partie 3' est localisée au niveau télomérique (Figure 14). Il contient 4 à 5 segments C λ fonctionnels précédés du segment J λ correspondant (Figure 15)^{3,28}. Les segments V λ sont regroupés en 3 clusters différents contenant des membres de différentes familles de V λ ³. En fonction des haplotypes existants, on recense environ 30 à 36 gènes V λ fonctionnels et un nombre équivalent de pseudogènes. Il existe également au niveau de ce locus des gènes codant des pseudo-chaînes légères ψ LC. Il s'agit des gènes *V_{préB}* et *λ 14.1* (ou *λ 5*). L'expression de ces 2 gènes entraîne la formation de pseudo-chaînes légères ayant une grande homologie avec les chaînes Lambda. Ces gènes s'expriment au cours des stades précoces du développement B³. Ces pseudo-chaînes légères s'associent avec les chaînes lourdes pour former le récepteur pré-B.

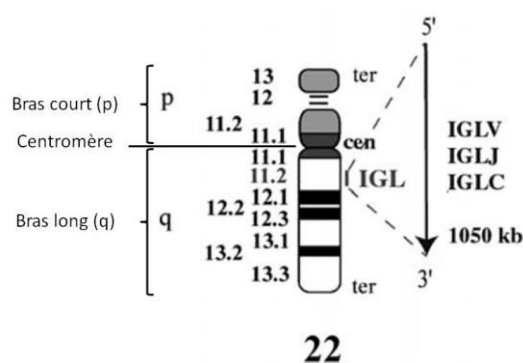


Figure 14 : Représentation du chromosome 22 avec localisation du locus IGL, adapté d'après Lefranc *et al.*²⁸.

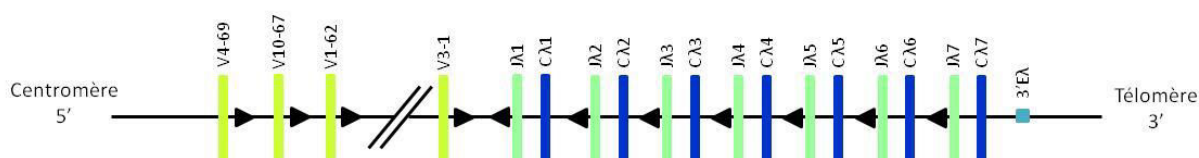


Figure 15 : Schéma simplifié du locus IGL en 22q11.2, adapté selon Schroeder *et al.* et Van Dongen *et al.*^{3,23}.

Les flèches représentent les RSS. Le carré bleu représente l'enhancer en 3'.

Il existe des orphons *IGL* en dehors du locus *IGL* en 22q11.2. Ils sont localisés en 8q11.2, 18p11 et 22q12^{28,29}.



4- Le récepteur de la cellule B ou BCR

a- Structure du BCR

Le lymphocyte B est caractérisé par l'expression à sa surface d'un récepteur spécifique capable de reconnaître et d'interagir avec l'antigène : le récepteur B ou BCR (pour B-Cell Receptor).

Il s'agit d'un complexe transmembranaire qui contrôle le devenir de la cellule depuis les stades précoces de la différenciation lymphocytaire B³⁰. Il va guider la maturation cellulaire, la survie, l'anergie et la production d'anticorps dans les plasmocytes.

Le BCR est constitué d'une immunoglobuline de surface issue du réarrangement des gènes des immunoglobulines, ce qui permet de produire un site de liaison à l'antigène unique.

Le BCR est connecté à un réseau de kinases et phosphatases qui ajustent et amplifient son activation. L'immunoglobuline de surface n'a pas d'activité kinase par elle-même, mais via son couplage avec l'hétérodimère composé des protéines Ig α et Ig β (CD79A et CD79B respectivement) (Figure 16). Chacune de ces sous-unités CD79A et CD79B contient un immunorécepteur à motif d'activation basé sur les tyrosines (ITAM pour Immunoreceptor Tyrosine-based Activation Motif) composé de deux résidus tyrosine conservés³⁰.

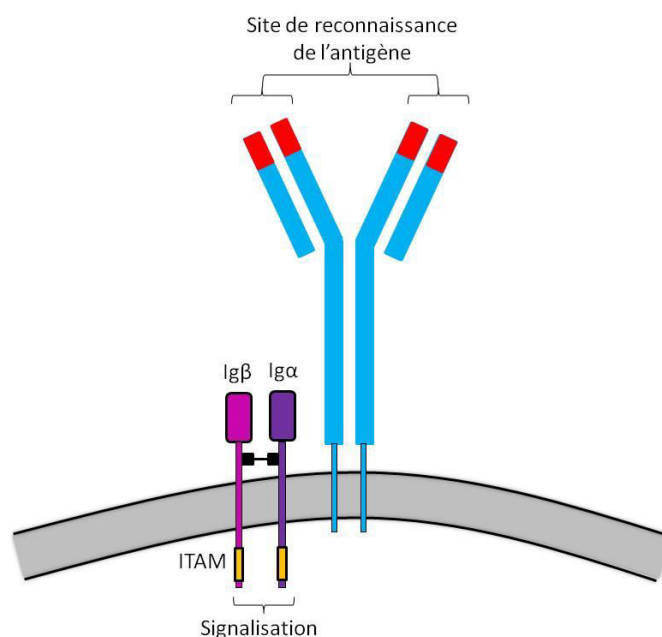


Figure 16 : Représentation du récepteur de la cellule B (BCR), adapté selon Immunologie, Janeway, De Boeck, 2009⁸.

BCR constitué de l'immunoglobuline de surface et des protéines Ig α et Ig β reliées entre elles par des ponts disulfures (en noir) et dont les domaines ITAM (en orange) permettent la signalisation.



b- Principales voies de signalisation induites par le BCR

L'agrégation du BCR et son changement de conformation induits par l'antigène entraînent la phosphorylation des ITAMs via la kinase LYN (pour Lyn proto-oncogene, Src family tyrosine kinase), provoquant la cascade d'activation du BCR (Figure 17)³⁰. Cette phosphorylation crée des sites d'ancrage pour les protéines contenant des domaines SH-2, telles que SYK (pour Spleen Tyrosine Kinase) et un recrutement accru de la kinase LYN. D'autres tyrosine kinases telles que FYN et BLK contribuent également à ce processus³⁰. On parle de kinases de la famille SRC (SFK pour SRC Family Kinases)³¹.

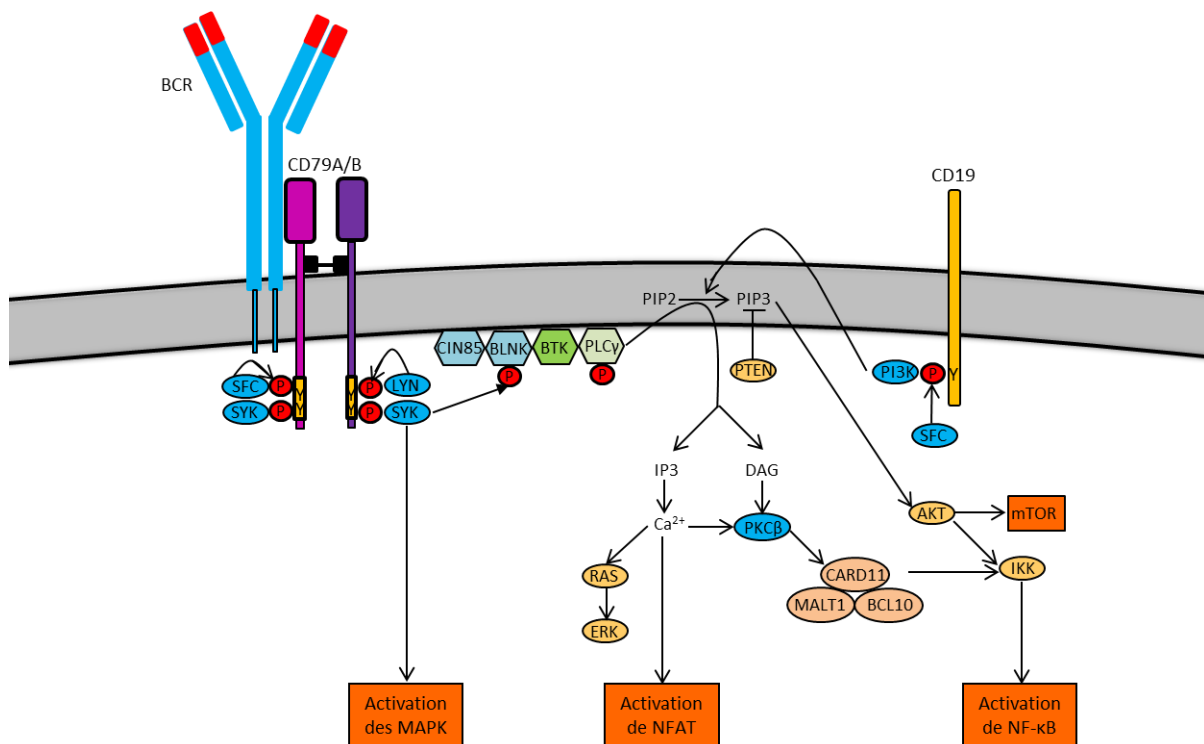


Figure 17 : Représentation schématique de la voie de signalisation du BCR adapté d'après Seda *et al.* et Young et Staudt^{30,31}.

Les protéines SYK et BTK (pour Bruton Tyrosine Kinase) sont essentielles pour la signalisation : SYK va recruter le complexe BLNK (pour B-cell linker protein) et CIN85 (pour Cbl Interacting protein of 85 kD) et activer BTK. L'absence de BTK ou une protéine non fonctionnelle va entraîner la mort du lymphocyte³⁰. L'activation de BTK dépend de la phosphorylation de la phosphoinositide 3-kinase (PI3K) par SYK. Cette phosphorylation est nécessaire pour la liaison de la PI3K au niveau de la membrane cellulaire, à proximité de ses substrats. Cette liaison de la PI3K est assurée par son association à différentes protéines qui serviront de sites



d'ancrage³⁰. CD19, BCAP (pour B-Cell Adaptor for PI3K) et GAB1/2 (pour GRB2-associated-binding protein 1 et 2) sont des protéines d'ancrage du PI3K phosphorylé. Elles permettent le recrutement et l'assemblage du signalosome du BCR, et l'amplification du signal. La phosphorylation de la PI3K peut court-circuiter SYK et passer directement par la liaison de LYN à la PI3K, mais dans ce cas l'activité de la PI3K n'est que faiblement augmentée³⁰. La PI3K activée va phosphoryler le phosphatidylinositide biphosphate (PIP2), créant un deuxième messenger (PIP3) servant de site d'ancrage pour recruter d'autres effecteurs tels que BTK, AKT, PLC γ (pour Phospholipase C γ) et PKC β (pour Protéine Kinase C β). La BTK joue un rôle crucial dans le développement lymphocytaire B et des mutations héréditaires du gène sont responsables de l'agammaglobulinémie de Bruton ou agammaglobulinémie liée à l'X. Il existe également des inhibiteurs de BTK administrables dans le traitement de certains lymphomes dont la LLC. La PLC γ va cliver PIP2 en diacylglycérol (DAG) et IP3 (inositol triphosphate, ce qui entraîne une augmentation du calcium intracellulaire. La combinaison de l'augmentation du calcium intracellulaire et la présence de DAG va permettre l'activation de la PKC β qui va phosphoryler différents substrats. L'activation du BCR va agir sur d'autres voies de signalisation comme la voie ERK/MAPK (pour Extracellular signal-regulated kinase et Mitogen Activated Protein kinases), NF κ B (pour Nuclear Factor-Kappa B), mTOR (pour mammalian Target Of Rapamycin), RAS et NFAT (pour Nuclear Factor of Activated T-cells)³⁰. Par exemple, le recrutement de la PKC β va entraîner l'activation de CARD11 (pour Caspase recruitment Domain-containing protein 11) par phosphorylation. Cette dernière va coordonner un complexe de signalisation aboutissant à l'activation de la voie NF κ B³¹.

LYN joue également un rôle régulateur en effectuant un rétrocontrôle négatif sur l'activation du BCR via des phosphatases : elle va phosphoryler des immunorécepteurs à motifs inhibiteurs basés sur des tyrosines ou ITIM (pour Immunoreceptor Tyrosine-based Motifs) au niveau du CD22 ou du Fc γ RIIb. Ces deux molécules vont servir de point d'ancrage pour PTEN (pour Phosphatase and Tensin homolog) et SHIP1/2 (pour SH2-containing Inositol 5' Polyphosphatase 1 et 2) qui vont enlever un 3'Phosphate à PIP3 et ainsi atténuer la voie de signalisation en aval de la PI3K. Cette voie inhibitrice est aussi activée en cas de contact prolongé avec des auto-antigènes et permet d'éliminer les B auto-réactifs ou de les maintenir sous forme anergique³¹. Par ailleurs, la molécule CD45 va contrôler l'activité de LYN en enlevant une phosphorylation inhibitrice (au niveau de la tyrosine 508), qui peut être rétablie par la CSK (pour C-terminal src tyrosine kinase).

Cette voie de signalisation peut être modulée selon le stade de maturation de la cellule et sa localisation (sang, ganglion...). Les lymphocytes B mémoires ayant réalisé la commutation de classe isotypique vers une IgG vont présenter une queue cytoplasmique plus longue de l'immunoglobuline, ce qui permet de propager la voie de signalisation du BCR et d'augmenter la réponse précoce à l'antigène³⁰. Dans les cellules B matures n'ayant pas encore rencontré l'antigène, la signalisation du BCR serait maintenue dans un état basal dit « tonique »³⁰.



5- Lymphopoïèse B normale

La maturation normale des lymphocytes B prend naissance au niveau des organes lymphoïdes primaires, à savoir la moelle osseuse pour la lignée B. Comme l'ensemble des cellules hématopoïétiques, les lymphocytes B dérivent des cellules souches hématopoïétiques après engagement de celles-ci vers la lignée B. Cette première phase va aboutir à la formation de récepteurs B spécifiques : le BCR¹⁰. La maturation finale aura lieu au niveau des organes lymphoïdes secondaires représentés par les ganglions, la rate et le tissu lymphoïde associé aux muqueuses (MALT pour Mucosa Associated Lymphoid Tissue). Cette phase aura lieu lorsque la cellule B aura rencontré l'antigène¹⁰. Les stades terminaux de maturation sont représentés par les lymphocytes B mémoire permettant de réagir plus rapidement face à un antigène connu, ou par les plasmocytes produisant les anticorps.

a- Phase indépendante de l'antigène

Les lymphocytes B sont produits dans la moelle osseuse à partir de cellules souches hématopoïétiques (CSH)^{2,32}. A partir d'une cellule souche hématopoïétique vont dériver des cellules précurseurs engagées dans le lignage lymphoïde. Ce sont les progéniteurs multipotents lymphoïdes (ou Lymphoid-primed Multipotent Progenitor ou LMPP) puis les progéniteurs lymphoïdes communs (ou CLP). Ces cellules vont ensuite pouvoir s'engager vers la lignée B. Le développement des stades précoces des lymphocytes B est régi par le réarrangement ordonné des gènes des immunoglobulines, et la sélection des cellules à différents stades de développement pour obtenir des récepteurs fonctionnels (Figure 18).

Les cellules stromales assurent le développement des lymphocytes B dans la moelle osseuse en établissant des interactions spécifiques à la surface des cellules B via des molécules d'adhérence et leurs ligands, et en produisant des facteurs de croissance tels que le SCF (pour Stem Cell Factor) reconnu par le récepteur Kit des cellules B immatures ou l'interleukine 7 (IL7) qui agit sur les cellules pro-B tardives et pré-B¹⁰. Les différents stades de maturation sont caractérisés par l'expression différentielle de marqueurs, permettant leur identification par cytométrie en flux (Figure 19).



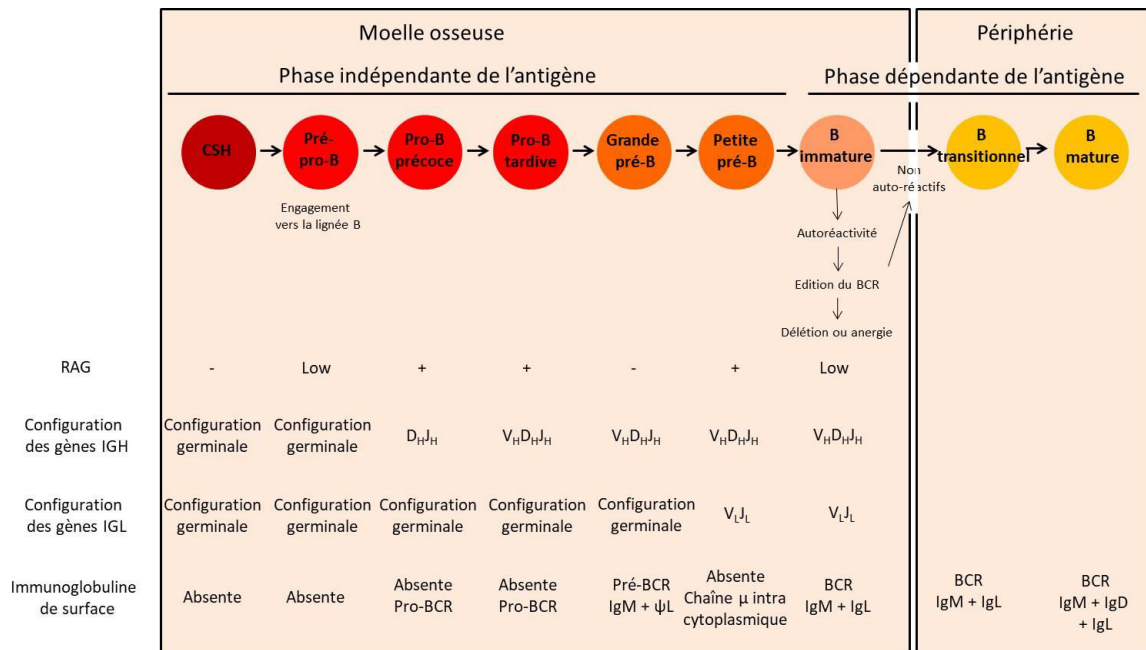


Figure 18 : Maturation lymphocytaire B précoce adapté selon Meffre *et al.* et Le système immunitaire, De Boeck^{33,10}.

Marqueur	Stade médullaire						Stade périphérique			
	CSH	pré-pro-B	pro-B	pré-BI	pré-BII	Cellule B immature	Cellule B mature			Plasmocyte
							naïve	activée	mémoire	
CD34	+	+	+	-	-	-	-	-	-	-
CD10	+	+	+	+	+	+	-	-	-	-
IL7Rα	+	+	+	-	-	-	-	-	-	-
CD19	-	-	+	+	+	+	+	+	+	+
CD79a	-	+	+	+	+	+	+	+	+	+
			(intra)	(intra)						(intra)
TdT	-	-	+	-	-	-	-	-	-	-
sIgM	-	-	-	-	-	+	+	+	+	-
sIgD	-	-	-	-	-	-	+	-	-	-
CD20	-	-	+/-	+	+	+	+	+	+	-
CD27	-	-	-	-	-	-	-	+	+	+
CD138	-	-	+	-	-	-	-	-	-	+
CD38	+	+	+	+	+	+	+faible	+	-	+

Figure 19 : Expression des marqueurs immunophénotypiques au cours des différents stades de développement de la cellule B, adapté d'après Armitage *et al.*, Ye et Graf, et LeBien et Fates^{2,34,35}.



i- Les stades précoces de cellule souche hématopoïétique et progéniteurs

Au stade de CSH, l'expression des facteurs de transcription Ikaros, PU.1 (pour Purine box factor 1) et E2A (pour Immunoglobulin enhancer binding factors E12/E47 ou TCF3 pour Transcription factor 3) participe au développement des stades précoces de la lignée B⁴. Ikaros joue un rôle dans le remodelage de la chromatine au niveau de régions spécifiques de l'ADN, permettant une conformation accessible au niveau des gènes impliqués dans le développement des cellules B. Le niveau d'expression de PU.1 permet d'orienter la cellule vers le lignage lymphoïde (en cas de faible expression) ou myéloïde (en cas de forte expression). Son niveau d'expression est régulé par le répresseur transcriptionnel Gfi1 (pour Growth factor independent 1 transcription repressor)⁴. E2A joue un rôle dans le maintien des CSH et régule leur cycle cellulaire. Les CSH expriment à leur surface la molécule c-kit (CD117).

Après le passage par le stade de progéniteur multipotent, les cellules vont pouvoir s'engager vers le lignage lymphoïde. C'est le stade LMPP. La cellule exprime alors le récepteur fms-related tyrosine kinase 3 (flt-3) capable de se lier à son ligand flt-3 retrouvé sur la membrane des cellules stromales. Cette interaction entraîne la production du récepteur à l'IL-7 (IL-7R ou CD127). Les LMPP vont commencer à exprimer RAG-1/2 et la TdT. La cellule va ensuite passer par le stade de progéniteur lymphoïde précoce dont une partie va migrer vers le thymus pour devenir des lymphocytes T, tandis que l'autre partie va rester dans la moelle osseuse pour devenir un progéniteur lymphoïde commun capable de produire des lymphocytes B, T ou NK⁴.

ii- Le stade pré-pro-B

Le premier stade de développement de la lignée B est appelé cellules pré-pro-B². Il débute avec l'expression du marqueur spécifique de la lignée B : le CD45R. Ces cellules sont également CD43³⁴. Le facteur de transcription EBF1 (pour Early B-cell Factor 1) est alors fortement exprimé⁴. A ce stade, les loci des immunoglobulines sont en configuration germinale. Il existe déjà des marqueurs de surface permettant de distinguer ces cellules (CD45R par exemple), mais elles n'expriment pas encore les éléments du BCR³³. Le niveau d'expression des enzymes RAG-1 et RAG-2 est faible. De nombreux facteurs de transcription constituent des éléments régulateurs importants à ce stade, parmi lesquels Pax-5 qui semble essentiel pour le maintien de l'engagement cellulaire vers la lignée B³³. E2A et EBF-1 vont se lier au gène *IGH*, facilitant l'accès à la région D_H-J_H, permettant la préparation à la première étape du réarrangement des gènes des immunoglobulines⁴. EBF1 est nécessaire pour l'expression de nombreuses protéines parmi lesquelles Igα et Igβ.



iii- Le stade pro-B

La cellule va ensuite exprimer de nouveaux marqueurs de surface : elle va commencer à exprimer les marqueurs CD79a (Ig α) et CD79b (Ig β), constituant le pro-BCR³³. Le complexe Ig α -Ig β est exprimé avec la protéine chaperonne Calnexine. La cellule est alors appelée cellule pro-B. Le facteur de transcription Pax-5, sous le contrôle d'EBF-1, va initier la transcription de gènes contrôles (dont le gène CD19), ce qui bloque l'expression des gènes non B et signe le passage vers le stade pro-B⁴.

La cellule va initier le réarrangement D_H-J_H, celui-ci pouvant se produire sur les 2 allèles *IGH*, puis le réarrangement V_H-D_HJ_H. Le réarrangement obtenu peut être productif ou non. Seuls les réarrangements productifs peuvent entraîner la formation d'une chaîne lourde protéique². Si le premier réarrangement V_HD_HJ_H n'est pas productif, le précurseur B peut faire une deuxième tentative avec l'autre allèle pour générer un réarrangement productif. La chaîne lourde fonctionnelle est alors produite en même temps qu'une pseudo-chaîne légère qui est composée de 2 protéines : $\lambda 5$ (ressemblant à la partie constante d'une chaîne légère) et V-preB (de structure proche de la partie variable d'une chaîne légère)^{2,8}. Les molécules de transduction du signal Ig α et Ig β (CD79a et CD79b) sont également exprimées à ce stade. Cette structure est appelée récepteur pré-B et le terme de cellule pré-B est alors utilisé. Les cellules pro-B expriment également les marqueurs CD34, CD10 et CD19, et pour la majorité d'entre elles la TdT³⁵.

iv- Le stade pré-B

Le stade pré-B est divisé en 2 : les cellules les plus immatures sont appelées grandes cellules pré-B ou pré-BI ou pré-B précoces. Les loci des chaînes légères sont en conformation germinale et les enzymes RAG-1 et RAG-2 sont faiblement exprimées³³. Les cellules les plus matures sont appelées petites cellules pré-B ou pré-BII ou pré-B ultérieur. Les enzymes RAG-1 et RAG-2 sont alors fortement exprimées et le réarrangement des chaînes légères peut avoir lieu. Le passage du stade pro-B au stade pré-B est caractérisé par la perte du CD34 et de l'expression de la TdT³⁵

Les cellules pré-BI sont sélectionnées sur l'expression du récepteur pré-BCR. Une cellule ne pouvant exprimer ce récepteur, à cause d'un réarrangement non productif, vont mourir par apoptose. On parle de contrôle pré-B⁴. Les réarrangements des gènes *IGH* sont alors stoppés afin d'éviter la formation simultanée de 2 chaînes IGH de spécificité différentes : on parle d'exclusion allélique^{2,32}. Il s'agit d'un processus par lequel un seul allèle d'un gène va être exprimé dans une cellule diploïde⁸.



Le pré-BCR va entraîner une signalisation favorisant l'exclusion allélique en réduisant l'expression de RAG-1/2, en phosphorylant RAG-2 au cours de la division cellulaire, favorisant sa dégradation et en réduisant l'accès du complexe recombinase au niveau du locus IGH⁸.

Les cellules pré-BI vont se diviser plusieurs fois, puis devenir des cellules pré-BII. A ce stade, le pré-BCR n'est plus exprimé et le réarrangement des chaînes légères débute alors, en général avec les loci IGK, puis, en cas de génération de réarrangements non productifs, une nouvelle tentative de réarrangement a lieu au niveau des loci IGL^{2,8}. Cette théorie est appuyée par le fait que les cellules B normales et malignes exprimant des chaînes légères Kappa, ont en général des loci IGL en configuration germinale, tandis que lorsque la cellule exprime une chaîne légère Lambda, les loci IGK ont généralement subi un réarrangement^{36,37}. De plus, du fait de la structure des gènes codant les chaînes légères des immunoglobulines, un nouveau réarrangement est possible sur le même allèle de chaîne légère lorsque la première tentative n'est pas productive et n'implique pas le premier segment variable (Figure 20).

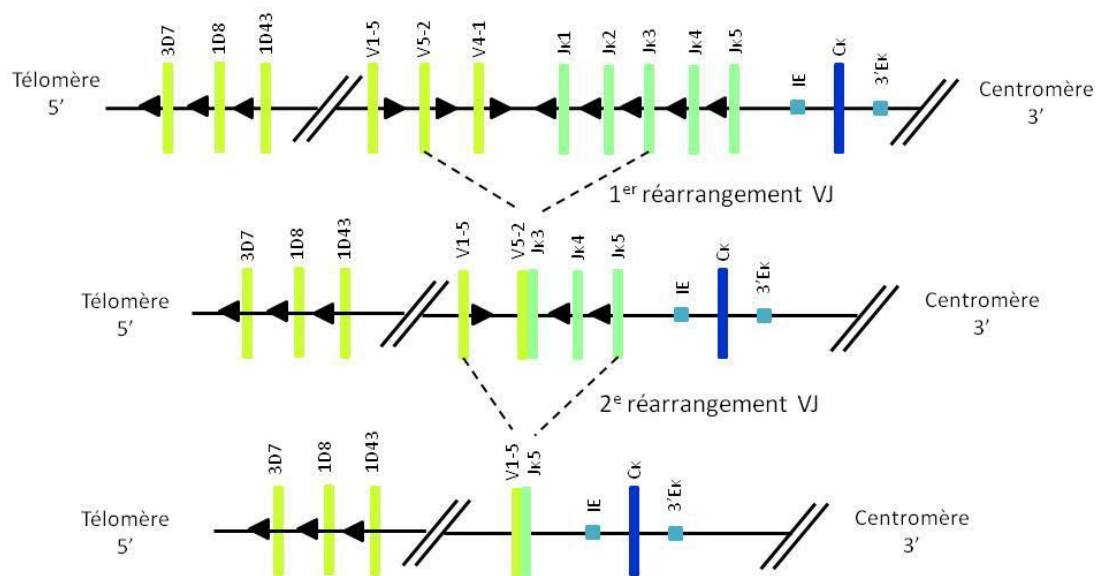


Figure 20 : Réarrangement VJ sur la chaîne légère Kappa adapté selon Le Système Immunitaire, De Boeck¹⁰. Deux recombinaisons VJ successives sur le même allèle IGK.

Le réarrangement productif de chaîne légère est alors exprimé. Si la chaîne légère produite ne peut pas s'apparier avec la chaîne lourde, alors un nouveau réarrangement débute (Figure 21). Si elle peut s'apparier, elle est alors exprimée à la surface de la cellule et on parle alors de cellule B immature. Ces cellules expriment une IgM de surface.



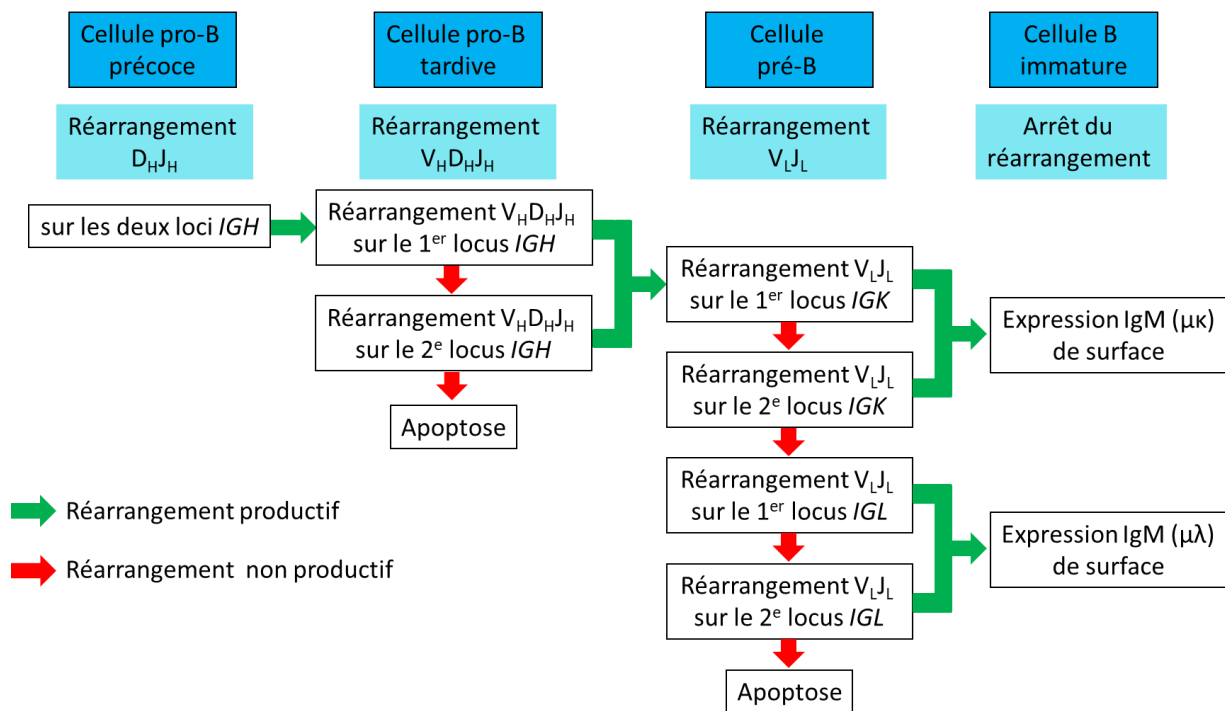


Figure 21 : Ordre de réarrangements des gènes des immunoglobulines au cours du développement lymphocytaire B précoce, adapté d'après Le système immunitaire, De Boeck¹⁰.

b- Phase dépendante de l'antigène : réponse T-dépendante

i- Les cellules B immatures

Les cellules B immatures produites par la moelle osseuse vont subir un nouveau processus de sélection afin de détecter les cellules dont le BCR reconnaît des antigènes du soi ou auto-antigènes : celles-ci vont alors soit mourir par apoptose, soit initier un nouveau réarrangement des parties variables lourdes ou légères (V(D)J) pour remplacer la chaîne initialement produite par une nouvelle^{2,38,39}. Cette modification de l'expression des chaînes variables est appelée édition du BCR (« receptor editing »). Cela permet d'éviter la génération de BCR auto-réactifs et d'améliorer la tolérance⁴⁰. Cette étape de sélection au niveau de la moelle osseuse est appelée tolérance centrale⁸. Les cellules B immatures sélectionnées par cette dernière étape deviennent des cellules B transitionnelles exprimant l'IgM de surface puis des cellules B matures qui vont quitter la moelle osseuse pour rejoindre le système immunitaire périphérique ou organes lymphoïdes secondaires. Ces cellules B matures co-expriment IgM et IgD (Figure 18)¹⁰.

Certaines cellules B auto-réactives peuvent survivre à cette étape de contrôle mais vont devenir anergiques, c'est-à-dire qu'elles ne répondront pas à tout autre stimulus antigénique⁴.



Elles peuvent également devenir ignorantes, c'est-à-dire que l'antigène du soi, pour lequel la cellule B a une affinité, ne sera pas détecté soit parce qu'il est séquestré ou en concentration faible, soit parce qu'il n'est pas capable d'induire une agrégation du BCR⁸. Cependant, ces cellules ignorantes peuvent devenir réactives dans certaines conditions (inflammation, augmentation de la concentration de l'auto-antigène notamment).

ii- Les cellules B matures

Les cellules B matures produites par la moelle osseuse vont transiter via le sang pour rejoindre les organes lymphoïdes secondaires. On parle de lymphocytes B naïfs, n'ayant rencontré que les antigènes du soi. Ces cellules vont former au niveau des organes lymphoïdes secondaires les follicules primaires, la zone du manteau ou des follicules secondaires.

Les follicules lymphoïdes primaires regroupent essentiellement des lymphocytes B, en contact avec des cellules stromales spécialisées appelées les cellules folliculaires dendritiques¹⁰. Ces dernières sont des cellules présentatrices d'antigène à l'état natif sous forme d'immun complexe. La cellule B naïve va rester transitoirement dans cette structure avant de rencontrer l'antigène. Lorsque la cellule B va rencontrer son antigène spécifique, elle va être retenue dans les zones thymo-dépendantes et être co-activée par les lymphocytes T CD4+ spécifiques du même antigène. Les lymphocytes B vont proliférer et continuer à se différencier.

Au cours de la réponse immune primaire, une partie des lymphocytes B activés va se différencier en plasmoblastes puis en plasmocytes. Ceux-ci sont spécialisés dans la production et la sécrétion d'anticorps. Les plasmoblastes sont un stade intermédiaire entre lymphocyte B activé et plasmocyte. La production d'anticorps débute alors mais la cellule continue à se diviser. Après quelques jours, la cellule peut mourir ou se différencier en plasmocytes⁸. Ces derniers vont perdre leur capacité à se diviser et leur durée de vie est de 4 semaines. A ce stade, les immunoglobulines de surface ne sont plus exprimées, entraînant une perte de la capacité de réponse face à l'antigène¹⁰. Cette première phase de différenciation plasmocytaire permet de produire rapidement des anticorps. Une partie de ces plasmocytes reste dans les organes lymphoïdes où leur durée de vie est courte, mais la majorité va migrer vers la moelle osseuse où ils vont pouvoir produire des anticorps de façon plus prolongée⁸.

Une autre partie des cellules B situées dans les follicules lymphoïdes primaires sont activées par l'antigène : on parle alors de follicules lymphoïdes secondaires avec un centre germinatif (Figure 22). Les cellules B résidentes au repos du follicule primaire sont repoussées en périphérie et forment la zone du manteau^{2,8}. Le centre germinatif va grandir avec la progression de la réponse immunitaire puis involuer lorsque l'infection est terminée (soit environ 3-4 semaines après le contact initial avec l'antigène)⁸. Les cellules B activées au niveau



du centre germinatif prennent le nom de centroblastes. Ces derniers, après avoir subi le processus d'hypermutation somatique des régions variables des gènes des immunoglobulines (obligatoire) et de commutation de classe (facultatif), mûrissent en centrocytes ayant perdu la capacité de se diviser¹⁰. Après quelques jours, deux zones différentes se dessinent au niveau histologique dans le centre germinatif : une zone sombre constituée essentiellement de centroblastes proliférants, exprimant faiblement l'immunoglobuline de surface (en particulier l'IgD) et assez peu de cellules folliculaires dendritiques, et une zone claire contenant principalement des centrocytes non proliférants, exprimant fortement l'immunoglobuline de surface, de nombreuses cellules folliculaires dendritiques et des lymphocytes Th (pour lymphocyte T helper)².

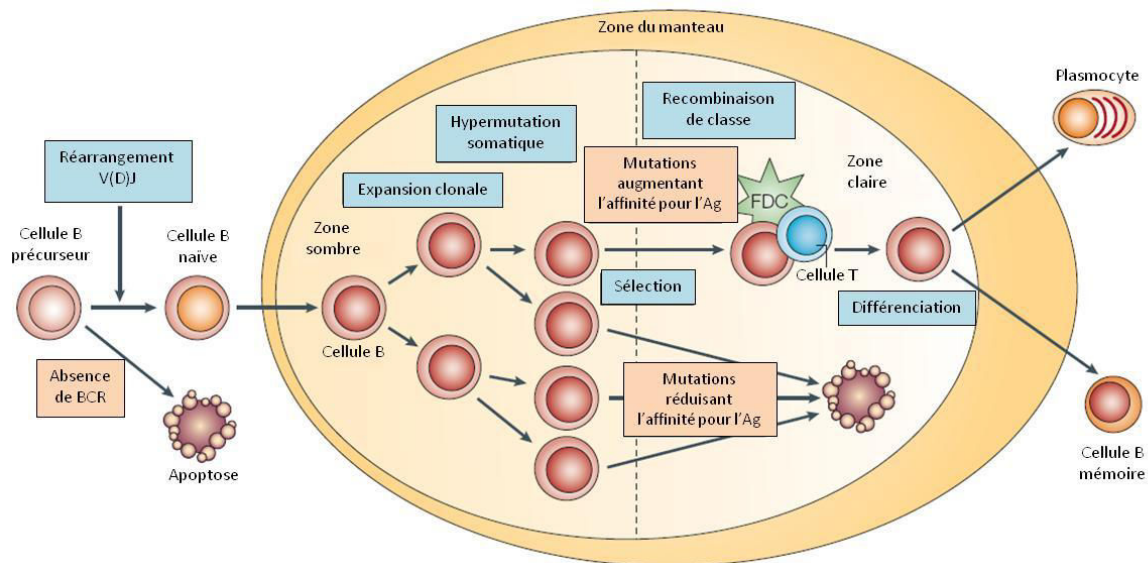


Figure 22 : Différenciation lymphocytaire B au niveau des centres germinatifs ganglionnaires, adapté d'après Küppers⁴¹. (FDC = Cellule Folliculaire Dendritique)

C'est donc au niveau du centre germinatif que les cellules B activées proliférantes vont subir le processus d'hypermutation somatique au niveau des gènes des régions variables des immunoglobulines avec une fréquence de l'ordre d'une paire de base pour 10^3 par division cellulaire (contre 1 pour 10^{10} dans les autres cellules). Ainsi chaque cellule fille va pouvoir acquérir de nouvelles mutations et il s'agit d'un processus par étape⁸.

Une sélection positive a lieu au niveau de la zone claire du centre germinatif ayant pour effet de sélectionner les cellules B exprimant un BCR de haute affinité pour l'antigène. Cette étape a lieu grâce à l'interaction avec les cellules folliculaires dendritiques qui présentent le même antigène sous forme d'immun complexe et les lymphocytes Th. De nombreuses cellules B vont acquérir des mutations défavorables et vont mourir par apoptose, puis être prises en charge par les macrophages présents dans les centres germinatifs : on parle alors de macrophages à corps tingibles à cause de la coloration des débris observés dans leur cytoplasme⁸. Les cellules B ayant acquis des mutations favorables vont alors se multiplier et pouvoir à nouveau subir



les processus d'hypermutation somatique puis de sélection. On parle de maturation d'affinité⁸. Plusieurs études ont montré que les cellules B du centre germinatif sont très mobiles et pourraient passer d'une zone à l'autre plusieurs fois^{42,43}.

Le processus de commutation de classe isotypique a lieu après activation par l'antigène et requiert la présence de cellules T et d'un certain nombre de cytokines⁸. La région constante utilisée dépend des cytokines produites par les cellules T auxiliaires et d'autres cellules immunitaires, opérant ainsi préférentiellement le switch vers un isotype d'immunoglobuline donné suivant l'environnement dans lequel se trouvent les cellules.

Les cellules B sélectionnées vont alors migrer vers d'autres zones du tissu lymphoïde secondaire ou dans la moelle osseuse. Les cellules B vont terminer leur différenciation en plasmocytes capables de sécréter des anticorps de haute affinité, issus de la commutation de classe isotypique. Les cellules B de haute affinité au niveau du centre germinatif peuvent également devenir des lymphocytes B mémoires de haute affinité. Ces cellules vont devenir quiescentes¹⁰. Elles ont une durée de vie plus longue et peuvent circuler dans l'organisme après stimulation dans les follicules lymphoïdes. Elles seront alors activées plus rapidement et pourront se différencier en plasmocytes, permettant la mise en place d'une réponse humorale plus rapide et plus puissante lors d'un contact ultérieur.

iii- Différenciation plasmocytaire

Les cellules B peuvent se différencier en plasmocytes grâce à la protéine régulatrice BLIMP-1 (pour B-Lymphocyte-Induced Maturation Protein 1). Celle-ci va inhiber dans les cellules B du centre germinatif la transcription des gènes impliqués dans la prolifération B, dans la commutation de classe et la maturation d'affinité⁸. Les cellules B engagées dans cette voie vont arrêter de proliférer et synthétiser et sécréter fortement l'immunoglobuline.

Plus précisément, BCL-6 (pour B-Cell Lymphoma 6) est fortement exprimé par les cellules du centre germinatif et inhibe BLIMP-1. La diminution de l'expression de BCL-6 va permettre l'expression de BLIMP-1 et l'engagement vers la différenciation plasmocytaire. Le facteur de transcription MITF (pour Microphthalmia-associated Transcription Factor) inhibe également la différenciation en plasmocyte⁴⁴. L'inhibition de MITF permet l'expression de BLIMP-1 et IRF4 (pour Interferon-Regulatory Factor 4) permettant la différenciation plasmocytaire. La levée d'inhibition de BLIMP-1 va entraîner la répression du facteur de transcription PAX-5, nécessaire au maintien en stades B, ce qui permet de lever son inhibition sur XBP1 (pour X-Box binding Protein 1). BLIMP-1 et XBP1 permettent l'expression de l'immunoglobuline. IRF4 active les enhanceurs au niveau des loci des chaînes légères. OBF1 (pour Octamer-binding transcription Factor (OCT)-Binding Factor 1) active un enhanceur du locus IGH et des promoteurs au niveau d'*IGK*⁴⁴.



iv- Différenciation en lymphocyte B mémoire

Les lymphocytes B mémoire ont une durée de vie longue. Leur capacité de division est très faible voire nulle. Ils expriment une immunoglobuline de surface mais ne sécrètent pas ou peu d'anticorps⁸. Ils vont permettre une réponse rapide avec prolifération en cas de second contact avec l'antigène⁴⁴.

Il existe des lymphocytes B mémoire de type IgM qui ont une capacité de prolifération plus faible et qui sont donc moins impliqués dans les réponses secondaires à l'antigène⁴⁵. Ces cellules sont IgM⁺, CD27⁺ et ont un profil mutationnel plus faible. Il s'agirait d'une population réservoir. Ainsi, en cas de réexposition à l'antigène, ces lymphocytes B mémoire IgM⁺ pourraient ré-entrer dans le centre germinatif et subir à nouveau le processus d'hypermutation somatique et la maturation d'affinité. Il existe également des lymphocytes B mémoire switchés, c'est-à-dire exprimant une IgG, IgA ou IgE, qui seront particulièrement efficaces dans les réponses immunes secondaires, mais dont le nombre va diminuer au cours du temps⁴⁵.

c- Réponse T-indépendante

La plupart des réponses immunitaires humores sont dépendantes des cellules T_H. Cependant, certains anticorps dirigés contre de nombreux antigènes bactériens peuvent être produits chez des sujets présentant des déficits en cellules T². Il existe deux classes d'antigènes T-indépendants (ou TI).

Les antigènes TI-1 entraînent une activation polyclonale de cellules B et peuvent être considérés comme des mitogènes des lymphocytes B. C'est le cas par exemple du Lipopolysaccharide capable de stimuler directement le BCR et par cosignal par liaison aux TLR (pour Toll-Like Receptor).

Les antigènes TI-2 sont caractérisés par des structures hautement répétitives, ce qui permet des interactions avec le BCR. Ces antigènes TI-2 sont typiquement des polysaccharides bactériens capsulaires. Il existerait des catégories spécifiques de cellules B impliquées principalement dans ces réponses T-indépendantes, à savoir les cellules B CD5⁺ et une fraction de cellules B de la zone marginale². On parle de cellules B-1. Celles-ci apparaissent de façon précoce au cours du développement. Cependant, chez le jeune enfant, la réponse aux antigènes polysaccharidiques n'est pas efficace avant l'âge de 5 ans. Les cellules B de la zone marginale apparaissent avec l'âge⁸. La réponse TI-2 est importante pour lutter contre les bactéries capsulées. En effet, la présence d'une capsule polysaccharidique permet d'échapper à la phagocytose et les cellules T ne pourront pas être activées par la présentation antigénique⁸.



B- Origine cellulaire des lymphomes B

D'une manière générale, les lymphomes sont séparés avec le lymphome de Hodgkin d'une part et les lymphomes non hodgkiniens (LNH) d'autre part. Parmi les LNH, le lignage B ou T permet de différencier les LNH-B, représentant 85 % des cas, des LNH-T². Dans ce travail, nous nous focaliserons sur les lymphomes B.

La classification OMS tend à différencier les différents types d'hémopathies selon leur contrepartie normale^{1,46}. Concernant les lymphomes B, la plupart correspondent à un stade de différenciation lymphocytaire donné (Figure 23). Cette approche est intéressante d'un point de vue physiopathologique. Cela permet d'identifier le stade cellulaire duquel dérive le clone malin et d'essayer d'élucider les mécanismes ayant entraîné la transformation de la cellule⁴¹. Ainsi, l'activation antigénique du BCR ou le microenvironnement peuvent jouer un rôle clé dans la transformation maligne. L'utilisation de techniques d'étude d'expression de gènes a permis de mettre en évidence des similitudes avec des stades de maturation donnés tels que les cellules B du centre germinatif ou les cellules B activées⁴¹. Par ailleurs, l'étude des mutations somatiques du gène de la chaîne lourde variable des immunoglobulines (*IGHV*) permet de différencier les lymphomes en 2 grands groupes, la présence de mutations somatiques étant un marqueur de transit dans le centre germinatif². Ainsi, un premier groupe correspond aux cas de lymphomes ne présentant pas d'hypermutation somatique du locus *IGHV* : c'est le cas des lymphomes pré-GC tels que la plupart des cas de lymphomes du manteau. Ce groupe comprend également les cas de lymphomes n'ayant pas ou peu subi l'hypermutation somatique au cours du passage dans le centre germinatif, ce qui est observé dans la leucémie lymphoïde chronique et le lymphome lymphocytaire². Le deuxième groupe comprend les cas de lymphomes pour lesquels une hypermutation du locus *IGHV* est mise en évidence. Dans ce cas, les lymphomes sont supposés dériver de lymphocytes au stade GC ou post-GC. Parmi ceux-ci, on retrouve notamment le lymphome de Burkitt, le lymphome folliculaire, ou le lymphome lymphoplasmocytaire².



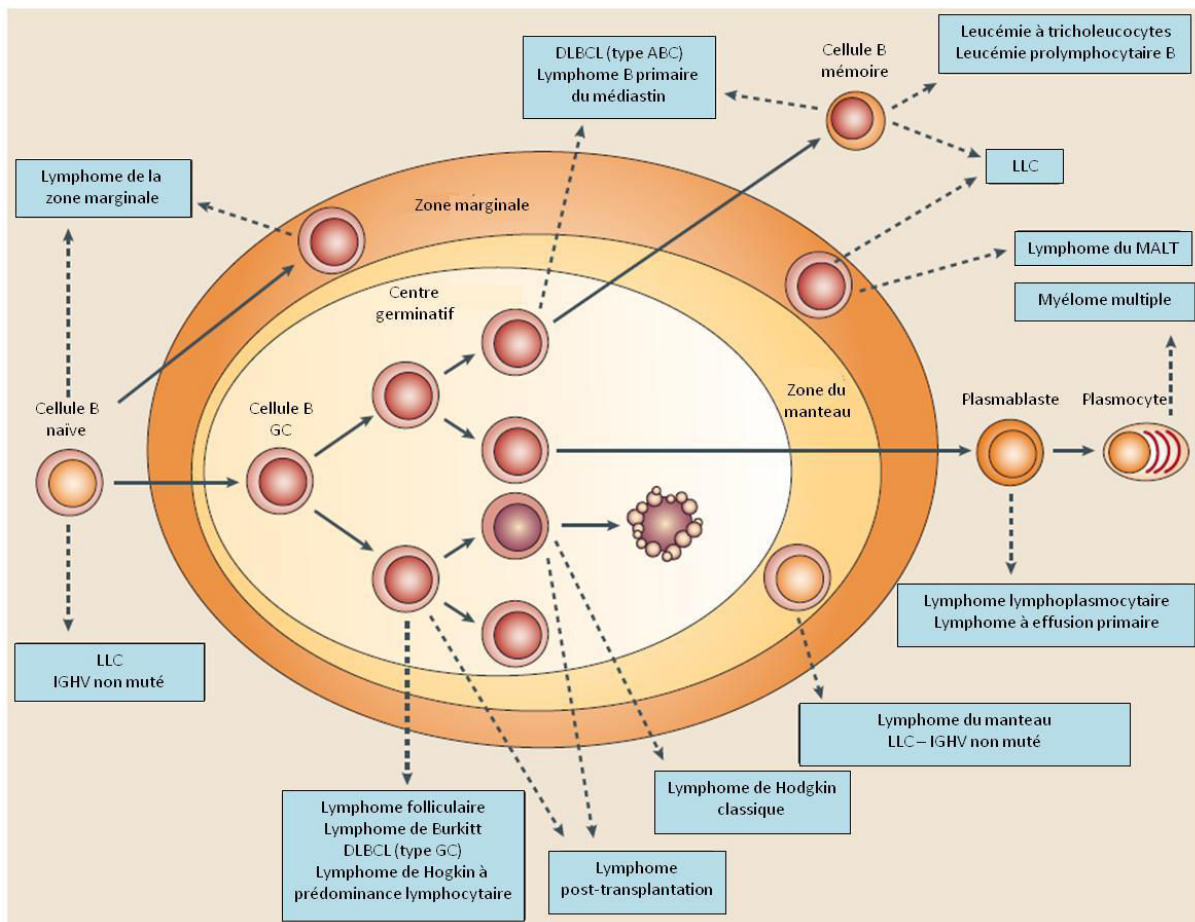


Figure 23 : Origine cellulaire des lymphomes B adapté d'après Küppers⁴¹.

Les flèches pleines indiquent la maturation lymphocytaire B normale. Les flèches en pointillés montrent les lymphomes B correspondant à un stade de maturation lymphocytaire donné. ABC = Activated B-Cell. DLBCL = Lymphome B diffus à grandes cellules. GC = centre germinatif. IGHV = gène variable de la chaîne lourde des immunoglobulines. LLC = Leucémie Lymphoïde Chronique. MALT = Mucosa Associated Lymphoid Tissue.

C- Voies d'oncogénèse dans les lymphomes B

Comme dans la plupart des cancers, la pathogénèse des lymphomes B est un processus multi-étapes impliquant l'accumulation de lésions génétiques affectant des proto-oncogènes et des gènes suppresseurs de tumeur². Cependant, le génome des lymphomes serait relativement stable selon les études cytogénétiques, comparé à celui de la plupart des tumeurs solides, et est caractérisé par quelques anomalies cytogénétiques récurrentes, en particulier les translocations chromosomiques².



On distingue les événements oncogéniques primaires tels que la translocation *IGH-BCL2* dans le lymphome folliculaire supposé survenir au cours du réarrangement VDJ dans la moelle osseuse. D'autres événements oncogéniques sont dits secondaires et apparaissent plus tardivement. C'est le cas par exemple des mutations de *TP53*.

Au niveau moléculaire, parmi les événements oncogéniques identifiés dans la pathogénèse des lymphomes B, on note l'activation d'oncogènes, soit par mutations introduites notamment lors du processus d'hypermutation somatique, soit par translocations réciproques, dont certaines sont caractéristiques d'un type de lymphome, ou l'inactivation de gènes suppresseurs de tumeurs par délétion chromosomique ou mutation². De plus, la lymphomagenèse peut comprendre l'introduction de gènes exogènes sous l'action de virus oncogéniques. Les translocations chromosomiques réciproques impliquent fréquemment les gènes des immunoglobulines (*IGH*, *IGK* ou *IGL*) et un proto-oncogène. Ce mécanisme permet de déréguler l'expression de l'oncogène qui est alors sous le contrôle d'éléments activateurs des gènes des immunoglobulines⁴¹.

1- Dérégulation oncogénique via translocation

Les translocations chromosomiques réciproques sont un mécanisme de dérégulation oncogénique bien connu dans les lymphomes B et constituent un événement primaire dans la lymphomagenèse⁴¹. Il s'agit généralement de translocations équilibrées, impliquant 2 chromosomes spécifiques, associés de façon récurrente dans certains types de lymphomes.

Il a été démontré que le processus de translocation est la conséquence d'erreurs se produisant soit sous l'action des enzymes RAG au cours du réarrangement VDJ, soit via AID lors de la recombinaison de classe au niveau d'*IGH* ou du processus d'hypermutation somatique des régions variables des gènes des immunoglobulines². En effet, ces processus génèrent des cassures double-brin de l'ADN, pouvant entraîner la formation de recombinaisons illégitimes à l'origine de ces translocations². Le processus de NHEJ permet la formation des jonctions dans ce type de translocation⁴⁷. Ceci entraîne la perte de quelques nucléotides, l'alignement de microhomologie terminale (de 0 à 3 pb) et l'insertion de nucléotides soit par la TdT, soit par les ADN polymérase μ ou λ .

Ces translocations impliquent typiquement un des loci des immunoglobulines et un proto-oncogène dont l'expression va alors être dérégulée lors de son rapprochement avec des éléments régulateurs du locus des immunoglobulines (Figure 24). Certaines translocations sont même une caractéristique de certains lymphomes. Par exemple, la translocation entre un locus d'immunoglobuline et le gène *MYC* dans le lymphome de Burkitt permet d'activer l'oncogène *MYC*. Il s'agit d'une caractéristique de ce type de lymphome, bien que non spécifique. De même, la translocation entre un locus d'immunoglobuline et le gène *CCND1* (pour Cycline D1) est une caractéristique du lymphome du manteau. Enfin, la translocation avec le gène *BCL2* (pour B-Cell Lymphoma 2) est retrouvée dans le lymphome folliculaire



notamment. Ce type de translocation peut être retrouvé à une faible fréquence dans les cellules B de sujets sains^{48,49,50}. Ceci suggère un processus de lymphomagénèse multi-étapes.

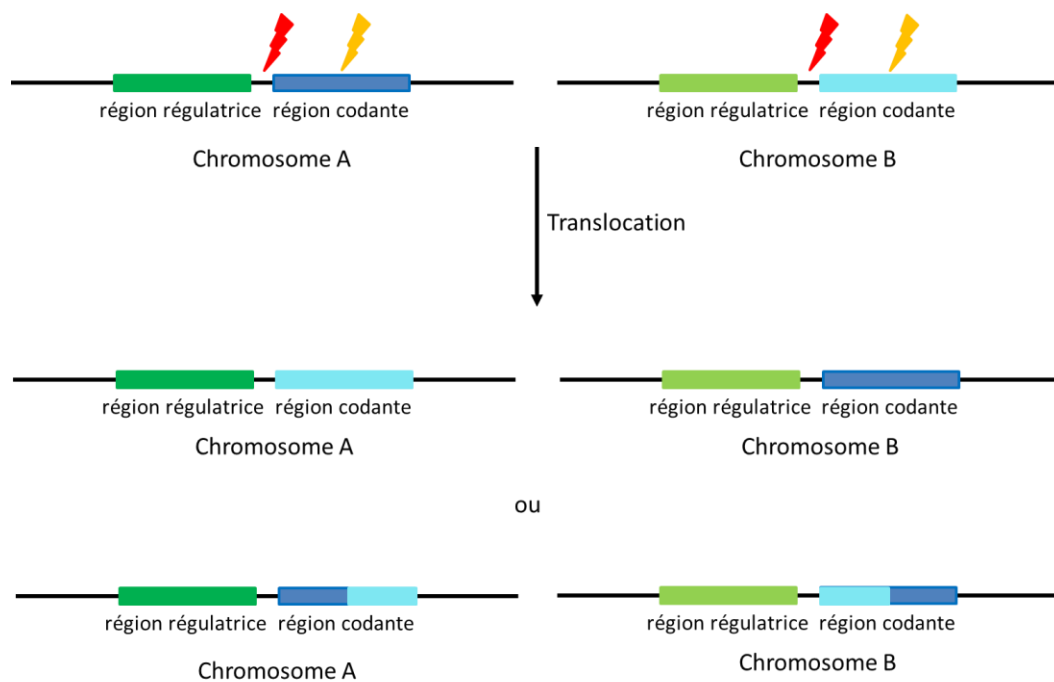


Figure 24 : Mécanismes de dérégulation génique via translocation dans les lymphomes B, adapté d'après Armitage *et al.*².

Les 2 grands mécanismes de dérégulation génique via translocation sont représentés. Dans le 1^{er} cas, la région régulatrice d'un gène situé sur le chromosome A va être rapprochée de la région codante du gène initialement présent sur le chromosome B, entraînant une dérégulation de l'expression de ce gène. Dans le 2^{ème} cas, la cassure a lieu au niveau de la région codante du gène et génère un gène de fusion, générant une protéine de fusion dont les propriétés biochimiques sont différentes de celles de la protéine native.

Un autre mécanisme de dérégulation via translocation consiste en la fusion des régions codantes de 2 gènes (Figure 24). Dans ce cas, une protéine chimérique est produite. Ce mécanisme est commun dans les leucémies aiguës et plus rare dans les lymphomes. Il est notamment retrouvé dans les lymphomes du MALT².



a- Translocations impliquant le gène des immunoglobulines

i- Points de cassure au niveau des gènes des immunoglobulines

Trois principaux types de cassures au niveau des gènes des immunoglobulines sont décrits.

Le premier point de cassure se situe au niveau des régions J ou D-J. Ce type de réarrangement entraîne fréquemment la perte ou l'addition de quelques nucléotides au niveau des extrémités D et J, ce qui suggère que ce type de translocations est dû à des erreurs au cours du mécanisme de recombinaison VDJ au niveau des stades précoces de la lymphopoïèse dans la moelle osseuse⁴¹. Ce cas est fréquemment retrouvé dans les translocations *IGH-BCL2* du lymphome folliculaire⁴¹.

Un autre type de translocations entraîne des points de cassure proches des segments VDJ réarrangés. Dans ce cas, les gènes des régions V présentent toujours des mutations somatiques⁴¹. Ces translocations pourraient donc se produire au moment du processus d'hypermutation somatique au cours duquel se produisent des cassures d'ADN.

Le troisième type de translocation entraîne des points de cassure au niveau des régions Switch des parties constantes des immunoglobulines. Des cassures de l'ADN se produisent physiologiquement à ce niveau au cours de la commutation de classe isotypique et ce type de translocation est donc supposé correspondre à une erreur au cours de ce processus⁴¹.

ii- Points de cassure au niveau du gène partenaire

Si les points de cassure au niveau des gènes des immunoglobulines, en particulier au niveau du locus *IGH*, sont bien décrits, l'origine de la cassure d'ADN au niveau de l'oncogène est en revanche moins claire. Il a été proposé que les régions concernées au niveau du gène partenaire pourraient contenir des séquences proches des RSS ou une structure similaire à un intermédiaire de recombinaison⁵¹.

Pour la translocation *IGH-BCL2*, la principale région de cassure au niveau de *BCL2* est confinée dans une région de 150 pb appelée la région majeure de cassure ou Mbr (pour Major breakpoint region)⁵². Il a été montré que le complexe RAG est capable d'aller provoquer des cassures au niveau de cette région y compris lorsqu'il n'y a pas de RSS. Ceci serait dû à la structure de cette région capable d'entraîner une stabilité à l'état simple brin⁵². Ainsi, le complexe RAG pourrait interagir à ce niveau et induire la recombinaison aberrante entre le locus VDJ et le Mbr, ce qui aboutirait à cette translocation. De plus, le complexe RAG possède



une activité de transposase^{53,54}. Un mécanisme de transposition pourrait expliquer de telles translocations ou entraîner l'insertion du fragment excisé au cours du processus VDJ.

Certains des gènes impliqués dans ces translocations impliquant les gènes des immunoglobulines subissent également le phénomène d'hypermutation somatique de façon aberrante, ce qui pourrait générer des cassures d'ADN⁴¹. C'est le cas par exemple pour le gène *BCL6* impliqué dans des translocations avec les loci des immunoglobulines⁴¹.

iii- Mécanismes et conséquences des translocations impliquant les gènes des immunoglobulines

Translocations résultant du processus de recombinaison V(D)J

De nombreuses translocations dans les lymphomes B présentent des caractéristiques structurales suggérant un rôle du processus de recombinaison V(D)J⁴⁸. Les deux exemples les plus typiques sont les translocations *IGH-BCL2* et *IGH-CCND1*. L'étude des points de cassure de ces dernières montre certaines particularités. En effet, pour les translocations *IGH-CCND1*, l'usage des JH est comparable aux cellules B normales, alors que dans les cas *IGH-BCL2* du lymphome folliculaire, les segments JH les plus en aval et les segments DH les plus en 5' sont préférentiellement impliqués^{48,55,56}. Les points de cassure de la translocation *MYC-IGH* du lymphome de Burkitt endémique sont également situés au niveau des régions JH et seraient liés à l'activité de RAG au cours de la recombinaison VDJ⁵⁷.

Ce type de translocation nécessiterait 3 cassures double brin : au niveau des segments DH et JH d'une part, et au niveau de l'oncogène d'autre part. Du fait de la perte de la portion entre les segments DH et JH impliqués, ce type de translocation n'est pas complètement équilibré⁴⁸.

Dans le cas de la translocation *IGH-BCL2*, les points de cassure sont proches de sites RSS au niveau de segments JH ou à l'extrémité 5' du RSS de segments D_HJ_H, ce qui implique que ce type de translocation se produit au moment où la cellule réalise le réarrangement DJ ou VDJ⁴⁸. Cette hypothèse est renforcée par le fait que des nucléotides ressemblant aux N-nucléotides sont ajoutés au niveau des points de cassure. Le fragment germinale entre le DH et le JH réarrangés étant délété, ces translocations ont probablement lieu au moment du processus de recombinaison D_HJ_H, au moment où les cassures double brin sont générées. De plus, le fait que le segment DH ait une séquence germinale est un argument supplémentaire, et en défaveur de l'hypothèse d'une translocation survenant sur un allèle ayant déjà opéré le réarrangement VDJ.

En parallèle, il avait été initialement proposé que les points de cassure au niveau de l'oncogène partenaire étaient liés au mécanisme de recombinaison V(D)J. Cependant, l'origine



de ce mécanisme est mal connu. En effet, des motifs ressemblant à des RSS ne sont généralement pas clairement identifiés au niveau des points de cassure, et le mécanisme ne serait alors pas guidé par ce type de séquence⁴⁸. Par ailleurs, des duplications de nucléotides sont fréquemment observées au niveau des points de cassure dans les gènes *BCL2* et *CCND1*. Ceci suggère des cassures de l'ADN décalées, ce qui n'est pas compatible avec un processus médié par RAG. De plus, au niveau des points de jonction *IGH-BCL2* et *IGH-CCND1* est observée l'insertion de nucléotides appelés les T-nucléotides (pour templated nucleotides)^{48,56}. Il s'agit de courtes copies de séquences JH, DH ou Mbr, pouvant être insérées directement ou inversées. Ces T-nucléotides ne sont généralement pas juste au niveau du point de cassure et présentent des mutations. Ceci suggère l'intervention d'un mécanisme de réparation de courtes séquences générant des erreurs^{48,55}.

Dans certaines translocations impliquant *IGH*, en particulier dans les leucémies aiguës lymphoblastiques T, il est rapporté des séquences appelées pseudo-signaux ou RSS cryptiques ou sites de mé-reconnaissance, qui correspondent à des régions de séquence proche de celle des RSS et qui fonctionnent de la même façon⁵¹. Des erreurs de recombinaison VDJ par les enzymes RAG peuvent également entraîner des délétions et des transpositions au niveau de gènes autres que les gènes des immunoglobulines ou du TCR⁴⁸. La séquence adjacente au point de cassure possède alors les caractéristiques heptamère/nonamère et le site de translocation est situé immédiatement en 5' d'un CAC (la séquence critique d'un heptamère)⁵¹. Cependant, ce type de séquence n'est généralement pas retrouvé au niveau du gène partenaire impliqué dans les translocations avec un gène des immunoglobulines.⁴⁷

Dans le cas des cassures au niveau de *BCL2*, un modèle basé sur la présence de séquences « Chi-like »^a au niveau des deux principales régions de cassures dans *BCL2* a été proposé^{48,58}. En effet, les translocations dans différents oncogènes se produisent à proximité d'une séquence consensus de 8 pb de type CC[A/T]CC[A/T]GC⁵⁸. En faveur de ce modèle, il a été démontré qu'une endonucléase clive spécifiquement au niveau du Mbr dans *BCL2* dans les cellules du lignage B^{48,56}. Une protéine de 45 kD (bp45) se lie à un motif Chi-like au niveau du Mbr (GGGAGGACGGGAGGAAGGCG). L'hypothèse est que cette protéine bp45 se lierait aux séquences Chi-like et servirait de molécule accessoire pour les nucléases endogènes. De plus, il existe des séquences Chi-like au niveau des segments JH et DH : la protéine bp45 pourrait jouer un rôle dans le rapprochement entre *IGH* et le gène partenaire et donc favoriser le mécanisme de translocation^{48,59}.

Il a également été démontré que RAG peut jouer un rôle de transposase, c'est-à-dire qu'il s'agit d'une enzyme capable de couper un fragment d'ADN au niveau de séquences de reconnaissance spécifiques et d'intégrer ce fragment au niveau d'une autre molécule d'ADN^{48,54,60}. Cette fonction des enzymes RAG permet d'intégrer les fragments d'ADN entourés de séquences RSS issus de la recombinaison VDJ au niveau d'une autre région d'ADN,

^a Les séquences Chi correspondent à une séquence de 8 nucléotides (octamère : CCACCAGC), décrite initialement chez *Escherichia coli*. Il s'agit de régions activatrices de recombinaison.



et potentiellement au niveau d'un oncogène. Dans ce cas, il s'agit d'une réaction de transesterification avec duplication de quelques paires de bases au niveau de l'insertion⁴⁸. Pour ce mécanisme, deux modèles sont proposés⁵⁴. Dans un premier modèle, la cassure en JH ou DH crée des extrémités franches capables de s'intégrer au niveau d'un autre fragment d'ADN. La structure branchée obtenue est alors à nouveau modifiée pour aboutir à une structure en épingle à cheveu et une jonction interchromosomique contenant le RSS⁵³. La translocation réciproque se produit lorsque les deux structures en épingle sont jointes (une au niveau du segment JH ou DH et une au niveau de l'autre chromosome) (Figure 25).

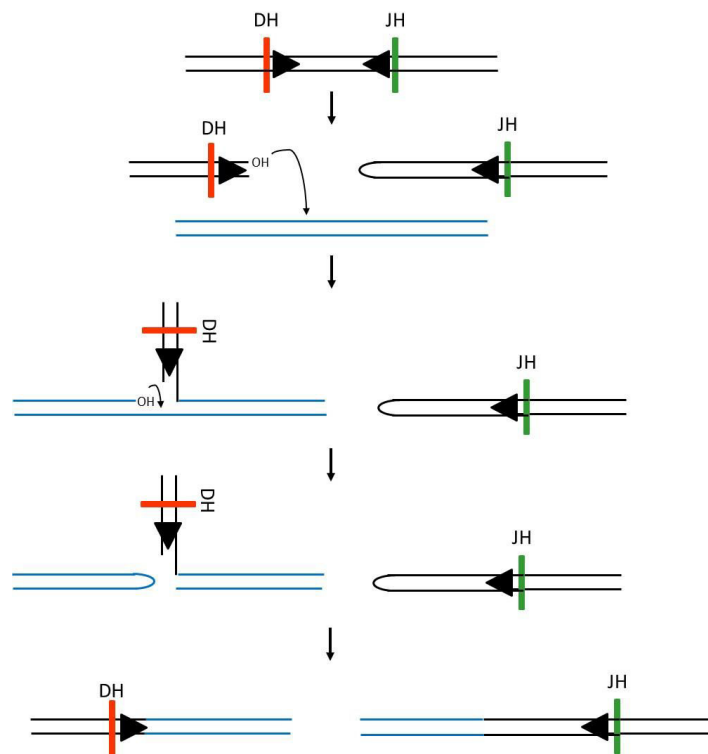


Figure 25 : Modèle de translocation via transposition médiée par RAG, faisant intervenir des extrémités franches, adapté d'après Roth et Craig⁵³.

Génération d'extrémités franches en VDJ (exemple en DH ici). La structure est alors capable de s'insérer sur un ADN cible (en bleu) pour générer une structure branchée. RAG coupe cette structure branchée et entraîne la formation de structures en épingles à cheveu, aboutissant à une translocation contenant alors le RSS au niveau de la jonction.

Dans ce cas, le RSS en 5' est retrouvé au niveau du site de translocation. Le deuxième modèle propose une insertion : les deux RSS génèrent des extrémités franches et permettent la transposition du fragment d'ADN excisé. Dans ce modèle, il est aussi possible d'obtenir une translocation lorsque la structure obtenue n'est pas insérée, via la formation d'une structure en épingle (Figure 26). Dans ce cas, un segment DH ou JH est retrouvé au niveau du site de translocation avec une large délétion du chromosome 14 lors du processus. La plupart des translocations *IGH-BCL2* et *IGH-CCND1* étudiées correspondent au second modèle⁴⁸.



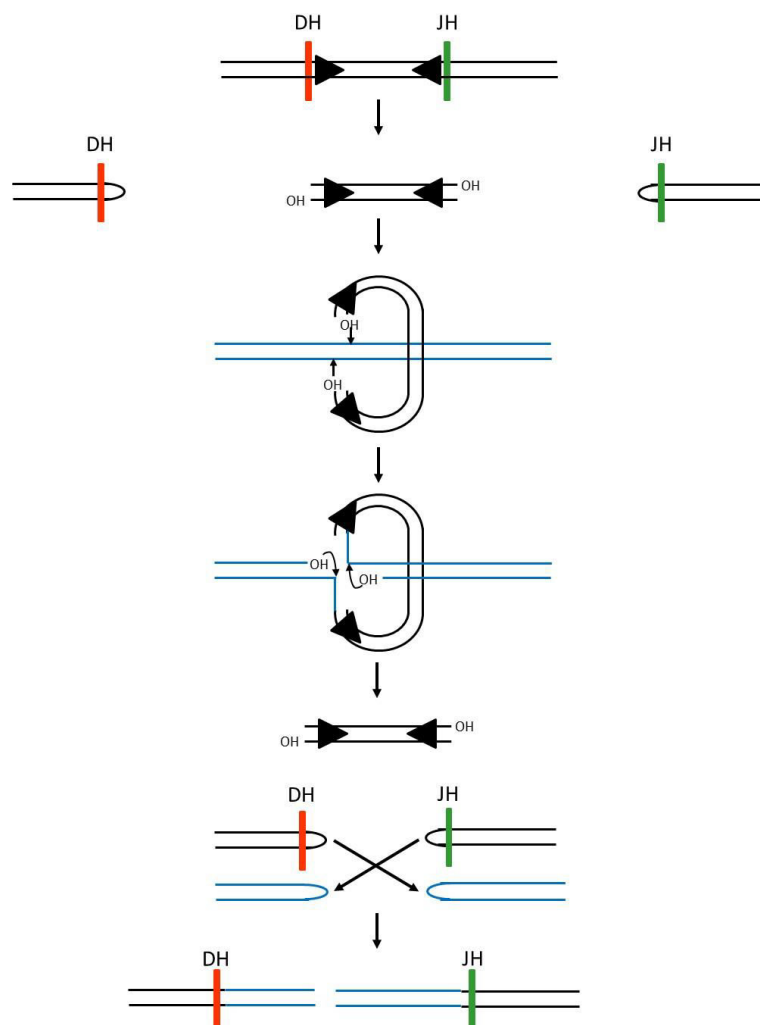


Figure 26 : Modèle de translocation par transposition via insertion médiée par RAG, adapté d'après Küppers et Dalla-Favera⁴⁸.

Le fragment excisé au cours de la recombinaison D_H - J_H est transposé au niveau d'un autre fragment d'ADN (traits bleus) et génère une structure branchée. RAG va à nouveau modifier cette structure, entraînant des cassures double brin au niveau de l'ADN cible, et des extrémités en épingle à cheveu. Enfin, le fragment d'ADN cible et les segments DH et JH sont joints, donnant lieu à une translocation.

En l'absence de RSS cryptique retrouvé, il a été proposé un rôle de la structure de l'ADN au niveau des points de cassure sur le gène partenaire⁵¹. Ainsi RAG2 pourrait être recrutée indirectement au niveau de régions chromatiniennes riches en histone 3 lysine 4 triméthylée (H3K4me3), RAG1 pourrait se lier directement aux histones, ou le complexe RAG pourrait interagir avec l'ADN de façon non spécifique⁶¹.



Un mécanisme mixte faisant intervenir une cassure double brin liée à la recombinaison V(D)J médiée par RAG au niveau du gène des immunoglobulines et une autre origine au niveau du gène partenaire est également possible⁴⁷. Une intervention d'AID au niveau du gène partenaire est susceptible d'être en cause. La co-expression de RAG et d'AID est possible à certains stades du développement. Au stade pro-pré-B, l'enzyme RAG est exprimée pour la génération de la recombinaison V(D)J et AID est également exprimée à un faible niveau comparé à celui des cellules du GC, mais ce serait suffisant pour la génération de translocations⁴⁷.

Enfin, la cassure au niveau de l'oncogène peut également être non spécifique et causée par des facteurs tels que les radiations ionisantes, des dommages causés par des radicaux libres oxygénés ou des réactions des topoisomérases II avortées⁴⁷.

La conséquence des translocations avec points de cassure au niveau des régions JH est une dérégulation liée à la présence de l'enhancer E_{μ} . La dérégulation a donc lieu au niveau du dérivé 14 (Figure 27).

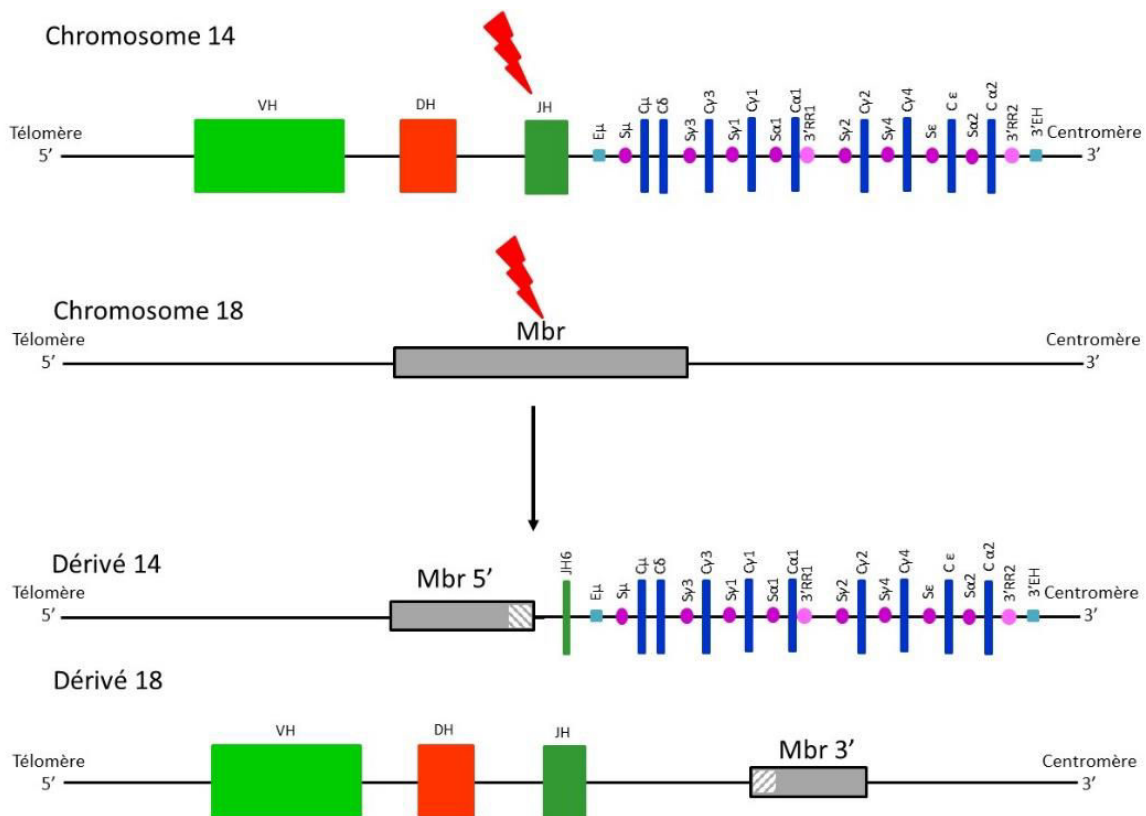


Figure 27 : Mécanisme de la translocation *IGH/BCL2* au cours du processus de recombinaison VDJ (exemple de cassure dans les régions JH pour *IGH* et au niveau de la *Mbr* pour *BCL2*). Les points de cassures sont représentés par les éclairs rouges au niveau des chromosomes 14 et 18.



Translocations résultant du processus de recombinaison de classe

De nombreuses translocations comprennent des points de cassure au niveau des régions switch d'*IGH*. C'est le cas par exemple pour la t(8;14) avec *MYC* dans le lymphome de Burkitt sporadique, la t(14;19) avec *BCL3* dans la leucémie lymphoïde chronique et la t(3;14) avec *BCL6* dans le DLBCL⁴⁸.

Du fait de la présence de points de cassure au niveau des régions Switch, ce type de translocation se produit probablement lors du processus de recombinaison de classe, au moment de la génération de cassure double brin de l'ADN. Généralement, les oncogènes impliqués sont alors situés sur le dérivé 14 et juxtaposés aux régions régulatrices 3'RR situées au niveau des parties constantes d'*IGH*, entraînant leur dérégulation (Figure 28)⁴⁸.

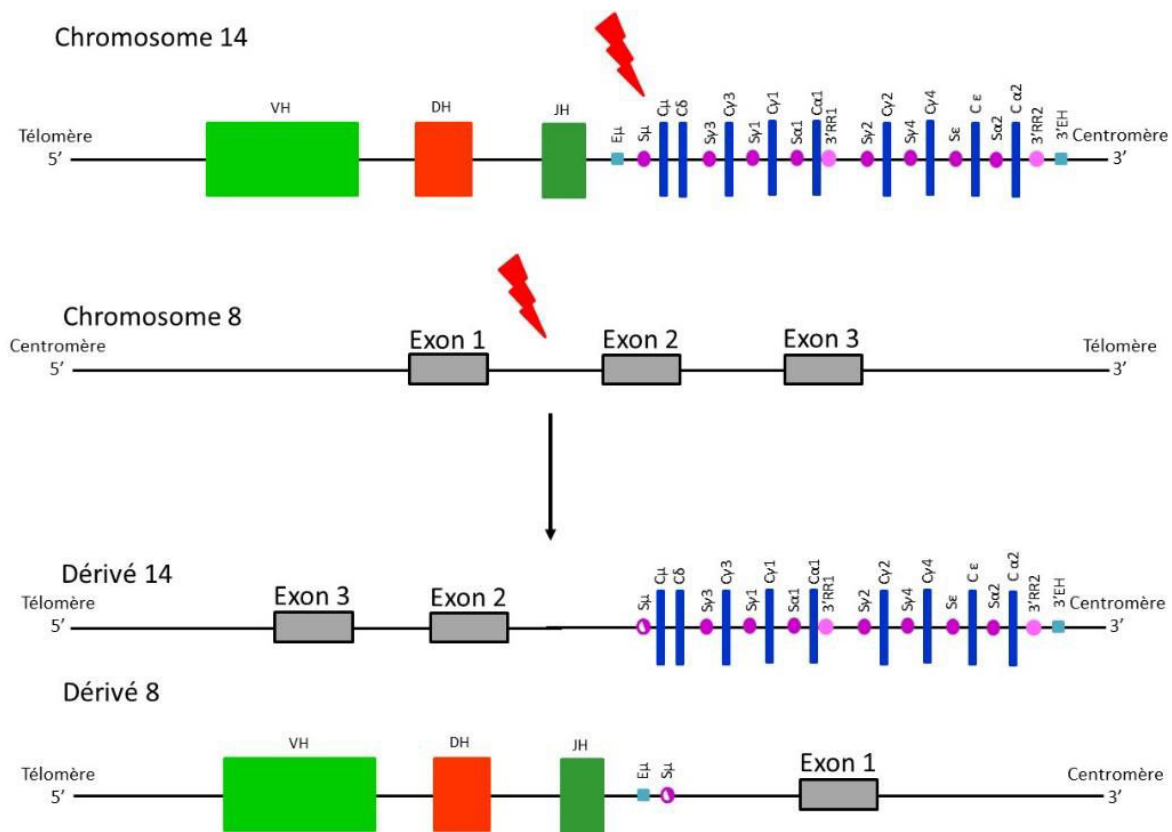


Figure 28 : Mécanisme de la translocation *IGH-MYC* faisant intervenir la recombinaison de classe, adapté d'après Blum *et al.*⁵⁷

Exemple de cassure au niveau de la région Sμ. Les points de cassures sont représentés par les éclairs rouges au niveau des chromosomes 14 et 8.



Cependant, ce type de translocation entraîne en parallèle la délocalisation du segment JH-intron $C\mu$ au niveau de partenaire de translocation et cela peut également entraîner une dérégulation de gènes au niveau de l'autre chromosome dérivé⁴⁸. C'est le cas par exemple pour la t(4;14) dans le myélome multiple⁶².

Bien que la recombinaison de classe entraîne deux cassures double-brin, ce type de translocation fait généralement intervenir une seule région switch. C'est le cas pour les translocations faisant intervenir $S\mu$. C'est également le cas pour de nombreuses translocations faisant intervenir $S\gamma$ ou $S\alpha$. Cela suggère qu'une structure dans le complexe de recombinaison pourrait favoriser l'interaction d'une seule région switch avec le chromosome partenaire au cours de la translocation⁴⁸.

De plus, il a été montré que les lymphocytes B sains peuvent effectuer une recombinaison entraînant une délétion de morceaux de régions $S\mu$ ⁴⁸. Il s'agirait d'un événement de stabilisation de l'isotype permettant de prévenir la recombinaison de classe. Ainsi, les translocations impliquant $S\mu$ au niveau des deux dérivés d'une translocation pourraient résulter de ce type d'événement et non d'une véritable recombinaison de classe.

La recombinaison de classe est également susceptible d'entraîner la dérégulation de gènes via des événements d'insertion⁴⁸. Ceci a été décrit dans un cas de myélome dans lequel une séquence délétée au cours de la recombinaison de classe a été insérée à proximité du gène *CCND1*. Comme ce fragment d'ADN contient l'enhancer $\alpha 1$, cette insertion entraîne la surexpression de la cycline D1⁶³.

La génération de translocations au cours du processus de recombinaison de classe pourrait être favorisée par la formation de boucles R (« R-loop »)^a ou par la présence de promoteurs entraînant des transcriptions divergentes ou convergentes, ceci favorisant la formation d'ADN simple brin⁴⁷. En effet, il a été démontré que les régions hautement actives sur le plan transcriptionnel sont des cibles d'AID dans des modèles murins^{64,65}.

Cependant, le mécanisme d'hypermutation somatique est également susceptible d'intervenir dans la génération de translocation impliquant les régions switch, par cassure au niveau du gène partenaire⁴⁸.

^a Boucle R ou « R-loop » : Séquences riches en G sur le brin non transcrit, générant un ARN riche en G pouvant s'hybrider sur le brin transcrit, tandis que l'ARN polymérase continue la transcription⁴⁷. Cela génère donc des hybrides ADN-ARN et une structure à 3 brins.



Rôle du processus d'hypermutation somatique dans les translocations et la lymphomagénèse

Le processus d'hypermutation somatique entraîne la modification de nucléotides, mais également des délétions et duplications, faisant ainsi intervenir des cassures double brin de l'ADN⁴⁸. Ceci a été prouvé au niveau des régions variables dans les cellules B^{12,66}. L'étude des points de cassure dans certaines translocations est en faveur de ce mécanisme. Dans le cas du lymphome de Burkitt, les points de cassures des translocations se situent fréquemment au niveau des régions variables réarrangées ou des introns des régions J en aval, soit les régions ciblées par l'hypermutation somatique⁴⁸. De plus, dans ces cas, des mutations somatiques ont été identifiées au niveau des régions variables réarrangées. Un autre élément important est l'absence de signal RSS identifiable à proximité des points de cassure. Dans ces cas, des motifs type CpG ou WGCW (ou W est un A ou un T) peuvent être retrouvés⁴⁷. Enfin, des mutations somatiques sont présentes en 5' des régions variables, alors que ces segments sont séparés de l'enhancer au cours de la translocation. La proximité de l'enhancer étant nécessaire pour l'hypermutation somatique, cela suggère que la translocation se produit donc après l'hypermutation du segment variable réarrangé⁴⁸.

Le processus d'hypermutation somatique touche d'autres gènes impliqués dans les translocations, tels que *BCL6*, impliqué dans les translocations *IGH/BCL6* dans les DLBCL.

Ce type de translocation entraîne le rapprochement de l'oncogène avec l'enhancer E μ sur le dérivé 14, entraînant sa dérégulation.

Autres mécanismes de translocation

Dans certains cas, aucun des 3 processus précédents ne semble compatible avec les points de cassure retrouvés. En particulier, cela concerne les translocations impliquant les gènes des chaînes légères des immunoglobulines dans lesquels les points de cassure se situent 2 à 5 kb en amont des segments J non réarrangés d'*IGK* ou *IGL*. Ceci a été décrit par exemple dans le lymphome de Burkitt⁶⁷. Dans ces cas, aucune séquence mimant un RSS ou un switch n'est retrouvé et de tels segments non réarrangés ne sont pas touchés par l'hypermutation somatique⁴⁸. Il est possible que la structure ouverte des gènes des immunoglobulines dans les cellules B puisse entraîner une augmentation du risque de cassure de l'ADN, indépendamment des processus de recombinaison et d'hypermutation somatique⁴⁸.

Les gènes des immunoglobulines sont aussi impliqués dans des translocation complexes impliquant parfois 3 partenaires. Ceci démontre l'instabilité génétique des loci des immunoglobulines dans les cellules B, et la variété des réarrangements aberrants pouvant être observés.



b- Translocations n'impliquant pas les gènes des immunoglobulines

Bien que la majorité des translocations observées dans les lymphomes B implique les gènes des immunoglobulines, il existe également des translocations impliquant d'autres gènes et pouvant entraîner des transcrits de fusion. Un exemple bien décrit est celui du lymphome du MALT dans lequel la $t(11;18)(q21;q21)$ est retrouvée dans environ 50 % des cas². Cette translocation implique les gènes *BIRC3* (pour baculoviral IAP repeat containing 3, anciennement *API2* pour Apoptosis Inhibitor 2, IAP signifiant Inhibitor of Apoptosis Protein) en 11q21 et *MALT1* (pour Mucosa-associated lymphoid tissue lymphoma translocation gene 1) en 18q21. *BIRC3* appartient à la famille des protéines inhibitrices de l'apoptose, tandis que *MALT1* induit la voie NFκB. Cette translocation entraîne la formation d'un transcrit de fusion, entraînant une activation du complexe NFκB, avec des effets anti-apoptotiques pouvant contribuer à la pathogénèse du lymphome du MALT².

Un autre exemple est celui du DLBCL dans lequel le gène *BCL6* (situé en 3q27) est fréquemment impliqué dans des translocations avec différents partenaires tels que *PIM1* (pour Pim proto-oncogene) en 6p21 ou *CIITA* (pour class II major histocompatibility complex transactivator) en 16p13^{2,68-70}. La protéine *BCL6* agit comme un potentiel répresseur transcriptionnel de séquences promotrices. Son expression est restreinte dans les cellules B au stade GC. Ces translocations entraînent la formation d'un transcrit de fusion dans lequel le promoteur et le premier exon non codant de *BCL6* sont remplacés par des séquences dérivées du gène partenaire. Le point commun entre ces promoteurs est un spectre plus large d'activité au cours du développement B par rapport au gène *BCL6* sauvage, en particulier dans les stades post-GC incluant les plasmocytes et immunoblastes². La translocation a pour effet de lever la sous-régulation normale de l'expression de *BCL6* associée à la différenciation en cellules post-GC.

Un dernier exemple est celui des gènes *PDL1* et *PDL2* (pour Programmed Death Ligand 1 et 2, situés en 9p24.1), fréquemment transloqués dans les lymphomes B primaires du médiastin⁷¹. Cela entraîne une augmentation de l'expression des *PDL1* ou *2* soit via l'échange de promoteurs hautement exprimés dans les cellules B, soit via la juxtaposition d'éléments enhancer de la même façon que pour les translocations impliquant les gènes des immunoglobulines⁷². Il a été montré que l'expression des *PDL1* et *2* modifie la réponse anti-tumorale dans plusieurs types de cancers⁷³. Certaines translocations entraîneraient un phénotype avantageux sur le plan immunologique dans un contexte d'activation des cellules T⁷². De plus, ces translocations pourraient également avoir comme deuxième effet d'annuler l'expression du gène partenaire ayant une fonction suppressive de tumeur. Enfin, certaines translocations pourraient entraîner la perte du site de liaison de micro-ARN, allonger la demi-vie du transcrit, et augmenter l'efficacité traductionnelle par la modification de la structure secondaire de l'ARN messenger⁷².



2- Mutations géniques

Le processus d'hypermutation somatique génère via AID des mutations au niveau des parties variables des immunoglobulines, mais peut aussi entraîner des mutations dans d'autres gènes tels que *BCL6* ou *CD95* (ou *FAS*)^{13,14,15}. Dans de rares cas, ces mutations peuvent contribuer à l'apparition de lymphomes. Par exemple, des mutations inactivatrices de *CD95* sont retrouvées dans environ 20 % des lymphomes B post-centre germinatif. *CD95* (ou *FAS*) est une protéine transmembranaire dont l'interaction avec son ligand va entraîner la mort cellulaire par apoptose. Son inactivation par mutation peut protéger la cellule contre cette voie d'apoptose⁴¹.

Une hypermutation somatique aberrante de *BCL6* et d'autres proto-oncogènes tels que *MYC* ou *PAX5* a été mise en évidence dans les lymphomes B diffus à grandes cellules (DLBCL)⁷⁴. Dans le cas de *MYC* par exemple, les mutations peuvent toucher aussi bien les régions non traduites que les régions codantes, entraînant des changements d'acides aminés, ayant de potentielles conséquences fonctionnelles².

Il peut également se produire des mutations ponctuelles indépendamment du processus d'hypermutation somatique². Ces mutations sont reconnues comme étant des événements transformants et peuvent toucher des gènes suppresseurs de tumeurs (mutations inactivatrices de *TP53* ou du gène codant pour IκBα (pour nuclear factor of kappa light polypeptide gene enhancer in B-cell inhibitor alpha)) ou des proto-oncogènes (mutations ponctuelles activatrices de *RAS*). Il a également été démontré que la voie NFκB peut être altérée dans les DLBCL via des mutations ponctuelles au niveau de gènes régulateurs positifs (*CARD11* par exemple) ou négatifs (*TNFAIP3* par exemple) du complexe NFκB, entraînant une activation constitutive de ses fonctions anti-apoptotiques et prolifératives^{2,75}.

Avec le développement du NGS, de nombreuses mutations ont pu être mises en évidence dans les lymphomes⁷⁶. Quatre grands types de mutations peuvent être distingués sur le plan fonctionnel :

- Modifications de la prolifération et de l'apoptose, parmi lesquelles des mutations gain de fonction ou perte de fonction, entraînant une dérégulation des voies de signalisation NFκB et/ou JAK-STAT. On note parmi celles-ci les mutations des gènes *CARD11*, *TNFAIP3*, *MYD88*, *CD79A/CD79B*, *PRDM1*, *STAT6* et *JAK2*.

- Modifications de la différenciation, comprenant également des mutations gain de fonction ou perte de fonction. Ce type de mutation entraîne un arrêt de différenciation à un stade spécifique ou entraîne une différenciation vers un lignage donné. Les gènes affectés codent des facteurs de transcription, parmi lesquels : *BCL6*, *NOTCH1*, *FOXO1*, *PAX5*, *IRF4*, *BTG1*, *ETS1* et *MEF2B*.

- Modifications épigénétiques. Ce type de mutation inclut des facteurs impliqués dans l'ajout ou la suppression de modifications post-traductionnelles au niveau de l'ADN et des



histones. Les gènes touchés sont par exemple *MLL2*, *EZH2*, *KDM2*, *EP300*, *CREBBP*, *JMJD2C*, *TET2* et *DNMT3A*. Ce type de mutation touche également les gènes impliqués dans l'épissage des ARN et l'export nucléaire. Par exemple, des mutations inactivatrices des acétyltransférases *EP300* et *CREBBP* empêchent l'activation via acétylation de *TP53*, contribuant à la lymphomagénèse⁷⁷.

- Modifications du microenvironnement et de la régulation immune, comprenant des mutations gain de fonction ou perte de fonction. Ce type de mutation peut entraîner des modifications d'interactions stromales ou supprimer la réponse immune en réaction aux cellules lymphomateuses. Les gènes impliqués sont notamment *B2M*, *CD58*, *CIITA*, *TNFRSF9*, *TNFRSF14*, *PDL1*, *PDL2* et *EPHA7*.

Ces différents types d'altérations peuvent être associés⁷⁶.

3- Autres mécanismes d'activation de proto-oncogènes

Il existe d'autres mécanismes d'activation de proto-oncogènes dans les lymphomes, tels que l'amplification génomique². Par exemple, dans les DLBCL peuvent être mis en évidence des amplifications du gène *cREL*. L'utilisation de techniques de cytogénétique moléculaire telles que la CGH array (pour Comparative Genomic Hybridization) ou la SNP array (pour Single Nucleotide Polymorphism) sont susceptibles de mettre en évidence de nouveaux loci concernés par l'amplification.

4- Inactivation de gènes suppresseurs de tumeur

L'inactivation de gènes suppresseurs de tumeurs peut survenir après mutation et/ou délétion. Le principal gène concerné est *TP53*. Comme dans les autres cancers, son inactivation provient fréquemment de la délétion d'un allèle et d'une mutation ponctuelle de l'autre allèle².

L'analyse cytogénétique des lymphomes révèle des délétions chromosomiques spécifiques récurrentes, suggérant la présence de gènes suppresseurs de tumeurs. Parmi elles, on retrouve la délétion du bras long du chromosome 6 (del(6q)), isolée ou non, généralement associée à un pronostic péjoratif². Il a été proposé que celle-ci entraîne la perte d'*HACE1* (pour HECT domain and Ankyrin repeat Containing E3 ubiquitin protein ligase 1), situé en 6q16 et potentiel gène suppresseur de tumeur⁷⁸. En effet, il est délété dans 40 % des LNH-B et une hyperméthylation de son promoteur est observée dans 60 % des cas. De plus, la sous-régulation de son expression dans des lignées cellulaires entraîne une diminution de l'apoptose et une accumulation en phase S et G2/M du cycle cellulaire. Le gène suppresseur de tumeur *FOXO3* (pour Forkheadbox O3) a également été proposé comme candidat⁷⁹. En



effet, il est situé dans la région minimale de délétion en 6q21 et son expression est significativement diminuée en cas de del(6q).

L'inactivation de gènes suppresseurs de tumeur peut aussi résulter de mécanismes épigénétiques, par hyperméthylation de régions promotrices, entraînant une inactivation transcriptionnelle. C'est le cas par exemple pour la perte d'expression de *p16* dans un certain nombre de LNH dont le lymphome folliculaire^{2,80}.

5- Transformation par les virus

Certains virus peuvent être impliqués dans la transformation des cellules B. En effet, les virus sont capables d'insérer leur génome dans leurs cellules cibles pouvant entraîner l'intégration d'un génome viral et/ou l'interruption de gènes endogènes².

Un exemple bien connu est celui de l'EBV (pour Epstein-Barr Virus) qui est retrouvé dans quasiment tous les cas de lymphome de Burkitt endémique, dans environ 40 % des lymphomes de Hodgkin classiques et dans de nombreux lymphomes post-transplantation⁴¹. La protéine LMP1 (pour Latent Membrane Protein 1) virale est capable d'activer la voie NFκB^{81,82}.

Le virus HHV8 (pour Human Herpes Virus 8) a également un pouvoir oncogénique via la protéine virale FLIP (pour FADD-like interleukine-1-β-converting enzyme-inhibitory protein) qui active le facteur de transcription NFκB⁸³.

D'autres agents pathogènes tels que le VIH (pour Virus de l'Immunodéficience Humaine), le VHC (pour Virus de l'Hépatite C) ont un rôle indirect sur le développement de LNH en altérant le système immunitaire et/ou en entraînant une stimulation antigénique chronique². Il a par exemple été montré que l'incidence des lymphomes liés à l'EBV chez les patients infectés par le VIH est élevée et que ceux-ci sont typiquement agressifs^{2,84}. Les lymphocytes T cytotoxiques spécifiques de l'EBV sont quantitativement normaux, mais ne fonctionnent pas correctement. En effet, ils perdent leur capacité à produire de l'interféron-γ en réponse aux peptides de l'EBV⁸⁵.



6- Activation du BCR

Les cellules B normales sont sélectionnées sur la présence d'un BCR fonctionnel et non autoréactif. Il semble que dans le cas des lymphomes B, la présence d'un BCR soit nécessaire pour la survie de la cellule. Ceci peut dépendre soit d'un signal autonome produit par le BCR, soit par une activation du BCR par l'antigène⁴¹. Dans la plupart des lymphomes B, le BCR est exprimé, même si parfois à des niveaux très bas comme dans la LLC. L'hypothèse selon laquelle un signal dépendant du BCR est nécessaire à la survie de la cellule B lymphomateuse est soutenue par le fait que les translocations impliquant les gènes des immunoglobulines affectent quasiment toujours l'allèle non productif⁴¹.

Plusieurs types de lymphomes dont le lymphome folliculaire et le lymphome de Burkitt, montrent l'apparition de mutations somatiques dites « ongoing », c'est-à-dire en cours au sein de la tumeur, sans pour autant perdre l'expression du BCR⁴¹. Pourtant, ces mutations entraînent dans environ 10 % des cas la perte de fonctionnalité du BCR par mutation, délétion ou duplication entraînant un décalage du cadre de lecture, dans des conditions sans sélection antigénique^{11,41}. Par conséquent, la persistance de l'expression d'un BCR dans ce type de lymphome suggère que ces cellules subissent une sélection. Ainsi, des signaux de survie seraient produits par l'expression du BCR de même que dans les cellules B normales⁴¹.

7- Inactivation du BCR

Dans de rares cas, les cellules du lymphome n'expriment pas de BCR. C'est le cas par exemple dans le lymphome de Hodgkin, le lymphome B post-transplantation, le lymphome à effusion primaire et le lymphome B médiastinal⁴¹.

Dans le lymphome de Hodgkin classique, des mutations inactivatrices des régions IGHV transforment des réarrangements des segments variables initialement fonctionnels en non fonctionnels dans environ 25 % des cas⁸⁶. Le nombre de mutations inactivatrices apparaissant dans les cellules B du centre germinatif est faible, ce qui suggère que les cellules de Hodgkin et de Reed-Sternberg dériveraient de cellules B du centre germinatif pré-apoptotiques⁴¹. Le mécanisme par lequel ces cellules arrivent à échapper à l'apoptose pourrait être expliqué par l'infection à l'EBV retrouvée dans environ 40 % des cas de Hodgkin classique. En effet les cellules infectées vont exprimer les protéines LMP1 et LMP2A (pour Latent Membrane Protein 2A). LMP1 est la principale protéine oncogénique de l'EBV. Elle mime la signalisation du CD40 et induit des voies de croissance et de survie, dont la voie NF- κ B.⁸⁷ LMP2A possède un domaine ITAM également retrouvé sur le CD79A et B au niveau du BCR. Il a été montré à partir de modèles murins transgéniques que l'expression de LMP2A peut remplacer la voie de signalisation médiée par le BCR⁸⁸. Ainsi, quand une cellule B du centre germinatif infectée par l'EBV va subir des mutations délétères pour le BCR, l'expression de LMP2A pourrait permettre



à la cellule d'échapper à l'apoptose. D'autres événements oncogéniques sont ensuite nécessaires pour permettre l'émergence du clone malin.

Dans les lymphomes B diffus à grandes cellules médiastinaux, l'expression du BCR est indétectable et les éléments de la cascade de signalisation du BCR sont régulés négativement⁴¹. L'activité de l'enhancer IgH est également diminuée, sans pour autant que des mutations inactivatrices des régions IGHV aient été mises en évidence⁸⁹. De même dans le lymphome à effusion primaire, l'expression du BCR est généralement faible à indétectable, sans mutation inactivatrice au niveau d'IGHV mise en évidence⁴¹. Dans ce type de lymphome, l'expression des facteurs de transcription PU.1, OCT2 (pour Organic Cation Transporter 2) et BOB1 (pour B-cell specific Octamer Protein 1), qui sont importants pour la régulation de la transcription de l'immunoglobuline, est diminuée⁹⁰. Ce type de lymphome perd généralement l'expression d'autres marqueurs B tels que le CD19 ou le CD20. Ceci indique une perte du phénotype B, de la même façon que dans le lymphome de Hodgkin classique. Cependant, d'autres aspects phénotypiques sont aussi compatibles avec une différenciation plasmablastique^{41,91}.

8- Activation par l'antigène

La stimulation par la liaison à l'antigène pourrait contribuer à la prolifération et la survie des cellules lymphomateuses⁴¹. Par exemple, dans la leucémie lymphoïde chronique (LLC), il a été montré que le BCR reconnaît fréquemment des auto-antigènes^{92,93}. Dans d'autres cas, les cellules de la LLC ont une spécificité pour des antigènes étrangers comme des protéines virales comme par exemple la protéine p24 d'HTLV1 (pour Human T-Lymphotropic Virus type 1)⁹⁴. L'étude des réarrangements des régions variables lourdes et légères dans la LLC a permis d'identifier plusieurs sous-groupes présentant des BCR hautement similaires. On parle de BCR stéréotypé. Il est possible que les cellules lymphomateuses dérivent d'un sous-groupe de cellules B non reconnues par les systèmes de contrôles et présentant un répertoire de segments V restreint, mais il est plus probable que cela soit le reflet d'une sélection et activation de ces cas par des types d'antigènes restreints^{95,96,97}. Cet aspect sera détaillé plus en détails dans le chapitre sur la LLC (Chapitre 2 : la leucémie lymphoïde chronique, page 79).

Un répertoire restreint des régions variables des immunoglobulines est une caractéristique du lymphome primitif du système nerveux central^{41,98}. Dans ce cas, seul le segment VH est concerné, ce qui indique que la liaison de l'antigène se produit exclusivement au niveau VH. Il s'agirait donc d'un superantigène, c'est-à-dire que le même antigène est capable d'interagir avec plusieurs immunoglobulines différentes. La stimulation de la cellule B par un superantigène pourrait donc être impliquée dans la pathogénèse de ce type de lymphome⁴¹.

Une activation du BCR par l'antigène serait aussi impliquée dans la pathogénèse du lymphome folliculaire⁴¹. De nombreux patients présentent des cellules lymphomateuses avec BCR autoréactif⁹⁹. De plus, elles présentent des mutations « ongoing » c'est-à-dire des mutations



successives provenant du processus d'hypermutation somatique, et ces mutations indiquent une sélection par l'antigène¹⁰⁰. Enfin, environ 80 % des cas de lymphomes folliculaires présentent des mutations somatiques des régions variables entraînant la formation de motifs liant les carbohydrates, dont la forte sélection suggère une implication dans la genèse du lymphome¹⁰¹.

Dans le cas de lymphomes B associés au VHC, il a été démontré que des antigènes viraux sont en cause dans la lymphomagénèse⁴¹. En effet, le BCR des cellules lymphomateuses interagit directement avec une protéine de l'enveloppe virale¹⁰². De plus, le traitement antiviral par interféron-alpha permet également une régression des cas de lymphomes B spléniques associés au VHC, alors que ce traitement n'entraîne pas de régression tumorale chez des patients atteints de lymphome B splénique VHC négatifs¹⁰³.

La stimulation des cellules de lymphome B par des antigènes du non soi peut également être indirecte via les lymphocytes T helper⁴¹. Ce serait le cas dans le lymphome du MALT gastrique associé à *Helicobacter pylori*. Dans ce cas, ce sont les lymphocytes T helper qui reconnaissent les antigènes bactériens et qui stimulent la prolifération des cellules lymphomateuses¹⁰⁴. Dans certains cas de lymphomes du MALT gastrique et salivaire, des auto-antigènes seraient reconnus par les cellules lymphomateuses, suggérant une potentielle synergie entre antigènes du non soi et auto-antigènes dans la lymphomagénèse^{41,105}.

9- Rôle du microenvironnement

Le microenvironnement tumoral semble important pour la survie et/ou la prolifération des cellules tumorales⁴¹. Dans le lymphome folliculaire, les lymphocytes tumoraux prolifèrent dans les centres germinatifs et sont en contact étroit avec les lymphocytes T helper et les cellules folliculaires dendritiques, comme dans un centre germinatif normal⁴¹. Certains lymphocytes tumoraux peuvent être retrouvés dans les régions interfolliculaires mais ils ont alors un faible index prolifératif, suggérant la nécessité du contact avec le microenvironnement cellulaire du centre germinatif pour proliférer¹⁰⁶. De plus, *in vitro*, ces cellules ne peuvent être cultivées qu'en présence de cellules T CD4+ ou de cellules stromales, avec un anticorps dirigé contre le récepteur CD40, dont l'activation dans les cellules du centre germinatif entraîne des signaux de survie^{107,108}.

Dans le cas de la LLC, les cellules tumorales circulantes ont un index prolifératif faible⁴¹. Les lymphocytes malins prolifèrent dans les ganglions et la moelle osseuse au sein de centres prolifératifs où ils sont en contact avec les lymphocytes T CD4+ et les cellules dendritiques.

Les cellules tumorales du lymphome de Hodgkin sont rares (< 1 %) et l'infiltrat cellulaire est principalement composé de lymphocytes T, de polynucléaires éosinophiles, de macrophages, de cellules B et de plasmocytes. Bien que cet infiltrat puisse traduire une réponse inflammatoire inefficace contre le lymphome, il semble qu'au moins une partie de ces cellules



soient attirées par les cellules tumorales. C'est le cas des lymphocytes T CD4+ recrutés par la production de la cytokine TARC (pour Thymus and Activation-Regulated Chemokine), normalement produite par les cellules dendritiques⁴¹. De plus, les cellules de Hodgkin expriment des molécules importantes pour l'interaction avec les lymphocytes T helper^{41,109}. Ceci semble indiquer un rôle de ces dernières dans la lymphomagénèse.

10- Rôle des ARN longs non codants

Les ARNs longs non codants (lncRNAs) constituent un groupe hétérogène d'ARNs de longueur supérieure à 200 nucléotides et représentant plus de la moitié du transcriptome non codant chez les mammifères¹¹⁰. Actuellement, plus de 100 000 lncRNAs sont répertoriés chez l'Homme, de nombreux loci entraînant la formation de différents transcrits¹¹⁰⁻¹¹³. Leur expression est régulée au cours du développement et peut être tissu- et cellule-spécifique. Ils permettent la régulation de l'expression génique aux niveaux transcriptionnel, post-transcriptionnel, ainsi que via la modification de la chromatine. Ils sont capables de recruter des complexes modifiant la chromatine, modulant la structure de l'hétérochromatine qui réprime l'expression génique. Ils peuvent également activer la transcription en renforçant des régions enhancer ou en induisant des modifications de conformation tri-dimensionnelle de la chromatine¹¹⁰. D'autres lncRNAs peuvent interagir avec des facteurs de transcription ou des protéines se liant à l'ARN pour réguler indirectement la transcription. De plus, ils peuvent entrer en compétition avec les microARNs (miARNs) en se liant aux ARNm, altérant ainsi indirectement l'expression génique. Enfin, les lncRNAs peuvent moduler les événements de transformation des ARNs tels que l'épissage, l'édition, la localisation, la traduction et la dégradation.

L'expression des lncRNAs est impliquée dans les processus d'oncogénèse ou de suppression de tumeur¹¹⁰. Des études montrent que la dérégulation de différents lncRNAs favorise la cancérogénèse, la progression de la maladie et les métastases. Dans les cancers solides, il a été démontré que les lncRNAs contribuent à toutes les caractéristiques du cancer, telles que la viabilité, la prolifération, la mobilité et l'angiogénèse¹¹⁴. Un exemple est celui de MALAT1 (pour Metastasis-Associated in Lung Adenocarcinoma Transcript 1), fortement exprimé dans différents types de tumeurs dont le cancer pulmonaire non à petites cellules et dont le rôle dans la migration cellulaire et la croissance tumorale a été démontré *in vitro*¹¹⁵. De plus en plus d'éléments suggèrent que les lncRNAs sont des acteurs importants dans l'hématopoïèse normale et maligne¹¹⁰. Une équipe a ainsi mis en évidence l'expression différentielle des lncRNAs au cours du développement des cellules B¹¹⁶. Par exemple, MYB-AS1 (pour MYB AntiSens RNA1, situé en 6q23.3), SMAD1-AS1 (pour Smad family member 1 AntiSens RNA1, situé en 4q31.21) et LEF1-AS1 (pour Lymphoid Enhancer binding Factor 1-Antisens RNA1, situé en 4q25) sont impliqués dans le développement lymphoïde B précoce. Il s'agit de transcrits antisens de facteurs de transcription impliqués dans les cellules B précoces (Figure 29)^{110,116}. RP11-132N15.3 (aussi appelé lnc-BCL6.3, situé en 3q27.3, environ 240 kb en amont de *BCL6*)



est exprimé dans les centroblastes et dans une moindre mesure dans les centrocytes. Il joue un rôle dans la modulation de la réaction GC et pourrait agir en régulant le répresseur transcriptionnel BCL6^{110,116}. De plus, l'étude du transcriptome dans les différents sous-types de cellules B au cours du développement montre que l'expression des lncRNAs est souvent restreinte à un subset, de façon plus spécifique que pour les gènes codant des protéines ou des récepteurs de membrane¹¹⁷. Les lncRNAs les plus spécifiques d'un type cellulaire n'étaient pas retrouvés dans des prélèvements de tissu lymphoïde entier ou de cellules mononucléées sanguines, probablement à cause de leur dilution parmi un grand nombre de cellules hétérogènes¹¹⁷. Il a également été démontré que l'expression différentielle des lncRNAs permet d'identifier l'engagement dès le stade de progéniteur vers les lignages lymphoïdes B ou T¹¹⁸.

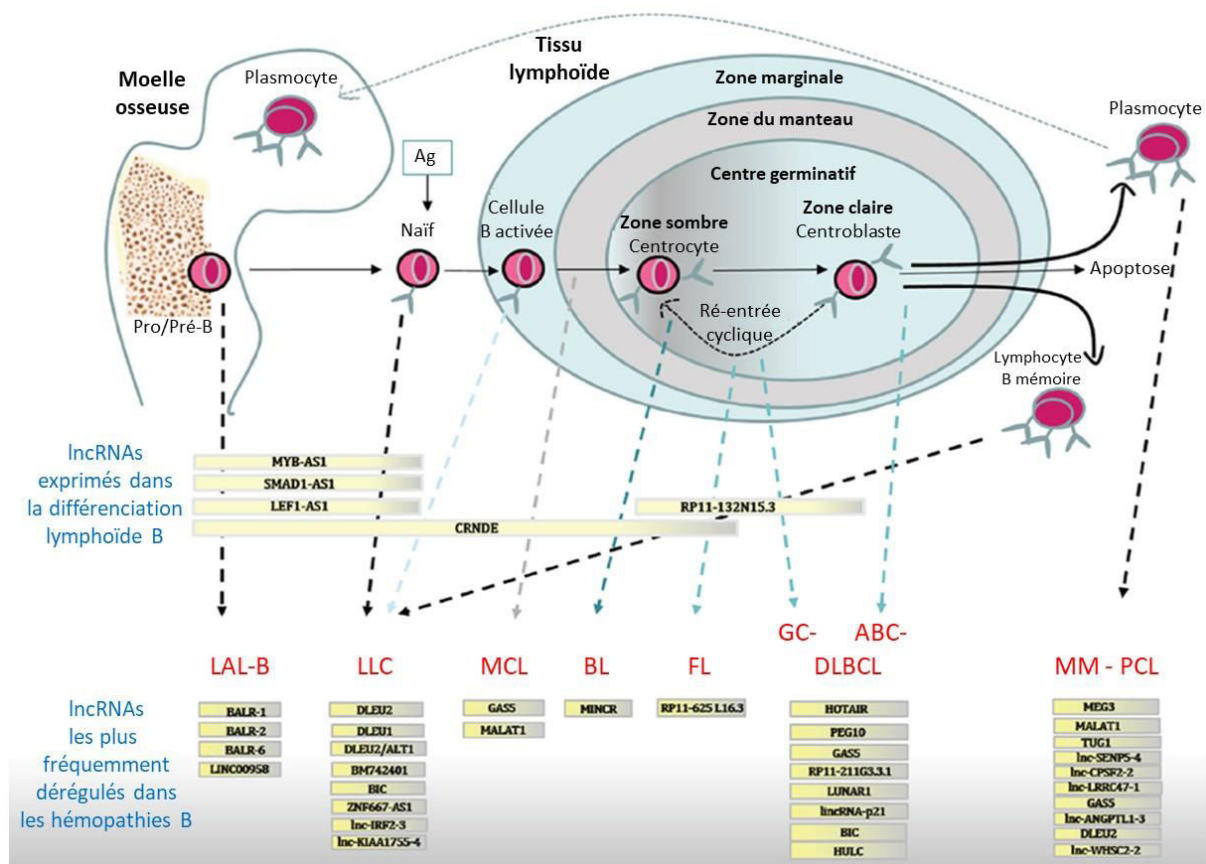


Figure 29 : lncRNAs exprimés au cours de la différenciation B et dans les hémopathies lymphoïdes B, adapté d'après Nobili *et al.*¹¹⁰.

LAL-B = Leucémie Aiguë Lymphoblastique B, MCL = Lymphome du manteau, BL = Lymphome de Burkitt, FL = lymphome folliculaire, MM = myélome multiple, PCL = Plasma Cell Leukemia (Leucémie à plasmocytes)



Actuellement, le rôle des lncRNAs dans les cancers est démontré et ils sont étudiés dans les hémopathies B. Certains sont déjà identifiés comme dérégulés dans les lymphomes B et leur nombre pourrait augmenter de façon exponentielle au cours des prochaines années¹¹⁰. Plusieurs équipes se sont intéressées à l'expression des lncRNAs dans différents types de lymphomes (Figure 29). Ceci sera développé dans le Chapitre 2 : la leucémie lymphoïde chronique (voir page 82 concernant la physiopathologie).

Les lncRNA peuvent également être dérégulés en cas de translocations. Un exemple dans le DLBCL est le lncRNA GAS5 (pour Growth Arrest Specific 5 situé en 1q25.1) retrouvé impliqué dans une translocation avec le gène *BCL6*¹¹⁹. Ce lncRNA joue un rôle dans l'arrêt du cycle cellulaire et l'apoptose avec une activité suppressive de tumeur¹¹⁰. La translocation décrite entraîne la formation d'un transcrite de fusion *GAS5-BCL6* exprimé à partir du dérivé 3 de façon préférentielle par rapport à *BCL6* de l'autre chromosome 3. Dans le lymphome du manteau, des essais ont été réalisés sur des lignées cellulaires en sous régulant l'expression de *GAS5*, connu comme étant un effecteur de la voie mTOR (pour mammalian Target Of Rapamycin). Ceci entraîne une diminution des effets des antagonistes de la mTOR sur la viabilité cellulaire, la synthèse d'ADN et la capacité à former des colonies^{110,120}. Ainsi, la modulation de l'expression de *GAS5* pourrait constituer une nouvelle cible thérapeutique¹¹⁰.

11- Modèles de lymphomagenèse

Plusieurs modèles de lymphomagenèse peuvent être proposés selon le stade auquel les altérations vont se produire (Figure 30)⁷⁶. Dans un premier modèle, les différentes altérations se produisent à un stade donné unique de différenciation lymphocytaire. Dans un deuxième modèle, les cellules acquièrent des modifications initiales lors des stades précoces de la différenciation lymphoïde qui peuvent être à l'origine d'une expansion clonale ou d'une résistance à l'apoptose, puis ces cellules vont développer des modifications additionnelles à des stades de différenciations plus tardifs. Dans le troisième modèle, les altérations initiales apparaissent au niveau de progéniteurs non engagés dans une lignée et ayant des capacités d'auto-renouvellement de façon inhérente ou par acquisition via altérations. Comme dans le deuxième modèle, ces cellules vont acquérir des altérations additionnelles au cours de la différenciation lymphoïde. Au cours des stades de développement ultérieurs, les cellules peuvent conserver des capacités d'auto-renouvellement et ainsi former des réservoirs pour diversification clonale.



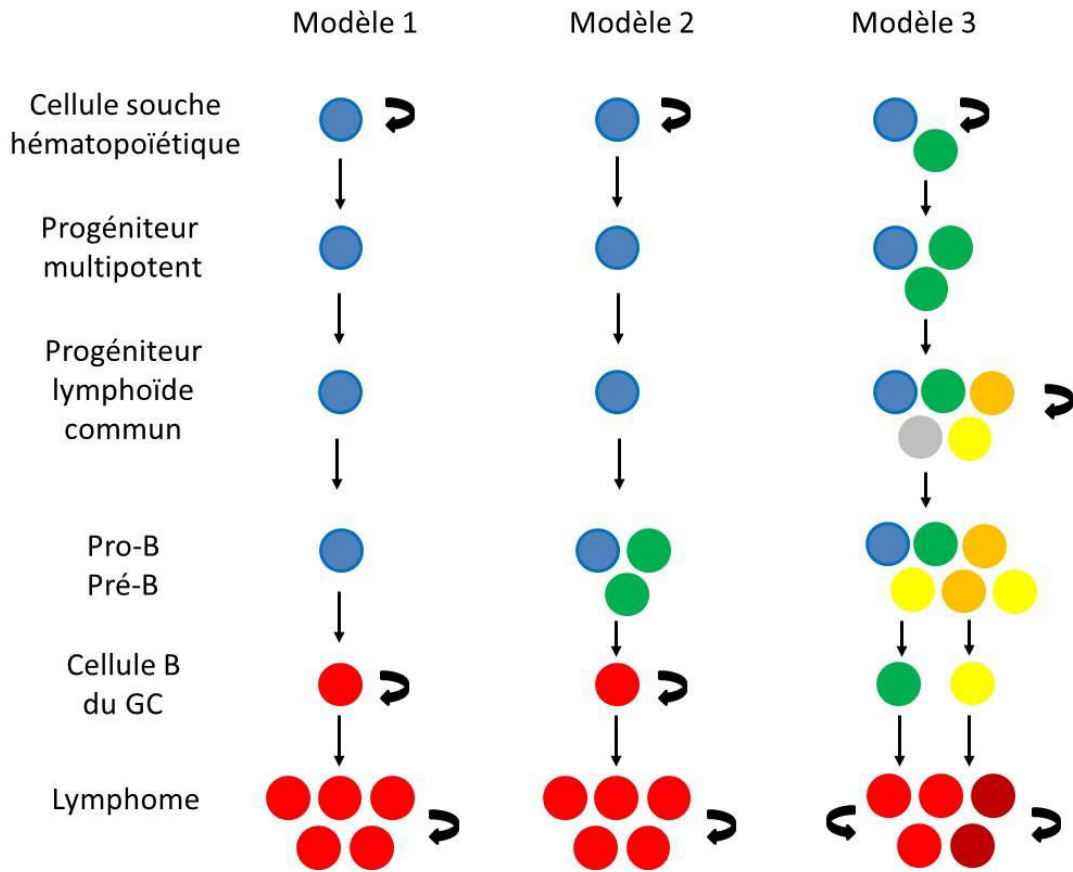


Figure 30 : Les 3 modèles de lymphomagénèse, adapté d'après Weigert et Weinstock⁷⁶.

Le lymphome de Burkitt sporadique pourrait constituer un exemple du modèle 1, l'origine cellulaire supposée étant les centroblastes du GC et les lésions génétiques primitives se produisant au niveau du GC⁷⁶. Le lymphome de Burkitt endémique serait un exemple du modèle 2, dans la mesure où l'apparition de la translocation *MYC-IGH* ferait intervenir RAG au cours du processus de réarrangement VDJ⁷⁶. Enfin, indépendamment du moment où se produisent les altérations génétiques au cours de la transformation lymphoïde, il semble exister des relations hiérarchiques : ainsi, des sous-clones à des stades de différenciation donnés et présentant des capacités d'auto-renouvellement peuvent entraîner l'émergence d'une population tumorale exprimant un phénotype différent. Ce serait le cas dans le myélome et le lymphome de Hodgkin dans lesquels une population de cellules B mémoires présente les mêmes réarrangements des gènes des immunoglobulines que le clone tumoral et pourrait constituer un réservoir de propagation de cellules tumorales, comme dans le modèle 3⁷⁶.



Chapitre 2 : la leucémie lymphoïde chronique

A- Définition

La leucémie lymphoïde chronique (LLC) est l'hémopathie lymphoïde B mature avec phase circulante la plus fréquente chez l'adulte. Elle est caractérisée par une expansion clonale de petits lymphocytes matures monomorphes, à noyau rond ou légèrement irrégulier, retrouvés dans le sang, la moelle osseuse, la rate et les ganglions¹. Des formes prolymphocytaires et des paraimmunoblastes peuvent également être observés, formant des centres de prolifération dans les tissus infiltrés¹.

Selon les classifications OMS 2008 et 2016^{1,46}, la LLC est caractérisée par la présence d'au moins 5 G/L de lymphocytes monotypiques.

Le terme de lymphome lymphocytaire est réservé aux cas avec atteinte ganglionnaire, avec un compte de lymphocytes monotypiques inférieur à 5 G/L dans le sang^{1,46}.

La notion de lymphocytose B monoclonale correspond à la présence d'une population lymphocytaire B monotypique circulante inférieure à 5 G/L, sans signe clinique associé^{1,46}.

B- Epidémiologie

Il existe des différences épidémiologiques selon les régions géographiques. En effet, la LLC est moins fréquente dans les populations africaines et asiatiques que dans la population caucasienne¹²¹. L'incidence de la LLC est plus de 10 fois inférieure dans la population asiatique et cela serait dû à des particularités génétiques dans la mesure où cette différence reste valable chez des sujets d'origine asiatique vivant dans les pays occidentaux, ce qui serait en faveur d'une prédisposition génétique de la LLC^{1,122}.

Le sex ratio homme/femme est de 1,5 à 2,1¹.

L'incidence est de 2 à 6 cas pour 100 000 habitants par an et augmente avec l'âge jusqu'à 12,8/100 000 habitants à l'âge de 65 ans¹.

En France, l'incidence de la LLC et du lymphome lymphocytaire a été estimée à 4464 nouveaux cas en 2012¹²³. L'incidence augmente avec l'âge et passe de 1,2 pour 100 000 de 40 à 44 ans à 52,5 pour 100 000 de 90 à 94 ans chez l'homme¹²³. Selon ces données, l'âge médian au diagnostic est de 71 ans chez l'homme et de 74 ans chez la femme. L'incidence observée entre 1980 et 2012 dans cette étude est en augmentation (1934 nouveaux cas répertoriés en



1980, 3014 nouveaux cas en 1990, 4017 en 2000 et 4481 en 2010), probablement en raison de l'augmentation de l'espérance de vie et d'un meilleur accès aux outils diagnostiques.

Aux Etats-Unis, l'incidence ajustée à l'âge de la LLC est estimée à 4,1 pour 100 000 habitants¹²⁴. En 2015, le nombre de nouveaux cas de LLC était de 14 620¹²⁵.

C- Etiologies

1- Facteurs de risque génétiques

Il existe probablement une prédisposition génétique à la LLC. Une prédisposition familiale est décrite dans 5 à 10 % des cas de LLC avec 2 cas ou plus décrits dans la même famille^{126,127}. Le risque global est augmenté de 2 à 7 fois chez les apparentés au 1^{er} degré d'un patient atteint de LLC¹. Des cas de familles avec une transmission semblant autosomique dominante de la LLC ont été décrits¹²⁸. Le risque est également augmenté en cas d'hémopathie lymphoïde dans la famille. L'âge de survenue en cas de LLC familiale a été décrit comme diminué¹²⁹. Cependant une étude à partir des bases de données du cancer en Suède ne trouve pas de différence significative¹³⁰.

Certaines équipes ont étudié la présence de polymorphismes génétiques pouvant constituer des facteurs de risque. Une étude par SNP-array, réalisée sur 206 familles présentant des cas de LLC issues d'Europe et d'Amérique, a permis d'identifier trois loci potentiellement intéressants situés en 2q21, 18q21 et 6p22¹³¹. En 2q21, le gène le plus pertinent identifié est *CXCR4*. Une équipe a séquencé ce gène dans 186 cas de LLC familiales, comparées à 213 cas contrôles, mettant en évidence 3 polymorphismes pouvant entraînant des conséquences fonctionnelles délétères sur le plan protéique¹³². Le locus 6p22 a également été étudié. Il concerne le locus du Complexe Majeur d'Histocompatibilité (CMH). Cependant son implication reste peu claire¹³³.

Certains polymorphismes ont été décrits comme facteurs prédisposants. Par exemple, des SNP dans les gènes *IRF4*, *LEF1* et *BCL2* ont été décrits¹³⁴. De plus, un SNP entraînant une diminution d'expression de mir15a et mir16-1 a été retrouvé dans des cas de LLC familiales. Cela entraînerait une augmentation de l'expression de ZAP70 et BCL2, aboutissant à une augmentation de la signalisation du BCR et des signaux anti-apoptotiques¹³⁴.

Cependant, l'étude de polymorphismes isolés a donné des résultats peu concluants. L'association de plusieurs variants de faible risque est possible¹³³. Des études d'association génétique pangénomique ou GWAS (pour Genome Wide Association Studies) ont été réalisées. Les SNP étudiés sont fréquents dans la population générale avec une fréquence d'au moins 5 %. Une équipe a ainsi réalisé en 2008 l'analyse de 299 983 SNP chez 1529 patients atteints de LLC, comparés à 3115 cas contrôles et a mis en évidence 6 loci à risque : 2q13,



2q37.1, 6p25 (proche du gène *IRF4*), 11q24, 15q23, 19q13¹³⁵. Actuellement, 30 loci ont été identifiés dans les études GWAS comme entraînant une susceptibilité à la LLC¹³⁶. Cependant, la fonction de ces différents loci reste encore méconnue. La plupart de ces SNP sont localisés dans des régions non codantes ou intergéniques, suggérant un rôle de régulation génique¹³³.

2- Facteurs de risque environnementaux

Il a été démontré que l'exposition professionnelle à certains toxiques peut augmenter le risque de développer un certain nombre de cancers². Cependant, pour la LLC, il n'a pas été prouvé de lien clair entre exposition à des toxiques et survenue de la maladie. En particulier, l'exposition aux radiations ionisantes a été étudiée, et aucune relation n'a pu être mise en évidence aussi bien dans le cadre d'une exposition professionnelle qu'après exposition aux armes nucléaires ou traitement dans le cadre de cancers solides^{137,138}. Quelques études rapportent cependant un risque augmenté, avec notamment une surmortalité liée à la LLC en cas d'exposition à de faibles doses de radiations ionisantes¹³⁹. De même, une équipe retrouve un lien entre exposition au radon et LLC^{140,141}. Cependant, ceci reste controversé et n'a pas été confirmé dans une autre étude chez des sujets travaillant dans des mines d'Uranium¹⁴².

L'exposition aux champs magnétiques et électriques a également été étudiée et ne retrouve pas d'association significative^{143,144}. Une étude norvégienne retrouve un risque augmenté de LLC chez les sujets vivant à proximité de lignes à haute tension, mais de façon non significative¹⁴⁵.

L'exposition aux ultra-violets a été peu étudiée et montrerait une relation inverse, avec effet protecteur sur l'apparition de la LLC et des LNH-B en général^{146,147}.

Plusieurs études retrouvent un risque augmenté de LLC chez les agriculteurs^{148,149,150,151,152,153,154}. Certains types de pesticides sont mis en évidence tels que les organophosphorés¹⁵³. L'exposition aux insecticides augmente le risque de LLC¹³⁴.

Cependant, l'impact de ces différents facteurs reste difficile à établir du fait du faible nombre de cas de LLC, l'âge d'apparition tardif, ou le regroupement des cas de LLC avec d'autres types de leucémies ou de lymphomes¹⁵⁵.



3- Facteurs liés au mode de vie

Le lien entre tabac et cancer est bien établi. Sa relation avec l'apparition de la LLC a été étudiée et il n'a pas été mis en évidence de façon claire d'impact dans le cas de la LLC¹⁵⁵. De même, l'effet de l'utilisation de teintures capillaires reste controversé.

L'impact de l'indice de masse corporelle ou du régime a également été étudié et ne met pas en évidence de lien clair avec la survenue de la LLC^{155,156}. Une étude récente en Espagne montre un risque augmenté en cas de régime occidental par rapport au régime méditerranéen¹⁵⁷.

4- Antécédents médicaux

Les études concernant les antécédents médicaux des patients atteints de LLC se sont particulièrement intéressées aux maladies auto-immunes et ne montrent pas de relation significative avec l'apparition de la LLC^{155,158}. L'anémie de Biermer serait associée à un risque plus élevé¹⁵⁹.

Les antécédents de transfusion sanguine pourraient avoir un rôle dans la survenue de la LLC, mais cet aspect reste controversé^{160,155}.

Certaines infections virales augmentent le risque de LLC, en particulier l'infection par le VHC ou les herpesvirus par exemple^{134,154,161}. Les antécédents de pneumonies constituent un facteur de risque^{161,162}. Cependant, il se peut qu'elles soient en fait une conséquence de l'immunodéficience liée à l'apparition de la LLC¹⁵⁵.

D- Physiopathologie

1- Origine cellulaire

L'origine cellulaire dans la LLC est une question majeure restant encore débattue. En effet, les classifications actuelles cherchent à mettre en relation le type d'hémopathie et sa contrepartie cellulaire normale. Les cellules de la LLC sont CD5+, CD23+. Ce phénotype diffère de la plupart des sous-populations B normales¹⁶³. Ainsi, la maladie dériverait de cellules B CD5+^{164,165}. Il a été proposé que la LLC dériverait d'une sous-population lymphocytaire B spécifique distincte présente dans les amygdales¹⁶⁶. D'autres auteurs ont montré que les cellules de la LLC présentent des similarités avec les cellules de la zone marginale splénique¹⁶⁷. Certains supposent qu'il s'agit de cellules B naïves conventionnelles du fait de cas présentant



un statut IGHV non muté avec des réarrangements stéréotypés communs¹⁶⁸. Par ailleurs, des études de profils d'expression génique montrent des caractéristiques communes quelque soit le statut mutationnel IGHV, suggérant une origine cellulaire commune^{169,170}. Le profil des cellules de LLC se rapproche de celui de cellules B CD5+ sanguines retrouvées chez des sujets sains, les cas IGHV non mutés correspondant à des cellules B matures CD5+, CD27- sanguines et les cas IGHV mutés étant proches de cellules post-GC CD5+, CD27+¹⁷¹.

2- Sélection par l'antigène

L'étude des réarrangements des gènes des immunoglobulines dans la LLC a mis en évidence l'utilisation récurrente de certaines régions variables chez des patients différents¹⁶³. On parle de BCR stéréotypé. Certains de ces réarrangements VDJ sont dans un très grand nombre de cas associés à un réarrangement donné d'*IGK* ou *IGL*^{95,172}. Suivant les caractéristiques des différents stéréotypes, l'immunoglobuline présente des profils distincts de réactivité à l'antigène^{173,174,175}. Il a été montré que certains auto-antigènes tels que le facteur rhumatoïde réagissent avec le BCR de cellules de la LLC¹⁷⁶. D'autres BCR de LLC réagissent contre des antigènes viraux tels que la protéine pUL32 du cytomégalo virus¹⁷⁷. Ceci suggère qu'un ou plusieurs antigènes pourraient être responsables de la sélection des différents répertoires exprimés dans la LLC. La présence de néo- ou auto-antigène ou de cellules apoptotiques, agissant en synergie avec des agents pathogènes bactériens pourraient favoriser l'émergence de progéniteurs de LLC ou des cellules malignes, via la production en continu de BCR avec des caractéristiques structurales particulières¹⁷⁸.

3- Altérations génétiques

a- Anomalies cytogénétiques

De nombreuses anomalies cytogénétiques récurrentes sont retrouvées dans la LLC. Les principales sont la délétion du bras long du chromosome 13 (del(13q)), la trisomie 12, la délétion du bras long du chromosome 11 (del(11q)), et la délétion du bras court du chromosome 17 (del(17p)). Ces anomalies sont associées à un pronostic différent. Ces différentes anomalies et leur impact pronostique seront détaillées dans le paragraphe sur les marqueurs pronostiques de la LLC (page 101).



b- Mutations somatiques

Avec l'apparition du NGS, de nombreuses mutations somatiques ont été mises en évidence. Les principales concernent les gènes *TP53*, *NOTCH1*, *BIRC3* et *SF3B1*. De nombreuses études se sont intéressées à l'impact pronostique de ces différentes mutations. Ceci est développé dans le paragraphe sur les marqueurs pronostiques de la LLC (G-3-Mutations récurrentes hors *IGHV*, page 111).

c- Modifications épigénétiques

Des dérégulations épigénétiques sont retrouvées dans les cancers et entraînent l'activation d'oncogènes ou la sous-régulation de suppresseurs de tumeur, aboutissant à la prolifération cellulaire¹⁷⁹.

Dès la fin des années 1980, une équipe s'est intéressée à la méthylation d'oncogènes tels que *MYC* dans la LLC et d'autres hémopathies en comparaison à des leucocytes de donneurs sains¹⁸⁰. La méthylation de l'ADN entraîne une inactivation génique. Ils ont montré une hypométhylation de ces différents gènes dans la LLC. Au début des années 1990, une autre équipe démontre une hypométhylation génomique globale dans la LLC¹⁸¹. Ces résultats ont été confirmés plus récemment^{182,183}. De plus, une hypométhylation des séquences satellites alpha serait un marqueur pronostique associé à une survie sans traitement plus courte¹⁸³. L'hypométhylation globale dans la LLC a été également mise en évidence avec les techniques de NGS en 2012, et toucherait principalement l'intérieur des gènes et les séquences enhancer¹⁸⁴. Ceci a été confirmé par une autre équipe, retrouvant une hypométhylation des enhancers et des motifs de liaison aux facteurs de transcription, alors que les régions transcrites étaient hyperméthylées¹⁸⁵. Les gènes affectés par cette hyperméthylation peuvent être de différents types : suppresseurs de tumeurs, codant des facteurs de transcription, ou des miARN ou des gènes impliqués dans la survie et la prolifération, ou associés à un pronostic donné¹⁸⁶. Par exemple, la méthylation d'un ou plusieurs sites CpG dans l'intron 1 de *ZAP70* est corrélée à une perte d'expression du gène, à un statut *IGHV* muté et un bon pronostic^{187,188,189}.

Le profil de méthylation a été étudié suivant les sous-groupes de LLC, en particulier suivant le statut mutationnel *IGHV*. Une étude retrouve 3265 sites CpG méthylés différemment¹⁸⁴. En se basant sur la méthylation des îlots CpG, il est possible de regrouper les cas de LLC en 3 catégories distinctes : les NBC-like LLC (pour Cellules B naïves), les MBC-like LLC (pour Cellules B mémoires) et les i-LLC (pour intermédiaires). Les NBC-like LLC regroupent majoritairement les cas *IGHV* non mutés et sont associées à un pronostic péjoratif ; les MBC-like LLC rassemblent les cas *IGHV* mutés et ont un pronostic favorable ; les i-LLC ont des caractéristiques intermédiaires et dériveraient d'un autre sous-type cellulaire, possiblement des cellules B exposées à l'antigène mais n'étant pas passées par le centre germinatif¹⁸⁴. Cependant, récemment, l'étude du méthylome dans différents stades de maturation



lymphocytaire B montre que certaines anomalies décrites dans la LLC sont en fait retrouvées dans des cellules normales^{185,190}. Il faudra donc revoir les anomalies décrites et les comparer aux stades de développement B pour définir les modifications réellement pathogéniques¹⁸⁶.

Dans la LLC, il a été montré que le profil de méthylation de l'ADN est relativement stable au cours du temps et hautement similaire dans les cellules sanguines (quiescentes) et nodulaires (proliférantes)¹⁹¹. Ceci implique donc que la méthylation aberrante retrouvée dans la LLC est un événement précoce dans la lymphomagenèse. Il existerait cependant des modifications subtiles et non récurrentes du profil de méthylation au cours du temps et en fonction des traitements à l'échelon individuel¹⁸⁴. Il a été montré qu'une augmentation de l'hétérogénéité de la méthylation est corrélée avec une complexité génétique sous-clonale avancée, suggérant une co-évolution de nouveaux profils de méthylation avec les altérations génétiques¹⁹². Cette hétérogénéité de la méthylation qui apparaîtrait au hasard est associée à une évolution clinique péjorative¹⁹³. Les auteurs parlent de méthylation désordonnée et suggèrent que cela constitue, comme l'instabilité génétique, un mode d'évolution et de survie pour les cellules cancéreuses.

Des modifications dans la régulation des enzymes modifiant les histones contribuent également à l'initiation et la progression des cancers¹⁹⁴. Dans la LLC, l'expression aberrante d'EZH2 (pour Enhancer of Zest homolog 2) et les mutations affectant les composants des complexe de remodelage de la chromatine sont impliqués dans la pathogénèse de la LLC¹⁸⁶.

EZH2 est une sous-unité de PRC2 (pour Polycomb Repressive Complex 2) qui catalyse la triméthylation de H3K27, entraînant une extinction transcriptionnelle¹⁸⁶. *EZH2* peut être muté et surexprimé dans différents types de cancers. Dans la LLC, seule la surexpression d'EZH2 a été rapportée¹⁸⁶. Elle touche principalement les cas IGHV non mutés¹⁹⁵. Elle est associée à une fréquence diminuée des altérations de *TP53* parmi les LLC IGHV non mutées. Elle entraîne une augmentation de la triméthylation de H3K27 et est liée à une augmentation de la viabilité cellulaire. A l'inverse, il a été montré que la sous-régulation d'EZH2, via l'utilisation d'inhibiteurs, provoque une apoptose¹⁹⁵. Les mécanismes responsables de cette surexpression d'EZH2 sont multiples. La sous-expression de miR-101 a été démontrée dans les cancers et dans la LLC^{196,197}. Celui-ci se lie à l'ARNm d'EZH2 et entraîne sa dégradation. Le microenvironnement pourrait également influencer l'expression d'EZH2¹⁸⁶. La stimulation des cellules de LLC par le CD40L entraîne l'activation de la voie NFκB non canonique et induit l'expression d'EZH2 responsable de la sous-régulation de gènes associés avec la sénescence cellulaire¹⁹⁸. Il a également été montré que c-Rel, qui fait partie des facteurs de transcription de la voie NFκB, induit l'expression d'EZH2 dans les lymphocytes activés et les cellules lymphoïdes malignes¹⁹⁹.

Les mutations de gènes codant des composants des complexes de remodelage de la chromatine pourraient influencer la pathogénèse de la LLC¹⁸⁶. Ceux-ci peuvent réguler la réparation de l'ADN, la réplication et la transcription¹⁸⁶. Des mutations de *ARID1A* (pour AT-Rich Interaction Domain 1A) codant une sous-unité d'un complexe de remodelage de la chromatine, sont retrouvées dans environ 2 % des LLC²⁰⁰. Dans d'autres types de cancers, des



mutations perte de fonction ont été retrouvées, laissant supposer qu'il s'agit d'un gène suppresseur de tumeur et d'un régulateur du cycle cellulaire^{201,202,203,204}. Le rôle d'ARID1A dans la pathogénèse de la LLC reste à établir¹⁸⁶. Des mutations de *CHD2* (pour Chromodomain Helicase DNA binding protein 2) ont été rapportées dans environ 5 % des cas de LLC^{205,206}. La majorité de celles-ci sont des mutations tronquantes ou touchant des résidus conservés de domaines fonctionnels, laissant supposer un potentiel rôle de suppresseur de tumeur²⁰⁵. Elles sont retrouvées principalement dans les cas IGHV mutés^{206,205}.

Les mutations touchant des gènes codant des enzymes régulatrices de l'épigénome sont rares dans la LLC. Parmi ces gènes, le plus fréquemment muté (environ 4 % des cas) est *SETD2* (pour SET domain containing 2) qui code une histone méthyl transférase responsable de la triméthylation de H3K36 associée avec des régions activement transcrites^{186,207}. Ce gène peut également être délété dans 3 % des cas²⁰⁸. Ces délétions sont associées avec les anomalies de *TP53*, la présence d'un caryotype complexe et le chromothripsis^a. Les anomalies de *SETD2* semblent être mono-alléliques et seraient de mauvais pronostic²⁰⁸.

d- Altérations de miARN

Les miARN sont de courts ARN simple brin non codants qui régulent l'expression de gènes via la répression traductionnelle ou la dégradation de transcrits²¹¹. Une dérégulation des miARN a été démontrée dans les cancers, et en particuliers les hémopathies. Leur rôle peut être celui de suppresseurs de tumeurs ou d'oncogènes (ou oncomirs), mais aussi de régulateurs des capacités invasives des cellules tumorales^{212,213}.

Dès 2002, l'implication des miARN dans la LLC a été décrite²¹⁴. Deux miARN situés en 13q14 sont décrits : miR15 et miR16. La délétion du bras long du chromosome 13 entraîne la perte de ces deux miARN. Ceci a été validé par SNP-array²¹⁵. Une étude rapporte la présence de mutation germinale dans le précurseur primaire mir16-1- miR15a, entraînant une diminution de l'expression de ces miARN, et souvent associée à la délétion de l'autre allèle²¹⁶. Une équipe a montré dans un modèle murin de LLC que des mutations ponctuelles de la région 3' flanquante de miR16-1 entraînent une diminution de l'expression de miR16-1²¹⁷. Un autre groupe a montré que les niveaux d'expression de ces deux miARN sont inversement corrélés à celui de BCL2²¹⁸. Ces miARN régulent également l'expression de BCL2 à un niveau post-transcriptionnel. Par ailleurs, les auteurs ont mis en évidence une homologie entre les

^a Le chromothripsis (du grec, chromos pour chromosome, et thripsis pour briser en éclats) est un phénomène caractérisé par de multiples réarrangements (de quelques dizaines jusqu'à plusieurs centaines) affectant un seul ou un faible nombre de chromosomes, survenu au cours d'un seul événement considéré comme catastrophe cellulaire^{209,210}. Cela entraîne une oscillation du nombre de copies autour de 2, et normalement pas plus de 3 niveaux différents^{209,210}.



extrémités 5' de ces deux miARN et la séquence de l'ARNm de BCL2. De plus, l'inhibition de BCL2 par miR15 et miR16-1 dans une lignée cellulaire leucémique entraîne l'apoptose²¹⁸. Une autre étude basée sur les profils d'expression de gènes prédits par informatique comme potentiellement cibles de ces deux miARN montre également une régulation de BCL2²¹⁹. Cependant, cela reste controversé. Une équipe ne retrouve une diminution significative d'expression de miR15a et miR16-1 qu'en cas de délétion du bras long du chromosome 13 homozygote, et ne met pas en évidence de relation avec l'expression de BCL2²²⁰. Il a par ailleurs été démontré que miR-16 entraîne la régulation négative de transcrits de gènes dont l'extinction entraîne un arrêt du cycle cellulaire en phase G0/G1.

Un autre miARN important dans la LLC est miR29. Il a été décrit comme un miARN suppresseur de tumeur ciblant plusieurs oncogènes dont *TCL1* (pour T-Cell Leukemia/Lymphoma 1)^{221,222}. Une expression élevée de *TCL1* est corrélée à un phénotype agressif dans la LLC et avec un statut IGHV non muté²²³. *TCL1* interagirait avec ATM (pour Ataxia Telangiectasia Mutated) et entraînerait une activation de la voie NFκB²²⁴. miR-181b serait également un régulateur de *TCL1*²²¹. Par ailleurs, une équipe a montré qu'une diminution de son expression serait un marqueur de progression de la maladie²²⁵.

La dérégulation des miARN par modifications épigénétiques est un événement important dans le développement de la LLC^{216,226}. La méthylation aberrante permet l'extinction transcriptionnelle de certains miARN^{227,228}. Par exemple, miR129-2 est un miARN suppresseur de tumeur fréquemment méthylé dans la LLC²²⁹. Cette méthylation est associée à une survie réduite. Un autre miARN suppresseur de tumeur, miR9-3 est méthylé fréquemment dans la LLC²³⁰. Il régulerait normalement négativement la voie NFκB.

Le profil d'expression des miARN dans la LLC a été étudié^{231,220}. Plusieurs équipes ont essayé de relier ces profils avec les marqueurs pronostiques connus tels que le statut mutationnel IGHV^{216,220}. De plus, l'expression de certains miARN tels que miR21 et miR181b sont des facteurs pronostiques indépendants²³². Enfin, miR34-a, cible de p53, est significativement sous exprimé en cas d'anomalie de *TP53*, mais son expression est très variable dans les autres cas²³³. Des niveaux bas d'expression de miR34-a sont associés à un temps de doublement de la lymphocytose réduit et une survie sans traitement plus courte.

e- Rôle des lncRNAs

Parmi les lncRNAs potentiellement impliqués dans la LLC, DLEU1 et DLEU2 (pour Deleted in Lymphocytic leukemia 1 et 2) ont été particulièrement étudiés¹¹⁰. Ils sont localisés en 13q14.1 et sont délétés dans plus de la moitié des cas de LLC²³⁴. Dans cette même région sont situés miR15a et miR16-1 dont le rôle a été décrit page 86. Il a été démontré dans des modèles murins que la délétion de l'ensemble du cluster Dleu2/miR15a/miR16-1 entraîne une lymphoprolifération semblable à la LLC et plus agressive qu'en cas de délétion de miR15a/miR16-1 seuls²³⁵. Il a été observé dans des lignées cellulaires cancéreuses que DLEU2



inhibe la croissance cellulaire et la capacité à former des colonies de façon dépendante de miR15a/miR16-1²³⁶. DLEU2 régule négativement le niveau d'expression des Cyclines D1 et E1 via miR15a/miR16-1. De plus, des modifications épigénétiques telles que la déméthylation de DLEU1 et de l'isoforme DLEU2/Alt1 est fréquente dans la LLC et entraîne une régulation négative de la transcription de DLEU2 et miR15-a/miR16-1, ainsi que de gènes suppresseurs de tumeurs proches régulant l'activité de NFκB²³⁷.

Deux lncRNAs intervenant dans la voie p53 sont également impliqués dans la LLC : NEAT1 (pour Nuclear Enriched Abundant Transcript 1) et lincRNA-p21 (lincRNA pour long intergenic non coding RNA)²³⁸. Ils sont induits en cas de dommage de l'ADN en présence de p53 fonctionnel. Ainsi, en cas d'anomalie de *TP53* dans la LLC, ces lncRNA seront sous exprimés. Il a également été montré que p53 se lie au promoteur de NEAT1²³⁸. L'utilisation de nutlin-3, un inhibiteur de MDM2, induit la transcription de NEAT1 et lincRNA-p21 sous la dépendance de p53. NEAT1 régulerait la rétention de certains ARNm, entraînant une répression traductionnelle²³⁹.

MIAT (pour Myocardial Infarction Associated Transcript) est un lncRNA surexprimé dans les lignées lymphoïdes avec un phénotype B mature, dont la LLC²⁴⁰. Cette surexpression est plus marquée dans les formes agressives (formes avec délétion du bras court du chromosome 17 ou du bras long du chromosome 11, ou trisomie 12) que dans les formes indolentes (délétion du bras long du chromosome 13) de LLC. De plus, MIAT produit une boucle de régulation avec OCT4, et ces deux molécules seraient essentielles pour la survie cellulaire²⁴⁰. MIAT s'accumule dans le noyau dans des corps nucléaires spécifiques et pourrait affecter l'épissage et exercer une régulation sur l'expression de gènes²⁴¹. MIAT favoriserait la prolifération des cellules B matures malignes¹¹⁰.

BM742401 est un lncRNA dont le promoteur est fréquemment méthylé dans la LLC, alors qu'il est déméthylé dans les cellules B normales sanguines²⁴². Il s'agirait d'un suppresseur de tumeur et sa méthylation dans la LLC serait corrélée avec un stade clinique avancé (Rai ≥ 2) et une lymphocytose plus élevée. Sa surexpression dans des lignées cellulaires entraîne l'inhibition de la prolifération cellulaire et favorise l'apoptose. Les mécanismes exacts de ses fonctions de suppresseur de tumeur sont encore à préciser.

Quelques études ont été réalisées sur l'expression globale des lncRNA dans la LLC. Une première étude de NGS des ARN met en évidence 127 lncRNA exprimés de façon différente entre des cellules normales et de LLC²⁴³. Une autre équipe a réalisé le profil d'expression des lncRNA chez 217 patients atteints de LLC en stade Binet A, en comparaison avec 26 sous-populations de cellules B normales sanguines et amygdaliennes²⁴⁴. Une signature de 24 lncRNA spécifiquement dérégulés dans la LLC a été mise en évidence. Quelques lncRNA étaient associés à certains groupes pronostiques tels que le statut IGHV non muté, la délétion du bras long du chromosome 11 ou du bras court du chromosome 17 et le statut *NOTCH1* muté). Enfin, l'expression de deux lncRNA (lnc-IRF2-3 et lnc-KIAA1755-4) permettait de différencier trois groupes pronostiques. Le groupe de haut risque présentait une expression élevée concomitante de ces deux lncRNA, tandis que le groupe de bas risque avait des niveaux bas d'expression. Les autres cas étaient classés dans un profil de risque intermédiaire.



Le rôle des lncRNA dans la LLC commence à émerger et de futures études permettront de mieux comprendre leur rôle dans la pathogénèse et probablement de les intégrer dans des scores pronostiques.

4- Microenvironnement

Le microenvironnement tissulaire joue un rôle primordial dans la pathogénèse de la LLC. Les événements majeurs de la pathogénèse ont lieu dans les organes lymphoïdes périphériques et la moelle osseuse où un microenvironnement favorable est établi et maintenu par la coévolution de cellules tumorales et normales²⁴⁵. Le microenvironnement de la LLC possède les mêmes caractéristiques que les microenvironnements dans le cancer c'est-à-dire de nouveaux vaisseaux (angiogénèse) pour l'apport de nutriments, la production locale de facteurs de croissance et la protection des cellules malignes contre le système immunitaire²⁴⁵.

Alors que la majorité des cellules de LLC sanguines ne se divisent pas, elles prolifèrent au niveau des tissus. Cette prolifération a lieu au niveau de centres de prolifération ou pseudo-follicules²⁴⁶. Les cellules sont alors en contact étroit avec des cellules accessoires telles que les lymphocytes T, des cellules stromales d'origine mésenchymateuse, et des cellules «nurse-like» dérivant des monocytes^{163,247,248,249}. Des études d'expression génique montrent que les cellules de LLC issues des ganglions ont une expression plus forte de E2F et MYC par rapport aux cellules sanguines²⁵⁰. Dans les organes lymphoïdes secondaires, les cellules de LLC ont une activation de la signalisation du BCR et des récepteurs de la famille du TNF (pour Tumor Necrosis Factor). Les cellules produites dans les tissus et échappant à l'apoptose s'accumulent et pourraient passer dans le sang, puis ré-entrer dans les tissus pour continuer à proliférer²⁴⁵.

a- Les acteurs du microenvironnement

i- Les cellules « nurse-like »

Il s'agit de cellules dérivées des monocytes. Ces cellules ont été décrites initialement *in vitro*²⁵¹. Les monocytes cultivés avec des cellules de LLC se différencient spontanément en cellules «nurse-like» qui pouvaient promouvoir la survie et la prolifération des cellules de LLC²⁵². Elles ont été identifiées *in vivo* dans les ganglions et la moelle osseuse de patients atteints de LLC, et elles sont corrélées à la viabilité des lymphocytes leucémiques^{251,253}. Une étude montre que l'activation des voies RAGE (pour Receptor for Advanced Glycation End-product) et du TLR9 par la libération d'HMGB1 (pour High Mobility Group Box 1) par les cellules de LLC entraîne la différenciation en cellules «nurse-like»²⁵⁴. Les cellules «nurse-like» expriment fortement les gènes correspondant aux molécules d'adhésion et interagissent avec les cellules de la LLC via des chimiokines et des cytokines dont BAFF (pour B-cell Activating



Factor) et APRIL (pour A Proliferation-Inducing Ligand)²⁵⁵. Elles sécrètent également du CD14 soluble sous l'action des cellules de la LLC, ce qui favorise la survie de ces dernières²⁵⁶. *In vitro*, cela entraîne une augmentation de l'expression de gènes anti-apoptotiques tels que BCL2²⁵⁷.

ii- Les cellules stromales mésenchymateuses

Elles sont considérées comme la partie nourricière pour les progéniteurs hématopoïétiques normaux²⁵¹. Dans la moelle osseuse, elles contribuent à l'architecture médullaire. Elles sont retrouvées dans les organes lymphoïdes secondaires des patients atteints de LLC²⁴⁸. Elles sécrètent de façon constitutive des chimiokines permettant le recrutement des cellules de LLC²⁵¹. De plus, par contact direct, elles induisent *in vitro* une sur-régulation de facteurs (dont ZAP70 et le CD38) qui favorisent la survie et la prolifération des lymphocytes de LLC²⁵⁸. Elles augmentent également la synthèse de glutathion et la glycolyse, entraînant la survie et une résistance au traitement²⁵⁹. Au niveau de la moelle, les cellules stromales favorisent la survie via ROR1 (pour Receptor tyrosine kinase like Orphan Receptor 1)²⁶⁰. Il s'agit d'un récepteur à domaine tyrosine kinase dont l'expression est fortement augmentée dans la LLC¹⁶⁹. Son expression est faible dans le sang normal²⁶¹. ROR1 serait impliqué dans la survie cellulaire²⁶².

iii- Les lymphocytes T

Les lymphocytes T sont en contact étroit avec les lymphocytes de la LLC²⁵¹. Ils fournissent des signaux de régulation dans la niche tissulaire²⁶³. Il a été mis en évidence une activation chronique et un épuisement, avec en particulier l'expression de marqueurs d'épuisement tels que CD244, CD160 et PD1, entraînant des défauts de cytotoxicité, de formation de la synapse immunologique avec les cellules présentatrices d'antigène, et de migration²⁶⁴.

iv- Les lymphocytes NK

Les cellules NK du microenvironnement de la LLC ont une capacité réduite à lyser les cellules de LLC avec un défaut en granules cytoplasmiques²⁵¹. Elles sont également déficientes dans la polymérisation de l'actine et la formation de la synapse immunologique²⁵¹. Les cellules de la LLC pourraient également inhiber les cellules NK par contact direct. Il a été montré que les patients atteints de LLC ont une hyperexpression plasmatique de la molécule HLA-G (HLA pour Human Leukocyte Antigen), ce qui induit l'apoptose des cellules NK et perturbe leur pouvoir cytotoxique²⁶⁵.



v- Les cellules suppressives d'origine myéloïde ou MDSC pour Myeloid-Derived Suppressor Cells

Les MDSC jouent un rôle dans les défauts mis en évidence dans les lymphocytes T et NK du microenvironnement de la LLC²⁵¹. Elles ont un phénotype myéloïde aberrant (CD14b+ et HLA-DR-). Les MDSC sont augmentées chez les patients atteints de LLC non traités. Elles suppriment les fonctions T dans le microenvironnement et induisent des lymphocytes T régulateurs.

vi- Les cellules endothéliales et les cellules folliculaires

Ces cellules sont essentielles pour le « homing » et la rétention des cellules de LLC dans les tissus²⁵¹. L'adhésion aux cellules endothéliales induit la survie, l'activation et la résistance au traitement des cellules de LLC²⁵¹. Cette interaction se ferait via des intégrines. Le contact avec les cellules folliculaires dendritiques passerait par le CD44 (récepteur membranaire de l'acide hyaluronique) exprimé sur les cellules de LLC, ainsi que via l'acide hyaluronique²⁵¹. Le CD44 interagit avec d'autres protéines de surface et cytoplasmiques exprimées par les cellules de LLC telles que le CD38 et ZAP70. L'inhibition du CD44 par un anticorps monoclonal entraîne une activité cytotoxique directe pour les cellules de LLC, tout en ayant peu d'effet sur les cellules B normales²⁶⁶. Ainsi, le CD44 pourrait être une cible thérapeutique. Le CD27 est également impliqué dans l'interaction avec le stroma. Son expression est corrélée avec l'expression de ZAP70 et est augmentée lors de la liaison du BCR²⁵¹.

b- L'activation des voies de signalisation du BCR

Les interactions entre les cellules de la LLC et du microenvironnement dépendent d'un BCR fonctionnel, permettant l'interaction avec l'antigène, et l'activation des voies de signalisation en aval²⁴⁵.

Les signalisations-réponses du BCR sont hétérogènes dans la LLC²⁶⁷. D'un côté, il existe des cas de LLC qui semblent non répondeurs à la stimulation de l'IgM de surface. Ceci est plus fréquent parmi les cas de LLC IGHV muté et est associé à une maladie indolente²⁶⁸. Cependant, parmi les cas répondeurs, il existe une hétérogénéité de la réponse *in vitro*²⁶⁷. Il semble que les niveaux d'expression de l'IgM de surface soient le principal déterminant de ces différences de réponses²⁶⁷. En général, l'immunoglobuline de surface est faiblement exprimée dans la LLC, mais avec des niveaux variables²⁶⁷. Cela pourrait être dû à une stimulation répétée par l'antigène ou une sous-modulation du récepteur. Une forte sous-modulation semble contribuer à la perte des signalisations-réponses de l'IgM, en particulier dans les cas IGHV



mutés²⁶⁸. La perte de la capacité de signalisation dans la LLC est associée avec une activation constitutive de ERK et du facteur nucléaire des cellules T activées, ce qui est en faveur du fait qu'une stimulation continue par l'antigène *in vivo* conduit à l'anergie²⁶⁹. Une sous-modulation moins importante, caractéristique plus fréquente dans les LLC IGHV non mutées, entraîne un maintien de la capacité de signalisation, se traduisant par une activation partielle des voies en aval et mimant les caractéristiques des cellules B anergiques murines²⁶⁷. Ce maintien plus fort de la capacité de signalisation dans les LLC IGHV non mutées pourrait être important pour le comportement de la tumeur²⁶⁷.

Des hétérogénéités dans la signalisation-réponse existent au sein du clone malin et pourraient être le reflet de temps différents de la stimulation par l'antigène parmi les cellules sanguines recirculantes²⁶⁷. Enfin, l'activation de différents isotypes d'immunoglobuline de surface pourrait également entraîner des réponses variables²⁶⁷. L'activation de l'IgM de surface entraîne *in vitro* soit l'absence de réponse, soit une activation relativement persistante de ERK1/2, tandis que l'activation de l'IgD entraîne une réponse transitoire dans quasiment toutes les cellules de LLC^{267,270}. L'activation persistante de ERK1/2 est critique pour la ré-entrée dans le cycle cellulaire, ce qui pourrait expliquer pourquoi la signalisation induite par l'IgD ne semble pas influencer le comportement de la maladie²⁶⁷. La différence de comportement entre l'IgM et l'IgD de surface pourrait être due à leurs niveaux d'expression relatifs ou à des différences dans l'association de cofacteurs de signalisation²⁶⁷. Les cas de LLC ayant effectué la commutation de classe isotypique sont relativement rares et ont donc été moins étudiés.

Bien que la modulation de l'IgM de surface joue un rôle important dans la détermination de la signalisation-réponse, d'autres facteurs doivent également intervenir²⁶⁷. La capacité de signalisation du BCR est corrélée avec l'expression de ZAP70 et la surexpression de ZAP70 augmente cette signalisation^{271,272}. ZAP70 pourrait augmenter indirectement l'activation de SYK et/ou séquestrer des inhibiteurs de SYK²⁷³. ZAP70 pourrait également moduler d'autres voies de signalisation dont celles contribuant à la migration cellulaire²⁶⁷. L'expression du CD38 est corrélée à la capacité de signalisation de l'IgM²⁷⁴. Le CD38 s'associe à l'IgM dans les cellules de LLC, mais sa capacité à moduler directement les signalisations-réponses n'est pas clairement identifiée^{267,275}. L'interaction CD38-CD31 semble contribuer aux voies impliquées dans la migration et le homing, et augmente la survie via l'induction de BCL2 et BCL-xL^{267,276,277}.

De nombreux autres facteurs sont exprimés de façon différente dans la LLC et peuvent dans certains cas moduler directement la capacité de signalisation de l'IgM²⁶⁷. Cependant, la perte de capacité de la signalisation-réponse est associée à l'absence de phosphorylation induite par SYK, ce qui est un événement précoce dans la formation du signalosome^{267,270}. Il semble donc qu'une étape régulatrice clé doit exister au niveau ou à proximité du récepteur²⁶⁷. Les réponses des lymphocytes de LLC *in vivo* sont potentiellement dépendantes des caractéristiques des cellules elles-mêmes et de l'antigène qui interagit (soluble versus lié au tissu/macromolécule), ainsi que de la présence potentielle de cellules supports²⁶⁷.



c- Le rôle des Toll-like receptors

Les Toll like receptors ou TLR sont des récepteurs de surface faisant partie du système immunitaire inné. Ils sont capables de se lier à des antigènes microbiens de structure conservée et d'activer ensuite les réponses immunitaires innées¹⁶³. Différents TLR sont exprimés dans la LLC, dont le TLR7 et le TLR9, et leur stimulation peut induire l'expression membranaire de marqueurs d'activation tels que CD25 et CD86²⁷⁸.

La voie de signalisation des TLR dans le microenvironnement ganglionnaire pourrait contribuer à l'activation de NFκB, l'expression de molécules co-stimulatrices et la régulation de la survie des cellules de LLC²⁷⁹. Il a été démontré que la stimulation des TLR peut diminuer la sensibilité des cellules à la fludarabine *in vitro* lorsqu'elles sont porteuses d'anomalies de mauvais pronostic telles qu'un statut IGHV non muté²⁸⁰. De plus, la perte de TIR8 (pour Toll Interleukine-1 Receptor 8) favorise la progression de la LLC dans un modèle murin²⁸¹. Cela suggère que la coopération entre BCR et TLR et/ou leur régulation pourraient impacter l'évolution du clone malin²⁴⁵.

d- Les chimiokines

Les récepteurs de chimiokines et les molécules d'adhésion exprimées par les cellules de LLC sont critiques pour le « homing » et leur rétention dans les tissus²⁸². Notamment, le contact avec les cellules stromales accessoires présentes dans le microenvironnement fournissent aux cellules de LLC des signaux de survie^{283,284}. Les cellules « nurse-like » dans le ganglion et les cellules stromales médullaires attirent les cellules de LLC via les récepteurs de chimiokines CXCR4 et CXCR5 qui sont fortement exprimés dans la LLC. Certaines chimiokines telles que CCL3, CCL4, CCL22 et l'interleukine 8 sont sécrétées par les cellules de LLC et peuvent recruter des lymphocytes T et d'autres cellules accessoires pour promouvoir la survie²⁴⁵. La régulation de CXCR4 par CXCL12 induit le trafic continu des cellules entre le stroma et le sang. L'expression de CXCR4 peut être utilisée comme marqueur pour identifier les cellules qui ont récemment quitté ou qui vont retourner dans le stroma²⁸⁵.

Les cellules de LLC ZAP70+ migrent plus rapidement sous l'effet de chimiokines *in vitro*²⁸⁶.



e- Le cytosquelette

La protéine HS1 (pour Hematopoietic lineage cell-specific protein 1) est un régulateur du remodelage du cytosquelette, contrôlant le trafic et le homing des lymphocytes²⁸⁷. Elle influence l'invasion tissulaire et l'infiltration dans la LLC. HS1 est un régulateur de l'actine au niveau de la synapse immunologique des cellules T, et elle influence dans les cellules NK et les cellules dendritiques différentes fonctions telles que la lyse de cellules cibles, l'adhésion cellulaire, le chimiotactisme et l'assemblage de l'actine au niveau de la synapse lytique²⁸⁸. Ceci suggère l'importance de l'organisation du cytosquelette dans le trafic et le « homing » des cellules de la LLC.

E- Clinique

1- Présentation clinique classique

Les caractéristiques cliniques sont très variables (présentation, évolution...). La plupart des patients sont asymptomatiques, mais certains patients présentent une asthénie, une anémie hémolytique auto-immune, des infections, une splénomégalie, une hépatomégalie, des adénopathies. Un petit pic monoclonal peut être retrouvé¹³⁴.

En général, la moelle est infiltrée, ainsi que les ganglions, le foie et la rate. D'autres sites extraganglionnaires peuvent parfois être impliqués¹. La peau est le principal organe non lymphoïde pouvant être infiltré, mais cela représente moins de 5 % des cas². Une atteinte gastro-intestinale ou méningée est extrêmement rare².

Dans environ 25 % des cas, le diagnostic de LLC a lieu à la suite de la découverte fortuite d'une hyperlymphocytose sur un bilan biologique de routine^{2,134}. Parfois, ce sont l'apparition d'adénopathies persistantes ou des signes cliniques en lien avec des cytopénies auto-immunes qui amènent le patient à consulter.

L'évolution clinique de la LLC est très hétérogène. Environ un tiers des patients ont une survie prolongée, supérieure à 20 ans, ne nécessitant aucun traitement et décèderont de causes sans rapport avec la LLC. Dans un tiers des cas, les patients vont présenter une maladie agressive et nécessiter immédiatement un traitement. Enfin, les autres vont connaître une phase initiale indolente suivie d'une progression de la maladie²⁸⁹.

Dans 3 à 10 % des cas, la LLC peut transformer en lymphome agressif (le syndrome de Richter)²⁹⁰. Le pronostic est alors sombre.



Parfois, la LLC peut évoluer vers une leucémie prolymphocytaire, associée à une évolution péjorative. Cette entité est principalement définie sur le plan cytologique avec la présence de plus de 55 % de formes prolymphocytaires¹.

2- Les variants

a- Le lymphome lymphocytaire

Le lymphome lymphocytaire est considéré comme étant la même entité que la LLC selon la classification OMS^{1,291}. Dans ce cas, le clone lymphocytaire est localisé au niveau des ganglions ou de la rate, la phase leucémique étant inférieure à 5 G/L. Peu d'études s'intéressent actuellement aux différences entre ces 2 entités. Une équipe a évalué différents paramètres selon la lymphocytose sanguine et retrouve des différences concernant les caractéristiques génétiques (anomalies cytogénétiques moins fréquentes et statut muté des gènes de la chaîne lourde des immunoglobulines dans le lymphome lymphocytaire) et immunophénotypiques (positivité du CD38 plus fréquente que dans la LLC)²⁹². La survie et le taux de réponse au traitement n'était pas significativement différents suivant la lymphocytose. Une autre équipe rapporte des différences cytogénétiques avec une fréquence plus élevée de trisomie 12 dans le lymphome lymphocytaire, et diminuée de délétion du bras long du chromosome 13²⁹³. Plus récemment, une étude portant sur 890 cas de LLC et lymphome lymphocytaire (dont 113 cas de lymphome lymphocytaire) confirme la fréquence plus élevée de trisomie 12 et de la positivité du CD38, mais rapporte un statut non muté des gènes de la chaîne lourde des immunoglobuline plus fréquent²⁹⁴. Des mutations de *NOTCH1* sont retrouvées dans de nombreux cas. Le lymphome lymphocytaire peut évoluer vers une LLC. La survie globale est moins bonne que dans la LLC, avec un taux de transformation plus élevé²⁹⁴.

b- La lymphocytose B monoclonale

La lymphocytose B monoclonale (ou MBL pour Monoclonal B-cell Lymphocytosis) correspond à la mise en évidence d'un clone lymphocytaire B sanguin représentant moins de 5 G/L (critère nécessaire pour le diagnostic de LLC) et en l'absence d'adénopathie ou de signe clinique (critères pour parler de lymphome lymphocytaire). La notion d'adénopathie est basée sur la palpation d'adénopathie avec un seuil de 1,5 cm, plutôt que sur l'imagerie²⁹⁵.

L'apparition de cette nouvelle entité en 2008 ainsi que la modification du seuil pour parler de LLC impliquent une reclassification de certains patients dans les études datant d'avant ces critères.



La classification OMS 2016 différencie les MBL « low count » dans lesquelles la lymphocytose clonale est inférieure à 0,5 G/L, des MBL « high count » dans lesquelles le clone est compris entre 0,5 et 5 G/L²⁹¹. Les cas de MBL « low count » ont des caractéristiques différentes de la LLC et ont un faible risque de progression.

On distingue également les MBL type LLC dont les caractéristiques immunophénotypiques sont celles de la LLC (CD5+, CD23+), des MBL non LLC (CD5-) et des MBL atypiques (CD5+, CD23- ou CD20fort)^{296,297}.

Les sujets présentant une MBL « low count », bien que persistante dans la majorité des cas, ont un risque très faible d'évoluer vers une LLC²⁹⁸. De plus, l'espérance de vie est similaire à celle de la population générale²⁹⁵.

Actuellement les MBL « high count » sont supposées être un stade précurseur de la LLC au cours duquel des anomalies génétiques sont acquises selon un processus multi-étapes²⁹⁵. De nombreuses études montrent un risque de progression vers la LLC ou un lymphome lymphocytaire estimé entre 1 et 2 % par an^{295,299,300,301}. La survie globale de ces sujets ne semble pas différente de celle de la population générale³⁰². Cependant, certains marqueurs pronostiques utilisés dans la LLC pourraient permettre de mieux stratifier ces patients. Les sujets porteurs de MBL « high count » ont un risque infectieux plus important³⁰³.

La fréquence des MBL chez des sujets de plus de 40 ans est estimée à environ 5 %^{295,304,305,301}. La fréquence augmente avec l'âge : elle est de 0,2 à 0,3 % chez des sujets de moins de 40 ans, de 3,5 à 6,7 % entre 40 et 60 ans et de 5 à 9 % chez les plus de 60 ans. L'utilisation de techniques plus sensibles entraîne l'augmentation de ces fréquences (plus de 20 % chez les sujets de plus de 60 ans et jusqu'à 75 % chez les plus de 90 ans)³⁰⁴.

3- Les complications

a- Le syndrome de Richter

Le syndrome de Richter est défini par l'apparition d'un lymphome agressif chez des patients atteints de LLC ou de lymphome lymphocytaire³⁰⁶. Deux variants pathologiques sont distingués : un variant type DLBCL et un variant type lymphome de Hodgkin¹.

Le variant type DLBCL est caractérisé histologiquement par des plages de grandes cellules B, soit des centroblastes (60 à 80 % des cas), soit des immunoblastes (20 à 40 %). Il faut différencier la LLC agressive : dans ce cas, les cellules de LLC sont de taille augmentée avec un index prolifératif plus important. Cependant, il n'existe pas de critère précis selon la classification OMS entre ces deux formes. La distinction repose alors sur l'interprétation et l'expertise de l'anatomopathologiste³⁰⁶.



D'un point de vue immunophénotypique, les cellules du syndrome de Richter variant DLBCL sont généralement CD20+ et ne sont que peu fréquemment CD5+ (seulement environ 30 % des cas)^{306,307}. L'expression du CD23 n'est positive que dans environ 15 % des cas. L'immunophénotype correspond à des cellules post-GC (IRF4 positif) dans 90 à 95 % des cas. Les 5 à 10 % restants ont un phénotype type GC (CD10+)^{306,307}. On retrouve une expression de MYC dans 30 à 40 % des cas, corrélée avec la présence d'une translocation impliquant MYC en cytogénétique³⁰⁸. Les cellules expriment BCL2, mais des lésions génétiques de ce gène (translocations ou mutations somatiques) ne sont que rarement mises en évidence. Une expression de PD-1 (pour Programmed cell Death 1) est retrouvée dans plus de 80 % des cas, alors qu'elle est rare dans la LLC et les DLBCL de novo³⁰⁶. La grande majorité sont EBV-négatifs. Enfin, dans environ 80 % des cas, l'analyse des gènes des immunoglobulines montre une clonalité en rapport avec celle de la LLC initiale, ce qui est en faveur d'une réelle transformation^{306,307,309}. Dans les autres cas, cela révèle une autre origine de clonalité.

Au niveau de la physiopathologie, le syndrome de Richter variant DLBCL ne présente pas les mêmes caractéristiques génétiques que les DLBCL *de novo* comme par exemple l'inactivation des gènes de l'acétyltransférase *CREBBP/EP300* et de *B2M*. Les altérations génétiques retrouvées dans ces cas de Richter sont des anomalies des gènes *TP53*, *NOTCH1*, *MYC* et *CDKN2A*^{306,309,310}. Les mutations de *TP53* apparaissent dans environ 60 à 80 % des cas et sont généralement acquises au moment de la transformation³⁰⁹.

Le pronostic du syndrome de Richter variant DLBCL est généralement péjoratif. Il est généralement meilleur en cas de DLBCL non relié génétiquement à la LLC initiale avec une survie médiane d'environ 5 ans, tandis qu'elle est de 8 à 16 mois en cas de réelle transformation³⁰⁶.

Dans le cas du variant type lymphome de Hodgkin, le diagnostic nécessite la présence de cellules de Reed-Sternberg de phénotype CD30+, CD15+, CD20-, avec un fond cellulaire de petits lymphocytes T, de cellules épithélioïdes, de polynucléaires éosinophiles et de plasmocytes³¹¹. La plupart de ces cas sont EBV-positifs et présentent un réarrangement des immunoglobulines différent de celui de la LLC de départ, suggérant l'apparition d'un lymphome *de novo* chez des patients atteints de LLC, en relation avec l'EBV³⁰⁹.

Le variant type lymphome de Hodgkin est rare. L'incidence à 5 et 10 ans est estimée à 0,25 et 0,5 % respectivement³¹². La survie globale médiane est de 4 ans³⁰⁶.



b- Cancers secondaires

Il a été démontré que les patients atteints de LLC ont un risque accru de développer des cancers secondaires. Ceci peut être expliqué par l'atteinte immunologique rencontrée dans la LLC, mais aussi par les chimiothérapies reçues par le patient, une prédisposition génétique aux cancers ou encore un biais dû au suivi clinico-biologique en lien avec la LLC^{313,314}. Ceci est particulièrement marqué pour la survenue de cancers de la peau non-mélanomes, représentant environ un tiers des cancers secondaires, avec un risque 3,66 fois plus important comparé à la population générale³¹⁴⁻³¹⁶. Une équipe s'est intéressée au profil cytogénétique des patients atteints de LLC développant un cancer secondaire³¹⁷. Ils retrouvent un caryotype associant plus de 2 anomalies et une délétion du bras long du chromosome 13 comme facteurs de risque de développer un cancer secondaire en cas de LLC.

c- Complications immunologiques

Des complications auto-immunes peuvent survenir au cours de la maladie. L'incidence est estimée entre 4 et 10 %^{318,319,320}. L'anémie hémolytique auto-immune est la plus fréquente, suivie par la thrombopénie. L'érythroblastopénie et la neutropénie auto-immune sont plus rares avec une incidence inférieure à 1 %³¹⁸. Dans ces deux derniers cas, on observe une destruction auto-immune des précurseurs granuleux et rouge respectivement au niveau de la moelle osseuse. Certains traitements de la LLC peuvent entraîner des cytopénies auto-immunes. Ainsi, la monochimiothérapie à base d'analogue de purine telle que la Fludarabine peut entraîner une anémie hémolytique auto-immune^{321,322}. Les patients ayant déjà reçu plusieurs lignes de traitement sont plus à risque de développer ce type de complication. Le mécanisme est mal connu. Cela pourrait être lié à un déséquilibre sanguin du ratio TH17/Treg³¹⁸. L'alentuzumab a été rapporté comme pouvant entraîner une thrombopénie auto-immune³²³. Les traitements à base de Bendamustine sont également associés à une fréquence augmentée de cytopénies auto-immunes, en particulier l'érythroblastopénie³¹⁸. Plusieurs études montrent une survie globale moins bonne en cas de cytopénie auto-immune^{318,324,325}. Cependant, cela reste controversé avec certaines études ne mettant pas en évidence de différence significative^{320,326}.

Une des principales complications de la LLC est la survenue d'un déficit immunitaire, entraînant un risque infectieux. En particulier, l'hypogammaglobulinémie est le déficit immunitaire le plus fréquent dans la LLC^{318,327}. Des anomalies des lymphocytes T, NK, monocytes et polynucléaires neutrophiles peuvent également contribuer au déficit immunitaire^{318,328,329,330}.



F- Diagnostic

1- Numération-Formule sanguine

La numération-formule sanguine (NFS) met en évidence une hyperlymphocytose. Sur le frottis, on observe des petits lymphocytes monomorphes matures, sans atypie. Ces lymphocytes sont généralement de petite taille, avec un rapport nucléocytoplasmique élevé, un noyau généralement régulier et une chromatine mottée (Figure 31). Il peut exister quelques atypies avec présence d'encoches ou encore quelques formes prolymphocytaires ou de rares para-immunoblastes. La présence d'ombres de Gümprecht est assez typique dans la LLC¹²⁴. Il s'agit alors de lymphocytes lysés au moment de l'étalement sur lame et cela traduit une fragilité cellulaire. Elles ne sont pas spécifiques de la LLC. La présence d'ombres de Gümprecht peut entraîner une sous-estimation de la lymphocytose lors d'une formule manuelle. Le pourcentage d'ombres de Gümprecht a été proposé comme marqueur pronostique^{331,332,333}.

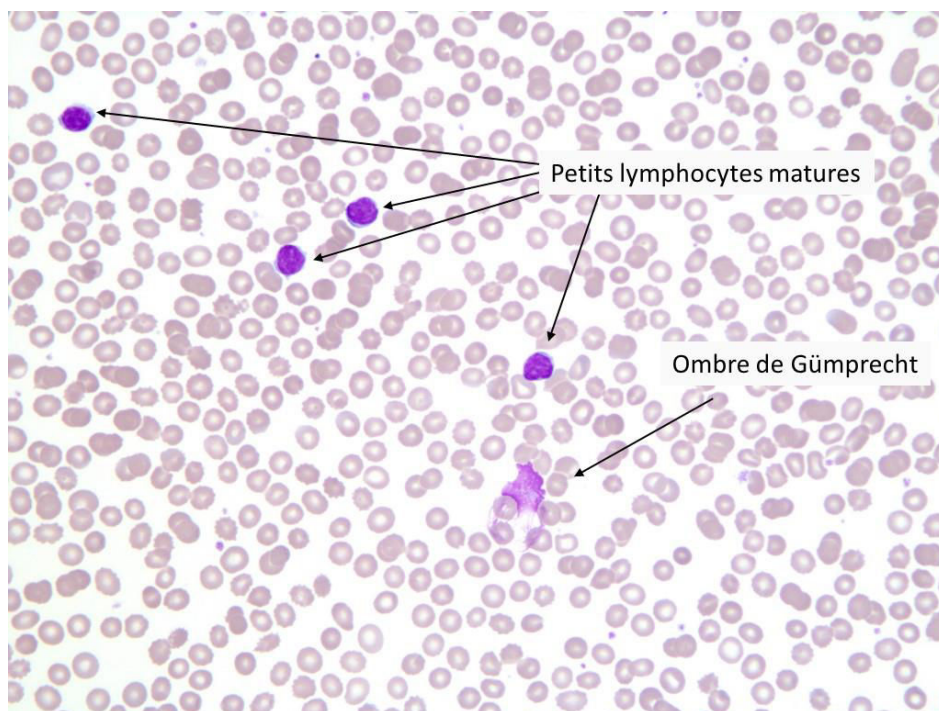


Figure 31 : Image de frottis sanguin coloré au MGG dans un cas de LLC.



2- Cytométrie en flux

La cytométrie en flux retrouve une expansion de lymphocytes B et permet d'établir le score de Matutes-Moreau selon le Tableau 1 : 5 points sont attribués en fonction de la positivité ou négativité de certains marqueurs. On parle de LLC en cas de score de Matutes-Moreau ≥ 4 ^{1,334,335}.

Marqueur	0 point	1 point
CD5	-	+
CD23	-	+
FMC7	+	-
CD22/CD79b	+	- ou faible
Chaîne légère monotypique	Intensité normale ou forte	Intensité faible

Tableau 1 : Calcul du score de Matutes-Moreau

L'expression du CD20 est généralement positive faible. Les cellules de LLC expriment également le CD200, ce qui permet d'identifier certains cas de LLC avec expression aberrante de quelques marqueurs, et de les distinguer par exemple d'un lymphome du manteau dans lequel ce marqueur est négatif³³⁶.

3- Autres

Un test de Coombs direct peut être réalisé : ce test permet de mettre en évidence la présence d'anticorps fixés sur les hématies. Le résultat peut être positif en l'absence d'anémie hémolytique auto-immune, mais dans ce cas, le risque d'en développer une est plus important¹³⁴.

Le myélogramme n'est pas indispensable au diagnostic. En effet la moelle est souvent infiltrée et n'apporte pas d'élément au diagnostic. En revanche, en cas de cytopénies non expliquées, il peut être réalisé afin d'en déterminer la cause (envahissement, autre hémopathie type syndrome myélodysplasique concomitant ou origine périphérique par exemple).



G- Facteurs pronostiques

1- Cytogénétique

a- Introduction

L'impact pronostique de l'analyse cytogénétique dans la LLC a été étudié dès la fin des années 1980^{337,338,339}. La réalisation du caryotype a connu de grandes avancées au cours de cette période. En effet, les cellules de la LLC ont un faible index mitotique rendant leur culture difficile. Jusqu'aux premiers ajouts de mitogènes en 1978 dans les conditions de cultures, très peu d'anomalies chromosomiques pouvaient être identifiées chez les patients atteints de LLC³³⁷. A partir des années 1980, des mitogènes tels que le 12-O-tétradécanoylphorbol-13-acétate (TPA), le lipopolysaccharide (LPS), ou encore le pokeweed mitogen (PKW), TPA+interleukine 2 (IL-2) ont été utilisés, permettant de détecter des anomalies chez 40 à 60 % des patients³⁴⁰. Au cours des années 1990, l'utilisation de l'Hybridation de Fluorescence *in situ* (ou FISH pour Fluorescence *in situ* Hybridation) interphasique a permis d'étudier les principales anomalies mises en évidence dans la LLC telles que la trisomie 12²³⁴. Dans les années 2000, de nouveaux progrès ont permis d'augmenter encore le nombre de patients chez lesquels une anomalie peut être détectée à partir de la réalisation d'un caryotype conventionnel. Ainsi, avec l'ajout de mitogènes tels que le CD40 ligand (CD40L) ou par des CpG-oligodéoxynucléotides (CpG-ODN), en association ou non avec des interleukines (IL-4 et IL-2 respectivement) permet d'identifier des anomalies chez 80 à 90 % des cas de LLC^{340,341,342,343}.

Des anomalies chromosomiques peuvent être mises en évidence dans la LLC. Quatre principales anomalies sont retrouvées. Il s'agit de la délétion du bras court du chromosome 17 (ou del(17p)), la délétion du bras long du chromosome 11 (del(11q)), la trisomie 12 ou la délétion du bras long du chromosome 13 (del(13q)). On peut retrouver au moins une de ces anomalies chez environ 80 % des patients^{134,234}. D'autres anomalies chromosomiques peuvent être retrouvées : délétion du bras long du chromosome 6 (del(6q)), gain 2p par exemple. La majorité des anomalies retrouvées sont des déséquilibres génétiques. On peut cependant retrouver des translocations impliquant les gènes des immunoglobulines avec une fréquence variable selon les études³⁴⁴ et les autres anomalies équilibrées sont très rares. Une évolution clonale avec hétérogénéité intraclonale n'est pas rare. Enfin, dans certains cas, le caryotype est complexe (c'est-à-dire comportant au moins 3 anomalies).



b- Délétion de TP53

La délétion du bras court du chromosome 17 (del(17p)) entraîne la perte du gène *TP53* qui est un gène suppresseur de tumeur. L'intégrité de ce gène est prépondérante pour maintenir la stabilité du génome³⁴⁵. La fréquence de cette anomalie est de 3 à 7 % au diagnostic^{234,346,347} et augmente avec l'évolution de la maladie, en particulier après traitement^{348,349}.

Cette délétion peut concerner l'ensemble du bras court du chromosome 17 ou être cryptique, avec une région minimale de délétion correspondant au locus 17p13.1³⁴⁷. Quelle que soit la présentation au caryotype, une FISH du locus *TP53* doit être réalisée.

Il s'agit d'une anomalie de mauvais pronostic avec une survie globale médiane de 32 mois contre 111 mois en cas de caryotype normal²³⁴.

La proportion de noyaux atteints par l'anomalie a un impact pronostique. Un cut-off à 25 % avait été proposé par l'équipe de Tam *et al.* en 2009 et montrait une survie globale significativement meilleure chez les patients présentant l'anomalie dans moins de 25 % des noyaux³⁵⁰. De même, une autre équipe montre un impact pronostique péjoratif sur la survie sans traitement (ou TFS) d'un nombre de noyaux délétés supérieur à 20 %³⁵¹.

La del(17p) est fréquemment associée à un caryotype complexe^{346,348,352}.

Dans plus de 80 % des cas (jusqu'à 90 %), la del(17p) est associée une mutation du gène *TP53*^{346,348,353}. Seulement 4 à 5 % des LLC sans del(17p) présentent une mutation monoallélique de *TP53*^{348,352,353}.

La délétion de *TP53* est importante car elle est prédictive de la réponse au traitement par analogues puriques tels que la Fludarabine. En effet, la délétion 17p entraîne une résistance à ce type de chimiothérapie^{354,348}. La survie après allogreffe de cellules souches hématopoïétique est également diminuée (survie globale à 3 ans de 25 % contre 59 % en l'absence de del(17p), et survie sans progression à 3 ans de 18 % contre 63 % lorsque l'anomalie n'est pas détectée)³⁵⁵. Sa présence ne constitue pas un critère de traitement mais devrait conduire à un suivi rapproché des patients. En général, la recherche de del17p n'est pas systématiquement réalisée chez les patients au diagnostic mais doit être étudiée au moment où un traitement est envisagé et à la rechute³⁵⁶.



c- Délétion du bras long du chromosome 11

La délétion du bras long du chromosome 11 (ou del(11q)) entraîne la perte du gène *ATM* (pour Ataxia Telangiectasia Mutated). Ce gène est essentiel pour la régulation du cycle cellulaire en activant P53 et en augmentant la réponse aux dommages de l'ADN³⁴⁶. Elle peut également entraîner la perte d'autres gènes situés sur le 11q tels que *BIRC3*. La région minimale de délétion est de 0.68 Mb située en 11q22.3³⁴⁷. La fréquence de l'anomalie est de 10 à 17 %^{234,346,347}.

Les patients présentant une del(11q) ont en général une LLC plus évoluée et plus agressive, avec une survie diminuée (survie globale de 79 mois contre 111 en cas de caryotype normal) et des adénopathies plus marquées²³⁴. La survie sans traitement est également diminuée avec une médiane estimée à 22 mois contre 35 mois en cas de caryotype normal³⁵¹.

Une mutation de l'autre allèle d'*ATM* est retrouvée dans 30 à 40 % des cas avec del11q^{357,358}. Dans ce cas, la survie des patients est moins bonne. Une étude montre que la survie sans progression après une première chimiothérapie est diminuée en cas de double inactivation d'*ATM* par délétion et mutation, et est comparable à la présence d'une anomalie de *TP53* (délétion ou mutation)³⁵⁹.

De plus, une délétion du bras long du chromosome 11 peut entraîner la perte du gène *BIRC3*. Ce gène peut également être muté. La présence d'anomalie de *BIRC3* est rare au diagnostic de LLC et augmente au cours de l'évolution (jusqu'à une fréquence de 24 % chez les patients réfractaires à la fludarabine)³⁶⁰.

d- Trisomie 12

La trisomie 12 est retrouvée à une fréquence de 11 à 16 % au diagnostic dans la LLC^{346,234,347}. Cette anomalie est associée à un pronostic intermédiaire avec une survie globale médiane de 114 mois contre 111 en cas de caryotype normal^{234,361}. La réponse au traitement est généralement bonne, et la fréquence de cette anomalie n'augmente pas dans les groupes de patients réfractaires à la chimiothérapie³⁴⁶.

Les cas de LLC avec trisomie 12 ont souvent une morphologie atypique avec une augmentation du nombre d'aspects prolymphocytaires sur le frottis sanguin, et sont associés à une expression plus forte du FMC7 et de l'immunoglobuline de surface^{362,363}.

La trisomie 12 est trouvée à une fréquence plus élevée dans le lymphome lymphocytaire avec un biais d'utilisation d'IGHV3-21²⁹³.



Cette anomalie est associée aux trisomies 18 et 19 dans 2 à 5 % des cas^{346,364}. La présence des trisomies 12 et 19 est associée aux cas de LLC ayant réalisé le switch IgG et avec réarrangement IGHV4-39³⁶⁵.

La trisomie 12 est associée à la présence de mutation du gène *NOTCH1* dans environ 20 à 30 % des cas et est fréquemment associée à des formes agressives de LLC, avec une survie diminuée^{366,367}. La délétion du bras long du chromosome 14 (del(14q)) est également associée à la trisomie 12 : la présence de ces deux anomalies combinées à une mutation de *NOTCH1* et un statut IGHV non muté (avec biais de représentation d'IGHV1-69) identifie un groupe de mauvais pronostic³⁶⁸.

e- Délétion du bras long du chromosome 13

La délétion interstitielle du bras long du chromosome 13 (del(13q)) est l'anomalie la plus fréquente. Elle est retrouvée dans plus de 50 % des cas de LLC. Cette anomalie peut être cryptique, c'est-à-dire non visible au caryotype. Elle peut apparaître après translocation visuellement équilibrée au caryotype mais entraînant en réalité une del(13q) alors visible par d'autres techniques telles que la FISH. Cette anomalie est de bon pronostic lorsqu'elle est isolée. La survie de ces patients montrait dès 2000 une médiane de survie globale de 133 mois contre 111 pour un caryotype normal²³⁴. La survie sans traitement a été revue sur une large cohorte en 2016 et retrouve une médiane de TFS à 72 mois contre 35 en cas de caryotype normal³⁵¹.

La zone de délétion est située au niveau 13q14, mais la taille est variable³⁶⁹. Il a été montré que cette délétion peut emporter différents gènes. D'une manière générale, elle entraîne la perte des gènes *DLEU2/MIR15A/MIR16-1*. Ce locus est considéré comme la région minimale de délétion. La perte de *MIR15A/MIR16-1* entraîne une levée de surexpression de *BCL2* qui a un rôle anti-apoptotique. De plus larges délétions concernent en plus le gène *RB1*. Dans ce cas, la valeur pronostique de la del(13q) est différente, la TFS étant significativement moins importante en cas de délétion concomitante de *RB1*^{370,371}. Il a été proposé de classer en différents groupes de del(13q) suivant les points de cassure et les gènes impliqués³⁷¹. Ainsi la del(13q) de type I entraîne la perte du locus *DLEU2/MIR15A/MIR16-1* avec des points de cassures pouvant correspondre à la région minimale de délétion (type Ia) ou plus larges (type Ib), alors que la del(13q) de type II entraîne en plus la perte de *RB1*.

La del(13q) peut être mono-allélique ou bi-allélique, sans que cela change l'impact pronostique. Les del(13q) de type II sont généralement mono-alléliques³⁷¹. En revanche, la quantité de noyaux atteints par la del(13q) a un impact pronostique péjoratif lorsqu'elle est très élevée : avec un seuil fixé à 65,5 % de noyaux atteints par l'anomalie, la TFS à 5 ans est de 79 % quand le pourcentage est inférieur, et seulement de 38 % lorsque plus de 65,5 % des noyaux sont touchés³⁵¹. Une autre étude montre le même impact pronostique péjoratif avec un seuil de noyaux atteints fixé à 80 %³⁷².



f- Délétion du bras long du chromosome 6

La délétion du bras long du chromosome 6 (ou del(6q)) est retrouvée dans environ 6 % des cas de LLC²³⁴. La fréquence varie selon l'évolution de la maladie et serait un marqueur de progression de la maladie^{347,373}.

g- Gain 2p

Les anomalies avec gain 2p sont principalement mises en évidence grâce aux nouvelles techniques de cytogénétique moléculaire telles que la CGH-array^{374,347}. Les gains 2p entraînent une amplification des gènes *MYCN*, *cREL*, *MSH2* ou *ALK*^{374,375}. La région minimale d'amplification se situe en 2p16.1-2p15³⁷⁶. La fréquence de cette anomalie varie selon les études de 2 à 8 %^{374,347,376,375}.

Les cas de LLC présentant cette anomalie sont associés à la présence d'une del(11q) et au statut IGHV non muté^{374,347}. Cette anomalie serait considérée comme un potentiel marqueur de progression³⁷⁴.

h- Caryotype complexe

Le caryotype complexe est défini comme la présence d'au moins 3 anomalies chromosomiques³⁷⁷. La fréquence du caryotype complexe dans la LLC au diagnostic est de 16 % selon Baliakas *et al.*³⁷⁸. Cette fréquence est variable et semble augmenter avec l'évolution de la maladie³⁷⁷.

Le caryotype complexe est associé aux délétions 17p et 11q, ce qui peut expliquer en partie le mauvais pronostic retrouvé et l'accumulation d'anomalies cytogénétiques^{379,378}. Il est également associé au statut IGHV non muté et à l'expression du CD38³⁷⁸. La valeur pronostique du caryotype complexe est encore débattue, mais semble selon plusieurs études avoir un impact pronostique péjoratif en termes de survie sans progression et de survie globale^{380,379,378}.



i- Translocations

Les translocations sont rapportées dans environ 20 % des cas de LLC³⁸¹. Ces translocations impliquent principalement les gènes des immunoglobulines, majoritairement *IGH*, et le locus 13q14.

Les translocations impliquant les gènes des immunoglobulines sont relativement rares dans la LLC (4 à 9 %) ^{377,344}. Les partenaires récurrents sont *CCND1*, *BCL2* et *BCL3*³⁸¹. Dans la majorité des cas, le caryotype présente d'autres anomalies associées. Les pertes 3q et gains 8p semblent être associés aux translocations impliquant les gènes des immunoglobulines. Actuellement entre 30 et 40 loci partenaires sont décrits, mais les gènes impliqués ne sont que rarement identifiés³⁴⁴. Ces cas présentent généralement une cytologie atypique avec des aspects lymphoplasmocytaires et une présence de prolymphocytes plus fréquente³⁴⁴. Le pronostic associé aux translocations impliquant les gènes de immunoglobulines est généralement péjoratif, excepté lorsque le partenaire est *BCL2* pour lequel le pronostic n'est pas modifié³⁴⁴. Ces translocations entraînent généralement une dérégulation du gène partenaire. Une exception est décrite avec la formation d'un transcrite de fusion au cours d'une translocation t(12 ;14)(q23;q32) ayant pour conséquence une anomalie de distribution cellulaire de la sulfotransférase impliquée³⁸². Egalement, des translocations impliquant deux gènes d'immunoglobulines ont été rapportés (t(14 ;22)(q32;q11))^{344,383}. Le mécanisme pathogénique de ce type de translocation est actuellement inconnu. La dérégulation d'un oncogène à proximité est possible.

Les translocations impliquant le locus 13q14 sont quasiment toujours associées à une délétion du locus au cours de la translocation³⁸¹. Il n'y a pas de partenaire récurrent retrouvé dans ces cas.

j- Autres

Les télomères sont des structures protectrices situées aux extrémités des chromosomes, composés de répétitions riches en G et associées à un complexe protéique appelé « shelterin »³⁸⁴. La longueur des télomères chez l'Homme varie de 10 à 15 kb³⁸⁵. A chaque division cellulaire, les télomères sont raccourcis avec perte de répétitions télomériques, à cause d'une réplication incomplète. Ceci est à l'origine de la sénescence et des cancers³⁸⁶. Des télomères raccourcis non protégés risquent d'être fusionnés par erreur par le système de réparation de l'ADN, entraînant des aneuploïdies ou des réarrangements chromosomiques structuraux³⁸⁴. Par conséquent, des anomalies des télomères peuvent entraîner une instabilité chromosomique au cours de la cancérogénèse. Les télomères peuvent être régénérés par la télomérase et sont protégés par la fusion des protéines « shelterin ».



Dans la LLC, des télomères raccourcis sont associés à un mauvais pronostic et une fréquence augmentée d'anomalies cytogénétiques péjoratives telles que la del(17p), la del(11q) ou un caryotype complexe^{387,388,389}. Ce serait également associé à la présence de mutations de *SF3B1* et *NOTCH1*³⁹⁰. L'identification de télomères courts au diagnostic serait associée à des survies globale et sans traitement diminuées³⁹⁰. L'érosion des télomères fait partie de l'évolution naturelle de la LLC et sa rapidité est un marqueur d'agressivité de la maladie³⁹¹. Les patients présentant une anomalie de *TP53* ont une dysfonction importante des télomères³⁹². La longueur des télomères est inversement corrélée à l'activité de la télomérase^{387,393}. Ainsi, une forte activité de la télomérase est associée à une survie raccourcie dans la LLC.

Un raccourcissement des télomères est associé à un statut IGHV non muté^{394,395,396,397}. Les patients IGHV non mutés avec des télomères courts auraient une survie comparable à celles des patients IGHV non mutés³⁹⁸.

Des mutations du gène *POT1* (pour Protection Of Telomeres 1), qui code un composant du complexe des protéines shelterin, entraînent des anomalies des télomères et chromosomiques³⁹⁹. Ces mutations sont retrouvées chez 3,5 % des patients atteints de LLC, et ont une fréquence augmentée chez les patients IGHV non mutés. Ces mutations ont également été décrites dans des cas de LLC familiales⁴⁰⁰.

k- Evolution clonale

En cytogénétique, l'évolution clonale a été étudiée par FISH principalement et montre l'acquisition de marqueurs de mauvais pronostic tels que la del(17p) et la del(11q) dans 65 % des cas⁴⁰¹. Certaines anomalies cytogénétiques sont considérées comme initiatrices telles que la del(13q) hétérozygote ou la trisomie 12, tandis que d'autres semblent acquises au cours de l'évolution (del(13q) homozygote et gain 2p par exemple)^{344,402}. La chimiothérapie ne semble pas être un facteur entraînant l'apparition de nouveaux clones. L'évolution clonale est associée à une survie globale diminuée, essentiellement en cas d'apparition de marqueurs de haut risque⁴⁰¹.



2- Génétique de la partie variable des chaînes lourdes des immunoglobulines (IGHV)

a- Statut mutationnel IGHV

A partir des années 1990, différentes équipes se sont intéressées au statut mutationnel IGHV. Cette caractéristique permet de mieux comprendre la cellule d'origine dans la LLC, un statut IGHV muté traduisant un passage dans le GC. Le séquençage des régions variables de la chaîne lourde des immunoglobulines peut mettre en évidence des polymorphismes⁴⁰³. Aussi, un seuil de 98 % d'homologie a été choisi pour séparer les cas.

Ce seuil de 98 % d'homologie reste débattu. Encore récemment, une publication rapporte que le taux de mutation en tant que variable continue est plus important que le seuil de 98 %⁴⁰⁴.

Les premières études rapportent un statut IGHV muté dans environ 50 % des cas^{95,405}.

Le statut IGHV s'est révélé être un marqueur pronostique important. Ainsi, un statut IGHV non muté est associé à une maladie plus agressive et un pronostic péjoratif⁴⁰⁶. A l'inverse, un statut IGHV muté semble prédire une maladie plus indolente. La médiane de survie était de 8-9 ans en cas de statut IGHV non muté contre 24 ans voire non atteinte dans le groupe IGHV muté^{406,407}. Une exception concerne le réarrangement IGHV3-21, généralement présentant à un statut muté, mais entraînant une survie semblable à celle du groupe IGHV non muté⁴⁰⁸. La fréquence de la trisomie 12 est également plus élevée dans ce groupe^{406,407,409}. Les patients présentant un réarrangement IGHV3-21 ont un CDR3 plus court et expriment généralement une chaîne légère lambda⁴¹⁰. De plus, l'étude du réarrangement du locus *IGL* montre une utilisation hautement restreinte de V λ 2-14/J λ 3. Ceci soutient l'hypothèse d'une implication de la reconnaissance de l'antigène dans la survenue de la LLC. Cependant, il semblerait que plus le réarrangement IGHV3-21, ce soit l'appartenance à un sous-groupe (ou subset) particulier qui modifie la survie⁴¹¹.

Par ailleurs, l'étude de la séquence IGHV dans différents clones ne montre généralement pas de différence notable⁴⁰⁶. De plus, l'étude au cours du temps chez un même patient ne montre pas de modification de la séquence IGHV. Ceci indique l'absence d'acquisition de mutations « ongoing » dans la LLC. Une exception est décrite dans un sous-groupe de cas dans lequel a été mise en évidence avec une grande diversité intraclonale, en faveur d'une interaction active continue avec l'antigène⁴¹².

L'expression du CD38 a été étudiée en parallèle du statut mutationnel IGHV : les patients IGHV non muté sont généralement CD38+⁴⁰⁷.



b- Réarrangements IGHV et définition de sous-groupes ou

« subset »

L'étude des réarrangements VDJ montre que 10 gènes VH sont utilisés par plus de la moitié des patients atteints de LLC, en particulier IGHV1-69, IGHV3-23 et IGHV4-34 sont les plus fréquents (environ 12 % chacun), suivis de IGHV1-02, IGHV3-30.3 et IGHV3-07⁴⁰⁶. Les fréquences d'utilisation de ces gènes peuvent correspondre pour certains à la fréquence d'utilisation observée dans des cellules B normales (IGHV3-23 et IGHV3-30.3 par exemple), tandis que d'autres sont surreprésentés dans la LLC tels qu'IGHV1-69. Concernant les gènes JH, la majorité des cas de LLC réarrange avec JH4 (plus de 40 % des cas) ou JH6 (plus de 35 %) ⁴⁰⁶. Ceci est comparable au profil observé dans des cellules B normales. Concernant le DH réarrangé, le plus fréquent est DH3-3, et aucune association avec des VH ou JH n'a été mise en évidence⁴⁰⁶.

Il existe des profils de réarrangements VDJ différents suivant le statut mutationnel IGHV, certains réarrangements étant associés à un statut mutationnel donné. C'est le cas par exemple d'IGHV1-69 qui est généralement associé à un statut non muté⁴¹³.

L'étude de réarrangements V(D)J des chaînes lourdes et légères chez les patients atteints de LLC montre l'expression d'une immunoglobuline identique dans environ 1 % des cas⁴¹³.

Des études aux Etats-Unis et en Europe sur les profils d'utilisation des gènes des immunoglobulines ont été réalisées^{414,96,415}. Elles révèlent des groupes de cas de LLC portant des BCR hautement similaires entre les patients : on parle de BCR stéréotypés. Les critères actuels pour parler de BCR stéréotypés sont un réarrangement VDJ de la chaîne lourde des immunoglobulines présentant un gène *IGHV* du même clan (I, II ou III), et des séquences du CDR3 de taille identique avec certains acides aminés conservés à des positions exactes^{416,417}. De plus, la séquence en acides aminés doit être identique à au moins 50 %, et avec des propriétés physicochimiques semblables à plus de 70 %. Le clan I regroupe les VH1, VH5 et VH7, le clan II les VH2, VH4 et VH6 et le clan III comprend les VH3⁴¹⁶. Le fait de retrouver de tels groupes ou « subsets » chez des sujets géographiquement éloignés et non apparentés soutient le rôle d'une sélection par l'antigène. Ainsi, il a été proposé que les cas avec BCR stéréotypés dériveraient de progéniteurs sélectionnés par l'antigène tandis que les autres cas, dits « hétérogènes » auraient pour origine des cellules B conventionnelles⁴¹⁸.

L'identification de BCR stéréotypés a permis d'identifier différents subsets. Une large étude menée en 2012, incluant 7596 cas de LLC, définit 19 subsets dits majeurs pour lesquels au moins 20 séquences sont retrouvées⁴¹⁷. Ces subsets majeurs regroupent 41 % des cas stéréotypés et 12 % de l'ensemble de la cohorte examinée. D'une manière générale, environ un tiers des cas expriment un BCR stéréotypé⁴¹⁷. Une étude en cours avec l'analyse des séquences d'environ 21 000 cas révèle des proportions similaires⁴¹⁹. La distinction de subsets sur la base de la séquence du CDR3 permet aussi de séparer les cas de LLC avec des caractéristiques clinico-biologiques comparables^{415,420}. Par exemple, le subset 4 est



caractérisé par un âge au diagnostic diminué et une évolution indolente, tandis que le subset 8 regroupe des patients à risque de transformation vers un syndrome de Richter (Tableau 2)^{420,421}.

	Subset 1	Subset 2	Subset 4	Subset 8
Fréquence dans la LLC	2,4 %	2,8 %	1 %	0,5 %
Réarrangements IGHV	IGHV1/5/7 IGHD6-19 IGHJ4	IGHV3-21 IGHD non identifié IGHJ6	IGHV4-34 IGHD5-18 IGHJ6	IGHV4-39 IGHD6-13 IGHJ5
Statut mutationnel IGHV	IGHV non muté	IGHV muté et non muté	- IGHV muté - SHM « ongoing »	IGHV non muté
Caractéristiques génétiques	Mutations récurrentes de <i>NFKB1E</i>	- Del(11q) - Mutations fréquentes de <i>SF3B1</i> - Mutations peu fréquentes de <i>TP53</i>	Peu d'anomalies génétiques	- Mutations de NOTCH1 - Trisomie 12
Evolution (survie sans traitement)	Très agressive (1,9 ans)	Très agressive (1,6 ans)	Très indolente (11 ans)	Très agressive (1,5 ans) Transformation en Richter

Tableau 2 : Caractéristiques principales des principaux subset dans la LLC, adapté d'après Stamatopoulos *et al.*⁴¹⁹.

Les subsets définissent des groupes pronostiques distincts. Ainsi, le subset 2, qui représente 2,8 % des LLC, comprend des cas à statut IGHV muté ou non muté, avec une fréquence élevée de del(11q), de mutations de *SF3B1* et peu d'anomalies de *TP53*, et est associé à une maladie agressive (survie sans traitement de 1 à 6 ans)⁴¹⁹. A l'inverse, le subset 8, représentant 1 % des LLC, est associé à un statut IGHV muté, peu d'anomalies génétiques, et une maladie indolente (survie sans traitement de 11 ans)⁴¹⁹.



3- Mutations récurrentes hors *IGHV*

a- *TP53*

Le gène *TP53* peut être altéré soit par délétion, soit par mutation inactivatrice. Les mutations de *TP53* sont retrouvées chez 90 % des cas avec del(17p)³⁵³. A l'inverse, environ 60 à 70 % des patients ayant une mutation de *TP53* ont également une del(17p). Environ 5 % des cas de LLC non traitée sans del(17p) ont une mutation de *TP53*^{348,353}.

Les anomalies de *TP53* sont associées à un mauvais pronostic, avec une survie diminuée et une résistance aux chimiothérapies à base d'analogues de purines tels que la Fludarabine^{348,422}. Dans ce contexte, il est important de rechercher la présence de délétion ou de mutation avant l'instauration de toute chimiothérapie, dans la mesure où un traitement conventionnel par Fludarabine risque d'entraîner la sélection du clone porteur d'anomalie de *TP53*^{423,424,425,353}. Dans certains cas, la présence de mutation de *TP53* est sous-clonale et peut ne pas être détectée par séquençage type Sanger par manque de sensibilité³⁵³. De plus, les mutations de ce gène sont réparties sur l'ensemble du gène, rendant difficile son séquençage par technique Sanger en routine pour tous les patients atteints de LLC. Le développement du NGS a permis de résoudre ces difficultés en permettant de diminuer le seuil de sensibilité et d'étudier un plus grand nombre de cibles en même temps. Il est actuellement recommandé de rechercher les mutations de *TP53* par NGS avec une profondeur suffisante (au moins 100 reads sur l'ensemble du gène)³⁵³. Certains variants introniques proches des exons (+2 ou -2) pouvant toucher les sites d'épissage sont considérés comme pathogéniques³⁵³. De plus, des mutations synonymes peuvent tout de même avoir des conséquences sur l'épissage et leur signification doit être vérifiée sur les bases de données (en particulier IARC TP53 et TP53 website)^{426,427}. En effet, ce type d'anomalie a été décrit dans certaines familles atteintes de Li-Fraumeni^{428,353}.

b- *ATM*

Le gène *ATM* est situé en 11q22.3. Il est de grande taille (plus de 146 kb) et comprend 69 exons dont 62 codants. Il s'agit d'un suppresseur de tumeur dont la fonction est de contrôler l'intégrité de la réparation de l'ADN et de réguler la progression dans le cycle cellulaire. *ATM* est la principale protéine activatrice de p53 en réponse aux cassures double brin de l'ADN. Il a été démontré dans la LLC que la réponse transcriptionnelle aux cassures double brin est entièrement dépendante d'*ATM* et seulement une partie de cette activité est transférée à la voie p53, aboutissant à un signal pro-apoptotique⁴²⁹. Etant donnée la taille du gène et l'absence de mutation récurrente dans la LLC, son séquençage n'était pas accessible par technique conventionnelle type Sanger. L'apparition du NGS permet dorénavant d'avoir des données mutationnelles chez un grand nombre de patients concernant ce gène.



Les mutations d'*ATM* sont retrouvées chez 10 à 16 % des patients atteints de LLC^{430,431,432,425}.

Comme pour la délétion, la présence de mutation d'*ATM* est de mauvais pronostic avec une survie globale et sans traitement raccourcies^{432,425}. Les mutations d'*ATM* sont associées à un statut IGHV non muté^{432,425}.

Comme pour *TP53*, ce gène peut être délété et/ou muté. La présence de mutation d'*ATM* en cas de del(11q) est retrouvée dans 30 à 40 % des cas^{357,358,430,431}. Les patients présentant une atteinte biallélique d'*ATM* (mutation et délétion ou double mutation) perdent la fonctionnalité d'*ATM*, tandis qu'une atteinte monoallélique permet une fonction conservée de la protéine^{432,357}. Il a été montré que les cas avec inactivation biallélique d'*ATM* sont réfractaires *in vitro* aux chimiothérapies cytotoxiques³⁵⁷. Les mutations d'*ATM* peuvent survenir au cours de l'évolution en cas de del(11q) préexistante³⁵⁷. Il a été démontré que la présence de mutation germinale hétérozygote d'*ATM* est plus fréquente en cas de del(11q), ce qui suggère un rôle dans la progression de la maladie mais pas dans son initiation⁴³³. L'inactivation biallélique est de plus mauvais pronostic avec une survie sans progression inférieure (7,4 mois en cas de del(11q) et mutation, contre 17,1 mois en cas de del(11q) seule et 30,8 mois pour mutation isolée)³⁵⁹. La survie globale est alors significativement diminuée et est comparable à celle des patients présentant une délétion de *TP53*⁴³⁴.

Différents types de mutations sont susceptibles d'avoir des impacts différents. Dans l'Ataxie-Télangiectasie, maladie génétique autosomale récessive dans laquelle *ATM* est muté, une hétérogénéité clinique est retrouvée et pourrait être attribuable aux différents types de mutations rencontrées⁴³⁰. En effet, les mutations tronquantes sont associées à un phénotype d'Ataxie-télangiectasie complet, tandis que les mutations faux-sens ou affectant les sites d'épissage entraînent une expression clinique plus modérée, avec une expression et une fonction conservées d'*ATM*^{435,436}. Il est encore actuellement parfois difficile de connaître l'impact d'une mutation d'*ATM* étant donnée la variété de mutations rencontrées, certaines pouvant être germinales ou somatiques, avec un impact différent sur la fonctionnalité de la protéine. Certaines mutations germinales peuvent correspondre à des SNP ou au contraire avoir un impact et les données restent vagues. Leur interprétation biologique reste difficile. Ainsi, des tests *in vitro* ont été mis au point pour essayer d'étudier l'impact des mutations sur la fonction d'*ATM*. Par exemple, plusieurs équipes utilisent des méthodes basées sur le suivi de l'accumulation de p53 et p21 après exposition des cellules à des radiations ionisantes^{357,437,438}. D'autres tests reposent sur l'utilisation d'étoposide et de nutlin-3a afin de mettre en évidence des défauts dans les voies ATM et TP53⁴³⁹. Le traitement des cellules par Doxorubicine ou Fludarabine avec suivi de l'expression de p21 a été proposé⁴³⁰. Enfin, une technique basée sur la mesure de l'expression de gènes sous la dépendance d'*ATM* ou TP53 par RT-MLPA (pour Reverse Transcription- Multiplex Ligation dependent Probe Amplification), après traitement par radiations ionisantes, a été proposée⁴⁴⁰. Ainsi, l'avènement du NGS, l'accumulation de données cliniques et l'utilisation de ces tests fonctionnels permettront à l'avenir de mieux connaître et interpréter ces mutations.



c- *NOTCH1*

NOTCH1 (situé en 9q34.3) code un facteur de transcription activé par ligand régulant différentes voies de signalisation impliquées dans la différenciation de progéniteurs hématopoïétiques en cellules T immatures et de lymphocytes B matures en cellules productrices d'anticorps^{441,442}.

Des mutations activatrices de *NOTCH1* sont retrouvées dans environ 10 % des cas de LLC au diagnostic et jusqu'à 15-20 % chez les patients évolutifs ou en rechute⁴⁴³. Dans les cas de Richter, la fréquence augmente jusqu'à 30 %^{443,444}. Les patients mutés pour *NOTCH1* présentent plus fréquemment des caractéristiques biologiques péjoratives telles qu'un statut IGHV non muté, des anomalies de *TP53* et une trisomie 12^{443,445,366,367,446}.

La principale mutation retrouvée est une délétion de 2 nucléotides (délétion CT : c.7541_7542delCT, p.P2514Rfs, NM_017617)⁴⁴⁴. Elle représente environ 77 % des mutations de *NOTCH1*¹⁶³. Il s'agit de mutations activatrices entraînant la génération d'un codon stop prématuré et la formation d'une protéine NOTCH1 sans le domaine C terminal, contenant un domaine PEST (une séquence riche en proline, acide glutamique, sérine et thréonine). Cela entraîne l'accumulation d'une isoforme protéique active⁴⁴⁴. Il a été montré que NOTCH1 est constitutivement exprimée dans la LLC, mais la présence de ce type de mutation génère une isoforme protéique plus stable et active⁴⁴⁷.

La présence de mutation de *NOTCH1* au diagnostic est un marqueur indépendant de mauvais pronostic⁴⁴³. Les patients mutés ont stade clinique plus avancé au diagnostic et une survie globale plus courte^{444,445}. La survie sans traitement est également raccourcie⁴⁴⁶. Le risque de transformation en Richter est significativement supérieur^{444,445}.

d- *SF3B1*

SF3B1 (pour Splicing Factor 3B subunit 1) est un gène situé en 2q33.1 et code une protéine importante dans la machinerie de l'épissage. Il s'agit de mutations faux-sens ou de délétion n'entraînant pas de modification du cadre de lecture⁴⁴⁸. La principale mutation récurrente retrouvée touche le codon 700 (mutation K700E)⁴⁴⁹. Le fait que ces mutations soient localisées dans un domaine spécifique (domaines HEAT) suggère un rôle fonctionnel spécifique⁴⁴⁹. Etant donné que SF3B1 est un régulateur de l'épissage de gènes impliqués dans la progression dans le cycle cellulaire et l'apoptose, il est probable que ces mutations jouent un rôle dans la prolifération et/ou la survie cellulaire¹⁶³.



Des mutations de *SF3B1* sont retrouvées chez environ 10 % des cas de LLC au diagnostic et 17 % en cas de maladie progressive^{163,449,448,206}. Elles semblent être acquises au cours de l'évolution de la maladie^{424,450}. La proportion de sous-clones porteurs de ce type de mutation augmente avec le temps, indépendamment de l'administration de chimiothérapies⁴⁵⁰.

Ces mutations sont associées à la délétion du bras long du chromosome 11 et sont mutuellement exclusives des mutations de *TP53*^{449,448,449}. Elles sont plus fréquemment associées à un statut IGHV non muté^{449,450,446}.

Les mutations de *SF3B1* sont associées à un pronostic péjoratif. Les patients mutés ont une maladie plus agressive et une survie globale diminuée²⁰⁶. La survie sans traitement est significativement réduite⁴⁴⁹.

e- BIRC3

BIRC3 est un gène situé en 11q22.2, à 6 Mb d'*ATM*, et code une protéine membre de la famille IAP, capable d'inhiber l'apoptose en se liant à TRAF1 et TRAF2 (pour Tumor necrosis factor receptor-associated factors 1 et 2). *BIRC3* peut réguler la voie NFκB. Elle active la voie de signalisation NFκB canonique, ce qui peut promouvoir la croissance et la survie des cellules de LLC, et inhibe la voie de signalisation NFκB non canonique¹⁶³. Il a été démontré que les cellules de LLC porteuses de mutation de *BIRC3* présentent une activation constitutive de NFκB et sont moins sensibles à la chimiothérapie³⁶⁰.

Les mutations retrouvées seraient inactivatrices et entraîneraient la génération de transcrits aberrants présentant des codons stop prématurés, responsables de l'élimination ou de troncature du domaine RING C-terminal³⁶⁰.

Ce gène peut également être altéré par délétion. Généralement, les anomalies de *BIRC3* sont monoalléliques³⁶⁰. L'inactivation bi-allélique a été retrouvée dans 7,5 % des cas avec délétion du bras long du chromosome 11⁴⁵¹. Dans ce cas, la survie sans traitement est significativement diminuée.

Les mutations de *BIRC3* sont peu fréquentes au diagnostic de la LLC, et la fréquence augmente chez les patients réfractaires à la chimiothérapie par Fludarabine³⁶⁰.

Une anomalie de *BIRC3* (soit mutation, soit délétion) est mutuellement exclusive d'une anomalie de *TP53*³⁶⁰. Elle est associée à la délétion du bras long du chromosome 11 (avec ou sans délétion de *BIRC3*) et un statut IGHV non muté⁴⁴⁶.

La présence d'anomalie de *BIRC3* est associée à un pronostic péjoratif indépendamment des autres marqueurs pronostiques tels que les anomalies de *TP53*³⁶⁰. La survie globale est réduite³⁶⁰.



f- Autres mutations

Les études de séquençage d'exome ont permis de mettre en évidence de nombreuses mutations récurrentes dans la LLC à des fréquences plus faibles (inférieures à 10 %)⁴⁴⁴.

Des mutations de *MYD88* (pour Myeloid Differentiation primary response 88), gène situé en 3p22.2, ont été mises en évidence dans 3 à 10 % des cas de LLC^{444,449}. Il s'agit le plus souvent d'une mutation ponctuelle récurrente (p.L265P). Cette mutation est également décrite dans différents types de lymphomes⁴⁵². En particulier, elle est retrouvée de façon récurrente dans la maladie de Waldenström^{453,454}. La protéine MYD88 participe aux voies de signalisation de l'interleukine-1 et des TLR. au cours de la réponse immunitaire⁴⁵⁵. Il s'agirait de mutations activatrices⁴⁵². Bien qu'initialement décrite dans des lymphomes sécrétant une immunoglobuline monoclonale d'isotype IgM, cette mutation n'est pas associée à la présence d'un pic dans la LLC^{456,457}. Les patients atteints de LLC porteurs de cette mutations sont plus jeunes, souvent âgés de moins de 50 ans^{444,458}. Cette mutation est associée à un statut IGHV muté^{444,458}. Il a été décrit un association avec un caryotype complexe⁴⁵⁹. Certains auteurs rapportent une maladie cliniquement plus avancée au diagnostic en cas de mutation, sans différence significative sur la progression ou la survie^{444,456,459}. De plus, parmi les cas de LLC IGHV muté, la présence de mutation MYD88 serait associée à une survie sans traitement raccourcie⁴⁵⁶. D'autres rapportent une survie globale comparable à celle de la population générale⁴⁵⁸. L'impact de cette mutation sur la survie reste controversé⁴⁶⁰.

Le gène *DDX3X* (pour DEAD-box helicase 3 X-linked), situé en Xq11, est muté chez environ 10 % des cas de LLC non traitées avec maladie progressive⁴⁶¹. Il code une hélicase impliquée dans différentes étapes de modification des ARN dont la transcription, l'initiation de la traduction, l'épissage et l'export des ARNm. Ces mutations sont plus fréquentes en cas de statut IGHV non muté. Elles sont de type tronquante avec des mutations non-sens ou des insertions-délétions entraînant un décalage du cadre de lecture. Elles entraînent une perte d'expression de la protéine. Ces mutations sont plus fréquentes chez les patients en rechute après traitement (jusqu'à 17 %)⁴⁶¹. Elles sont associées à une survie sans progression raccourcie⁴⁶².

Des mutations inactivatrices de *NFKBIE* (pour NFκB Inhibitor Epsilon), situé en 6p21 ont été décrites comme récurrentes dans la LLC. La principale mutation retrouvée est une délétion de 4 pb⁴⁶³. Ces mutations sont susceptibles d'entraîner la formation d'une protéine tronquée et entraîneraient une activation de la voie NFκB⁴⁶⁴. Elles sont retrouvées à une fréquence d'environ 7 %, avec une augmentation dans les cas de stade avancé^{463,465,466}. Elles sont associées à un pronostic péjoratif⁴⁶³.



MED12 (pour Mediator complex subunit 12), situé en Xq13.1, a récemment été identifié comme muté dans la LLC à une fréquence d'environ 7 %⁴⁶⁷. Ces mutations sont mutuellement exclusives avec celles de *NOTCH1*. De plus, l'expression de *NOTCH1* est augmentée en cas de mutation de *MED12*. Ces mutations sont associées à un statut IGHV non muté.

Le gène *EGR2* (pour Early Growth Factor 2), situé en 10q21, a été retrouvé muté dans différentes études^{465,468}. Ces mutations sont retrouvées dans environ 4 % des cas de LLC⁴⁶⁸. Leur fréquence augmente dans les cas de LLC de stade avancé⁴⁶⁵. Elles sont associées à un âge au diagnostic plus faible et à une survie sans traitement et une survie globale raccourcies^{465,468}. Elles entraînent un pronostic défavorable comparable à celui conféré par les altérations de *TP53*⁴⁶⁸. Elles ont été retrouvées dans des précurseurs B de patients atteints de LLC, ce qui suggère un rôle initiateur de ces mutations⁴⁶⁵. Elles sont associées à un statut IGHV non muté, des altérations d'*ATM*, de *TP53* et *NOTCH1*⁴⁶⁸.

Le gène *XPO1* (pour Exportin 1 gene), situé en 2p15, a également été retrouvé muté dans la LLC. Il code une protéine impliquée dans l'export nucléaire de protéines et d'ARNm. Les mutations retrouvées touchent un codon particulier (p.E571) et sont situées dans une région hautement conservée, ce qui suggère que ces mutations ont un impact sur l'activité de la protéine⁴⁴⁴. Elles sont retrouvées dans 1 à 5 % des cas de LLC^{444,366,469}. Elles sont associées à un statut IGHV non muté^{444,366}. Elles sont également associées aux mutations de *NOTCH1*, *SF3B1* et *TP53*^{444,470}. Il ne semble pas y avoir d'impact sur la survie des patients⁴⁷⁰. Les mutations de ce gène sont intéressantes dans la mesure où il existe un inhibiteur d'*XPO1* (le selinexor) qui pourraient être utilisables en thérapeutique chez ces patients.

Le gène *BRAF* est un proto-oncogène fréquemment muté dans les mélanomes et certains cancers de la thyroïde. Il est également quasiment constamment muté dans la leucémie à tricholeucocytes, un type de LNH-B⁴⁷¹. Des mutations activatrices ont également été décrites dans la LLC. La principale mutation retrouvée est la V600E au niveau de l'exon 15. Quelques autres mutations sont rapportées dans l'exon 11. Ces mutations entraînent la perte de la conformation inactive de la protéine, ce qui entraîne une activation de la voie des MAPK (pour Mitogen Activated Protein Kinases)⁴⁷². La fréquence de ces mutations est d'environ 3-4 % dans la LLC^{473,465}. Elles seraient retrouvées à une fréquence plus élevée dans les syndrome de Richter et pourraient constituer une cible thérapeutique avec l'utilisation des inhibiteurs de *BRAF*^{V600E}, déjà utilisés dans le traitement du mélanome et de la leucémie à tricholeucocytes réfractaire⁴⁷⁴.



Le gène *FBXW7*, situé en 4q31, code une ubiquitine ligase et est un suppresseur de tumeur. Ces mutations entraîneraient une activation de la voie Notch⁴⁷⁵. Il a été retrouvé muté dans 3-4 % des cas de LLC^{449,459}. Ces mutations sont associées à la trisomie 12⁴⁴⁹. Elles sont mutuellement exclusives des mutations de *NOTCH1*^{449,476,469}. Elles sont retrouvées de façon clonale et ne semblent pas sélectionnées par la chimiothérapie⁴⁷⁷. De plus, elles ont été identifiées au niveau de progéniteurs hématopoïétiques dans la LLC, suggérant un rôle initiateur⁴⁷⁸. Elles seraient associées à un risque de progression plus élevé⁴⁷⁹.

Le gène *POT1*, situé en 7q31, intervient dans la protection des télomères. Des mutations de ce gène ont été retrouvées chez 3,5 à 5 % des patients atteints de LLC^{399,480}. Elles sont associées à la présence d'anomalies télomériques et chromosomiques, ce qui suggère son rôle dans l'acquisition d'événements oncogéniques³⁹⁹. De plus, il s'agit d'une anomalie clonale avant traitement, ce qui indique un rôle précoce dans l'initiation de la maladie⁴⁷⁷. Une étude montre un impact pronostique péjoratif des mutations de *POT1*⁴⁸¹. Certaines de ces mutations ont été rapportées dans les formes familiales de LLC⁴⁰⁰.

SAMHD1, situé en 20q11, a été retrouvé muté de façon récurrente dans la LLC⁴⁸². Ces mutations sont retrouvées à une fréquence faible (inférieure à 5 %)^{482,483}. Leur fréquence augmente dans les cas de LLC réfractaires⁴⁸². *SAMHD1* régulerait la prolifération cellulaire et la survie et serait impliqué dans des interactions protéiques spécifiques en réponse aux dommages de l'ADN⁴⁸². Il s'agirait d'un gène suppresseur de tumeur⁴⁸³. Ces mutations seraient un événement fondateur dans la LLC^{484,424}.

RPS15 (pour Ribosomal Protein S15) est un gène situé en 19p13 et code une protéine membre de la sous-unité ribosomale 40S. Les mutations retrouvées dans la LLC sont de type faux-sens et sont restreintes à une région de 7 acides aminés conservée au cours de l'évolution⁴⁶⁶. Elles entraîneraient un défaut de régulation de p53⁴⁶⁶. La fréquence de ces mutations est d'environ 4-5 % dans la LLC avec un enrichissement dans les formes agressives et réfractaires (jusqu'à presque 20 %)^{402,466}. Un tiers des cas mutés pour *RPS15* présente également des anomalies de *TP53*⁴⁶⁶. L'impact pronostique des mutations de *RPS15* reste à définir avec certains auteurs ne mettant pas en évidence d'impact pronostique indépendant, tandis que d'autres retrouvent une survie sans progression raccourcie^{466,402}.

Des mutations de *KLHL6* (pour Kelch-like family member 6) ont été identifiées. Ce gène est situé en 3q27 et code une protéine impliquée dans la formation du centre germinatif au cours de la maturation lymphocytaire B⁴⁸⁵. Les mutations retrouvées sont situées à proximité du site de début de transcription⁴⁴⁴. Ceci pourrait indiquer un rôle de l'hypermutation somatique



dans l'apparition de ces mutations. De plus, elles sont associées avec un statut IGHV muté⁴⁴⁴. Ces mutations sont retrouvées dans moins de 1 % des cas de LLC et leur impact clinique n'est pas connu⁴⁸⁰.

D'autres gènes sont également décrits comme mutés de façon récurrente dans la LLC avec de faibles fréquences (inférieures à 5 %). L'impact clinique et leur utilisation en pratique restera à définir au cours de prochaines études.

g- Evolution clonale

Chaque cellule cancéreuse d'un individu est caractérisée par une combinaison d'événements mutationnels comprenant des mutations initiatrices (ou « driver »), des mutations accompagnatrices (ou « passagers »), ainsi que d'éventuels variants héréditaires prédisposants⁴⁸⁴. Les cellules cancéreuses se propagent et se diversifient au cours de la croissance tumorale sous l'effet de facteurs intrinsèques (microenvironnement tumoral) ou extrinsèques (chimiothérapie par exemple). Cela aboutit à une population hétérogène de sous-clones différents aux niveaux génétique et phénotypique, reliés selon un lignage hiérarchique⁴⁸⁴. La population tumorale évolue selon une sélection Darwinienne⁴⁸⁴.

Les premières données permettant d'évoquer l'évolution clonale dans la LLC sont issues des études cytogénétiques par FISH et SNP-array, avec notamment la mise en évidence de l'acquisition sous-clonale d'une del(13q) au niveau du second allèle^{215,486}.

Le NGS et les approches de séquençage du génome entier permettent de visualiser l'architecture sous clonale impliquée dans l'évolution de la LLC¹⁶³. Ces techniques permettent de détecter environ 2000 mutations par échantillon issu de patients atteints de LLC⁴⁸⁷. En utilisant les fréquences alléliques, il est possible de hiérarchiser l'apparition des différentes mutations.

Une étude en 2012 a ainsi séquencé des échantillons de 3 patients à différents moments au cours de la maladie (au diagnostic, après traitement et au cours de rechutes, ainsi qu'un prélèvement buccal)⁴⁸⁴. Les mutations détectées dans toutes les cellules et dans tous les échantillons étaient considérées comme des événements fondateurs, tandis que les autres mutations étaient considérées comme secondaires. Cette approche a permis de mettre en évidence l'évolution des mutations et des anomalies du nombre de copies géniques. Ainsi, 2 patients ont développé des anomalies cytogénétiques au cours du temps, tandis qu'un autre est resté stable. Les profils mutationnels variaient au cours du temps. Au moins une mutation initiatrice possible était détectable pour chaque patient, et dans chaque cas, une mutation fondatrice différente a été mise en évidence (*SF3B1*, *SAMHD1* et *MED12*)⁴⁸⁴. L'évolution des sous-clones était différente suivant les cas. Chez un premier patient, un premier sous-clone a



émergé après une première ligne de chimiothérapie, puis persiste à la suite d'un second traitement anti-leucémique, avec la coexistence d'un autre sous-clone prépondérant, ce dernier devenant majoritaire par la suite. Chez un autre patient, l'évolution sous-clonale était relativement stable tant que la maladie était aussi cliniquement stable. Puis au moment de la progression, un sous-clone a émergé. Enfin, le dernier patient présentait les mêmes sous-clones de façon stable tout au long de sa maladie. Au total, des sous-clones étaient mis en évidence au diagnostic dans tous les cas. Leur évolution peut varier et de nouveaux sous-clones peuvent apparaître après traitement. L'intérêt clinique de l'évaluation systématique de l'architecture sous-clonale au cours de l'évolution de la maladie reste à définir.

Une autre équipe a analysé des échantillons successifs chez 12 patients⁴⁸⁸. Ils rapportent la compétition entre 2 clones ou plus au cours de l'évolution chez 70 % des patients, tandis que pour les autres, la dynamique clonale reste stable. Ils ont également mis en évidence pour deux patients l'acquisition d'anomalies génétiques comparables au cours du temps, ce qu'ils appellent l'évolution clonale convergente.

L'identification de sous clones chez les patients atteints de LLC est désormais possible grâce au NGS et a un intérêt clinique, en particulier pour *TP53* : il a été montré qu'une mutation de *TP53* dans un sous clone minoritaire au départ entraîne le même pronostic péjoratif qu'en cas de clone majoritaire⁴⁸⁹. De plus, après traitement, ces clones mutés *TP53* émergent et deviennent majoritaires.

Il apparaît donc important d'identifier dès le diagnostic des mutations sous-clonales d'intérêt afin d'apprécier l'évolution de la maladie.

4- Marqueurs sériques

α- Béta-2-microglobuline

La bêta-2-microglobuline est une protéine extracellulaire faisant partie du CMH de classe I⁴⁹⁰. Elle est utilisée comme marqueur de l'activité de différentes maladies telles que certains cancers, des affections auto-immunes ou les infections⁴⁹⁰. Cette protéine est présente en faible quantité dans le sérum chez l'homme et sa concentration augmente notamment en cas d'infection (en particulier les infections virales par l'EBV ou le CMV)⁴⁹⁰. Dès les années 1980, il a été montré que la bêta-2-microglobuline est augmentée dans la leucémie lymphoïde chronique de façon corrélée avec la quantité de cellules tumorales et leur renouvellement^{491,490,492}. En particulier, il a été démontré une corrélation entre l'augmentation de la bêta-2-microglobuline et le stade clinique Binet, ainsi que l'infiltration tumorale⁴⁹². Il s'agit également d'un marqueur pronostique au sein du groupe de stade Binet A⁴⁹³. La bêta-2-microglobuline permet de prédire les survies sans progression et sans traitement^{494,495}.



Cependant, sa concentration sérique dépend largement de la fonction rénale des patients⁴⁹⁶. En effet, elle est filtrée à 95 % par le glomérule rénal et réabsorbée à 99,9 % par le tubule proximal où elle est catabolisée⁴⁹⁶. Ainsi, sa concentration augmente considérablement en cas d'insuffisance rénale. Ainsi, il a été proposé d'ajuster le taux de béta-2-microglobuline en fonction du taux de filtration glomérulaire pour avoir un meilleur reflet de l'augmentation due à la masse tumorale⁴⁹⁶.

Malgré l'apparition de nouveaux marqueurs moléculaires, la béta-2-microglobuline reste un marqueur pronostique utilisé en pratique courante dans la LLC et intégré dans des scores cliniques récents tels que le CLL-IPI⁴⁹⁷.

b- Lactate déshydrogénase

La lactate déshydrogénase (LDH) est une enzyme impliquée dans la glycolyse. Elle permet la conversion du pyruvate en lactate. Il existe cinq isoenzymes correspondant à des tétramères constitués de 2 types de sous-unités : H (pour Heart) et M (pour Muscle)⁴⁹⁸. Les différentes isoformes sont exprimées préférentiellement dans différents tissus. Les LDH3 (H₂M₂) et LDH4 (H₁M₃), représentant normalement environ 20 et 8 % respectivement de l'ensemble des LDH circulantes, sont préférentiellement produites par les lymphocytes et sont augmentées dans les lymphopathies. Ces isoformes ne sont pas dosées séparément et c'est donc le dosage global des LDH qui est utilisé.

Il a été montré que le taux de LDH est un marqueur prédictif de la survie sans traitement dans la LLC⁴⁹⁹. Il s'agit d'un marqueur de mauvais pronostic associé à un risque de développement de syndrome de Richter³⁴⁶.

c- Electrophorèse des protéines sériques

L'électrophorèse des protéines sériques permet d'étudier les différentes fractions des protéines dans le sérum et de mettre en évidence des anomalies quantitatives et qualitatives. En particulier, dans la LLC, la recherche d'hypogammaglobulinémie ou la présence d'une fraction anormale d'immunoglobuline appelée pic monoclonal est recherchée. La suspicion de pic monoclonal est ensuite étudiée par immunofixation afin de déterminer la clonalité (chaîne lourde et légère) impliquée. La recherche de chaînes légères libres peut également être réalisée car il s'agit d'une technique plus sensible dans la détection d'un composant monoclonal.

Une anomalie est mise en évidence dans environ un tiers des cas de LLC, avec une hypogammaglobunémie chez environ 25 % des patients et un pic monoclonal dans environ 11



% des cas^{327,500}. L'utilisation de la quantification des chaînes légères libres permet d'augmenter le pourcentage de patients présentant une fraction monoclonale jusqu'à 42 %⁵⁰¹. Un ratio anormal est retrouvé plus fréquemment au cours de l'évolution qu'au diagnostic ou avant traitement (39 versus 29 % respectivement)⁵⁰².

La présence d'un composant monoclonal est un facteur pronostique péjoratif dans plusieurs études^{457,501,502,503,504}. En particulier, la présence d'un pic d'isotype IgM a été rapportée comme de plus mauvais pronostic^{504,457}.

L'impact pronostique de l'hypogammaglobulinémie est plus controversé⁵⁰⁵. Certaines études rapportent un risque augmenté d'infections, en particulier en cas de déficit en sous-classes d'IgG⁵⁰⁶. D'autres ne retrouvent pas d'association entre hypogammaglobulinémie et survenue d'infection⁵⁰⁷. De plus, l'impact concernant la survie n'est pas clairement tranché⁵⁰⁵. Certaines études ne retrouvent pas d'impact péjoratif en analyse multivariée^{327,508}. D'autres décrivent une survie globale raccourcie⁵⁰⁹. Enfin, un impact péjoratif sur la survie sans traitement a été démontré⁴⁵⁷.

Actuellement, l'électrophorèse des protéines sériques est réalisée au diagnostic et au cours de l'évolution de la LLC afin de détecter essentiellement une hypogammaglobulinémie, qui pourra être prise en charge par l'administration d'immunoglobulines polyvalentes lorsqu'elle est significative.

d- CD23 soluble

Le CD23 est le récepteur de faible affinité de l'IgE⁵¹⁰. Il est exprimé à la surface des cellules de LLC. Il a également été retrouvé sous forme soluble dans le sérum de patients atteints de LLC, à des niveaux corrélés avec l'expression de l'antigène de surface, la lymphocytose, la masse tumorale et le stade clinique^{511,512}. Son augmentation est également corrélée au temps de doublement de la lymphocytose et permet de suivre la progression de la maladie⁵¹³. Sa détermination au diagnostic chez les patients de stade Binet A permet de prédire les cas à risque de progression^{510,514}. Il s'agirait d'un marqueur pronostique associé à une survie globale raccourcie^{515,510}. Son temps de doublement permet de distinguer les patients ayant une forme indolente ou agressive⁵¹⁶. Il s'agit également d'un marqueur pronostique en terme de survie sans traitement⁵¹⁷.

Une étude récente a démontré que les taux de CD23 soluble élevés sont un marqueur prédictif d'évolution vers une LLC chez les patients présentant une lymphocytose B monoclonale⁵¹⁸.

Actuellement, ce marqueur n'est pas utilisé en routine. Il est mesuré classiquement par technique ELISA (pour Enzyme-Linked ImmunoSorbent Assay). Une équipe a mis au point une méthode de dosage par cytométrie en flux utilisant des billes (cytometry bead array)⁵¹⁹. Cette



technique est parfaitement corrélée à l'ELISA et permettrait une meilleure accessibilité au dosage de ce marqueur.

e- Thymidine kinase

La thymidine kinase est une enzyme catalysant la phosphorylation de la désoxythymidine en thymidine monophosphate, un précurseur essentiel pour la synthèse de l'ADN⁵²⁰.

La thymidine kinase sérique est un marqueur permettant de prédire la survie sans progression et la survie globale^{494,521}. Il s'agit d'un marqueur prédictif de la réponse au traitement par Fludarabine⁵²¹. Ce marqueur permet de séparer en deux groupes pronostiques les patients en stade Binet A et de forme clinique non indolente⁵²². Il a été proposé comme marqueur prédictif du statut mutationnel IGHV^{523,524}. Ainsi, des taux élevés sont associés à un statut IGHV non muté et la présence d'anomalies cytogénétiques de mauvais pronostic^{525,524}. Enfin, une étude montre que le taux de thymidine kinase sérique diminue significativement avec l'âge et propose d'interpréter ce marqueur de façon ajustée à l'âge dans la mesure où les patients atteints de LLC sont généralement âgés⁵²⁶.

5- Marqueurs immunophénotypiques

a- CD38

Le CD38 est une glycoprotéine transmembranaire. Son expression par les cellules de LLC est un marqueur de mauvais pronostic, indicateur d'activation et de prolifération^{407,527,528}. Le seuil de positivité utilisé varie entre 20 et 30 % de cellules exprimant ce marqueur. L'expression du CD38 est associée à la présence d'adénopathies, un taux d'hémoglobine plus faible, une hépatomégalie et un taux de bêta-2-microglobuline élevé⁵²⁷. Il avait alors été proposé d'utiliser ce marqueur en routine pour détecter les patients de mauvais pronostic^{407,527}. En particulier, il avait été étudié à la fin des années 1990 comme alternative à l'étude du statut mutationnel IGHV, alors difficile à réaliser en routine de laboratoire⁴⁰⁷. Cependant, il a été montré une discordance entre ces deux marqueurs dans environ 30 % des cas⁵²⁹. De plus, l'expression du CD38 peut varier au cours de l'évolution de la maladie dans environ 25 % des cas⁵²⁹. Il s'agit d'un marqueur pronostique indépendant.



b- ZAP70

ZAP70 (pour ζ -Associated Protein) est une tyrosine kinase normalement exprimée dans les lymphocytes T et NK^{346,530}. Elle est exprimée dans les lymphocytes B normaux et malins à différents stades de maturation⁵³¹. Il a été démontré qu'elle joue un rôle dans la voie de signalisation du BCR^{532,533}. Elle est recrutée au niveau du complexe de signalisation du BCR sous l'effet de la liaison à l'antigène, d'une façon similaire à SYK. SYK et ZAP70 ont des structures homologues.

Dans la LLC, une forte expression de ZAP70 a été mise en évidence par étude de profils d'expression génique en cas de statut IGHV non muté par rapport au statut muté¹⁷⁰. L'utilisation de l'expression de ce gène mesurée par RT-PCR (pour Reverse Transcription-Polymerase Chain Reaction) a donc été proposée comme alternative à l'étude du statut mutationnel IGHV, alors peu répandue et difficile à réaliser dans les laboratoires. Ceci a été confirmé par différentes équipes, avec une mesure proposée par cytométrie en flux^{534,535}. Le seuil de positivité retenu dans les différentes études varie entre 10 et 20 %^{534,535}. Cependant, des cas discordants existent suivant le type de réarrangement IGHV et certains autres marqueurs pronostiques tels que les anomalies cytogénétiques³⁴⁶. La mesure par cytométrie en flux nécessite une fixation des cellules et une perméabilisation, et des problèmes de standardisation ont été relevés, rendant difficile son utilisation en pratique courante^{346,536,537}.

L'impact péjoratif de la positivité de ZAP70 a été démontré sur les survies globale, sans traitement et sans progression^{534,538}.

c- CD49d

Le CD49d est une molécule de surface dont l'expression induit la prolifération des cellules de LLC sous l'effet du microenvironnement⁵³⁹. Le seuil de positivité est généralement fixé à 30 %^{540,541,539}. Son expression est associée aux formes progressives de LLC, et à une survie globale et sans traitement raccourcie^{540,542}. Elle est également associée à l'expression du CD38 et de ZAP70^{540,542,541}. Une étude en 2014 retrouve le CD49d comme étant le marqueur immunophénotypique le plus fort pour prédire la survie globale⁵³⁹. Ce marqueur pourrait constituer une cible thérapeutique⁵⁴².



6- Temps de doublement de la lymphocytose

Le temps de doublement de la lymphocytose est un reflet de la progression de la maladie. Il consiste à suivre la lymphocytose dans le temps et est donc facile d'utilisation en pratique. Dès les années 1980, il a été démontré que ce paramètre est un facteur pronostique de la survie globale^{543,544,545}. Actuellement, ce paramètre est pris en compte pour parler de maladie progressive¹²⁴. Ainsi, un temps de doublement de la lymphocytose inférieur à 6 mois est un critère à prendre en compte pour l'initiation du traitement. Cependant, lorsque la lymphocytose est inférieure à 30 G/L, une période d'observation plus importante est nécessaire. La lymphocytose en elle-même ne constitue pas un critère de traitement actuellement, dans la mesure où les hyperlymphocytoses observées dans la LLC ne sont que rarement responsables de signes cliniques, contrairement à ce qui peut être observé avec les hyperleucocytoses des leucémies aiguës¹²⁴.

7- Classifications pronostiques

L'utilisation de classifications pronostiques est un enjeu majeur en cancérologie. Cela permet de distinguer différentes catégories de patients les plus homogènes possibles. Le but est de pouvoir proposer un traitement adapté à chaque profil et de pouvoir comparer les réponses dans le cadre d'études cliniques.

Dans la LLC, les classifications de Rai et Binet sont les principales classifications pronostiques utilisées en clinique. La classification de Rai est largement utilisée aux Etats-Unis d'Amérique, tandis que la classification de Binet est plutôt utilisée en Europe. Ces 2 méthodes de stratification ont été décrites dans les années 80 et restent encore aujourd'hui valables malgré l'apparition de nombreux marqueurs pronostiques.

a- Classification de Rai

La classification de Rai distingue 5 stades différents selon la présence ou non d'adénopathie, d'hépatome et/ou splénomégalie et de cytopénies (anémie définie par un taux d'hémoglobine inférieur à 11 g/dL, thrombopénie définie par une numération plaquettaire inférieure à 100 G/L) (Tableau 3)⁵⁴⁶.



Stade	Critère	Médiane de survie en 1975
Stade 0	Lymphocytose seule	> 150 mois
Stade I	Lymphocytose + Adénopathies	101 mois
Stade II	Lymphocytose + Splénomégalie et/ou hépatomégalie	71 mois
Stade III	Lymphocytose + Anémie	19 mois
Stade IV	Lymphocytose + Thrombopénie	19 mois

Tableau 3 : Classification de Rai selon Rai *et al.*⁵⁴⁶.

Le seuil pour l'anémie étant fixé par un taux d'hémoglobine inférieur à 11 g/dL et la thrombopénie par une numération plaquettaire inférieure à 100 G/L.

Actuellement, les stades I et II d'une part, et III et IV d'autre part sont regroupés⁵⁴⁷.

b- Classification de Binet

La classification de Binet distingue 3 stades différents selon la présence d'adénopathies au niveau de différentes aires ganglionnaires et la présence de cytopénies (Tableau 4)⁵⁴⁸. Une adénopathie est définie comme une augmentation du volume d'un ganglion d'au moins 1 cm⁵⁴⁷. Les aires ganglionnaires étaient les régions axillaires, cervicales et inguinales, unilatérales ou non, ainsi que la rate et le foie. Les seuils pour parler de cytopénie sont fixés à 10 g/dL pour l'hémoglobine et 100 G/L pour les plaquettes.

Stade	Critère	Médiane de survie en 1981
Stade A	< 3 aires ganglionnaires Absence de cytopénies	Identique à la population générale
Stade B	≥ 3 aires ganglionnaires Absence de cytopénies	7 ans
Stade C	Anémie et/ou Thrombopénie	2 ans

Tableau 4 : Classification de Binet selon Binet *et al.*⁵⁴⁸.

Le seuil pour l'anémie étant fixé par un taux d'hémoglobine inférieur à 10 g/dL et la thrombopénie par une numération plaquettaire inférieure à 100 G/L.



c- Score de Rossi

En 2013, Rossi *et al.* proposent une classification pronostique intégrant le statut mutationnel de 4 gènes mutés de façon récurrente dans la LLC (*TP53*, *NOTCH1*, *SF3B1* et *BIRC3*), ainsi que les anomalies cytogénétiques principales (del(17p), del(11q), trisomie 12 et del(13q))³⁶¹. Cette classification distingue 4 groupes pronostiques différents. Le groupe de haut risque correspond aux patients présentant une anomalie de *TP53* (mutation et/ou délétion) ou mutation de *BIRC3*. La survie à 10 ans est estimée à 29 %. La catégorie de risque intermédiaire correspond aux patients présentant une mutation de *SF3B1* ou *NOTCH1* ou une del(11q). La survie à 10 ans est alors estimée à 37 %. Le groupe de risque faible correspond aux patients avec caryotype normal ou avec une trisomie 12. Dans ce cas, la survie à 10 ans est estimée à 57 %. Enfin, la catégorie de risque très faible regroupe les patients présentant une del(13q) isolée. La survie à 10 ans est alors de 69 % et ne diffère pas significativement de la population générale. Cette classification reste valable au cours de l'évolution de la maladie.

d- CLL-IPI

En 2016, le CLL-IPI (pour Chronic Lymphocytic Leukemia- International Prognosis Index) a été établi par un groupe de travail international en étudiant le poids pronostique de 27 facteurs au diagnostic sur une cohorte de 3472 patients naïfs de tout traitement⁴⁹⁷. Les variables indépendantes retenues étaient le statut de *TP53* (mutation ou délétion), le statut mutationnel IGHV, la concentration sérique en bêta-2-microglobuline, le stade clinique selon Rai ou Binet, et l'âge (seuil fixé à 65 ans) (Tableau 5).

Critère	Pris en compte si	Point attribués
Statut TP53	Muté ou délété	4
Statut mutationnel IGHV	Non muté	2
Béta-2-microglobuline	> 3,5 mg/L	2
Stade clinique	Rai I à IV ou Binet B ou C	1
Age	> 65 ans	1

Tableau 5 : Critères et calcul du CLL-IPI selon The International CLL Working Group⁴⁹⁷.



La présence d'une altération du gène *TP53* (mutation et/ou délétion) est pondérée de 4 points. Le statut IGHV non muté et une concentration sérique en bêta-2-microglobuline supérieure à 3,5 mg/L attribuent 2 points chacun. Le stade clinique Rai non 0 ou Binet non A entraînent 1 point chacun. Un score allant de 0 à 10 est calculé en additionnant chaque critère. A partir de ce score, 4 catégories de risque sont distinguées.

Catégorie de risque selon le CLL-IPI	Score CLL-IPI	Survie globale à 5 ans
Bas	0-1	93,2 %
Intermédiaire	2-3	79,3 %
Elevé	4-6	63,3 %
Très élevé	7-10	23,3 %

Tableau 6 : Catégories de risque selon le score d'après le CLL-IPI⁴⁹⁷.

Ce score a été vérifié dans plusieurs cohortes^{549,550}. Il permet de discriminer correctement les patients pour la survie globale (OS) mais aussi pour la survie sans traitement (TFS). L'application du CLL-IPI n'est pas suffisante pour décider de mettre en place un traitement selon les recommandations actuelles¹²⁴.

H- Traitement

1- Critères de mise en place thérapeutique

Les patients ayant une LLC de stade précoce asymptomatique ne nécessitent pas de mise en place thérapeutique⁵⁴⁷. Il a été montré que le fait de traiter ces cas n'apporte pas de bénéfice en terme de survie⁵⁴⁷.

Les patients de stade Binet C (ou Rai III ou IV) doivent être traités⁵⁴⁷. En revanche, les cas avec risque intermédiaire (stade Binet B ou Rai I ou II) ne sont pas systématiquement traités hormis en cas de maladie progressive ou symptomatique. Autrement, un suivi seul est recommandé.

Les critères pour parler de maladie progressive ou active sont définis par l'iwCLL 2018 (pour international workshop on Chronic Lymphocytic Leukemia)⁵⁴⁷. Il s'agit de :

- la mise en évidence d'une insuffisance médullaire progressive se traduisant par l'aggravation d'une anémie et/ou d'une thrombopénie. Bien que les seuils de 10 g/dL pour



l'hémoglobine et de 100 G/L pour les plaquettes sont généralement utilisés (ils définissent le stade Binet C en particulier), certains patients peuvent avoir des taux de plaquettes bas mais stables au cours du temps et cela ne constitue pas forcément un critère de traitement ;

- la présence d'une splénomégalie massive (≥ 6 cm de débord costal) ou progressive ou symptomatique ;

- la présence d'adénopathie massive (≥ 10 cm) ou progressive ou symptomatique ;

- une progression de la lymphocytose avec une augmentation de plus de 50 % en 2 mois ou un temps de doublement de la lymphocytose inférieur à 6 mois. Ce dernier peut être extrapolé à partir de mesures de lymphocytose à deux semaines d'intervalle pendant une période de 2 ou 3 mois. En cas de lymphocytose inférieure à 30 G/L, une période d'observation plus longue est nécessaire. Les autres causes d'augmentation de la lymphocytose, telles que les infections ou certains médicaments, doivent être exclus. Il faut noter qu'une lymphocytose élevée n'est pas un critère de traitement dans la mesure où elle est rarement symptomatique dans la LLC ;

- la survenue de complications auto-immunes répondant peu aux corticoïdes ;

- la découverte de localisations extra-ganglionnaires symptomatiques ou avec impact fonctionnel (par exemple au niveau de la peau, le rein, le poumon ou la colonne vertébrale) ;

- la présence de signes cliniques liés à la maladie tels qu'une perte de poids d'au moins 10 % au cours des 6 derniers mois, ou une fatigue importante, ou une fièvre de plus de 38,0°C depuis au moins 2 semaines sans foyer infectieux, ou des sueurs nocturnes depuis au moins 1 mois sans cause infectieuse.

La présence d'une hypogammaglobulinémie ou d'un composé oligoclonal ou monoclonal n'est pas un critère de mise en place thérapeutique. En revanche, si le patient est traité, il convient de suivre l'évolution de ces anomalies.

Pour les patients ayant déjà reçu une ou plusieurs lignes de traitement, la réapparition de la LLC n'est pas un critère de traitement en dehors d'une maladie active telle que définie précédemment.



2- Evaluation de la réponse au traitement

L'évaluation de la réponse au traitement est généralement réalisée 2 mois après son arrêt ou, en cas de traitement en continu ou comportant une phase de maintenance, 2 mois après l'obtention de la réponse maximale⁵⁴⁷. Les modalités thérapeutiques sont précisées page 130.

Plusieurs niveaux de réponses sont définis.

La réponse complète correspond à une normalisation de la lymphocytose (< 4 G/L), l'absence d'adénopathie ou d'hépatosplénomégalie, l'absence de symptômes liés à la maladie, l'absence de cytopénie (Hémoglobine \geq 11 g/dL, Plaquettes \geq 100 G/L et Polynucléaires neutrophiles supérieurs à 1,5 G/L)⁵⁴⁷. Une évaluation de la maladie résiduelle doit être réalisée (voir 130)⁵⁴⁷.

La réponse partielle correspond à la présence d'une amélioration d'au moins 2 paramètres parmi les suivants : adénopathie, hépatosplénomégalie, symptômes cliniques et une diminution de la lymphocytose sanguine d'au moins 50 %⁵⁴⁷. De plus, au moins 1 cytopénie (parmi anémie ou thrombopénie) doit être corrigée.

Une maladie progressive au cours du traitement ou après celui-ci est définie par l'apparition de nouvelles localisations tumorales ou l'augmentation d'au moins 50 % d'une ou plusieurs adénopathies ou d'une hépatosplénomégalie, ou un doublement de la lymphocytose (au-dessus de 5 G/L), ou l'apparition d'un syndrome de Richter, ou la présence de cytopénie (anémie ou thrombopénie) en rapport avec la LLC (mais non auto-immunes)⁵⁴⁷.

Une maladie stable correspond à l'absence de réponse complète ou partielle, mais sans signe de progression⁵⁴⁷.

Un échec de traitement est défini par l'absence d'obtention d'un bénéfice, c'est-à-dire l'absence de réponse complète ou partielle⁵⁴⁷.

La rechute est définie par l'apparition d'une progression de la maladie après l'obtention d'une réponse complète ou partielle⁵⁴⁷.

Une maladie réfractaire est définie par un échec de traitement ou une progression dans les 6 mois après l'arrêt de traitement⁵⁴⁷.



3- Mesure de la maladie résiduelle

La mesure de la maladie résiduelle peut être réalisée avec différentes techniques sensibles telles que la cytométrie en flux, la PCR ou le NGS⁵⁴⁷. Ces techniques permettent de descendre à une sensibilité de 10^{-4} (soit 1 cellule/10 000) à 10^{-5} ^{547,551}. Le seuil retenu par l'iwCLL 2018 est de 10^{-4} ⁵⁴⁷. La maladie résiduelle peut être réalisée dans le sang ou la moelle. Elle doit généralement être réalisée dans le sang, mais il existe certaines drogues permettant l'élimination dans le sang mais pas la moelle : une évaluation médullaire est alors souhaitable⁵⁴⁷.

L'obtention d'une maladie résiduelle indétectable est toujours souhaitable dans la mesure où elle est associée à une évolution clinique meilleure selon différentes études cliniques^{552,553}.

4- Les différentes options thérapeutiques disponibles

a- Les traitements de la LLC

i- Les monochimiothérapies

Les agents cytostatiques

Ils comprennent les agents alkylants et les analogues de purine.

Parmi les agents alkylants, le Chlorambucil en monothérapie a longtemps été la thérapeutique utilisée¹²⁴. Cette molécule est peu toxique, de faible coût et administrable par voie orale. Cependant, elle ne permet pas l'obtention de réponse complète et peut entraîner après utilisation prolongée des cytopénies persistantes, des myélodysplasies et des leucémies aiguës secondaires¹²⁴. Il a par la suite été démontré l'intérêt de l'adjonction d'anticorps anti-CD20. Actuellement, l'administration de Chlorambucil en monochimiothérapie est limitée aux patients âgés fragiles ou ayant des comorbidités et constitue une thérapeutique palliative¹²⁴.

Il existe plusieurs analogues de purine utilisés dans la LLC. Le principal est la Fludarabine. Elle permet en monochimiothérapie d'obtenir de meilleures réponses globales que les autres traitements alkylants ou les corticoïdes^{124,554}. Elle permet également d'obtenir de meilleurs taux de rémission par rapport à d'autres chimiothérapies conventionnelles tels que le CHOP (pour Cyclophosphamide, Doxorubicine, Vincristine, Prednisone), mais sans permettre d'améliorer la survie globale^{555,556}.

Enfin, la Bendamustine est un autre agent alkylant utilisé dans la LLC. Son efficacité a été comparée au Chlorambucil : de meilleures réponses globales et survie sans progression était obtenues avec ce traitement, mais une toxicité plus importante était observée^{557,558}. La



survenue d'infections sévères était également plus fréquente avec la Bendamustine⁵⁵⁷. Son efficacité en seconde ligne en comparaison à la Fludarabine a également été étudiée et montre une meilleure efficacité en terme d'obtention de réponse complète et de survie sans progression⁵⁵⁹.

Les anticorps monoclonaux

Les anticorps monoclonaux utilisés dans la LLC comprennent essentiellement les anti-CD20, une glycoprotéine membranaire exprimée à la surface des lymphocytes B.

Le premier anti-CD20 utilisé est le Rituximab. Il est peu utilisé en monochimiothérapie dans la LLC mais son efficacité en chimio-immunothérapie a été largement démontrée¹²⁴.

L'Ofatumumab est un anti-CD20 humanisé avec une affinité augmentée pour le CD20 par rapport au Rituximab¹²⁴. Il peut être utilisé en monothérapie dans la LLC dans les cas réfractaires à la Fludarabine et l'Alemtuzumab¹²⁴.

L'Obinutuzumab (GA101) est un anti-CD20 humanisé avec une affinité augmentée pour le CD20 permettant d'induire une cytotoxicité supérieure par rapport aux autres anti-CD20. Il a été étudié chez des patients réfractaires ou en rechute et montre une bonne efficacité malgré des effets secondaires notables (notamment la neutropénie)¹²⁴. Il est actuellement évalué dans des études cliniques en combinaison avec d'autres chimiothérapies.

A côté des anti-CD20, un anti-CD52 est utilisé dans la LLC. Il s'agit de l'Alemtuzumab. Il s'agit d'un anticorps monoclonal humanisé. Il a été démontré une bonne efficacité de cette molécule en monothérapie chez des patients réfractaires ou en rechute après une première ligne de traitement à base d'agent alkylant et une deuxième ligne de traitement incluant la Fludarabine^{124,560,561}. Des infections sévères ont cependant été observées^{560,561}. Cet anticorps est également efficace en cas d'anomalies de haut risque telles que les anomalies de *TP53* ou la del(11q)^{124,562,563}. L'Alemtuzumab peut être utilisé en première ligne de traitement¹²⁴.

Les inhibiteurs du BCR

La voie de signalisation du BCR semble jouer un rôle important dans la survie des cellules de LLC^{124,267}. Ainsi, son inhibition constitue une nouvelle cible thérapeutique. Il existe des inhibiteurs de tyrosine kinases impliquées dans cette voie.

L'Idelalisib est un inhibiteur de la PI3K entraînant l'apoptose des cellules de LLC de façon temps et dose dépendante¹²⁴. Il inhibe le chimiotactisme des lymphocytes malins via CXCL12



et CXCL13, ainsi que la migration parmi les cellules stromales. Il inhibe les signaux de survie dérivés du BCR et des cellules « nurse-like ». Il est administrable par voie orale. Les essais réalisés chez des patients réfractaires ou en rechute et présentant des facteurs de mauvais pronostic ont montré de bons résultats en terme de survie globale et sans progression⁵⁶⁴. Les effets secondaires sévères comprenaient essentiellement la survenue de pneumopathie, de neutropénie fébrile et de diarrhée. Cette molécule est actuellement évaluée en combinaison avec d'autres agents. Elle a obtenu l'autorisation de mise sur le marché dans la LLC en association avec un anti-CD20 soit en traitement de première intention en cas d'altération de *TP53* chez les patients non éligibles à un autre traitement, soit en traitement de deuxième ligne.

L'Ibrutinib est un inhibiteur de BTK induisant l'apoptose des cellules B lymphomateuses ou de LLC. Son administration par voie orale chez des patients réfractaires ou en rechute a démontré un bénéfice significatif⁵⁶⁵. Les principaux effets secondaires relevés étaient des infections virales. Les essais suivants ont montré des effets secondaires plus modérés (diarrhée transitoire, fatigue et infections respiratoires hautes essentiellement)⁵⁶⁶. De plus, de très bons résultats en terme de survie globale et sans progression étaient relevés, indépendamment du stade clinique avancé, du nombre de lignes de traitement antérieures et de la présence de del(17p). Cependant, il existe l'apparition de résistances à l'Ibrutinib liées à la présence d'une mutation (C481S) au niveau de son site de fixation sur la BTK, entraînant une inhibition seulement réversible⁵⁶⁷. Des mutations de PLC γ 2 (R665W et L845F) ont également été mises en évidence et entraîneraient un gain de fonction avec une activité autonome du BCR. Ces mutations peuvent être présentes avant la mise sous traitement ou acquises au cours de l'évolution avant la rechute^{568,569}. Leur recherche permettrait de proposer un traitement additionnel par thérapie ciblée¹²⁴. Ce médicament a obtenu une autorisation de mise sur le marché dans la LLC en première ligne en monothérapie ou en deuxième ligne seul ou en association au Rituximab et à la Bendamustine.

L'Acalabrutinib (ACP-196), un inhibiteur de BTK irréversible et potentiellement plus sélectif est actuellement en cours d'évaluation¹²⁴.

Les inhibiteurs de BCL2

Actuellement, un inhibiteur de BCL2 est utilisé dans la LLC. Il s'agit du Venetoclax (ABT-199). Les premiers essais réalisés chez 3 patients réfractaires ont montré l'apparition d'une lyse tumorale en 24h⁵⁷⁰. Les essais suivants proposent l'augmentation progressive des doses en 4-5 semaines¹²⁴. Un essai de phase I/II a montré de bonnes réponses y compris chez des patients ayant reçu de multiples lignes de traitement et présentant des facteurs de mauvais pronostic tels que la del(17p) ou un statut IGHV non muté⁵⁷¹. L'ajustement des doses a permis d'éviter la survenue de syndrome de lyse tumorale. Les effets indésirables rencontrés étaient principalement une diarrhée modérée, l'apparition d'infection respiratoire haute et de nausée chez environ la moitié des patients. Une neutropénie sévère est également fréquemment



observée. Une autre étude chez des patients réfractaires ou en rechute avec del(17p) a montré des résultats encourageants⁵⁷². Cette molécule a obtenu une autorisation de mise sur le marché dans la LLC en seconde ligne soit en cas d'anomalie de *TP53* et en cas d'échec ou d'inéligibilité à un traitement par un inhibiteur du BCR, soit en l'absence d'altération de *TP53* et en cas d'échec à une chimio-immunothérapie et à un inhibiteur du BCR.

Les immunomodulateurs

Le Lenalidomide a été étudié chez les patients atteints de LLC de haut risque. Ce traitement entraîne une réaction appelée « tumor flare » (ou embrasement tumoral) avec sensation de chaleur et de brûlure au niveau des ganglions¹²⁴. Il permet d'obtenir des réponses à long terme avec un temps jusqu'à l'échec de traitement non atteinte avec un suivi médian de 4 ans⁵⁷³. Des essais ont été réalisés en combinaison avec le Rituximab avec de meilleures réponses sans augmentation de la toxicité. D'autres tests combinant Lenalidomide-Rituximab et Fludarabine ou Bendamustine ont été stoppés à cause de leur toxicité¹²⁴. La combinaison avec le Flavopiridol, un inhibiteur de kinase cycline-dépendant, a été testée et permet d'augmenter le taux de réponse sans entraîner de toxicité plus importante⁵⁷⁴. Enfin, le Lenalidomide est testé en traitement de maintenance chez les patients de haut risque et permettrait d'améliorer significativement la survie sans progression^{575,576}.

Le Pembrolizumab est un inhibiteur de PD-1. Il s'agit d'un anticorps monoclonal humanisé. Il a été testé chez des patients atteints de LLC réfractaires et des cas de transformations en Richter⁵⁷⁷. Il a montré un effet dans le deuxième groupe de patients.

Les CAR-T cells (pour cellules T porteuses d'un Récepteur à l'Antigène chimérique ou Chimeric Antigen Receptor) est une approche utilisée dans les leucémies aiguës lymphoblastiques. Leur utilisation a été testée dans la LLC avec un récepteur modifié dirigé contre le CD19⁵⁷⁸. Un patient avec forme réfractaire a bénéficié de cette approche. Il a développé un syndrome de lyse tumoral décalé dans le temps par rapport à l'injection. La rémission complète a été atteinte et a persisté 10 mois après le traitement. Ces résultats ont été partiellement confirmés dans de plus grandes séries de patients et la place de cette thérapeutique dans la LLC reste à établir¹²⁴.

ii- Les poly-chimiothérapies

Les avancées thérapeutiques dans la LLC reposent sur les combinaisons de différentes molécules. Dans la mesure où les agents alkylants et les analogues de purine ont des mécanismes d'action et des profils de toxicité partiellement différents, leur combinaison a semblé logique¹²⁴. L'utilisation de la Fludarabine et du Cyclophosphamide a montré une cytotoxicité synergique⁵⁷⁹. Les combinaison Fludarabine-Cladribine et Fludarabine-



Chlorambucil-Prednisone ont également été testées et ont montré respectivement une efficacité moindre et une toxicité plus importante¹²⁴. D'autres combinaisons ont également été testées mais restent moins utilisées que le tandem Fludarabine-Cyclophosphamide.

iii- Les chimio-immunothérapies

Les études précliniques ayant démontré une synergie entre le Rituximab et la Fludarabine, des essais cliniques se sont ensuite intéressés à ce type de chimio-immunothérapie¹²⁴. Les résultats ont démontré une meilleure survie sans progression et globale par rapport à un traitement à base de Fludarabine seule⁵⁸⁰. La combinaison Fludarabine-Cyclophosphamide-Rituximab (ou FCR) s'est révélée encore plus efficace, excepté dans le cas des patients avec del(17p)^{581,582}.

Plus récemment, la combinaison Rituximab-Bendamustine a été étudiée. Elle a montré de bons résultats chez des patients réfractaires à la Fludarabine⁵⁸³. Cependant, en première ligne de traitement, l'efficacité était moins bonne que la combinaison FCR⁵⁸⁴. Néanmoins, les complications infectieuses étaient plus sévères chez les patients de plus de 65 ans traités par FCR. Ainsi, le FCR reste le traitement de choix pour les patients en forme, tandis que les patients plus âgés peuvent bénéficier d'un traitement par Rituximab-Bendamustine¹²⁴.

La combinaison Obinutuzumab-Chlorambucil a été étudiée chez des patients ayant des comorbidités⁵⁸⁵. Cette approche est intéressante car elle a une efficacité supérieure à la combinaison Rituximab-Chlorambucil. Il a également été démontré que la co-administration d'Ofatumumab-Chlorambucil est plus efficace qu'un traitement par Chlorambucil seul chez le même type de patients⁵⁸⁶.

La chimio-immunothérapie Alemtuzumab-Fludarabine a été testée chez des patients en rechute et semble être une option intéressante¹²⁴. En revanche, les essais associant Fludarabine-Cyclophosphamide et Alemtuzumab ont été stoppés à cause de leur toxicité et d'une surmortalité⁵⁸⁷.

L'utilisation combinée de l'Alemtuzumab et du Rituximab a été étudiée dans des cas de LLC réfractaires ou en rechute avec de bons taux de réponse⁵⁸⁸. Ces résultats doivent être confirmés avec un plus grand nombre de patients¹²⁴.

iv- Les combinaisons d'agents ciblés

L'utilisation combinée de l'Idelalisib et du Rituximab a été comparée à l'utilisation du Rituximab seul chez des patients atteints de LLC ayant des comorbidités et son efficacité a été



validée⁵⁸⁹. Cette approche peut être proposée aux patients en rechute et ne pouvant recevoir d'autres chimiothérapie à cause de comorbidités¹²⁴. L'idelalisib a également été testé avec l'Ofatumumab chez le même type de patients et permet d'obtenir une survie sans progression plus longue qu'en cas d'utilisation d'Ofatumumab seul, mais avec une toxicité et des décès liés au traitements plus élevés⁵⁹⁰.

La combinaison Rituximab-Ibrutinib a été testée chez des patients de haut risque et montre une bonne efficacité avec l'obtention de rémissions durables et une toxicité modérée^{591,592}. La bithérapie Ofatumumab-Ibrutinib a été administrée chez des patients en échec après au moins 2 lignes de traitement. Cette approche est efficace, avec une bonne tolérance, y compris dans les cas de transformation en Richter⁵⁹³.

L'addition de l'Ibrutinib a été testée chez des patients recevant un traitement à base de Bendamustine-Rituximab⁵⁹⁴. L'efficacité est meilleure avec une survie sans progression allongée, sans toxicité supplémentaire observée.

La combinaison Venetoclax-Rituximab a été étudiée chez des patients réfractaires ou en rechute. L'efficacité est très bonne avec l'obtention de survie sans progression à 2 ans dans 82 % des cas et un fort taux de réponse profonde⁵⁹⁵. L'association Venetoclax-Obinutuzumab a été testée chez 12 patients non précédemment traités mais ayant des comorbidités⁵⁹⁶. Les résultats sont impressionnants avec une survie sans progression à 15 mois de 100 %, et l'obtention d'une réponse complète chez 7 patients, et une MRD négative chez 11 patients. Ces résultats doivent être confirmés et pourraient modifier la future prise en charge des patients¹²⁴.

v- L'allogreffe de cellules souches hématopoïétiques

L'allogreffe de cellules souches hématopoïétiques a été proposée chez les patients en bonne condition physique et ayant une maladie de haut risque. Il s'agit de la seule thérapie curative dans la LLC, permettant l'éradication de la maladie⁵⁹⁷. Cependant, cette thérapeutique ne peut être proposée qu'à des patients jeunes, ayant un donneur compatible. Elle est associée à une mortalité précoce élevée (entre 15 et 30 %) au cours des deux premières années après la procédure. De plus, la réaction du greffon contre l'hôte affecte la qualité de vie des patients dans environ 25 % des cas, au moins au cours des premières années.



b- Les traitements de support

i- Facteurs de croissance

L'administration de facteur de croissance granulocytaire (G-CSF pour Granulocyte-Colony Stimulating Factor) est recommandée en cas de chimiothérapie myélosuppressive ou en cas de cytopénie prolongée après traitement par Alemtuzumab⁵⁴⁷. L'érythropoïétine peut être administrée en cas d'anémie. Cependant, généralement, les cytopénies se normalisent grâce au traitement anti-leucémique adapté.

ii- Prévention des infections

Les infections sont une des principales complications au décours de la LLC. La prévention par vaccination est généralement proposée, bien qu'aucune étude randomisée démontrant l'effet de la vaccination sur le nombre d'infections et leur évolution dans la LLC n'ait été réalisée⁵⁴⁷. La vaccination doit avoir lieu si possible avant l'initiation de la chimiothérapie dans la mesure où l'immunosuppression occasionnée entraîne une moindre efficacité des vaccins^{547,598}. Le vaccin contre la grippe est recommandé chez les patients atteints de LLC. Les vaccins vivants sont contre-indiqués dans la mesure où ils peuvent entraîner des complications sévères voire fatales.

En cas d'hypogammaglobulinémie, il est actuellement recommandé d'administrer des immunoglobulines polyvalentes uniquement dans les situations où elle est associée à des infections répétées⁵⁴⁷. Il a été démontré que cette thérapeutique substitutive permet de diminuer le taux d'infections bactériennes et allonge le temps avant la 1^{ère} infection, même si elle ne modifie pas la survie ou l'évolution^{547,599}.

iii- Prise en charge des cytopénies auto-immunes

Le purpura thrombopénique immunologique et l'anémie hémolytique auto-immune en tant que seule anomalie de la LLC doivent être traités en premier lieu par corticothérapie et non chimiothérapie⁵⁴⁷. Les traitements de seconde ligne sont ceux utilisés en dehors de la LLC (Rituximab, immunoglobulines polyvalentes par exemple). En revanche, la présence de cytopénies auto-immunes réfractaires à ces thérapeutiques est une indication de traitement de la LLC⁵⁴⁷.



5- Choix du traitement en fonction des patients

a- Les paramètres à prendre en compte

Il existe actuellement un grand nombre de possibilités thérapeutiques. Le choix doit prendre en compte le stade clinique, les symptômes, la condition physique et les maladies concomitantes, les facteurs de risque génétique de la LLC, et la situation thérapeutique (1^{ère} ligne versus seconde ligne et réponse ou non au traitement précédent)¹²⁴.

b- Traitement de première ligne

Il est actuellement recommandé de débiter un traitement chez les patients ayant un stade clinique avancé (stade Binet C ou Rai III-IV) ou une maladie active, symptomatique. Dans ces cas, le premier paramètre à prendre en compte est l'état général du patient, c'est-à-dire sa condition physique et la recherche de comorbidités¹²⁴. Le deuxième paramètre à prendre en compte est la présence d'anomalie de *TP53* pouvant entraîner une résistance au traitement.

Ainsi, pour les patients en bonne condition physique et sans anomalie de *TP53*, le traitement de référence reste le FCR. Eventuellement, chez les patients de plus de 65 ans et en bonne condition physique, la bithérapie Rituximab + Bendamustine peut être envisagée. Cependant, en cas d'anomalie de *TP53*, une thérapie à base d'Ibrutinib doit être proposée, sauf contre-indication. Les deux autres options possibles sont le Venetoclax ou l'Idelalisib+Rituximab.

Chez les patients en condition physique altérée et sans anomalie de *TP53* sera proposée une thérapie à base de Chlorambucil en combinaison avec un anti-CD20 (Obinutuzumab ou Rituximab ou Ofatumumab), l'Obinutuzumab étant le plus efficace. L'Ibrutinib en monothérapie peut également être envisagé. En cas d'anomalie de *TP53*, l'Ibrutinib ou l'Idelalisib + Rituximab ou le Venetoclax peuvent alors être proposés⁶⁰⁰. D'autres alternatives possibles sont l'Alemtuzumab ou de fortes doses de Rituximab ou Ofatumumab¹²⁴.

c- Les traitements de deuxième ligne

En cas de rechute, le premier paramètre à prendre en compte est la durée sans progression¹²⁴. Si celle-ci est inférieure à 3 ans, il convient de changer de traitement par rapport à la première ligne. Si elle est supérieure à 3 ans, la même stratégie que la première ligne peut être utilisée. Le statut de *TP53* doit être vérifié.



Les thérapeutiques pouvant alors être proposés sont l'Idelalisib+Rituximab ou l'Ibrutinib en monothérapie, le Venetoclax, les combinaisons d'Ibrutinib ou Venetoclax avec un anti-CD20, l'Alemtuzumab seul ou en association ou le lénalidomide^{124,600}.

Le choix repose alors sur la condition physique du patient, la disponibilité des différentes options thérapeutiques, le profil de risque moléculaire de la LLC et les effets secondaires des nouvelles drogues suivant les caractéristiques des patients (risque de saignement et Ibrutinib, maladie cardiaque et Ibrutinib, maladie pulmonaire chez les patients CMV positifs ou avec antécédent de pneumopathie à *Pneumocystis* pour l'Idelalisib par exemple).

Actuellement, il est recommandé de proposer une allogreffe de cellules souches hématopoïétiques aux patients en bonne condition physique réfractaires ou avec une del(17p), s'ils ont rechuté après un traitement par inhibiteur de kinase et répondent à un deuxième type de traitement^{124,601}.

Enfin, les patients ayant une maladie réfractaire doivent être inclus à chaque fois que cela est possible dans des protocoles cliniques¹²⁴.



Chapitre 3 : Problématiques et objectifs

D'une façon générale, nous posons la question de l'impact des anomalies moléculaires et en particulier des gènes des immunoglobulines dans la lymphomagenèse B et plus particulièrement dans la LLC.

D'une part, afin de rechercher de nouveaux oncogènes, nous nous sommes intéressés aux remaniements chromosomiques (translocations) impliquant les gènes des immunoglobulines dans les lymphomes B, et en particulier dans la LLC et le lymphome lymphocytaire. Nous avons sélectionné les cas pour lesquels le partenaire n'est pas connu, dans la mesure où ce type de translocation entraîne la dérégulation du gène partenaire via le rapprochement d'un élément régulateur du gène des immunoglobulines. L'objectif a donc été de mettre en évidence les points de cassures dans trois cas de LLC/lymphome lymphocytaire pour lesquels ce type de translocation avait été identifié. Nous avons voulu caractériser le gène partenaire dans l'optique de découvrir de nouveaux oncogènes. Enfin, nous avons tenté de comprendre le mécanisme de survenue de ces translocations et le mode de dérégulation du gène partenaire sous l'effet de la translocation.

D'autre part, dans la suite d'un travail sur les LLC avec pic monoclonal d'immunoglobuline, nous avons étudié l'impact pronostique d'une électrophorèse des protéines sériques normale (immunoglobulines polyclonales sans hypogammaglobulinémie) dans l'ère du NGS. Pour cela, nous avons mis au point un panel Ampliseq appliqué aux lymphomes, permettant de détecter les mutations d'intérêt à visée diagnostique et/ou pronostique, ainsi que les principales variations du nombre de copies de la LLC, à savoir la del(17p), la del(11q), la trisomie 12 et la del(13q). Dans le cadre de la détection des variations du nombre de copies, nous avons collaboré avec l'équipe EA6309 de la faculté de médecine de Limoges afin d'améliorer un outil dédié, le logiciel Cov'Cop, en particulier pour son application dans les pathologies somatiques. Nous avons ensuite établi les performances de cette technique dans la détection des principales anomalies de nombre de la LLC. Enfin, nous avons évalué sa place dans la stratégie diagnostique au laboratoire, ainsi que son apport pour la détection de certaines autres anomalies de nombre et de disomies uniparentales (ou UPD pour UniParental Disomy).



Chapitre 4 : Matériels et méthodes

A- Patients et échantillons

1- Etude de translocations impliquant le gène des immunoglobulines

a- Patient n°1

Le premier patient est un homme chez lequel a été diagnostiqué en 2012, à l'âge de 56 ans, un lymphome lymphocytaire. Cliniquement, le patient se présentait alors avec de multiples adénopathies cervicales de 1 à 2 cm. Après biopsie ganglionnaire cervicale gauche en Septembre 2012, le diagnostic a été établi par examen anatomo-pathologique (lymphome lymphocytaire de cytologie atypique). Des prélèvements de cette biopsie ganglionnaire avaient également été transmis en cytologie, cytométrie en flux et cytogénétique. En cytologie, l'aspect était plutôt celui de grandes cellules d'aspect immunoblastique. Un immunophénotypage lymphocytaire avait alors été réalisé et montrait la présence d'un clone B représentant environ 30 % des cellules de l'échantillon et de profil CD5+, CD20+ d'expression modérée et de monotypie Kappa d'intensité faible. Une phase leucémique *a minima* était également retrouvée (lymphocytose à environ 2,5 G/L avec 1/3 de lymphocytes B monotypiques en cytométrie en flux, avec score de Matutes à 3, de profil CD5+, CD23+, FMC7+ fort, CD22+fort, monotypie Kappa d'intensité faible, CD11c+, compatible avec la cytologie atypique mise en évidence en anatomo-pathologie).

La présence en cytogénétique d'une translocation impliquant le gène *IGH* et un locus situé en 17q24 inconnu en font un cas intéressant. Ces résultats obtenus sur ganglion ont été confirmés sur sang.

Par ailleurs, l'étude du statut mutationnel *IGHV* retrouve un réarrangement *IGHV5-51/IGHD6-6/IGHJ6* de profil muté (taux de mutations de 3,5 %). L'électrophorèse des protéines sériques réalisée en Octobre 2012 retrouvait la présence de micropics multiples de composantes IgG Kappa et IgG Lambda.

Le cas de ce patient avait été discuté en réunion de concertation pluridisciplinaire, décidant d'une surveillance simple avec la réalisation de bilans sanguins tous les 4 mois et consultation médicale annuelle.

Nous avons sélectionné ce cas pour étude du remaniement entre *IGH* et le locus 17q24. Nous disposons de matériel en cytogénétique et biologie moléculaire sur ganglion et sur sang.



b- Patient n°2

Le deuxième patient est une femme chez laquelle le diagnostic de LLC a été réalisé en 2007 à l'âge de 75 ans. Elle est alors classée en stade Binet A et est restée en abstention thérapeutique. Un immunophénotypage lymphocytaire avait alors été réalisé : le score de Matutes était à 5 (monotypie Lambda), avec un CD38 positif. La cytologie ne retrouve pas d'atypie notable. L'électrophorèse des protéines sériques retrouvait une hypogammaglobulinémie à 4,1 g/L, et l'immunofixation ne mettait pas en évidence de composant monoclonal.

Début 2012, devant la survenue d'infections à répétition, une substitution par immunoglobulines polyvalentes est initiée.

Un bilan biologique complet est alors réalisé. Le statut IGHV est muté (réarrangement IGHV4-61/IGHJ4, taux de mutations à 4 %). En cytogénétique, une translocation impliquant les chromosomes 17 et 22 est mise en évidence dans 19 mitoses sur 20. Le locus au niveau du chromosome 17 est situé en 17q25. Il n'y a pas de del(17p) mise en évidence.

En Décembre 2012, devant des signes de progression de la maladie (augmentation de la lymphocytose et des adénopathies, et splénomégalie volumineuse), un traitement par Chloraminophène est initié. A partir du mois de Mars 2013, du Rituximab est également instauré pour 6 cures. Devant une bonne réponse, le traitement est arrêté en Août 2013 avec surveillance.

Début 2016, une reprise de la LLC est objectivée avec une lymphocytose à 111 G/L (14,2 G/L en Juin 2015), et une volumineuse splénomégalie. La cytogénétique réalisée alors ne met pas en évidence d'anomalie supplémentaire à la t(17;22). Un traitement par Rituximab-Bendamustine est alors mis en place en Avril 2016. Après 6 cures et l'obtention d'une bonne réponse (correction de la lymphocytose et absence de splénomégalie palpable), la patiente est en abstention thérapeutique. Les dernières nouvelles de Janvier 2018 rapportent une maladie stable.

Devant la présence d'une translocation impliquant le locus 22q11 (soit potentiellement *IgL*) et le locus 17q25 (soit proche de celui du patient n°1), ce dossier a été sélectionné. Nous disposons de matériel disponible en cytogénétique et biologie moléculaire.



c- Patient n°3

Le troisième patient est une femme chez laquelle a été diagnostiqué une LLC en 2009 au Royaume-Uni, à l'âge de 50 ans. Devant la présence de volumineuses adénopathies progressives, elle a bénéficié en 2010 d'un traitement par FCR (1 seule cure devant la survenue d'effets indésirables digestifs majeurs ainsi que de complications infectieuses).

En Janvier 2013, la patiente est vue en consultation au CHU de Limoges. Devant un stade Binet B progressif, un traitement par Rituximab-Bendamustine a été mis en place (6 cures). Une rémission complète avait alors été obtenue et une surveillance simple a alors été débutée.

Fin 2016, une reprise de la maladie avec un temps de doublement de la lymphocytose inférieur à 6 mois (Lymphocytes à 0,42 G/L en Septembre, 74 G/L en Décembre 2016), a motivé l'initiation d'un nouveau traitement. Depuis Janvier 2017 la patiente est sous Ibrutinib. En Mai 2017, devant la survenue d'infections de la sphère ORL et la présence d'une hypogammaglobulinémie majeure (1,7 g/L ; valeurs normales locales de 6,0 à 11 g/L), une substitution à base d'immunoglobulines polyvalentes a été mise en place.

Le score de Matutes est à 4 (chaîne légère de surface Lambda d'intensité medium), profil IGHV non muté (réarrangement IGHV3-21/ IGHJ6, taux de mutations somatiques à 0 %, assignée au subset 2). La cytologie de départ était commune, avec apparition au cours de l'évolution de prolymphocytes. La cytogénétique a mis en évidence un caryotype complexe avec notamment une translocation entre les chromosomes 8 et 14, impliquant *IGH* et n'entraînant pas de remaniement de *MYC* sur le chromosome 8 (vérification effectuée en FISH). Ce dossier a donc été sélectionné. Nous disposons de matériel en cytogénétique et biologie moléculaire sur sang. Les électrophorèses des protéines sériques réalisées depuis le diagnostic en 2012 ne mettent en évidence de pic monoclonal, mais une hypogammaglobulinémie.

2- LLC et électrophorèse des protéines sériques

Les critères de sélection et les caractéristiques des patients pour l'étude de l'impact pronostique de l'électrophorèse des protéines sériques sont détaillés dans l'article 1 (page 194).

3- Evaluation du logiciel Cov'Cop dans la LLC

Dans cette étude, 82 cas de LLC avec annotation cytogénétique complète ont été sélectionnés. Nous avons récupéré les données cliniques concernant le stade Binet et la survie sans



traitement. Les caractéristiques des patients au diagnostic sont résumées dans le Tableau 7. L'échantillon du diagnostic était disponible pour 50 patients.

		Total (n = 82)
Sexe		Hommes : 58,5 % (n = 48) Femmes : 41,5 % (n = 34)
Age (ans)		64,9 +/- 10,7
Stade Binet (n = 78)	A	76,9 % (n = 60)
	B	14,1 % (n = 11)
	C	9,0 % (n = 7)
Statut mutationnel IGHV (n = 77) (M-CLL)		42,9 % (n = 33)
Lymphocytose (G/L)		30,6 +/- 39,5
Hémoglobine (g/dL)		13,8 +/- 1,7
Plaquettes (G/L)		204,0 +/- 91,7
Cytogénétique (Caryotype +/- FISH et/ou QMPSF)	- del(17p)	11,0 % (n = 9)
	- del(11q)	20,7 % (n = 17)
	- trisomy 12	11,0 % (n = 9)
	- del(13q)	53,7 % (n = 44)
	- del(13q) isolée	25,6 % (n = 21)
Mutations	<i>NOTCH1</i>	9,8 % (n = 8)
	<i>SF3B1</i>	24,4 % (n = 20)
	<i>TP53</i>	20,7 % (n = 17)
	<i>BIRC3</i>	4,9 % (n = 4)
	<i>ATM</i>	21,7 % (n = 18)
Score de Rossi	Risque très faible	25,6 % (n = 21)
	Risque faible	15,9 % (n = 13)
	Risque intermédiaire	34,1 % (n = 28)
	Risque haut	24,4 % (n = 20)

Tableau 7 : Caractéristiques des patients de l'étude sur l'évaluation du logiciel Cov'Cop dans la LLC
La moyenne +/- l'écart-type sont indiquées pour l'âge, la lymphocytose, le taux d'hémoglobine et la numération plaquettaire.



B- Techniques conventionnelles

Les techniques conventionnelles utilisées pour cette thèse (extraction d'ADN, QMPSF (pour Quantitative Multiplex PCR of Short Fluorescent Fragments), Mutations somatiques caryotype et FISH) sont décrites dans l'article 1 (page 194).

Les caractéristiques des sondes de routine utilisée pour la recherche de translocation sont détaillées ci-après.

Concernant le locus *IGH*, deux sondes différentes ont été utilisées. La première est fournie par Kreatech (Leica, Wetzlar, Germany). Il s'agit d'une sonde de séparation, constituée de 2 signaux. Le signal rouge cible les gènes de jonction et de diversité, l'ensemble de parties constantes d'*IGH* ainsi que la région en aval du gène, tandis que le signal vert ne couvre qu'une partie de la région *IGHV* et la région en amont de celle-ci (Figure 32).

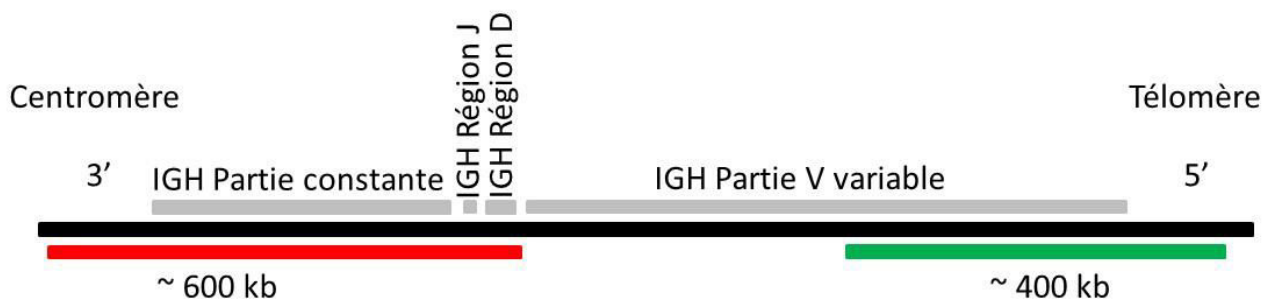


Figure 32 : Schéma de la sonde IGH Kreatech adapté selon les indications du fournisseur.

La seconde sonde utilisée ciblant *IGH* est la sonde Vysis LSI *IGH* Dual Color Break Apart Probe, fournie par Abbott (North Chicago, IL). Il s'agit également d'une sonde de séparation constituée de deux signaux : un signal vert (fluorochrome Spectrum Green) couvre l'ensemble de la région *IGHV* et un signal rouge-orange (fluorochrome Spectrum Orange) couvre la région en aval des régions constantes d'*IGH* (voir Figure 33).



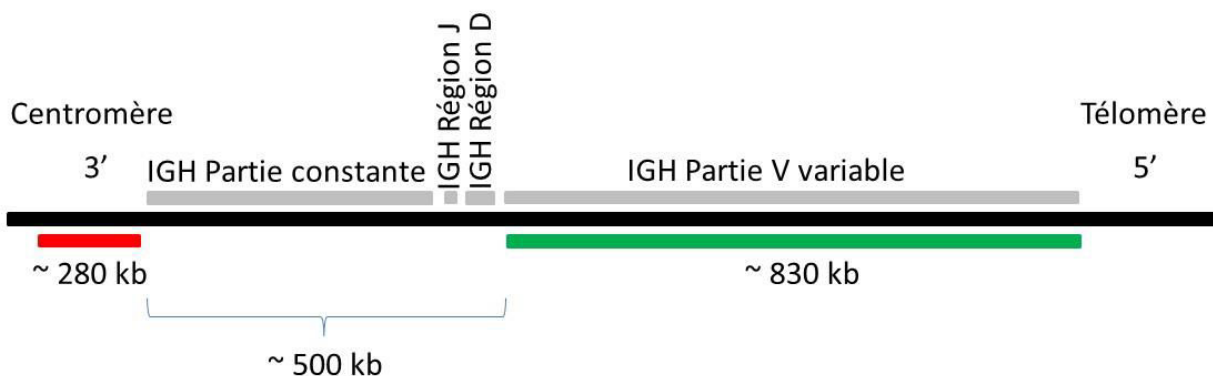


Figure 33 : Schéma de la sonde IGH LSI Vysis, adapté selon les indications du fournisseur

Une sonde ciblant le locus *IGL* a été utilisée. Il s'agit d'une sonde fournie par Dako (Agilent technologies, Santa Clara, CA). Il s'agit d'une sonde de séparation constituée d'un signal rouge ciblant la région télomérique d'*IGL* (segments de jonction et de régions constantes) ainsi que la région en aval du gène (Figure 34). Deux sondes marquées par de la fluorescéine (signal vert) ciblent *IGLV*.



Figure 34 : Schéma de la sonde IGL Dako, adapté selon les indications du fournisseur.

Une sonde ciblant *MYC* a également été utilisée. Il s'agit de la sonde *MYC* FISH DNA Probe, Split signal de chez Dako (Agilent technologies, Santa Clara, CA). Cette sonde est constituée de deux signaux (vert et rouge), ciblant les régions de part et d'autres de la principale région de cassure (breakpoint cluster region) retrouvée dans le gène *MYC*, entre les exons 1 et 2 et concernant les cas de lymphome de Burkitt sporadique (Figure 35). Les deux sondes couvrent également les régions de cassure en 5' (retrouvées dans le lymphome de Burkitt endémique) et 3' (retrouvées dans certaines translocations impliquant les gènes des chaînes légères des immunoglobulines) de la breakpoint cluster region⁵⁷. Cela permet théoriquement de détecter l'ensemble des réarrangements impliquant *MYC*.



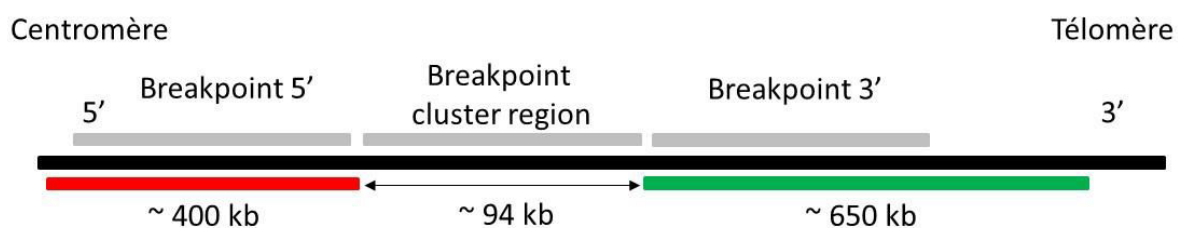


Figure 35 : Schéma de la sonde MYC Dako, adapté selon les indications du fournisseur.

Une seconde sonde MYC, fournie par Metasystems (Altlußheim, Germany) a été utilisée (Figure 36). Elle est constituée de deux signaux, un rouge et un vert situés de part et d'autre du gène MYC, espacées de 1,48 Mb.

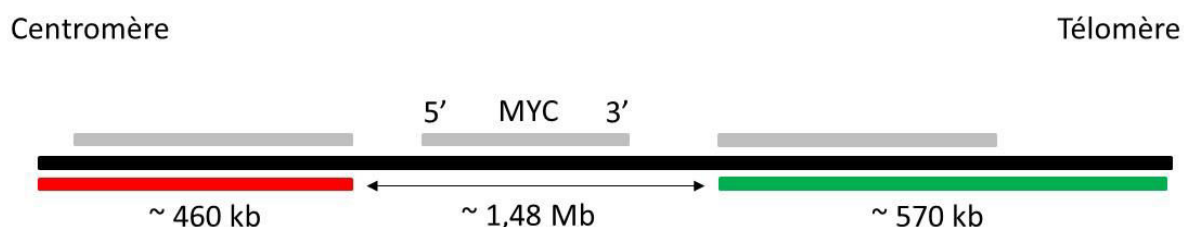


Figure 36 : Schéma de la sonde MYC Metasystem, adapté selon les indications du fournisseur.

C- Techniques spécialisées

1- FISH à façon

Afin de préciser les points de cassure retrouvés chez nos patients, nous avons dessiné des sondes de FISH à façon.

Nous avons pour cela utilisé des BACs (pour Bacterial Artificial Chromosome) et des fosmidés. Ceux-ci ont été choisis à partir des bases de données afin de cibler différentes régions au niveau des chromosomes 14 et 17. La méthode utilisée est dérivée du protocole décrit par Bertrand *et al.*⁶⁰² et consiste à interroger 2 bases de données différentes afin de limiter les erreurs de localisation possibles. La spécificité des séquences est vérifiée avec NCBI BLAST⁶⁰³. Les BACs et fosmidés sont ensuite commandés à BACPAC Ressources (Children's Hospital Oakland Research Institute, Oakland, CA). En parallèle, des amorces sont dessinées et commandées pour chaque sonde afin de vérifier par PCR la conformité de la séquence.



Une fois reçus, les BACs et fosmidés sont préparés et mis en culture avec des bactéries (*Escherichia coli*) sur milieu LB (pour Lysogeny Broth) pendant une nuit à 37°C. La suspension bactérienne est ensuite centrifugée et le culot obtenu est utilisé pour l'extraction d'ADN avec le kit « Wizard Plus SV Minipreps DNA Purification System » (Promega, Fitchburg, WI).

Les BACs ou fosmidés sont ensuite marqués à l'aide d'un fluorochrome (Spectrum Orange ou Spectrum Green ou Aqua, Abbott (North Chicago, IL)) par technique de Nick translation (kit Nick Translation, Abbott (North Chicago, IL)). Cette technique utilise une DNase entraînant des cassures simple brin de l'ADN cible (Figure 37). L'ADN polymérase I répare la brèche obtenue au niveau de l'extrémité 3'OH libre et continue l'élongation dans le 5'→3', le long du brin. Cela permet l'incorporation de dUTP fluorescents par l'ADN polymérase I. La suspension obtenue est ensuite évaporée et reprise dans 10 µL d'eau pour utilisation en FISH.

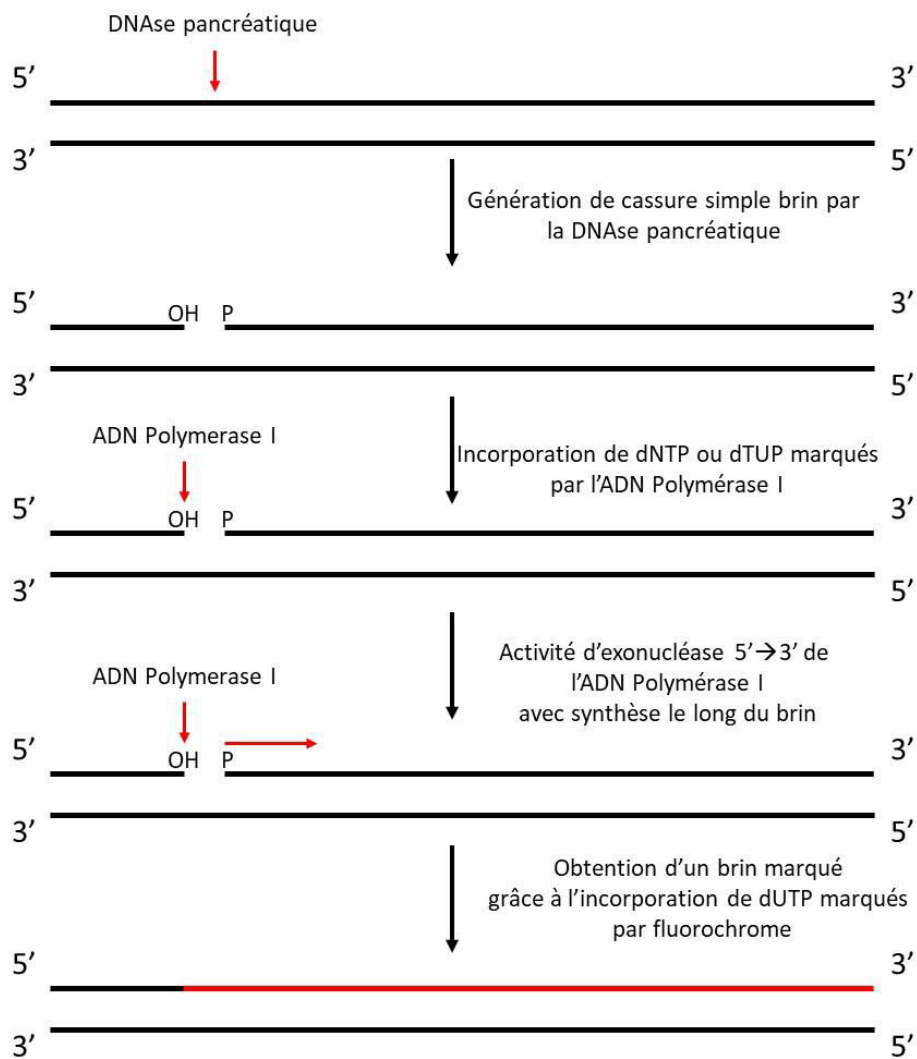


Figure 37 : Principe de la technique de Nick translation



Dans une première approche, des BACs ciblant le chromosome 17 ont été choisis. Ils ciblaient les régions 17q23 (RP11-661B17), 17q24 (RP11-387O17) et 17q25 (RP11-467J3).

Dans un deuxième temps, nous avons sélectionné des BACs ciblant la région IGHV (RP11-47P23, RP11-379F22 et RP11-72N10) (Figure 38).

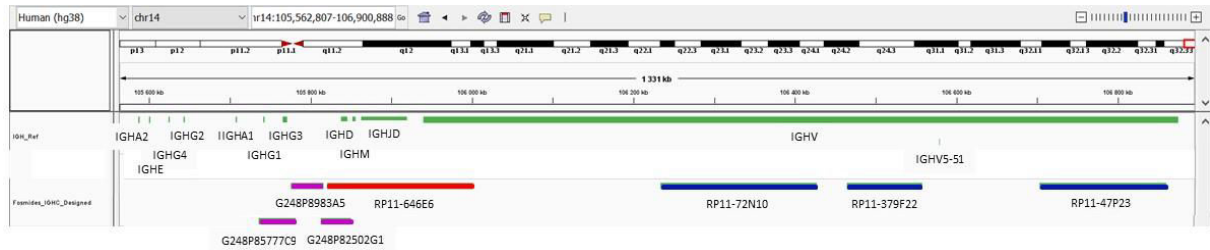


Figure 38 : Schéma de la localisation des BAC et fosmidés utilisés au niveau d'IGH.
Avec en bleu les BAC utilisés au niveau d'IGHV, en violet les fosmidés ciblant les régions IGHC et enfin en rouge un BAC intermédiaire.
Le segment IGHJD représente les régions de jonction et diversité.

Enfin, nous avons utilisé des fosmidés ciblant les régions constantes d'IGH (G248P85777C9, G248P8983A5 et G248P82502G1) et nous avons complété avec un BAC intermédiaire ciblant la région allant d'IGHD à la portion 5' d'IGHV (RP11-645E6).

2- LM-PCR et séquençage

Pour compléter l'approche cytogénétique, nous avons opté pour une technique de génétique moléculaire capable de détecter les points de cassure des translocations d'intérêt : la LM-PCR (pour Ligation Mediated-PCR). Cette technique a été utilisée pour la recherche de partenaires impliqués dans des translocations du TCR⁶⁰⁴.

Cette technique consiste à digérer l'ADN à tester avec des enzymes de restriction générant des bouts francs (Figure 39). Les enzymes utilisées et les sites de restriction correspondant sont indiqués dans le Tableau 8.



Enzyme de restriction	Séquence reconnue
PvuII	CAG-CTG
EcoRV	GAT-ATC
SmaI	CCC-GGG
StuI	AGG-CCT
SspI	AAT-ATT
DraI	TTT-AAA

Tableau 8 : Enzymes de restriction utilisées pour la technique de LM-PCR et séquence reconnue correspondante (sens 5'→3'). Le tiret dans la séquence indique le site de restriction.

Après digestion enzymatique, des adaptateurs sont ligués. Ensuite deux PCR successives sont réalisées, la deuxième étant une PCR nichée. Cette approche permet d'améliorer la spécificité de la technique. Les amorces utilisées sont, d'une part des amorces spécifiques du locus connu d'intérêt, dessinées par nos soins (dans notre cas, les gènes des immunoglobulines), nommées GSP1 et GSP2 respectivement pour la PCR1 et la PCR2, et d'autre part une amorce ciblant l'adaptateur ligué (nommée AP1 pour la première PCR et AP2 pour la PCR nichée). Les tailles attendues en l'absence de remaniement sont recherchées par alignement sur le génome. La révélation se fait par migration électrophorétique sur gel d'agarose à 1,5 % contenant du Bromure d'Ethidium, avec visualisation sous lampe à Ultra-Violets. Toute bande de taille différente doit être découpée pour purification et séquençage. Par sécurité, j'ai découpé toutes les bandes visibles. Enfin, chaque bande est purifiée à l'aide du kit Nucleospin Gel and PCR clean-up (Macherey Nagel, Düren, Germany) et une réaction de séquence est réalisée en sens et antisens (amorces GSP2 et AP2) avec du BigDye Terminator v1.1 (Applied, Thermo Fisher, Waltham, MA). La purification est alors faite à l'aide de colonnes de Sephadex G50 (VWR, Radnor, PA). Enfin, les séquences sont passées sur séquenceur capillaire ABI-Prism 3130xl (Thermo Fisher, Waltham, MA).

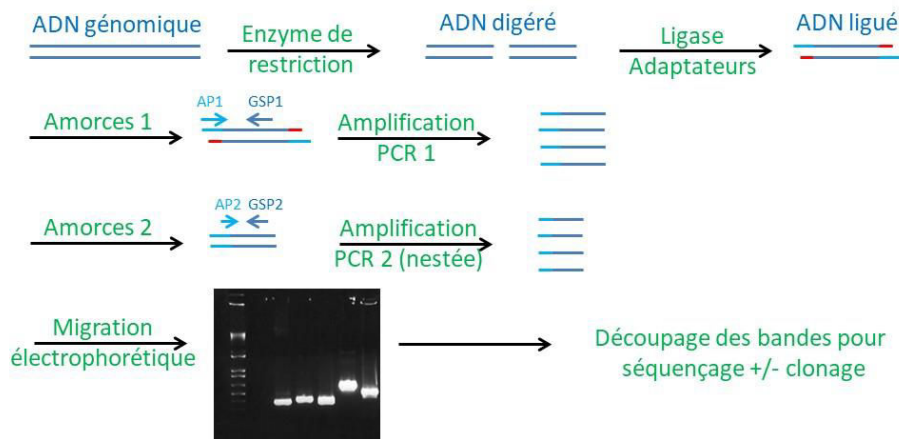


Figure 39 : Principe de la LM-PCR

AP1 et AP2 (pour Adaptator Primer 1 ou 2) sont les amorces ciblant l'adaptateur ligué ; GSP1 et GSP2 (pour Gene Specific Primer 1 ou 2) sont les amorces spécifiques du locus d'intérêt connu.



Le kit utilisé pour la réalisation de la LM-PCR est le Universal Genome Walker 2.0 (Takara Bio, Kyoto, Japan).

Nous avons dessiné des amorces au niveau des points de cassure potentiels tels que décrits dans la littérature. Concernant *IGH*, il s'agit principalement des régions IGJ et IGHM (au niveau de l'enhancer et du Switch μ ou à proximité), mais aussi des régions switch des autres régions constantes. Concernant *IGL*, les points de cassure concernent les différents gènes de jonction. Nous avons ainsi dessiné plusieurs centaines d'amorces correspondant aux spécifications de la technique, à savoir :

- une longueur de 26 à 30 nucléotides ;
- un pourcentage en GC compris entre 40 et 60 % ;
- l'absence d'hybridation côté 3' de l'amorce avec l'adaptateur ;
- un maximum de 3 G/C dans les 6 dernières positions en 3' de l'amorce.

De plus, dans la mesure du possible, l'amorce nichée ne doit pas être chevauchante avec la première amorce. Si le chevauchement est obligatoire, le côté 3' de l'amorce nichée doit avoir une séquence la plus unique possible.

3- CGH-array

La CGH-array (pour Comparative Genomic Hybridization-array) est une technique de cytogénétique moléculaire. Elle repose sur l'hybridation d'un ADN test marqué par un fluorochrome (en général vert) en comparaison à un ADN normal (en général marqué par un fluorochrome rouge) sur une puce contenant un grand nombre de sondes ciblant les différents chromosomes ou régions d'intérêt dans certaines pathologies. L'hybridation est réalisée sur lame. Différentes résolutions existent.

Nous avons réalisé une analyse de CGH-array pour le patient n°1 de l'étude des remaniements atypiques impliquant les gènes des immunoglobulines. Cette technique a été réalisée avec la technologie Agilent (Agilent Technologies, Santa Clara, CA). Une puce 400k a été utilisée.



4- Séquençage haut débit (NGS)

Dans ce travail, nous avons utilisé deux technologies différentes pour la réalisation du NGS.

Concernant l'étude des translocations impliquant les gènes des immunoglobulines, une approche de séquençage du génome entier « paired-end » a été réalisée sur HiSeq 4000 Illumina (San Diego, CA). La réalisation de ce séquençage a été sous-traitée à Integragen (Evry, France).

Pour l'étude réalisée dans le cadre de l'étude de l'impact pronostique de l'électrophorèse des protéines sériques dans la LLC, deux panels ciblés de gènes ont été mis au point pour la technologie IonProton (ThermoFisher, Waltham, MA).

a- Technologie Illumina

i- Principe

La technologie de NGS Illumina repose sur un séquençage par synthèse, utilisant l'incorporation de nucléotides marqués par un fluorophore, suivie en temps réel par caméra CCD. L'utilisation de « paired-end », c'est-à-dire d'un séquençage dans les deux sens des fragments d'ADN, permet d'améliorer la détection de réarrangements.

ii- Application au séquençage « whole-genome »

Le séquençage « whole genome », c'est-à-dire du génome entier, a été réalisé pour les trois patients d'intérêt retenus dans l'étude de réarrangements atypiques impliquant les gènes des immunoglobulines. L'ADN génomique des patients a été utilisé pour la réalisation de bibliothèques avec obtention de fragments d'environ 400 pb.

Pour les patients n°2 et 3, nous disposons de prélèvements largement infiltrés (plus de 90 % de cellules porteuses de l'anomalie), et une profondeur de lecture de 15-20x a été appliquée.

Concernant le patient n°1, nous avons utilisé de l'ADN extrait à partir de ganglion qui était plus infiltré que le sang. Sur ce prélèvement, environ 50 % des cellules étaient porteuses de l'anomalie. Après discussion avec Integragen, une profondeur de 15-20x a semblé suffisante pour la détection du remaniement chez ce patient.

Les données générées par le séquençage ont été alignées sur le génome de référence GRCh38.



iii- Analyse MANTA

L'analyse des données générées par le séquençage du génome entier pour la détection des remaniements chromosomiques a été réalisée grâce au logiciel MANTA. Il s'agit d'un logiciel développé par Illumina (San Diego, CA) permettant la détection de variants structuraux dont les insertions-délétions et les translocations chromosomiques.

Le logiciel MANTA utilise les fichiers bam (pour Binary Alignment Map) générés par le séquençage et procède en deux étapes⁶⁰⁵. Dans un premier temps, un graphique de toutes les associations de cassure dans le génome est construit. Dans un second temps, les composants de ce graphique sont travaillés pour la génération d'hypothèses de variants, l'assemblage, la génération d'un score et d'un rapport sous format vcf (pour variant call format). Ce logiciel permet d'obtenir un rapport en un temps limité (jusqu'à 4-5h contre jusqu'à 4-5 jours avec d'autres logiciels)⁶⁰⁵.

b- Technologie Ion Proton

i- Principe

La Technologie IonProton a été développée par Life Technologies (ThermoFisher, Waltham, MA). Il s'agit d'une technique de NGS par synthèse, avec détection en temps réel de l'incorporation des bases entraînant la libération d'un proton. C'est la modification du pH qui est alors mesurée.

Dans notre cas, le séquençage nécessite la préparation de bibliothèques par amplification clonale des régions d'intérêt à étudier. Pour cela, nous avons mis au point deux panels Ampliseq.

La réalisation des bibliothèques a été faite avec le « Ion AmpliSeq Library Kit 2.0 » (ThermoFisher, Waltham, MA) selon les recommandations du fournisseur.

Les échantillons ont ensuite été analysés sur un séquenceur IonProton (ThermoFisher, Waltham, MA) au niveau de la plateforme de séquençage du CHU de Limoges.



ii- Définition des gènes d'intérêt

Nous avons mis au point un premier panel de NGS pour la détection des mutations d'intérêt diagnostique et/ou pronostique dans les lymphomes. Nous avons fait une revue de la littérature et suivi les recommandations des groupes de travail afin de retenir les gènes et régions d'intérêt.

Les gènes d'intérêt dans la LLC intégrés dans ce premier panel sont *TP53*, *ATM*, *NOTCH1*, *SF3B1*, *BIRC3*, *BRAF*, *FBXW7*, *MYD88*, *XPO1*, *POT1*, *EGR2*, *NKBIE*, *DDX3X*, *RPS15*, *PLCG2* et *BTK* (Tableau 9).

Le deuxième panel dessiné ciblait uniquement *TP53* et *ATM* pour la recherche de mutations.

iii- Définition d'amplicons de référence pour l'analyse de déséquilibres génétiques

Dans l'optique d'utiliser notre panel de NGS pour détecter les variations du nombre de copies dans la LLC, des régions génomiques spécifiques ont été intégrées dans le panel Ampliseq lymphoïde (Tableau 9). Le séquençage de certains gènes tels que *TP53* ou *ATM* permet de couvrir ces régions d'intérêt (détection de la del(17p) et de la del(11q)) avec suffisamment d'amplicons. D'autres gènes ont été ajoutés pour assurer la détection de la del(13q) (*DLEU1*, *DLEU2*, *MIR16-1*, *MIR15A* et *RB1*) et de la trisomie 12 (*POU6F1*, *MDM2*, 2 cibles utilisées en QMPSF). D'autres régions ont été ajoutées au niveau du chromosome 8 pour détecter les gains 8q (*EIF3H*, *MYC*), et au niveau des chromosomes 18 (*MYOM1*, *DCC* et *BCL2*) et 19 (*CACNA1A* et *CCNE1*) pour la détection de ces trisomies. Enfin, des régions contrôles supposées normales ont été intégrées (chromosome 3 avec *EPHA3*, *PROS1* et *TBL1XR1*, chromosome 4 avec *ZAR1*, chromosome 7 avec *SUMF2* et *ASL*, chromosome 9 avec un amplicon ciblant *JAK2*, et chromosome 22 avec un amplicon ciblant *CECR1*).



Chromosome	Gène	Région	Référence	Cible pour CNV seulement	Cible pour mutation récurrente dans la LLC	Cible pour mutation récurrente dans autres lymphomes
Chr1	SPEN	E11	NM_015055			X
Chr1	ARID1A	from E1 to 20	NM_006015			X
Chr1	NOTCH2	from E26 to 28 + E34	NM_024408			X
Chr2	REL	from E1 to E10(with UTR)	NM_001291746			X
Chr2	XPO1	E14, E15	NM_003400		X	
Chr2	CXCR4	E1, E2	NM_003467			X
Chr2	SF3B1	from E1 to E25	NM_012433		X	
Chr3	MYD88	from E3 to E5	NM_001172567		X	
Chr3	EPHA3	E1, E3, E5, E7, E9, E11, E13, E15, E17	NM_005233	X		
Chr3	PROS1	E2, E4, E6, E8, E9, E10, E11, E13, E15	NM_000313	X		
Chr3	TBL1XR1	E3, E6, E7, E8, E12	NM_024665	X		
Chr3	KLHL6	E6, E7	NM_130446			X
Chr4	ZAR1	from E1 to E4	NM_175619	X		
Chr4	RASSF6	from E1 to E11	NM_201431			X
Chr4	PTPN13	from E2 to E48	NM_080685			X
Chr4	FAT4	E1	NM_024582			X
Chr4	PCDH10	E1	NM_032961			X
Chr4	FBXW7	from E2 to E12	NM_033632		X	
Chr5	EGR1	E1, E2	NM_001964			X
Chr6	IRF4	from E2 to E9	NM_002460			X
Chr6	NFKBIE	from E1 to E6	NM_004556		X	
Chr6	PRDM1	from E1 to E7	NM_001198			X
Chr6	TNFAIP3	from E2 to E9	NM_006290			X
Chr7	CARD11	from E2 to E9	NM_032415			X
Chr7	SUMF2	from E1 to E9	NM_015411	X		
Chr7	ASL	E1, E2, E4, E6, E8, E10, E12, E14, E15, E16, E17	NM_000048	X		
Chr7	POT1	from E5 to E19	NM_015450		X	
Chr7	BRAF	E11, E15	NM_004333		X	
Chr7	EZH2	E16, E18	NM_004456			X
Chr8	IKBKB	from E2 to E22	NM_001556			X
Chr8	EIF3H	from E1 to E8	NM_003756	X		
Chr8	MYC	from E1 to E3	NM_002467			X
Chr9	JAK2	E7	NM_004972	X		
Chr9	NOTCH1	from E25 to E28 + E34	NM_017617		X	
Chr10	EGR2	E3, E4	NM_001136178		X	
Chr11	SWAP70	from E1 to E3	NM_015055			X
Chr11	BIRC3	from E6 to E9	NM_001165		X	
Chr11	ATM	from E1 to E62	NM_000051		X	
Chr12	SOX5	from E1 to E15	NM_006940			X
Chr12	KMT2D	from E1 to E54	NM_003482			X
Chr12	POU6F1	from E2 to E6	NM_002702	X		
Chr12	MDM2	E1, E3, E5, E7, E9, E11	NM_002392	X		
Chr12	DTX1	E2	NM_004416			X
Chr13	RB1	E2, E4, E6, E8, E10, E12, E14, E16, E18, E20, E22, E24, E26	NM_000321	X		
Chr13	DLEU2	E1, E2, E4, E5, E7, E9, E10		X		
Chr13	Mir16-1	E1		X		
Chr13	Mir15A	E1		X		
Chr13	DLEU1	from E1 to E6		X		
Chr14	TRAF3	E3, E9, E10, E12	NM_145725			X
Chr15	SIN3A	E4, E8	NM_001145357			X
Chr16	CREBBP	from E1 to E31	NM_004380			X
Chr16	PLCG2	from E18 to E23	NM_002661		X	
Chr17	ATP2A3	E2, E5, E8, E11, E14, E17, E20, E23	NM_174953	X		
Chr17	TP53	from E2 to E11	NM_000546		X	
Chr17	MAP3K14	E12	NM_003954	X		
Chr17	CD79B	from E1 to E6	NM_001039933			X
Chr18	MYOM1	E2, E6, E8, E10, E12, E14, E16, E20, E22, E24, E26, E28, E30, E32, E34, E36, E38	NM_003803	X		
Chr18	DCC	E1, E3, E5, E7, E9, E11, E13, E15, E17, E19, E21, E23, E25, E29	NM_005215	X		
Chr18	BCL2	E2, E3	NM_000633			X
Chr18	RPS15	from E1 to E4	NM_001018		X	
Chr19	CACNA1A	E1, E4, E7, E10, E13, E16, E19, E41, E44, E47	NM_000068	X		
Chr19	KLF2	from E1 to E3	NM_016270			X
Chr19	CCNE1	from E1 to E11	NM_001238	X		
Chr19	CD79A	from E1 to E5	NM_001783			X
Chr22	CECR1	E2	NM_001282228	X		
Chr22	MAPK1	from E1 to E8	NM_002745			X
Chr22	XBP1	from E1 to E5	NM_005080			X
ChrX	DDX3X	from E1 to E17	NM_001356		X	
ChrX	ZMYM3	from E2 to E25	NM_001171162			X
ChrX	BTK	from E14 to E16	NM_000061		X	

Tableau 9 : Liste des cibles incluses dans le panel NGS lymphoïde.

Concernant le panel réduit à *TP53* et *ATM*, le but était la détection de la del(11q) et de la del(17p) uniquement. Des gènes contrôles ont été ajoutés : au niveau du chromosome 11 (*SWAP70*, gène situé en 11p, permettant de différencier une del(11q) d'une monosomie 11), au niveau du chromosome 17 (*MAP3K14*, *CD79B*, gènes situés en 17q). Les gènes contrôles



normaux sont *JAK2* sur le chromosome 9, *CECR1* sur le chromosome 22. De plus, un gène situé sur le chromosome Y (*SRY*) a été intégré afin d'améliorer la normalisation, y compris pour des puces déséquilibrées en terme de sex-ratio, et de détecter automatiquement les hommes et les femmes.

iv- Mise au point de panels NGS

Le panel Ampliseq lymphoïde commandé a été développé sur la plateforme en ligne AmpliSeq Designer. Il couvre 221,6 kb. 70 gènes ont été intégrés (soit exome, soit régions ciblées). Il s'agit d'un design en 2 pools avec un total de 1206 amplicons. Afin d'améliorer la couverture obtenue, en particulier au niveau de certains gènes d'intérêt, nous avons choisi une option de spécificité relâchée. La couverture théorique obtenue est de 97,68 %.

Un premier test avait été réalisé avec 5 patients par puce. La profondeur obtenue au niveau de *TP53* semblait suffisante avec cette configuration (minimum de 100 lectures nécessaires d'après les recommandations³⁵³) : la profondeur moyenne obtenue par amplicon était alors d'environ 9000X. Dans cette configuration, les régions d'intérêt les moins bien couvertes (telles que l'exon 5 de *TP53*) ont été séquencées avec une profondeur minimale d'environ 300X. Nous avons ensuite testé nos différents échantillons avec une configuration avec 7 à 8 patients par puce, permettant d'obtenir une profondeur de lecture moyenne d'environ 7000X par amplicon, avec un minimum de 100X pour l'exon 5 de *TP53*.

Le panel Ampliseq réduit à *TP53* et *ATM* a été développé de la même façon. Il couvre 30,22 kb. Il s'agit d'un design en 2 pools avec un total de 172 amplicons. La couverture théorique obtenue est de 96,39 %. La spécificité retenue est haute, la couverture des régions d'intérêt étant théoriquement suffisante. 50 patients ont été passés simultanément par puce. La couverture au niveau de *TP53* était suffisante (minimum de 100 lectures).

v- Logiciel CovCop

Le logiciel CovCop a été développé par les Drs Paco Derouault et Anne-Sophie Lia de l'EA6309 de la faculté de Médecine de Limoges. Son principe a été publié dans la revue *Bioinformatics* en 2016⁶⁰⁶. Il a été spécialement développé pour la détection des variations du nombre de copies après NGS basé sur la réalisation de bibliothèques par amplicons (amplification de gènes ou régions géniques ciblées) telles que celles obtenues avec nos panels Ampliseq. Le principe de



ce logiciel est basé sur l'analyse du nombre de lectures par amplicon et par patient au sein d'une même puce de séquençage. Plusieurs étapes de correction et de normalisation sont réalisées. Tout d'abord, les nombres de lectures par amplicon de chaque échantillon sont corrigés par le total des amplicons de chaque librairie. Une correction par rapport au pourcentage en GC et la longueur de chaque amplicon est ensuite réalisée. En effet, ces paramètres influencent l'efficacité d'amplification par PCR. L'étape de normalisation est ensuite appliquée : les amplicons ayant la variance la plus faible sont sélectionnés automatiquement comme amplicons de référence. Un ratio normalisé est alors généré pour chaque amplicon en divisant leur nombre de lectures par la médiane des lectures des amplicons de référence sélectionnés automatiquement. En l'absence d'anomalie de nombre, le ratio normalisé est centré sur la valeur 1. Ce ratio prendra la valeur théorique de 1,5 en cas de trisomie ou de gain d'un allèle et tendra vers la valeur de 0,5 en cas de délétion mono-allélique.

Il est possible d'assigner à chaque patient le sexe (homme ou femme) afin d'éviter de mauvaises interprétations d'amplicons situés sur le chromosome X. Il est également possible d'intégrer des amplicons situés sur le chromosome Y pour permettre une détection automatique. CovCop corrige alors les valeurs des amplicons situés sur le chromosome X en multipliant par 2 les valeurs obtenues chez les hommes. Ceci peut induire un biais qui reste limité si le nombre d'amplicons situés sur le chromosome X est faible. Une autre façon de s'affranchir de ce biais est de réaliser des puces ne contenant que des échantillons provenant d'hommes ou de femmes. Le seuil de détection des variations du nombre de copies peut être ajusté par l'utilisateur. Un minimum de 4 amplicons successifs variant dans le même sens est nécessaire pour considérer un locus comme potentiellement délété ou dupliqué.

Cov'Cop a été initialement développé pour la recherche d'anomalies constitutionnelles. Un minimum de 3 échantillons par puce comprenant au moins 100 amplicons chacun est nécessaire pour un fonctionnement optimal de Cov'Cop dans cette indication.

vi- Utilisation de la fréquence allélique des variants pour la détection des CNV

Le NGS permet d'identifier de nombreux SNP. La fréquence allélique de ces variants (ou VAF ou Variant Allelic Frequency) dans l'échantillon est d'environ 50 % en cas d'hétérozygotie, et de 100 % en cas d'homozygotie. Une valeur déviante de 50 % peut être observée en cas de CNV. Nous avons donc utilisé la VAF déterminée par NGS pour la détection des CNV. Le panel NGS utilisé est le panel de 70 gènes.

Nous avons mis au point et validé un algorithme (script R) permettant de sélectionner automatiquement les SNP détectés par NGS à partir des fichiers vcf comprenant les différents variants rapportés par NGS, et annotés à partir des bases de données des SNP (1000Genomes



et dbSNP). Le critère de sélection des SNP est une fréquence allélique dans la population générale supérieure à 1 %. Certains SNP ayant des valeurs aberrantes (c'est-à-dire des valeurs éloignées systématiquement de 50 ou 100 %, en lien avec des difficultés de séquençage) ont été éliminés. Les données issues des 82 patients de l'étude sont utilisées pour calculer la moyenne et l'écart-type des différents SNP situés sur un même amplicon. Des courbes ROC (pour Receiving Operator Characteristic) ont été construites afin de déterminer le seuil de normalité pour les amplicons. Celui-ci a été choisi afin d'obtenir une bonne spécificité et a été fixé à $\pm 2,5$ écart-types par rapport à la moyenne. Pour les SNP représentés moins de 5 fois dans la base de données, la moyenne et l'écart-type des SNP de tous les échantillons est appliquée. Le choix d'un seuil à $\pm 2,5$ écart-type a été fixé afin de privilégier la spécificité du test par rapport à la sensibilité de détection d'une anomalie.

Dans ce script nous avons inclus une représentation graphique des VAF avec les bornes de normalité par amplicon et l'analyse visuelle rapide des résultats, facilitant le repérage des anomalies présentes dans l'échantillon. Un package R est en cours de réalisation.



Chapitre 5 : Résultats

A- Etude de translocations atypiques dans les lymphomes B

1- Patient 1

a- Résultats préliminaires : Caryotype et FISH de routine

Le premier patient étudié a bénéficié d'un caryotype avec stimulation par IL2/DSP30 à partir d'une biopsie ganglionnaire. Celui-ci met en évidence un caryotype complexe comprenant des trisomies 12 et 19 ainsi qu'une translocation entre les bras longs des chromosomes 13 et 14 (Figure 40).

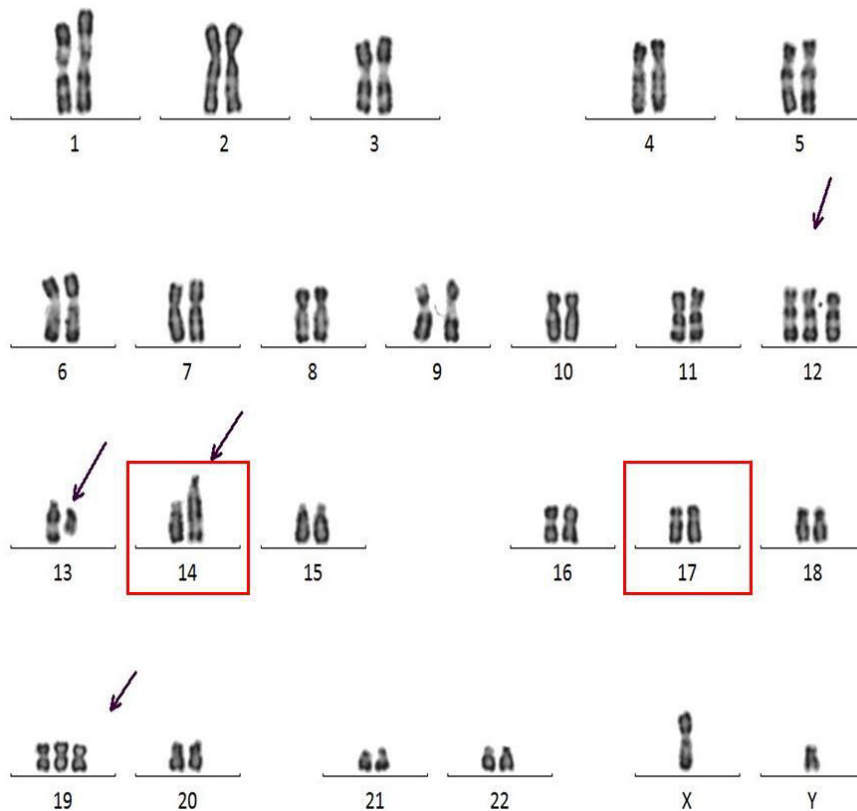


Figure 40 : Caryotype du patient n°1.

Il met en évidence une translocation entre les bras longs des chromosomes 13 et 14, ainsi que des trisomies des chromosomes 12 et 19. Le deuxième chromosome 14 (non impliqué dans la t(13;14)) et les deux chromosomes 17 paraissent normaux au caryotype.



Devant la présence d'une translocation impliquant le locus 14q32, avec potentiellement un remaniement d'*IGH*, une analyse en FISH a été réalisée ciblant *IGH* (Figure 41). La sonde utilisée est celle fournie par Kreatech (Figure 32, page 144).

A



B

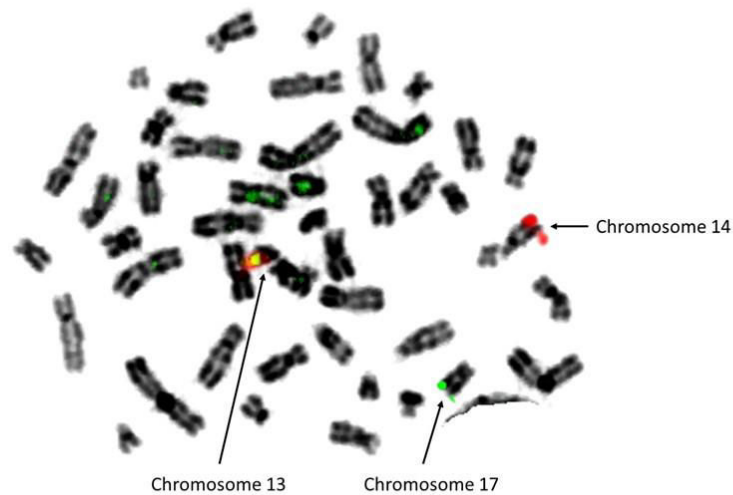


Figure 41 : Image de FISH IGH (sonde Kreatech), concernant le patient n°1.

A- Représentation schématique à partir d'IGV de la sonde Kreatech ciblant *IGH*. La sonde utilisée cible en rouge la région 3' du gène *IGH* (côté centromère, régions constantes jusqu'aux gènes de diversité) et en vert la partie télomérique de la région variable du gène *IGH*.

B- Image de FISH métaphasique avec la sonde Kreatech ciblant *IGH* concernant le patient n°1. Un signal vert est retrouvé sur un chromosome 17, un signal rouge est observé sur le chromosome 14 (non impliqué dans la t(13;14)) et une fusion est visualisée sur le dérivé 13 issu de la translocation entre les chromosomes 13 et 14.



Le résultat montre que la translocation entre les chromosomes 13 et 14 entraîne une délocalisation de l'ensemble du locus *IGH* sur le dérivé 13 (signal de fusion observé sur celui-ci), sans remaniement d'*IGH* (Figure 41). Elle met également en évidence un remaniement concernant l'autre chromosome 14, non vu au caryotype, avec délocalisation d'un signal vert sur un chromosome 17.

La formule chromosomique rendue sur ce prélèvement était alors :

48,XY,+12,+19[17]/48,sl,t(13;14)(q13;q32)[3]

.ish+12(TEL+)[2],t(13;14)(IGH+;IGH-)[5];t(14;17)(3'IGH+;5'IGH+)[5]

.nuc ish(TELx3,AML1x2)[65/110],(IGHx2)(5'IGHsep3'IGHx1)[30/50]

Au final, le patient présente donc un remaniement d'*IGH* cryptique avec implication d'un locus en 17q24. De plus la t(14;17) est présente dans le clone initial (proportion identique à la trisomie 12 sur FISH interphasique) tandis que la t(13;14) est sous-clonale. La translocation impliquant *IGH* est donc un événement précoce dans la lymphomagénèse de ce cas.

L'étude de la littérature retrouve un cas de lymphome B diffus des séreuses dans lequel une translocation t(14;17)(q32;q23), également cryptique, avait été identifiée⁶⁰⁷. Le gène partenaire n'avait pas été mis en évidence. Une étude concernant l'impact pronostique des translocations impliquant les gènes des immunoglobulines dans la LLC rapporte un cas impliquant un partenaire en 17q non identifié³⁴⁴. Les auteurs retrouvent un impact pronostique péjoratif des translocations impliquant les gènes des immunoglobulines, à l'exception de la t(14;18). Nous avons également repris les gènes situés au niveau des locus 17q23 à 17q25 sans identifier de gène connu comme impliqué dans les lymphomes B. Enfin, l'atlas de cytogénétique répertorie 43 partenaires différents identifiés dans des translocations impliquant *IGH* mais le locus 17q24 n'y figure pas (dernière mise à jour en Février 2010)⁶⁰⁸. Il en est de même pour les partenaires de translocation impliquant *IGL* et *IGK*. Une translocation *IGH/IGF2BP1* en 17q21 a été rapportée en 2014 dans un cas de leucémie aiguë lymphoblastique B⁶⁰⁹, mais le point de cassure semble plus télomérique dans notre cas.



b- Résultats par LM-PCR et cytogénétique

Nous avons vu précédemment que les translocations impliquant *IGH* entraînent, dans la majorité des cas, des cassures au niveau des régions de jonction ou autour de l'enhancer E μ . Ces points de cassure sont compatibles avec les résultats obtenus en FISH avec la sonde *IGH* Kreatech. Les régions d'intérêt étant ciblées, nous avons opté pour la technique de LM-PCR. Nous avons mis au point des amorces permettant de cibler ces régions d'intérêt. Cependant, au cours de cette approche, le remaniement impliquant *IGH* n'a pu être identifié. La technique semble valide car elle a permis de retrouver le réarrangement VDJ supposé fonctionnel mis en évidence par clonalité B.

Puisque les translocations impliquant *IGH* peuvent également impliquer les régions switch des différentes parties constantes, nous avons ensuite dessiné des amorces permettant de cibler par LM-PCR les régions switch des parties constantes *IGHG1*, *IGHA1* et *IGHE* qui sont les plus fréquemment impliquées dans le lymphome de Burkitt⁶¹⁰. Néanmoins, la FISH avec la sonde *IGH* Kreatech, ciblant en rouge les régions constantes, de jonction et de diversité, devrait entraîner une délocalisation du signal rouge sur le chromosome 17 en cas de cassure dans les parties constantes les plus en 3', ce qui n'est pas le cas. Les résultats obtenus en LM-PCR n'ont pas permis de mettre en évidence de remaniement pour ces régions.

c- Résultats de l'étude en FISH à façon

i- Confirmation des résultats négatifs de LM-PCR sur les régions constantes d'IGH

L'absence de remaniement impliquant les régions constantes d'*IGH* selon la technique de LM-PCR a également été vérifiée par FISH avec des fosmides ciblant les régions constantes les plus en 5', et en aval d'*IGHM* (G248P85777C9, G248P8983A5 et G248P82502G1). Les 3 fosmides restent sur le chromosome 14 (Figure 42).



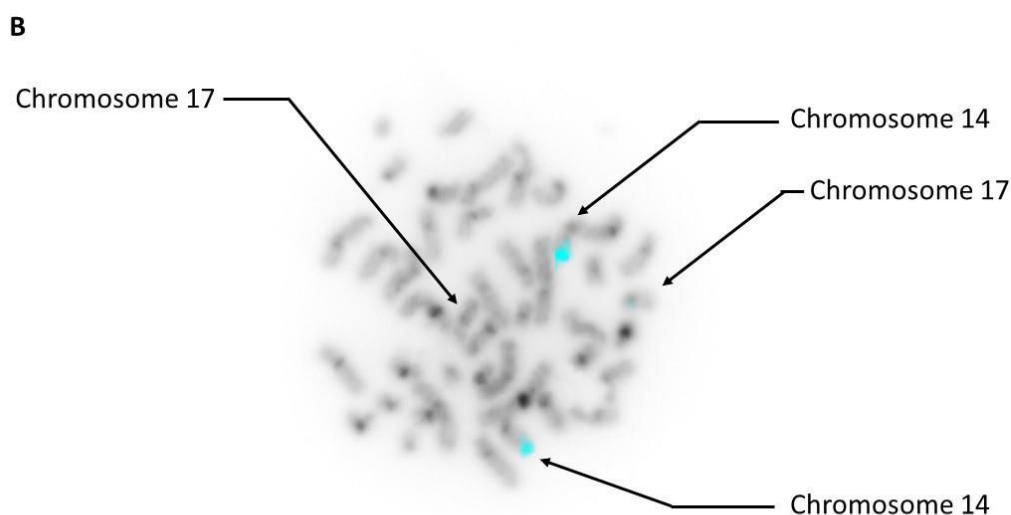
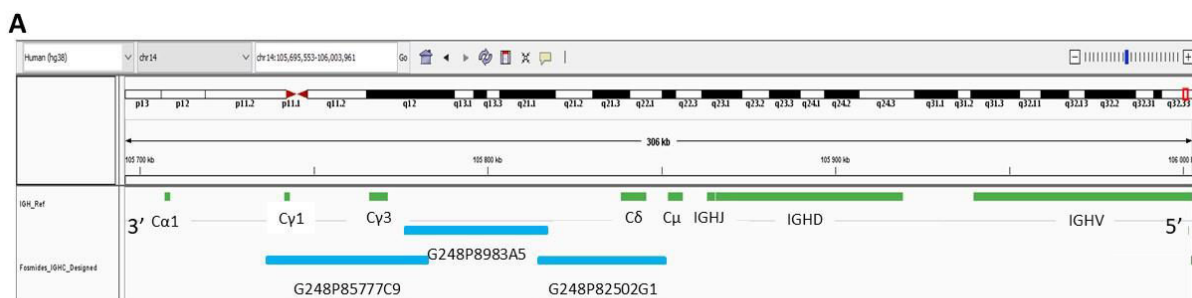


Figure 42 : A- Représentation schématique à partir du logiciel IGV du locus IGH (en vert) et des fosmides ciblant les régions constantes d'IGH (en bleu). B- Image de FISH métaphasique avec ces fosmides marqués en Aqua.

On observe l'absence de délocalisation des signaux bleus. La clonalité de cette mitose a été mise en évidence avec un BAC ciblant les régions IGHV (RP11-47P23), et les chromosomes 17 ont été marqués avec les sondes centromérique et télomérique pour repérage.

ii- Etude du dérivé 17

En parallèle, nous avons déployé une analyse en FISH ciblant le chromosome 17 afin de localiser le point de cassure sur ce chromosome. Nous avons sélectionné des BACs ciblant les régions 17q22, 17q23, 17q24 et 17q25. Aucune de ces sondes n'était délocalisée sur le chromosome 14. Nous avons ensuite utilisé une sonde télomérique du bras long du chromosome 17, montrant elle aussi l'absence de délocalisation du signal. Nous avons vérifié que ces résultats concernaient des mitoses clonales (48 chromosomes, c'est-à-dire porteuses des trisomies 12 et 19). De plus, une sonde centromérique du chromosome du 17 a été réalisée afin de confirmer la concordance télomère-centromère. Enfin, une sonde IGH a



également confirmé la présence du remaniement entre les chromosomes 14 et 17, avec le maintien de la région télomérique 17q sur le chromosome 17.

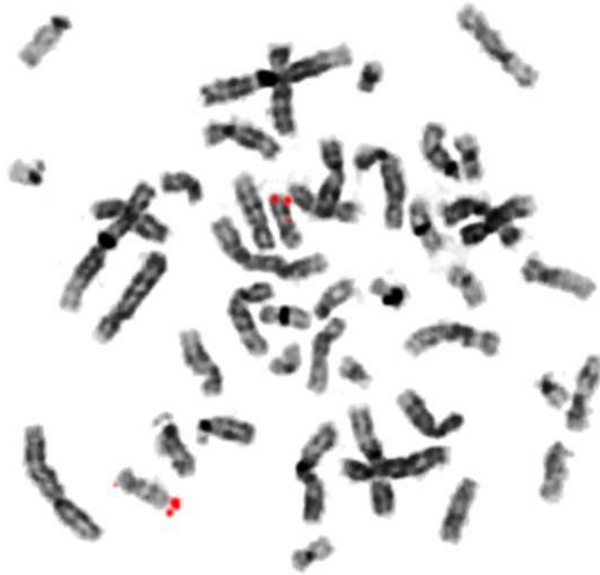


Figure 43 : Image de FISH métaphasique avec une sonde télomérique du bras long du chromosome 17. On observe l'absence de remaniement (deux signaux rouges sur deux chromosomes 17).

En conclusion, l'étude en FISH montre que remaniement n'est pas une translocation mais d'une insertion cryptique d'*IGH* au niveau du chromosome 17.

D'après la littérature, une insertion d'*IGH* est possible et pourrait impliquer le fragment d'excision issu du remaniement VDJ ou du switch des parties constantes^{48,63,611}. Dans cette hypothèse, selon la première analyse en FISH avec la sonde *IGH* fournie par Kreatech, un remaniement du fragment issu du switch entraînerait une séparation du signal rouge couvrant l'ensemble des régions constantes ainsi que les gènes de jonction et diversité. Etant donné que le signal observé sur le chromosome 17 est un signal vert, cela implique la délocalisation d'un fragment d'*IGHV*. Dans l'hypothèse d'une insertion issue de la recombinaison VDJ, nous savons que nous avons un réarrangement VDJ clonal de type *IGHV5-51, IGHD6-6* unique. La clonalité n'a pas mis en évidence d'autre remaniement VDJ. Dans ce cas, une grande partie de la partie variable des immunoglobulines est tronquée et sa délocalisation pourrait être visualisée. D'après le schéma de la sonde de FISH Kreatech, un tel remaniement entraînerait une cassure située à peu près au milieu de la sonde verte, avec un signal couvrant le fragment d'excision et un signal résiduel sur le chromosome 14 (Figure 32 et Figure 44). Malgré l'absence de signal résiduel vert visible sur le chromosome 14, cette hypothèse ne peut être



totalemment exclue étant donnée l'incertitude de la localisation exacte des bordures de la sonde verte. Il se peut également qu'un autre réarrangement VDJ non productif ait été initié avec un segment VH proche du télomère, entraînant un signal résiduel faible sur le chromosome 14, et donc invisible en FISH.

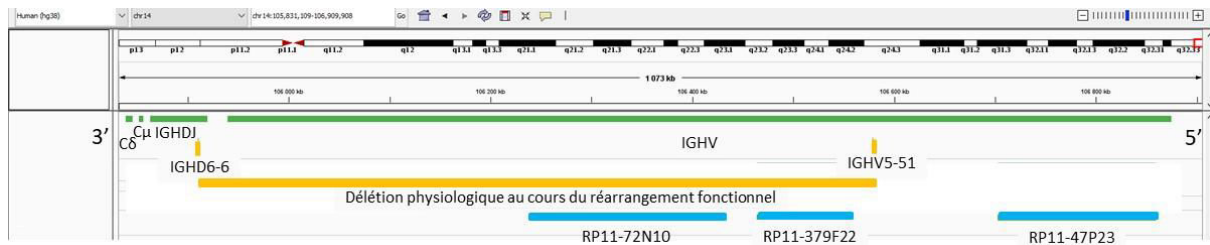


Figure 44 : Représentation schématique à partir du logiciel IGV du locus *IGH* (en vert) et du fragment d'excision selon le remaniement VDJ mis en évidence (en orange), ainsi que les 3 BACs utilisés pour vérifier la délocalisation du fragment d'excision (en bleu).

Dans l'hypothèse d'une insertion du fragment d'excision se produisant au cours du réarrangement VDJ, nous avons utilisé des BACs situés en amont et au niveau du fragment d'excision (Figure 44). Les résultats montrent que tous les BACs sont délocalisés sur le chromosome 17, ce qui infirme cette hypothèse (Figure 45). De plus, on observe une disparition du signal ciblant le fragment d'excision issu du réarrangement VDJ soit au niveau d'un chromosome 14 normal (en l'absence de translocation entre les chromosomes 13 et 14), soit au niveau du dérivé 13 issu de la t(13;14). Ceci indique que le réarrangement VDJ productif mis en évidence a lieu sur le chromosome 14 non impliqué dans le remaniement avec le chromosome 17.

Nous avons continué les explorations avec la réalisation de FISH au niveau de la région de cassure supposée. Pour cela, nous avons utilisé les BAC RP11-645E6, RP11-965B13 et RP11-14117. Cependant, les résultats obtenus étaient ininterprétables (réalisation sur lames déjà hybridées plusieurs fois avec déshybridation incomplète et dégradation du matériel génétique).

L'ensemble de ces résultats indique une région de cassure théorique située entre IGHDJ et IGHD6-6 ou IGHD6-6 et IGHD6-6.

La zone de cassure restante est alors d'environ 470 kb.



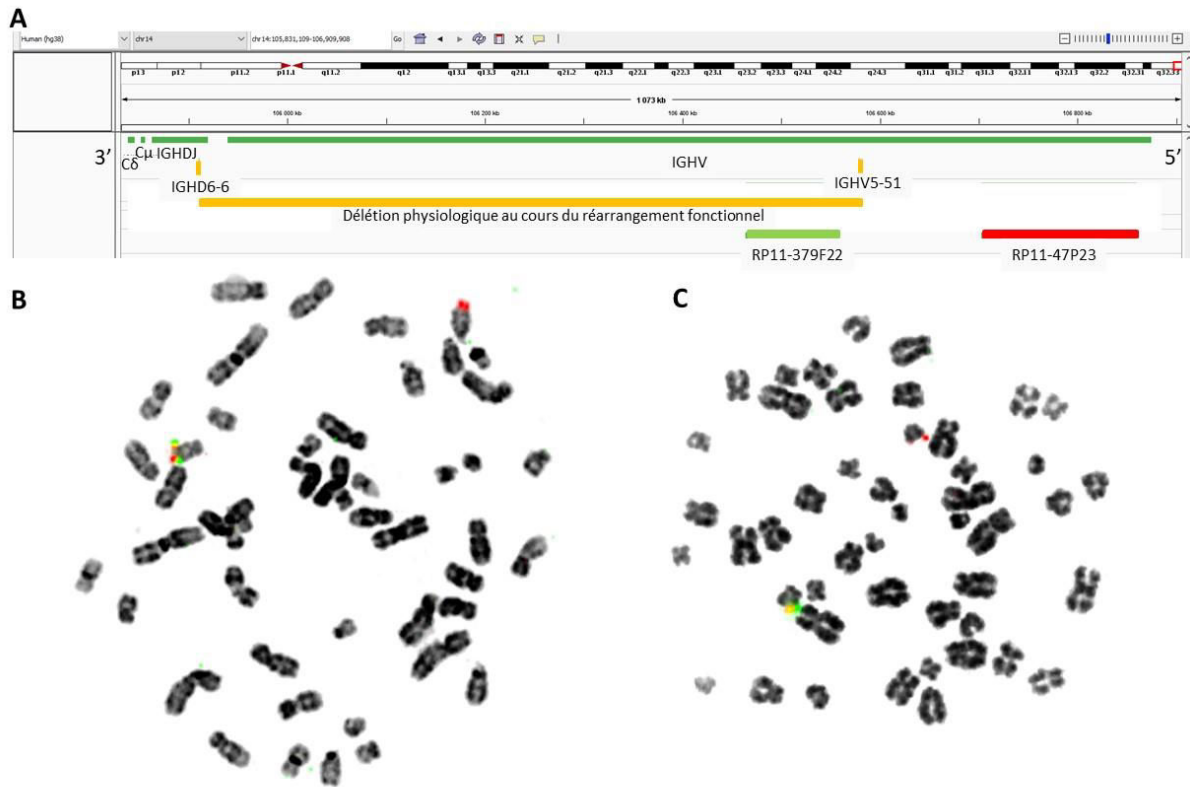


Figure 45 : Image de FISH métaphasique avec les BACs RP11379F22 et RP11-47P23 marqués respectivement en vert et rouge.

A- Représentation schématique à partir du logiciel IGV du locus IGH (en vert) et des BACs RP11-379F22 (vert) et RP11-47P23 (rouge).

B- Images de FISH métaphasique du clone sans t(13;14). On observe un signal rouge sur un chromosome 14 non remanié et une fusion sur le dérivé 17. Il manque un signal vert, ce qui peut être expliqué par la perte lors du réarrangement physiologique IGHV. C : Image métaphasique du clone présentant la t(13 ;14). On observe un signal rouge sur le dérivé 13 issu de la t(13;14) et une fusion sur le dérivé 17. Il manque un signal vert pouvant être expliqué par la perte lors du réarrangement physiologique IGHV avec délocalisation d'IGH sur le chromosome 13.

d- Etude en Multi-FISH

Nous avons également réalisé une analyse en multi-FISH pour ce patient qui ne permet pas de mettre en évidence le remaniement entre les chromosomes 14 et 17 (Figure 46 et Figure 47). En effet, l'idéogramme du dérivé 14 correspondant ne permet pas de déceler la présence de matériel issu du chromosome 17. Concernant le dérivé 17, on note un décrochage des couleurs en 17q sur une mitose, mais non reproductible sur d'autres mitoses et probablement artéfactuel. La translocation entre les chromosomes 13 et 14 est visible avec une partie du chromosome 13 clairement délocalisée sur le 14. A l'inverse, la portion de chromosome 14



délocalisée sur le chromosome 13 n'est pas visible, probablement à cause de la taille du fragment et du marquage faible des extrémités chromosomiques. La résolution de cette technique est donc insuffisante pour mettre en évidence le remaniement d'intérêt.

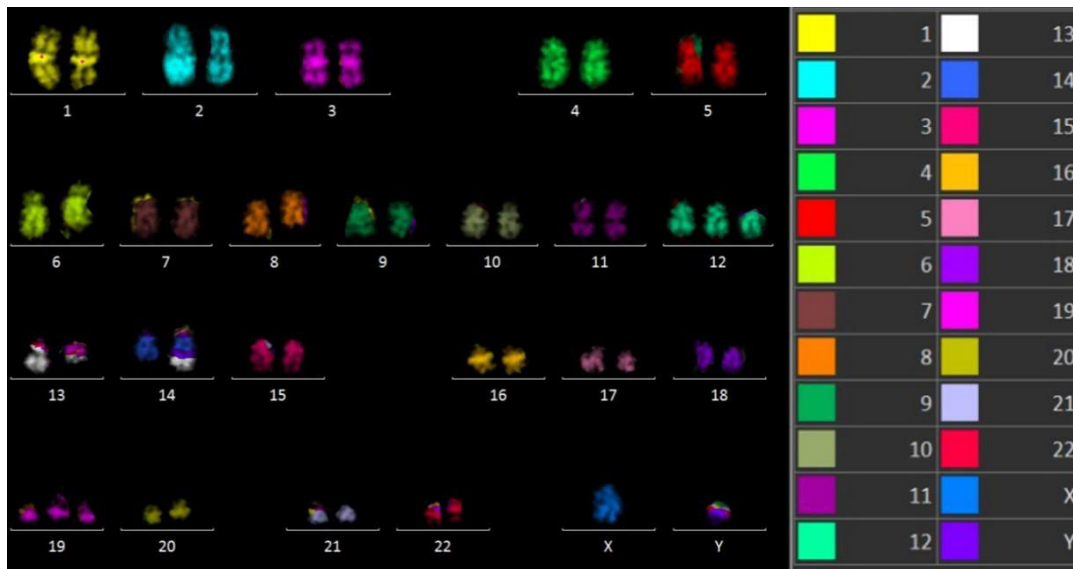


Figure 46 : Image de M-FISH réalisée chez le patient n°1. On visualise la translocation entre les bras longs des chromosomes 13 et 14. En revanche, le remaniement entre les chromosomes 14 et 17 est invisible. La mitose observée est porteuse des trisomies 12 et 19.

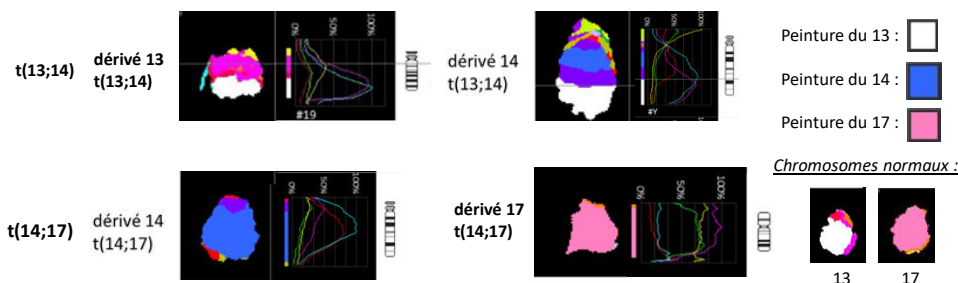


Figure 47 : Idéogrammes des chromosomes 13, 14 et 17 (multiFISH). Le dérivé 13 issu de la t(13;14) est identifiable du fait de sa taille réduite, mais la portion de chromosome 14 délocalisée n'est pas visible car sa taille est trop petite. En revanche, le dérivé 14 issu de la t(13;14) est bien visualisé. La t(14;17) n'est pas visible. On note un décrochage en 17q terminal des couleurs bleu et rouge (associées au chromosome 14) et une chute des couleurs jaune, vert et violet (associées au chromosome 17) sur le deuxième chromosome 17. Ces signaux restent trop faibles pour être interprétés et ne sont pas répétés sur les autres mitoses clonales analysées.



e- Etude en CGH-array

Enfin, une approche en CGH-array a été réalisée. En effet, certaines translocations apparemment équilibrées au caryotype entraînent néanmoins une perte de matériel visible en CGH-array⁶¹².

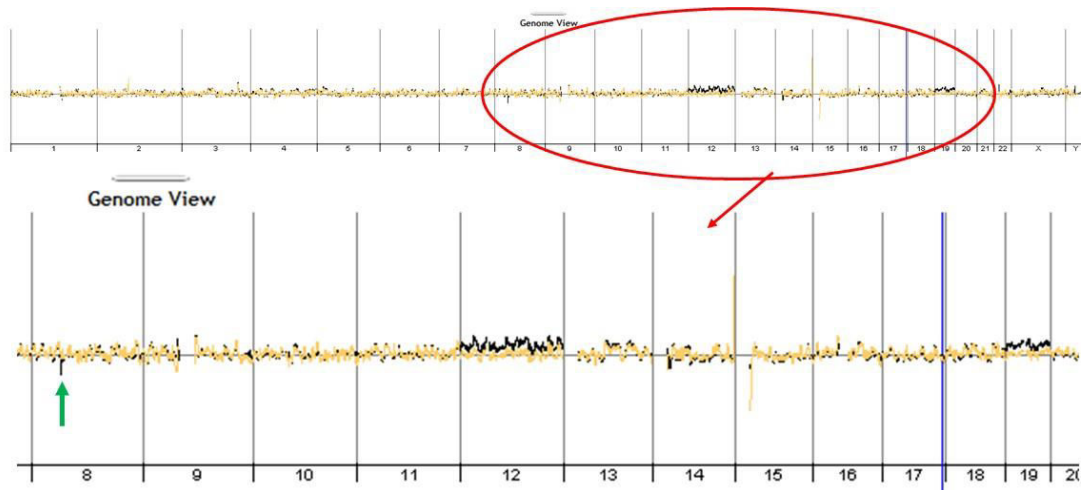


Figure 48 : Profil en CGH-array du patient n°1.

Le remaniement entre les chromosomes 14 et 17 n'est pas visible. On visualise faiblement en noir les trisomies 12 et 19. La flèche verte pointe la délétion 8p11.

Le résultat de la CGH-array permet de visualiser faiblement les trisomies 12 et 19 (Figure 48). Le fait qu'elles soient faiblement visibles est expliqué par la faible infiltration du prélèvement utilisé (ADN issu du sang). De plus, il s'agit d'une anomalie sous-clonale. Elles demeurent néanmoins visibles. En revanche la translocation entre les chromosomes 14 et 17 n'est pas visualisable alors qu'elle est autant représentée que les trisomies 12 et 19. En particulier, cette technique n'a pas permis de mettre en évidence de perte de matériel au niveau du chromosome 17. Le locus au niveau du chromosome 14 n'est pas interprétable en CGH en raison de l'utilisation d'un témoin ADN présentant un réarrangement VDJ physiologique de taille importante. La translocation entre les chromosomes 13 et 14 n'est pas visualisée non plus. Soit elle est sous clonale, soit elle est équilibrée et donc non visualisable. L'analyse d'un échantillon plus infiltré tel que le ganglion serait souhaitable. En revanche, on détecte une délétion 8p11.



f- Analyse du séquençage du génome entier

L'ensemble des différentes techniques réalisées étant peu informatives devant la petite taille de ce remaniement déséquilibré et face à l'épuisement du matériel pour la réalisation de FISH, nous avons décidé de réaliser un séquençage du génome entier à partir d'ADN de ganglion. Cette approche a permis de mettre en évidence le remaniement. Les points de cassures identifiés, annotés sur le génome de référence GRCh38, sont :

- chromosome 14, position 105864259 – chromosome 17, position 82971132 ;
- chromosome 14, position 105894509 – chromosome 17, position 82971126.

Le premier point de cassure identifié sur le chromosome 14 est situé au niveau d'IGHJ4. Ce point de cassure peut être visualisé à l'aide du logiciel IGV à partir des fichiers bam (Figure 49).



Figure 49 : Visualisation à partir du logiciel IGV des séquences obtenues par séquençage du génome entier au niveau du chromosome 14, position 105864259.

Les alignements colorés en gris correspondent aux séquences germinales. Ceux en violet représentent les séquences dont le début est aligné sur le chromosome 14 et la fin est localisée sur le chromosome 17. Le locus IGHJ4 est représenté en vert. Le premier point de cassure au niveau d'IGH est indiqué avec un trait pointillé rouge et est situé au niveau d'IGHJ4. Les deux cadres violets au niveau de la séquence indiquent l'heptamère (CACATTG) et le nonamère (ACAAAACC) du RSS jouxtant IGHJ4 et le réarrangement (voir Courte discussion et conclusion, page 173).

Le deuxième point de cassure identifié sur le chromosome 14 est situé au niveau d'IGHD4-17. Ce point de cassure a également été visualisé sur IGV (Figure 50).



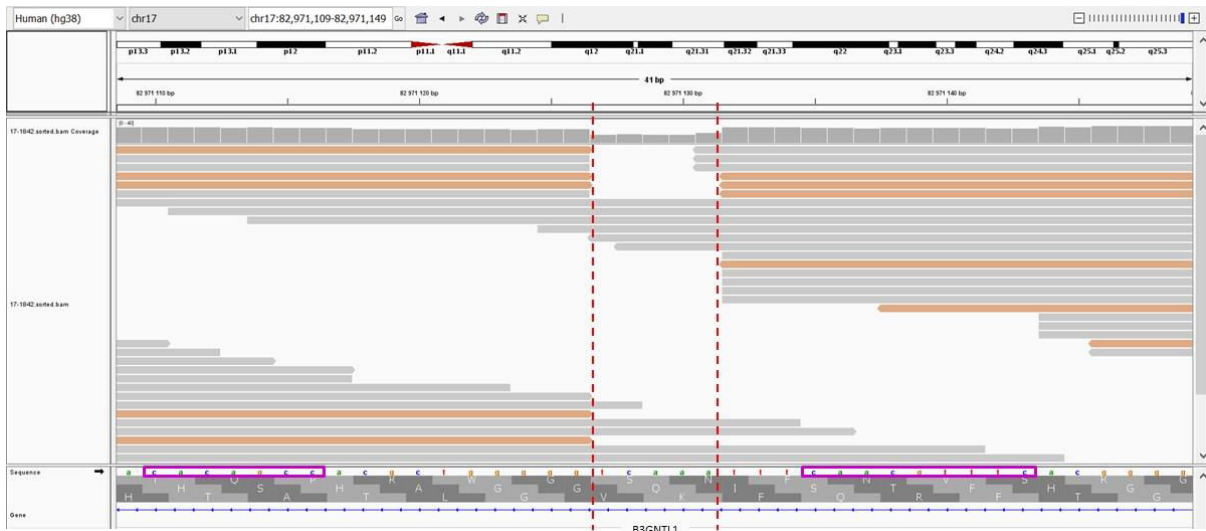


Figure 51 : Visualisation à partir du logiciel IGV des séquences obtenues par séquençage du génome entier au niveau du chromosome 17, centrée sur les positions 82971126 à 82971132. Les alignements colorés en gris correspondent aux séquences germinales. Ceux en orange représentent les séquences dont le début est aligné sur le chromosome 17 et la fin est localisée sur le chromosome 14. Les deux points de cassure au niveau du chromosome 17 sont indiqués par les traits pointillés rouges. Ils sont espacés de 6 pb et touchent le gène *B3GNTL1*. Les deux cadres violets au niveau de la séquence indiquent les séquences mimant un heptamère (CACAGCC) et un nonamère (CAACGTC) d'un RSS cryptique détecté par le site RSS database^{613,614} (voir Courte discussion et conclusion, page 173).

A partir de ces résultats, nous avons récupéré les séquences concernées. Ainsi, la séquence issue du premier point de cassure identifié, impliquant IGJH4 et la position 82971132 sur le chromosome 17 est :

ATGCAGCAAACCCTTCAGAGTTAAAGCAGGAGAGAGGTTGTGAGGACTCACCTGAGGAGACGGTG
 ACCAGGGTCCCTGGCCCCAGTAGTCAAAGTTTACGATCTTCAACGTTTCACGGGGACAAACCACA
 TCCAAACCACGGCAATTATATCCAGGCCGGTCTGTGCTCACGGGACGTTTCTATCGGACAG

Avec en rouge la portion issue du chromosome 14, en noir l'insertion de 9 nucléotides et en bleu la portion issue du chromosome 17.

La séquence au niveau du deuxième point de cassure identifié impliquant IGHD4-17 et la position 82971126 sur le chromosome 17 est :

CTGCGAGGGCTCACTCAGCGCTGTGGGAGGCATTATCTGTGCACAAGGGGTCCCCTCGAGACCCCA
 GCACCTCCCACACAGCCACGCTGGGGGCCCCCTCGGGGTAGTACCGTAGTCACACAGTAGGAG
 GGGCCTTCAAAAAAGCCCCTGAGTGTGCCAGCGCGCTTCCACGCGGGCCCTGGTCCAGG

Avec en bleu la portion issue du chromosome 17, en noir l'insertion de 13 nucléotides et en rouge la portion issue du chromosome 14.



A partir de ces résultats, nous avons mis au point des amorces pour vérifier ces points de cassure par PCR standard suivie d'un séquençage de type Sanger. Ceux-ci ont été validés (Figure 52 et Figure 53).

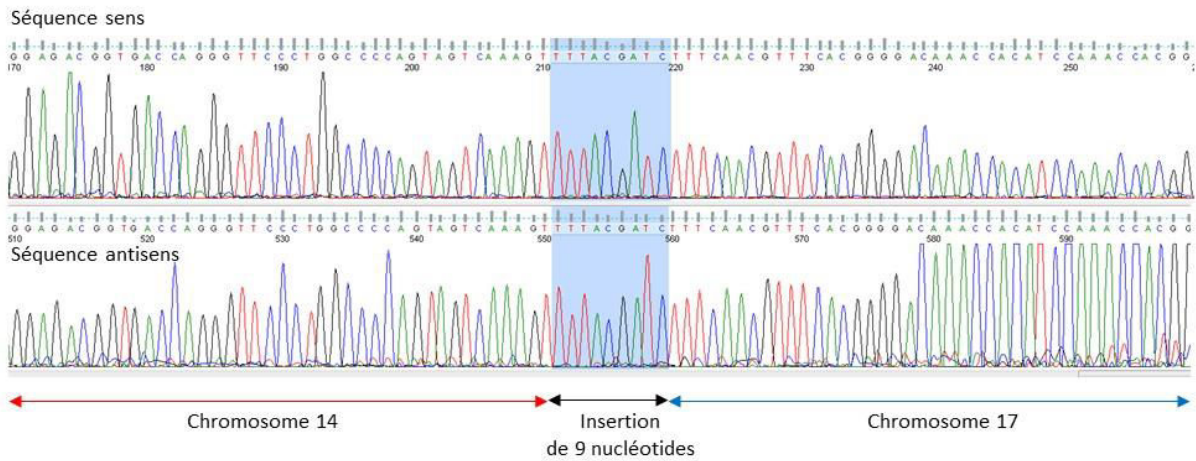


Figure 52 : Résultat du séquençage par technique Sanger du premier point de cassure impliquant IGJ4 et la position 82971132 sur le chromosome 17.

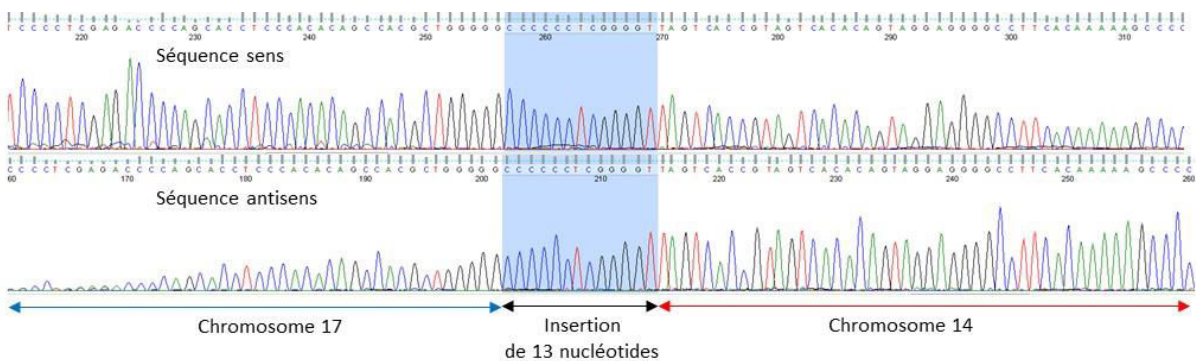


Figure 53 : Résultat du séquençage par Sanger du deuxième point de cassure impliquant IGHD4-17 et la position 82971126 sur le chromosome 17.

Les résultats obtenus en Sanger sont cohérents avec ceux du séquençage du génome entier et confirment l'insertion d'une portion d'*IGH* dans le chromosome 17, avec au niveau des jonctions l'insertion de 9 et 13 nucléotides, associée à une délétion de 6 paires de bases au niveau du chromosome 17. Ceci est compatible avec la CGH-array qui n'identifie pas de délétion au niveau du chromosome 17, celle-ci étant trop petite pour être visualisée par cette technique.



Cependant, le remaniement n'entraînerait pas l'insertion simple du fragment correspondant à la région située entre IGHJ4 et IGHD4-17. En effet les séquences obtenues correspondent à la région allant d'IGHJ4 vers IGHJ6 d'une part (et non IGHJ1), et d'IGHD4-17 vers IGHD3-16 (et non IGHD5-18) (Figure 54).

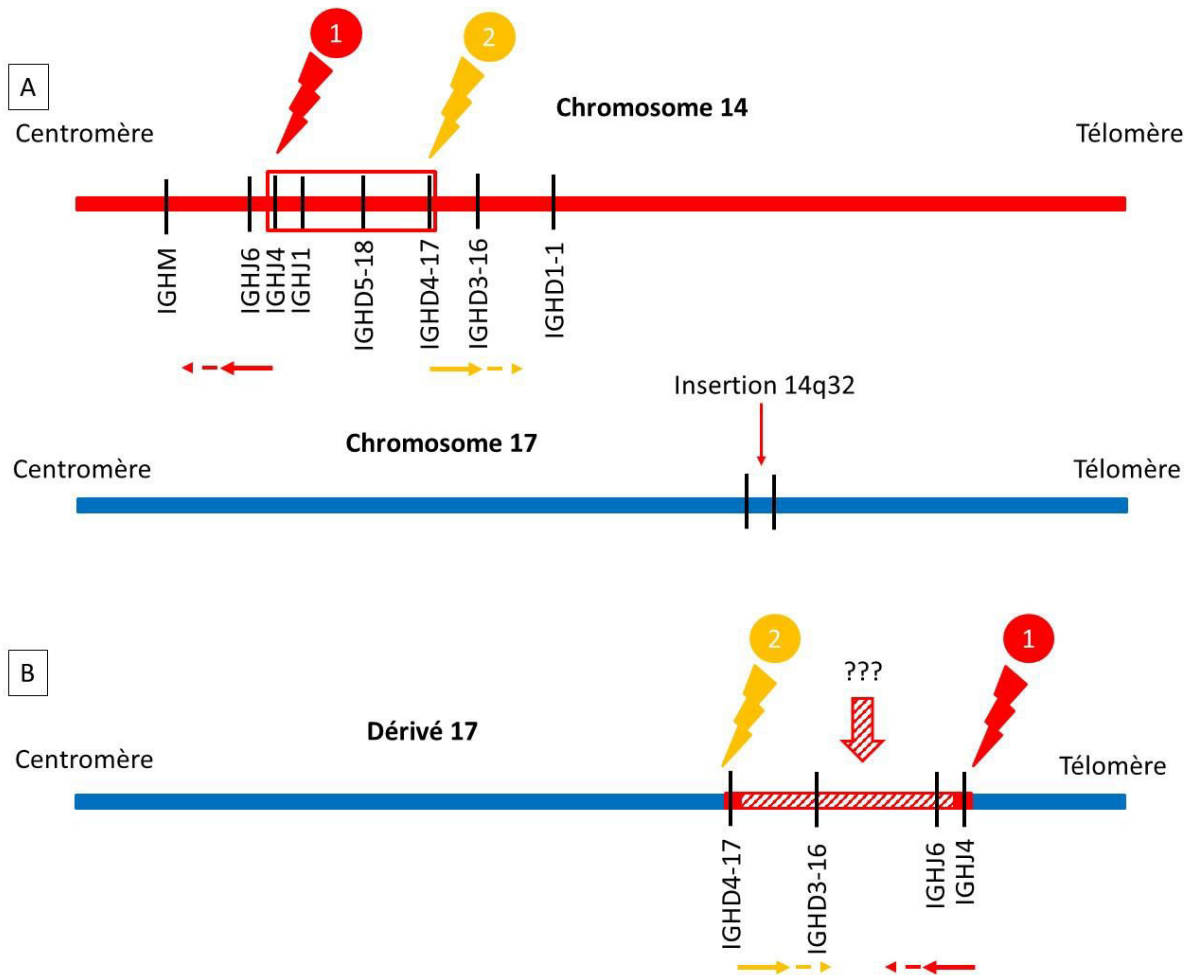


Figure 54 : Schémas représentant le réarrangement identifié chez le patient 1.

En A : la représentation des points de cassure sur les chromosomes 14 et 17 ; en B : le dérivé 17 obtenu avec les jonctions retrouvées par séquençage (flèches rouge et orange pointillées). La flèche rouge hachurée indique la séquence insérée de taille et séquence inconnue.

Ce résultat suggère donc un mécanisme plus complexe d'insertion. Ceci est compatible avec les résultats obtenus avec la FISH IGH (sonde Kreatech, Figure 32) ciblant la région située entre IGHJ4 et IGHD4-17 avec le signal rouge, et dont le résultat indiquait la délocalisation d'un fragment comprenant la région IGHV (signal vert). Ceci est également conforté par la FISH réalisée avec les fosmidés ciblant IGHV (Figure 45).



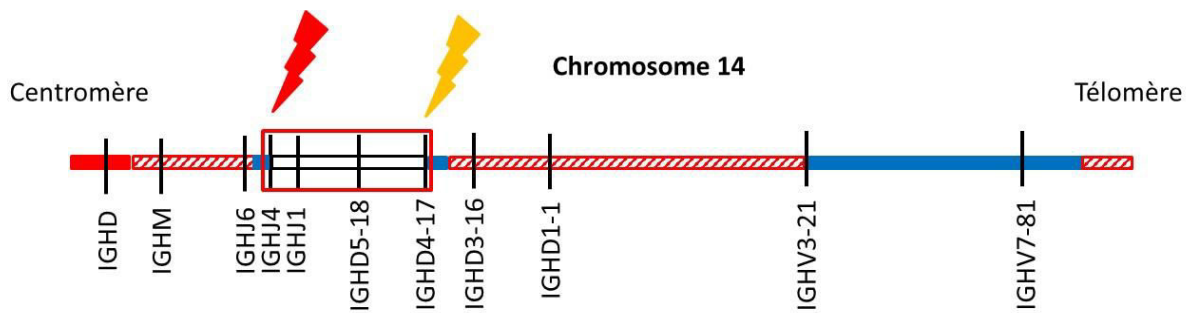


Figure 55 : Schéma récapitulatif des différents résultats obtenus concernant le remaniement entre les chromosomes 14 et 17 chez le patient n°1.

Le schéma représente un chromosome 14, avec en rouge, la partie restant sur le dérivé 14 (confirmation par FISH) ; en rouge hachuré, les régions incertaines sur leur localisation ; en bleu, les régions délocalisées sur le dérivé 17 (résultats obtenus par FISH et séquençage du génome entier) ; en blanc la région délétée au cours du remaniement.

L'ensemble des résultats obtenus indique donc une insertion d'IGH avec des points de cassure impliquant IGJ4 et IGHD4-17. Le fragment inséré comprend une grande partie d'IGHV. Le segment entre IGJ4 et IGHD4-17 semble délété au cours du réarrangement.

g- Courte discussion et conclusion

Nous avons utilisé des amorces ciblant IGJ6 en LM-PCR sans mettre en évidence le remaniement. L'étude du réarrangement montre que seules les enzymes SmaI et StuI coupent l'ADN sur le chromosome 17 après le remaniement et auraient généré des fragments d'environ 3000 et 3400 pb respectivement. Ces enzymes n'avaient pas été testées avec ces amorces dans un premier temps. Ceci a donc été réalisé. Les résultats obtenus ne permettent pas de mettre en évidence le remaniement. L'amplification de fragments de taille située entre 3000 et 4000 pb avait été réalisée dans d'autres tests de LM-PCR et est possible avec cette technique. Il semblerait donc y avoir un remaniement entre IGJ4 et IGJ6. L'utilisation d'amorces situées au niveau d'IGJ4 avec les différentes enzymes de restriction pourrait permettre de tester cette hypothèse. Cependant, avec le problème persistant d'épuisement du matériel biologique, nous allons dans un premier temps privilégier l'analyse plus approfondie des résultats de NGS.

Les résultats obtenus par séquençage du génome entier mettent en évidence deux points de cassure qui pris isolément pourraient suggérer une translocation d'IGH. Les résultats obtenus par FISH permettent d'identifier une insertion complexe.

Etant donnée la localisation des points de cassure identifiés au niveau d'IGH, une implication du complexe de recombinaison VDJ et en particulier des enzymes RAG dans la survenue de cette translocation est fortement probable. Nous avons analysé les séquences en IGJ4 et



IGHD4-17 : elles mettent en évidence des points de cassure au niveau des RSS (RSS avec espaceur 12 au niveau d'IGHD4-17 et espaceur 23 au niveau d'IGHJ4) (Figure 49). Ceci indiquerait un mécanisme faisant intervenir les enzymes RAG et le mécanisme de recombinaison VDJ dans l'apparition de ce remaniement. De plus, nous retrouvons au niveau des points de cassure l'insertion de nucléotides au hasard, ressemblant à des N-nucléotides.

L'origine du point de cassure au niveau du chromosome 17 est incertaine. Nous avons recherché la présence de RSS cryptique au niveau du chromosome 17 grâce au site RSS database⁶¹³. Celui-ci détecte une séquence pouvant correspondre à un RSS cryptique au niveau du point de cassure. Cependant, le point de cassure se situe à l'intérieur de cette séquence et non à l'extrémité ou à proximité de l'élément mimant l'heptamère (Figure 51). Les enzymes RAG peuvent se lier à l'ADN également via les histones ou encore de façon non spécifique. Leur implication pourrait être étayée par analyse ChIP-Seq (pour Chromatin ImmunoPrecipitation Sequencing) afin de mettre en évidence une fixation des enzymes RAG à ce niveau. Un autre mécanisme responsable d'une cassure double brin au niveau du chromosome 17 est possible. L'implication d'AID pourrait également être mise en cause. Cependant la séquence au niveau du chromosome 17 ne présente pas de motifs type CpG ou WGCW.

L'idéal pour affiner les régions délocalisées serait une analyse par FISH avec marche chromosomique. Cependant, le matériel biologique et en particulier les mitoses analysables disponibles sont à présent faibles. Nous allons réaligner les résultats du séquençage du génome entier sans référence (alignement *de novo*). Ceci risque d'être difficile étant donnée la séquence du gène *IGH* riche en GC et en séquences répétées. De plus, nous allons étudier plus précisément les différents résultats produits par l'analyse MANTA afin de déceler des remaniements au sein du chromosome 14. Une meilleure connaissance du fragment inséré nous permettra d'avancer dans l'interprétation du mécanisme de survenue du réarrangement.

Une hypothèse est qu'il s'agit d'un processus en plusieurs étapes avec une première translocation entre les chromosomes 14 et 17 au cours du réarrangement D_HJ_H avec les points de cassure identifiés par séquençage du génome entier puis un deuxième événement avec nouvelles cassures en 3' d'IGHJ4 sur le dérivé 14 et en aval d'IGHV sur le dérivé 17 (Figure 56).



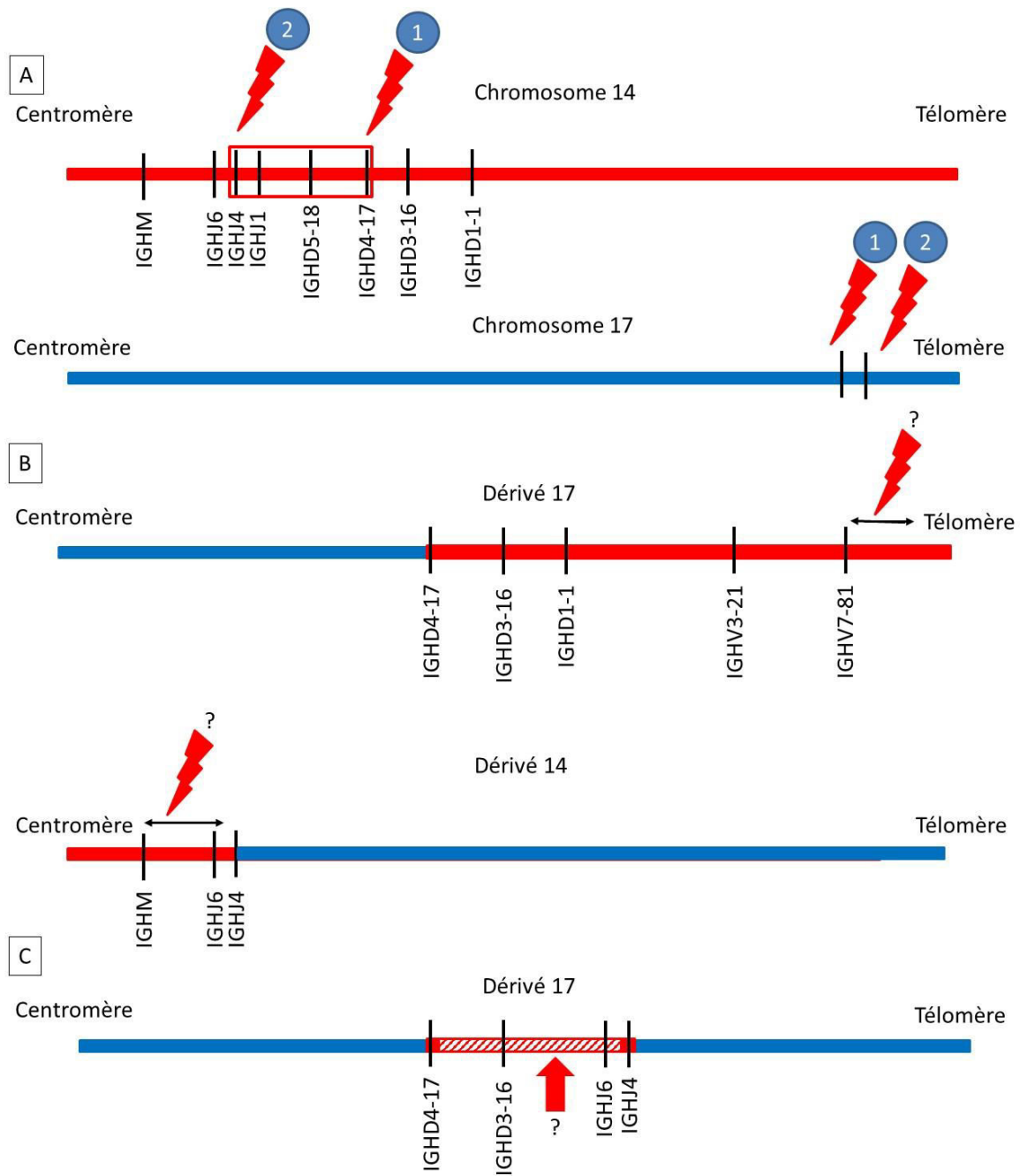


Figure 56 : Schéma de l'hypothèse en deux étapes pour le remaniement observé chez le patient n°1.
A- Points de cassure identifiés par séquençage du génome entier (éclairs rouges) au niveau des chromosomes 14 et 17.

B- Schéma de la translocation générée au cours d'une première étape selon les points de cassure identifiés par séquençage du génome entier. Les éclairs rouges indiquent les probables régions touchées au cours d'une deuxième étape par de nouvelles cassures double brin pouvant aboutir au remaniement complexe sur la base des données de FISH.

C- Schéma de l'insertion complexe obtenue au niveau du dérivé 17. La région hachurée en rouge correspond à la portion de chromosome 14 insérée de taille et séquence inconnues.



Une autre hypothèse est la génération de cassures sur le chromosome 14 avec circularisation du fragment généré, puis réouverture lors du réarrangement D_{HJH} et insertion dans le chromosome 17.

Un phénomène de chromothripsis peut également être en cause. Il s'agit d'un phénomène limité dans le temps, au cours duquel de nombreuses cassures (de 10 à des milliers) apparaissent de façon localisée sur un chromosome ou un bras de chromosome. Cette hypothèse semble moins plausible étant donné les points de cassures observés au niveau de RSS sur le chromosome 14, et non au hasard.

Par ailleurs, nous ne connaissons pas la région d'IGH exacte insérée dans le chromosome 17. En particulier, nous ne savons pas si l'enhancer E_{μ} est emporté, pouvant alors entraîner une surexpression du gène *B3GNTL1*. En l'absence de délocalisation de l'enhancer E_{μ} , on peut imaginer une disruption du gène entraînant une perte d'expression. Il existe des promoteurs en amont et en aval du point de cassure : la translocation, si elle implique un enhancer d'*IGH* (en particulier E_{μ}), est susceptible d'entraîner une dérégulation de l'expression du gène. La vérification de l'impact sur l'expression du gène sera faite par séquençage haut débit des ARN (RNA-Seq). Ceci nous permettra d'étudier une éventuelle dérégulation de ce gène et des autres situés à proximité.

Le gène *B3GNTL1* code une protéine impliquée dans la glycosylation et exprimée à différents niveaux de façon ubiquitaire. Des translocations impliquant ce gène ont été rapportées dans des cancers solides avec des partenaires situés sur les chromosomes 1, 3, 17 et X^{615,616}. Ces données sont en faveur de l'hypothèse d'un rôle oncogénique de la translocation retrouvée chez ce patient.



2- Patient 2

a- Résultats préliminaires

La deuxième patiente a bénéficié d'un bilan complet en 2011. Le caryotype alors réalisé montre une translocation entre les bras longs des chromosomes 17 et 22 (Figure 57). Il s'agit de la seule anomalie mise en évidence au caryotype.

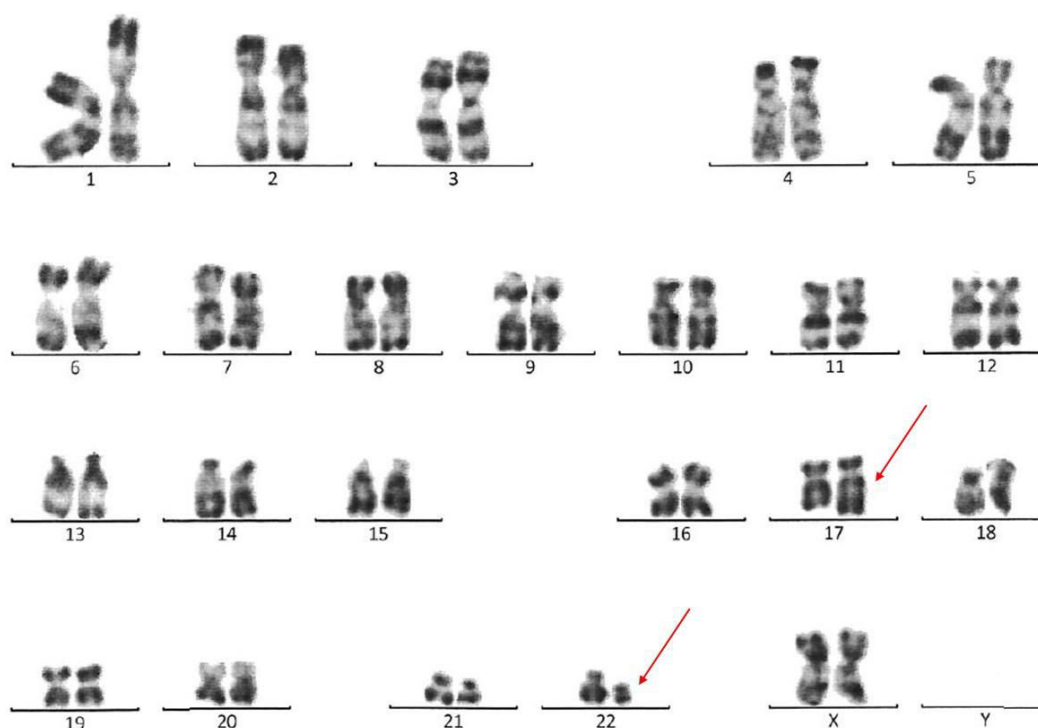


Figure 57 : Caryotype de la patiente n°2
Celui-ci met en évidence une translocation entre les bras longs des chromosomes 17 et 22.

La formule chromosomique était alors la suivante : 46,XX,t(17;22)(q25;q11)[19]/46,XX[1].

Le caryotype réalisé en 2016 montre la persistance de cette anomalie isolée. La formule chromosomique rendue était : 46,XX,t(17;22)(q25;q11)[15].

Devant la présence de l'implication du locus 22q11, une FISH pour IGL a été réalisée et montre un remaniement du gène (Figure 58). En particulier, on observe une délocalisation du signal rouge ciblant la région IGLC sur le chromosome 17.



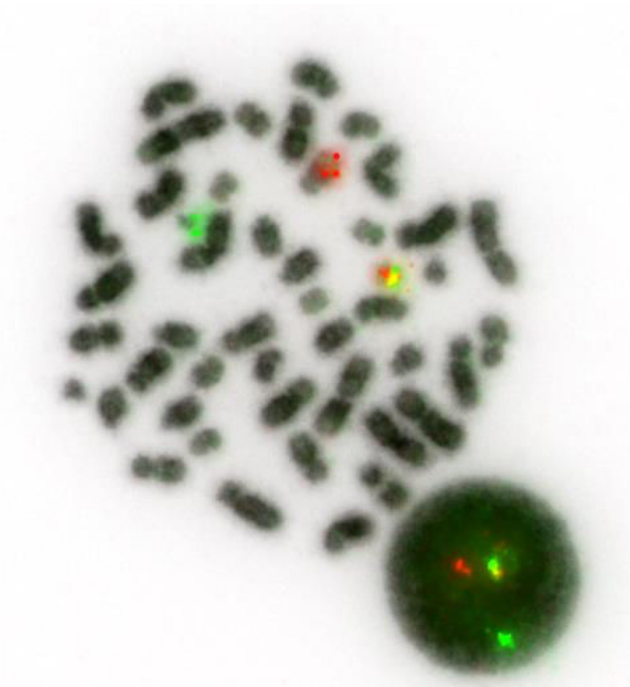


Figure 58 : Image de FISH avec la sonde IGL Dako concernant la patiente numéro 2. On observe une mitose avec un signal de fusion sur le chromosome 22 non remanié, un signal vert sur le dérivé 22 et un signal rouge sur le dérivé 17. On observe également un noyau dans lequel un remaniement d'IGL est mis en évidence (séparation du signal).

b- Résultats moléculaires

Les points de cassure rapportés dans les translocations impliquant *IGL* sont situés au niveau des régions de jonction et constantes^{617,618}. Nous avons mis au point des amorces pour réalisation de LM-PCR au niveau des régions IGLJ-C 1 à 5. Les différentes expériences menées n'ont pas permis de mettre en évidence le remaniement.

Un séquençage du génome entier a été réalisé et a permis de mettre en évidence les points de cassure suivants :

- chromosome 17, position 83175424 (gène AC139099.4) (Figure 59)
- chromosome 22, position 22850521 (au niveau d'IGLV3-6) (Figure 60)



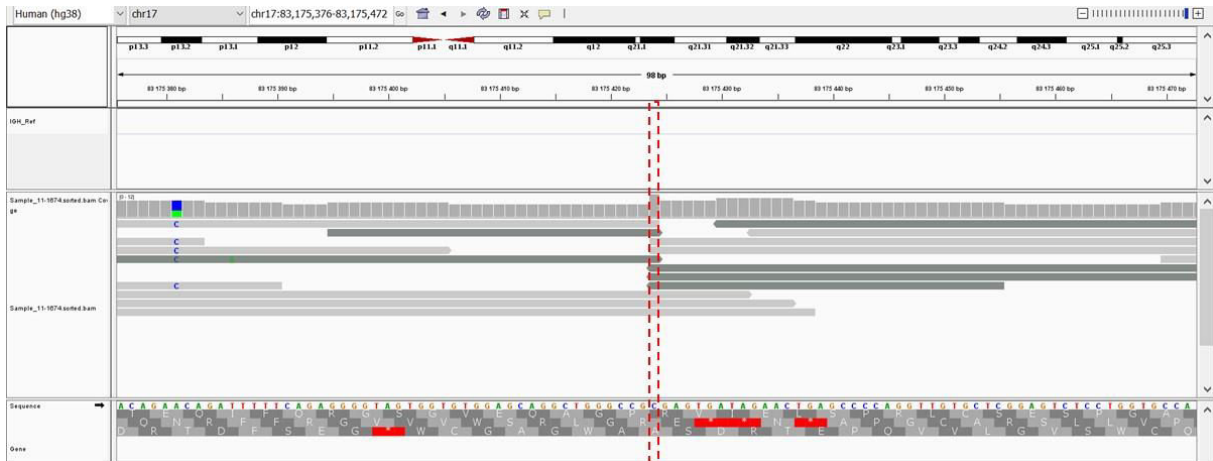


Figure 59 : Visualisation à partir du logiciel IGV des séquences obtenues par séquençage du génome entier au niveau du chromosome 17, position 83175424 concernant la deuxième patiente. Les alignements colorés en gris clair correspondent aux séquences germinales. Ceux en gris foncés correspondent aux séquences dont le début de la séquence est aligné sur le chromosome 17 et la fin est localisée sur le chromosome 22. Le point de cassure est entouré par un cadre pointillé rouge.



Figure 60 : Visualisation à partir du logiciel IGV des séquences obtenues par séquençage du génome entier au niveau du chromosome 22, position 22850521 concernant la deuxième patiente. Les alignements colorés en gris correspondent aux séquences germinales. Ceux en rouge indiquent une délétion. Ceux en bleu foncé correspondent aux séquences pour lesquelles le début est aligné sur le chromosome 22 et la fin est localisée sur le chromosome 17. Les points de cassure retrouvés d'après les séquences sont indiqués par les traits pointillés rouge. Les 9 nucléotides intermédiaires semblent délétés au cours du remaniement. A noter 4 bases homologues à ce niveau (encadrées en rose). L'étoile bleu indique une mutation G>A.

A partir de ces résultats, les séquences au niveau des points de cassure ont été récupérées. Ainsi, la séquence du dérivé 22 est la suivante :



TCATCATTTACAAGCCTCCTCTCTGTCCTCTCTTGCAAGCTCCTATGAGCTTACACAGCCACCCTCAGT
 GTCAGTGTACCAAGGACAGGCAGCCATGATCACCTGCTCTTGAGATAACCTCAAAGATGAGTATGTT
 TACTGGTTCTGGCAGAAGCCAGACCAGGCCATACTGGTGATATATGAAGGCAACAAGAGTGATAG
 AACTGAGCCCCAGGTTGTGCTCGGAGTCTCCTGGTGCCAGAACACTGGGCTTCCGTGGGACCTGAG
 GGGCTGAAACGGTGCTGGACCTGGACCCTTCTCACAGTACCCTCAGGTAGCCCATCTCAGACAGG
 AGGGGACCCCTTGTCCCCATGTTTGGAGCGTGCTGAGCTGCACACAGTCTCCTGGGTCTGAGAC
 CGCTGGTCTGCCCCACCCAGCTCTCTCCCTCTGG

avec en rouge la séquence issue du chromosome 22 et en bleu la séquence issue du chromosome 17. Le nucléotide A en gras souligné correspond à une mutation (G>A).

La séquence dérivée du chromosome 17 est la suivante :

CGAACGACTCGGCAGGCTCCAGAGTTTGGCTCATTTTCTGCTTGGTCCCTTTTTGATGCCTGTGATTTA
 TTTATAACAATAGTTTGCCTAATTCCAGATGTGGTGACACAGCCTGCAGACGCTGAGTGAATGAAG
 ACAGACCAGATTTTTCAGAGGGGTAGTGGTGTGGAGCAGGCTGGCCGCCAGGAATTTCTGATTTTC
 TGAGTCCAGCTCAGGGAACATGGCCA

avec en bleu la séquence issue du chromosome 17, et en rouge la séquence issue du chromosome 22. Les nucléotides soulignés et en gras correspondent à une homologie de séquence au niveau du point de cassure.

Ces résultats ont été confirmés par séquençage Sanger.

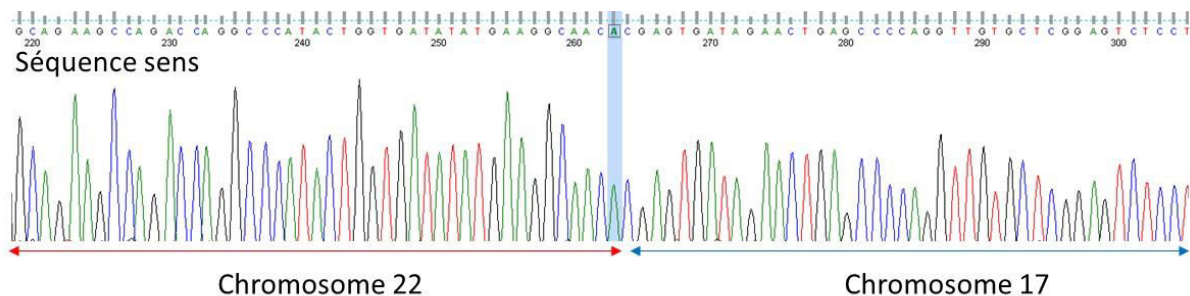


Figure 61 : Séquence sens obtenue par séquençage Sanger confirmant la jonction du der(22)t(22 ;17).

Concernant le dérivé 22, seule la séquence sens est lisible. La séquence antisens s'éteint après un poly-G. Une autre amorce située après cette région devrait permettre d'obtenir un résultat correct.

Le dérivé 17 n'a pas été mis en évidence. L'amorce utilisée au niveau du chromosome 22 semble trop éloignée du point de cassure. Nous allons mettre au point de nouvelles amorces plus proches afin de confirmer le dérivé 17.



Ce remaniement entraîne donc une délocalisation de l'enhancer 3'Eλ sur le dérivé 17 (Figure 62).

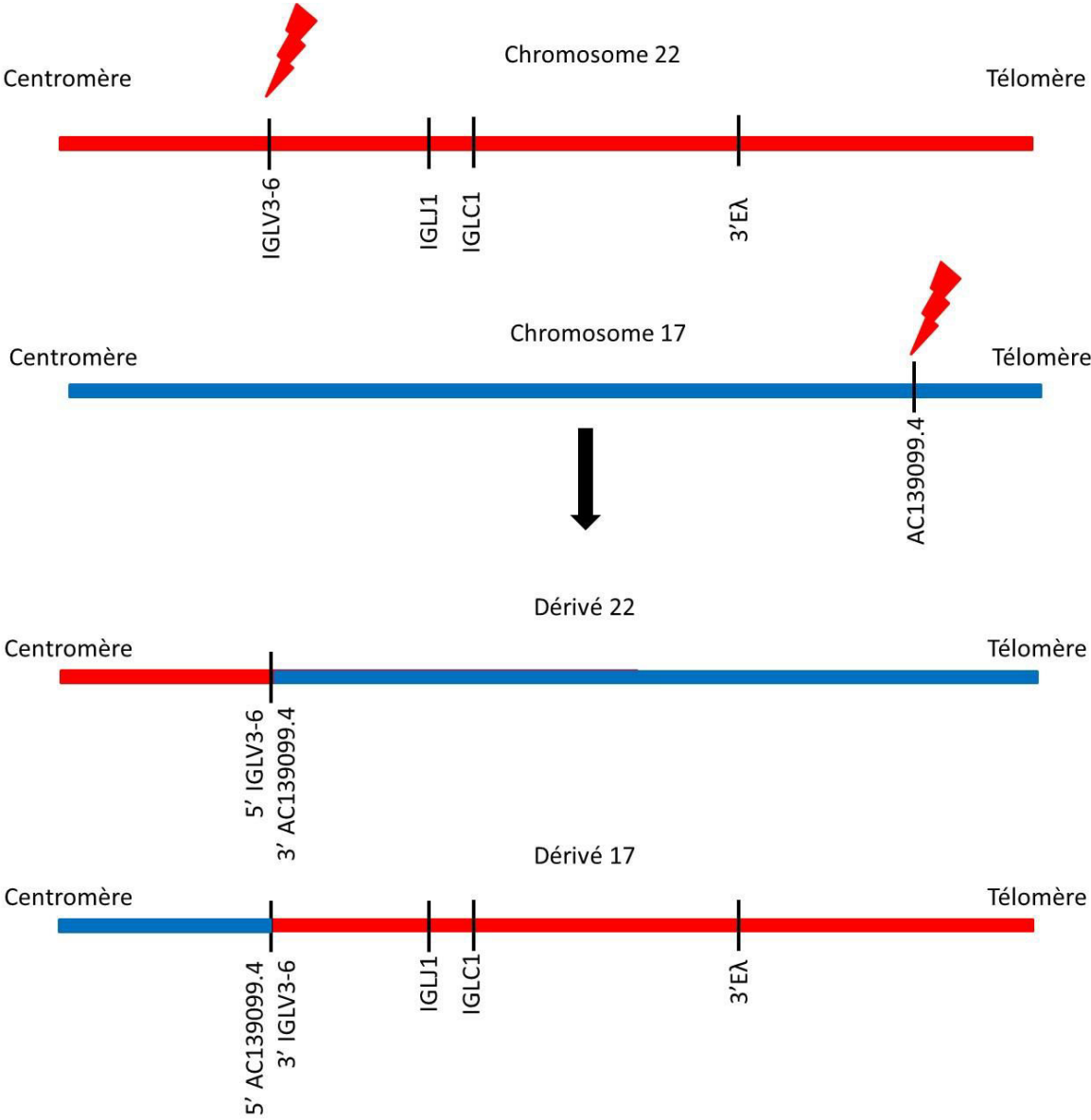


Figure 62 : Schéma représentant la translocation entre les bras longs des chromosomes 17 et 22 mise en évidence chez la patiente n°2.



c- Courte discussion

Le point de cassure au niveau du chromosome 22 est situé au niveau des régions variables à l'intérieur d'IGLV3-6. Il s'agit d'un gène non fonctionnel²³. Étant donnée la localisation au niveau des régions variables d'IGL, un mécanisme médié par RAG, au cours de la recombinaison VJ de la chaîne légère, peut être évoqué. Cependant, le point de cassure se situe à l'intérieur du gène IGLV3-6 et un peu à distance du RSS (133 pb).

En parallèle, la recherche de RSS cryptique au niveau du chromosome 17 a été réalisée à l'aide du site RSS Database et montre la présence de séquences pouvant correspondre à de potentiels RSS cryptiques à proximité du site de cassure. La première région retrouvée peut correspondre à un RSS avec espaceur 12 (CACAGCATGCAGACGCTGAGTCAAATGA), située 51 pb en amont du point de cassure, ou à un RSS avec espaceur 23 (CACAGCATGCAGACGCTGAGTCAAATGAAACAG) situé alors à 40 pb en amont du point de cassure. Son orientation ne semble pas compatible avec un réarrangement médié par RAG. Un autre RSS cryptique avec espaceur 23 est identifié 109 paires de bases en aval du point de cassure (CACAGTCACCCTCAGGTAGCCCATCTCAGACAGGAGGGG) avec une orientation compatible avec la translocation. Cependant, leur localisation semble trop éloignée du point de cassure pour être mise en cause. Certains auteurs ne les considèrent que si une séquence mimant un heptamère (CACNNNN) est présente dans les 5 nucléotides à proximité du point de cassure⁶¹⁹. D'autres ont pris en compte les 20 nucléotides adjacents⁶²⁰. Une dernière étude a pris un seuil à 21 paires de bases⁶²¹.

En revanche, le point de cassure au niveau du chromosome 17 est situé au niveau d'un motif CGCG. Ceci suggère un rôle d'AID au cours de l'hypermutation somatique. L'analyse de la séquence en amont du point de cassure dans la région IGL révèle la mutation d'un nucléotide dans 3 séquences sur 4, y compris dans des séquences non impliquées dans le remaniement entre les chromosomes 17 et 22. Il peut s'agir d'une mutation dans le cadre de l'hypermutation somatique.

Dans ce cas, un rôle d'AID dans le mécanisme de survenue de cette translocation au cours de l'hypermutation somatique est privilégié.

Concernant le gène partenaire, il s'agit d'*AC139099.4*. Le point de cassure est situé au niveau de l'intron 2 qui est le plus long, le gène étant constitué de 3 exons. Il est répertorié dans NCBI gene sous le nom LOC101929650. Il s'agit d'un lncRNA. Il est répertorié dans la base de données LNCipedia sous le nom lnc-METRNL-2:1 ou LNC-TSPAN10-103:2¹¹¹. Le transcrit fait 641 pb. Ce gène est situé 80 kb en aval du gène *METRNL* (pour Meteorin-like, glial cell differentiation regulator), qui est impliqué dans l'adipogénèse⁶²².

Le point de cassure au niveau du chromosome 17 est situé dans ce cas environ 200 kb en aval du point de cassure mis en évidence chez le patient n°1. Il se peut qu'une cible commune soit en cause. Nous prévoyons de réaliser du RNA-Seq pour étudier l'expression des gènes situés dans cette région, dont *B3GNTL1*, *AC139099.4* et *METRNL*.



3- Patient 3

a- Résultats préliminaires

Concernant le cas de la troisième patiente, la cytogénétique réalisée en 2012 puis en 2016 sur sang avec stimulation par IL2/DSP30 met en évidence un caryotype complexe comprenant notamment une translocation entre les chromosomes 8 et 14 (Figure 63). Les FISH réalisées montrent une implication d'*IGH* (Figure 64), sans réarrangement de *MYC*. Il met également en évidence une translocation entre les chromosomes 15 et 22 n'impliquant pas *IGL* (Figure 65) et une délétion du bras long du chromosome 11. C'est donc la t(8;14) qui nous intéresse car elle implique un gène des immunoglobulines et un partenaire inconnu en 8q24.

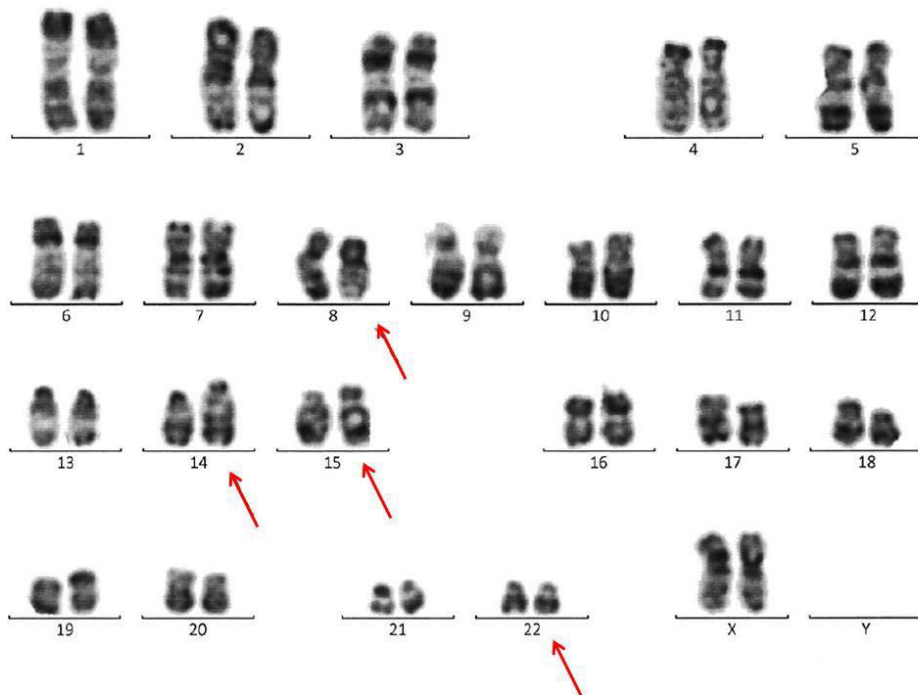


Figure 63 : Caryotype du patient n°3.

On observe la translocation entre les bras longs des chromosomes 8 et 14 d'une part et entre le bras court du chromosome 15 et le bras long du chromosome 22 d'autre part.



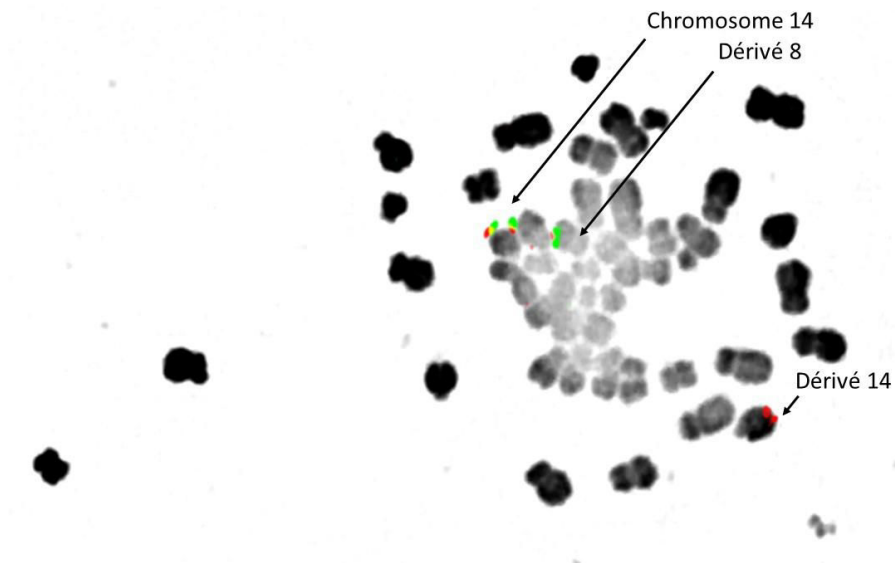


Figure 64 : Image de FISH métaphasique avec la sonde IGH Kreatech. On observe une fusion sur le chromosome 14 normal, une fusion avec un signal rouge semblant diminué sur le dérivé 8 et un signal rouge sur le dérivé 14.

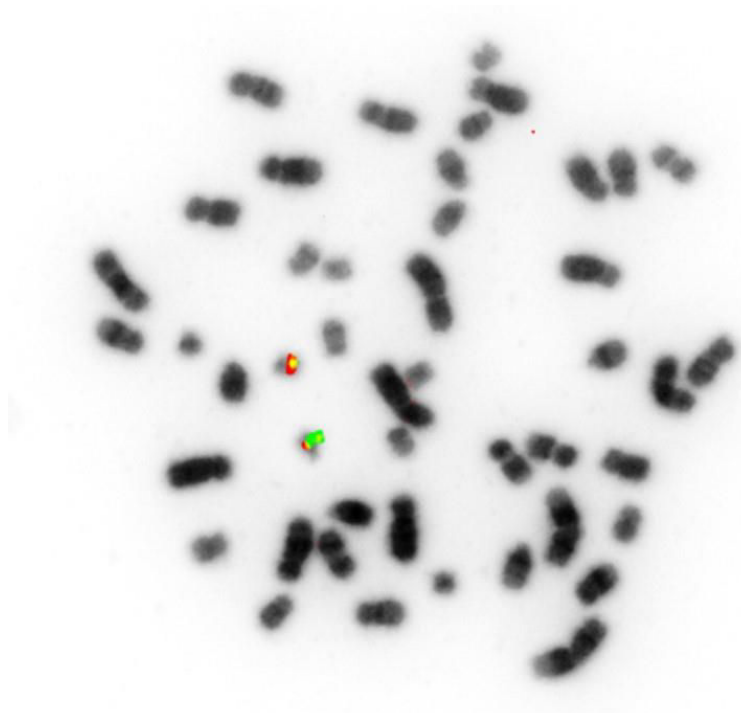


Figure 65 : Image de FISH métaphasique avec la sonde IGL Dako. On observe deux signaux de fusion situés sur les chromosomes 22.



La FISH du locus MYC avait été réalisée avec la sonde DAKO en 2012. Les images obtenues sur mitoses étaient de mauvaise qualité. Une seule mitose normale avait été vue. Les résultats de FISH interphasique montrent l'absence de séparation du signal, mais une amplification du locus avec 3 signaux de fusion est mise en évidence dans 13,5 % des noyaux analysés, compatible avec la présence d'un isochromosome 8q sous-clonal.

La formule cytogénétique rendue en 2012 est la suivante :

46,XX,del(11)(q14)[6]/46,XX,i(8)(q10),t(15;22)(p13;q12)[4]/46,XX,t(8;14)(q24;q32),t(15;22)(p13;q12)[9]/46,XX[1]

.ish t(8;14)(CMYC+,IGHdim;CMYC-,3'IGH+)[1]

.nuc ish(CMYCx3)[54/400],(IGHx2)(3'IGHsepIGHdimx1)[13/139],(D17Z1,P53)x2[150],(IGLx2)[238]

En 2016, le caryotype rendu était le suivant :

46,XX,t(8;14)(q24;q32),t(15;22)(p13;q12)[13]/46,sl,del(11)(q21q24)[7].ish 17p13(P53x2),17p11.1-q11.1(D17Z1x2)[20].nuc ish(P53,D17Z1)x2[206]

On observe en 2016 une augmentation de la proportion du clone comprenant la t(8;14) par rapport à 2012, probablement sélectionnée après traitement.

b- Résultats complémentaires

Le résultat de la FISH avec la sonde IGH de chez Kreatech montre une délocalisation du signal vert et d'une partie du signal rouge sur le chromosome 8. Ceci indique une région de cassure au niveau soit des régions constantes, soit des régions de jonction. Le plus probable, étant donnée la configuration de la sonde, est une cassure dans les régions constantes car le signal rouge en cas de cassure dans la zone de jonction semble assez faible pour pouvoir être visualisé. Dans cette hypothèse, nous avons réalisé de la FISH avec trois fosmidés ciblant les régions constantes les plus en amont du gène (côté télomérique) en aval d'IGHM : G248P85777C9, G248P8983A5 et G248P82502G1 initialement marqués tous les trois en Aqua (Figure 66).



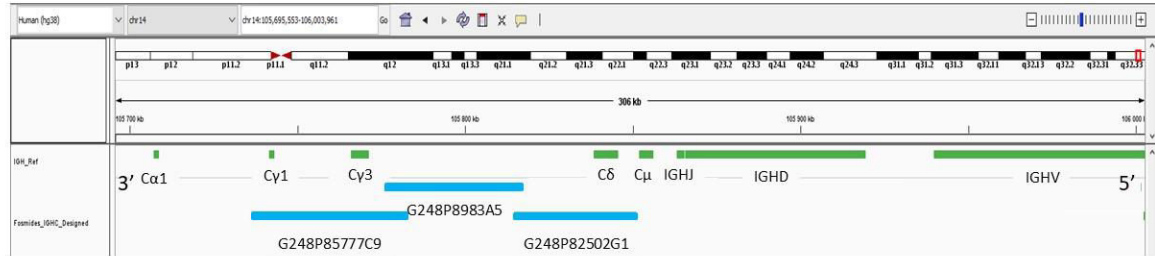
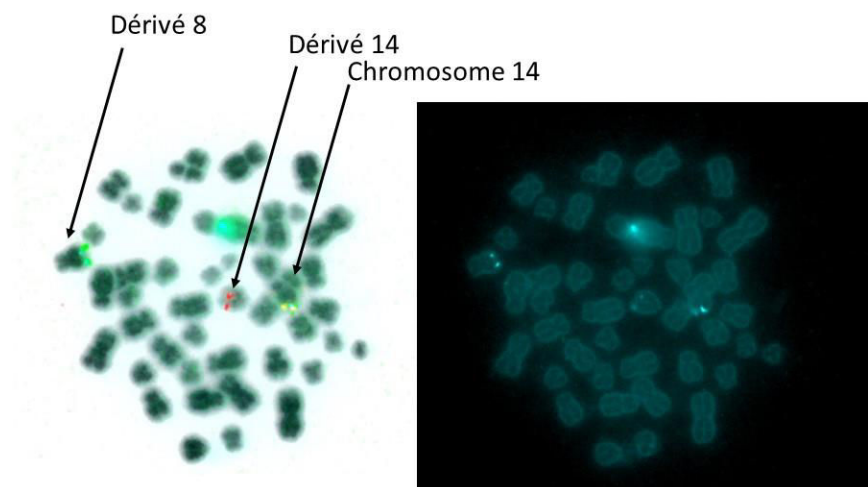
A**B**

Figure 66 : Image de FISH métaphasique avec les fosmides G248P85777C9, G248P8983A5 et G248P82502G1 marqués tous les trois en Aqua, ainsi que la sonde IGH Vysis.

A- Représentation schématique à partir du logiciel IGV du locus IGH (en vert) et des fosmides ciblant les régions constantes d'IGH (en bleu).

B- Trois signaux Aqua sont visibles : ils marquent le chromosome 14 normal et les dérivés 8 et 14. Le dérivé 8 semble être marqué plus fortement. La sonde IGH Vysis permet de visualiser une fusion sur le chromosome 14 normal, un signal rouge sur le dérivé 14 et un signal vert sur le dérivé 8.

Le résultat montre une séparation du signal Aqua avec un marquage résiduel sur le dérivé 14 faible et une délocalisation sur le dérivé 8 semblant plus marqué (Figure 66). Ceci semble indiquer un point de cassure au niveau des fosmides G248P85777C9 ou G248P8983A5. Pour aller plus loin, une deuxième FISH a été réalisée avec ces deux fosmides marqués respectivement en rouge et vert (Figure 67). Le résultat montre deux signaux de fusion marquant le chromosome 14 normal et le dérivé 8. Un signal rouge résiduel est retrouvé sur le dérivé 14. Ceci prouve une région de cassure située au niveau du fosmide G248P85777C9 ciblant les régions IGHG1 et IGHG3.



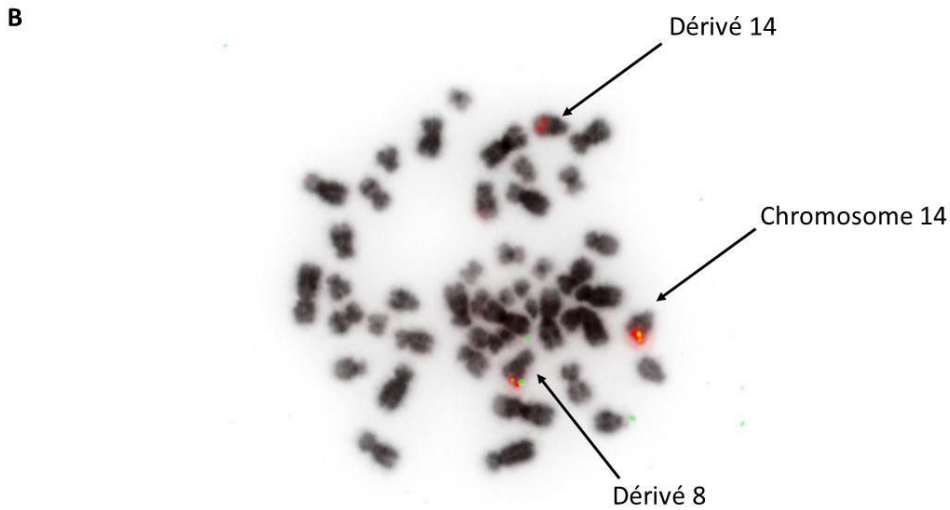
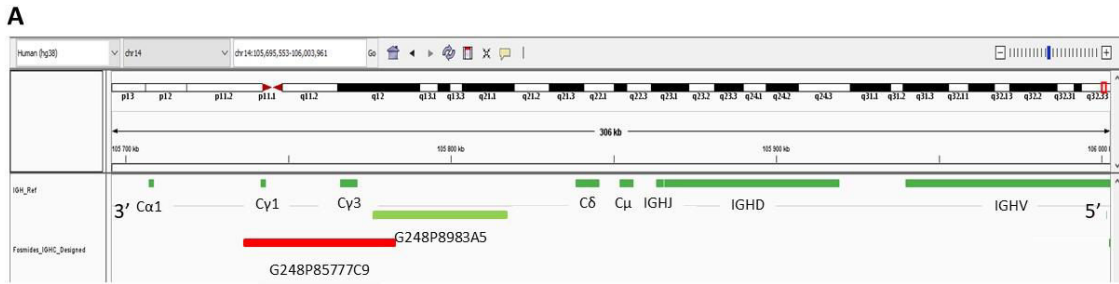


Figure 67 : Image de FISH métaphasique avec les fosmides G248P85777C9, G248P8983A5 marqués respectivement en rouge et vert.

A- Représentation schématique à partir du logiciel IGV du locus IGH (en vert) et des fosmides ciblant les régions constantes d'IGH : G248P85777C9, G248P8983A5 marqués respectivement en rouge et vert.

B- Deux fusions sont observables : l'une sur le chromosome 14 normal et l'autre sur le dérivé 8. Un signal rouge résiduel marque le dérivé 14.

Des expériences de LM-PCR ont alors été mises en œuvre avec des amorces ciblant IGHG1 et n'ayant pas permis de mettre en évidence le remaniement.

Une approche par séquençage du génome entier a alors été décidée. Celle-ci a permis de mettre en évidence le remaniement entre les chromosomes 8 et 14 avec les points de cassure suivants :

- Chromosome 14, position 105647500 (Figure 68)
- Chromosome 8, position 127196845 (Figure 69)





Figure 68 : Visualisation à partir du logiciel IGV des séquences obtenues par séquençage du génome entier au niveau du chromosome 14, position 105647500 concernant la patiente n°3. Les alignements colorés en gris correspondent aux séquences germinales. Ceux en violet représentent les séquences dont le début est aligné sur le chromosome 14 et la fin est située sur le chromosome 8. Le trait rouge représente le point de cassure identifié au niveau d'*IGH*. Il est situé 2711 pb en amont d'*IGHG2*. Une séquence de 7 nucléotides est commune aux deux chromosomes au niveau de la jonction des points de cassure et est encadrée ici en rose.

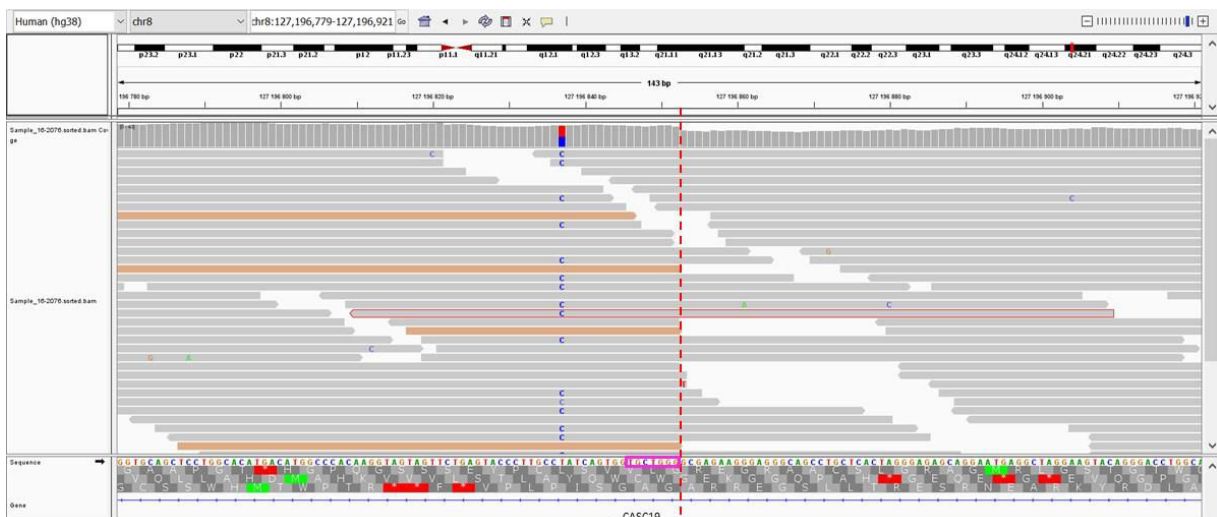


Figure 69 : Visualisation à partir du logiciel IGV des séquences obtenues par séquençage du génome entier au niveau du chromosome 8, position 127196845. Les alignements colorés en gris correspondent aux séquences germinales. Ceux en orange représentent les séquences pour lesquelles le début est aligné sur le chromosome 8 et la fin sur le chromosome 14. Le trait rouge représente la fin de lecture des séquences correspondant au remaniement au niveau du chromosome 8. Une séquence de 7 nucléotides est commune aux deux chromosomes au niveau de la jonction des points de cassure et est encadrée ici en rose. Le remaniement touche le gène *CASC19* au niveau du chromosome 8.



A partir de ces résultats, nous avons récupéré les séquences concernées. La séquence identifiée concernant le der(8)(t(8;14)) est la suivante :

ATTAAGTCCTGGGTGCAGCTCCTGGCACATGACATGGCCCACAAGGTAGTAGTTCTGAGTACCCTT
 GCCTATCAGTGGT**GCTGGGTCAGCTCCCTGAACGCACGGAGCCTCACCCCTCCCCTGCCCCAGGCCT**
GCTGCACTCTGGGCCTTTCTGGGCCTCCCTGGACTCTCCCTCCTCCCATCTGTGCACTCAGCACAGCT
CTCCCCTCCACTCCGCTGCTGA

avec en bleu la séquence issue du chromosome 8, en vert une séquence de transition commune aux deux chromosomes et en rouge la séquence issue du chromosome 14.

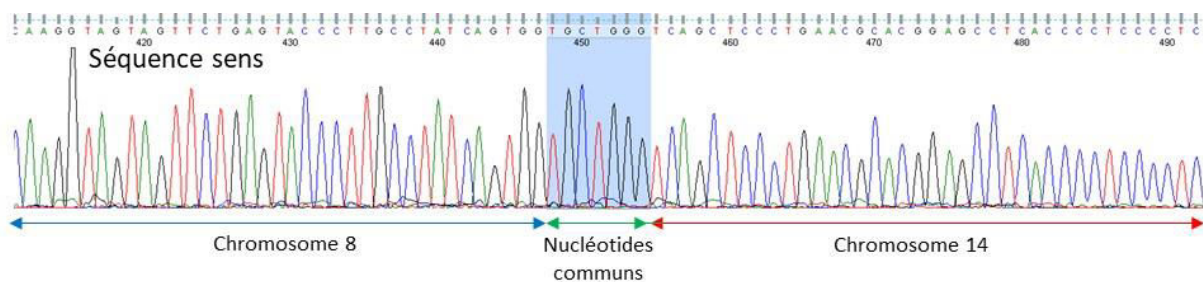


Figure 70 : Séquence sens obtenue par séquençage Sanger pour le der(8)t(8;14).

Ce résultat a été confirmé par séquençage Sanger. Seule la séquence sens est exploitable. En antisens, la présence d'une région répétée (CCCCTCCCCT) entraîne un arrêt prématuré de la séquence.

Ce point de cassure ne permet de mettre en évidence que le dérivé 8. La vérification par séquence du dérivé 14 issu des mêmes points de cassure ne permet pas d'amplification.

c- Courte discussion

Les résultats ont permis de mettre en évidence les points de cassure concernant la translocation entre les chromosomes 8 et 14. Le segment IGH impliqué est IGHG2 d'après le séquençage du génome entier. Ce résultat est discordant avec le résultat obtenu par FISH qui indiquait un point de cassure au niveau d'*IGHG3*. Il se peut, étant donnée la grande similarité entre ces régions, qu'il y ait une hybridation non spécifique de la sonde de FISH (fosmide).

Le point de cassure au niveau du chromosome 8 se situe dans le gène *CASC19* (pour Cancer Susceptibility 19), environ 538 kb en amont du gène *MYC*. La FISH réalisée avec la sonde *MYC* Dako aurait donc dû montrer une délocalisation du signal entier sur le chromosome 14.



Cependant, une seule mitose exploitable avait été visualisée et était normale. Il se peut que cette mitose n'ait pas été porteuse de la translocation. Les résultats de FISH interphasique sont compatibles avec le point de cassure retrouvé, n'entraînant pas de séparation du signal. Les résultats n'avaient pas été photographiés en 2012 et n'ont pas été reproduits avec la sonde DAKO. La sonde actuellement utilisée au laboratoire pour le locus MYC est la sonde Metasystem (Figure 36). Elle montre une séparation du signal (Figure 71). Cette sonde est constituée de 2 signaux vert et rouge plus espacés que la sonde DAKO et permet la détection d'un plus grand nombre de remaniement impliquant le locus 8q24, mais certains peuvent ne pas impliquer *MYC*, comme dans notre cas. C'est le cas par exemple des translocations impliquant *PVT1* (pour Plasmocytoma Variant Translocation gene 1), situé environ 52 kb en aval de *MYC*, et décrit dans des translocations variantes dans le lymphome de Burkitt⁶²³.



Figure 71 : Image de FISH métaphasique avec la sonde MYC Metasystem pour la patiente n°3. On observe une fusion sur un chromosome 8 normal, un signal rouge sur le dérivé 8 et un signal vert sur le dérivé 14.

Par séquençage du génome entier, nous n'avons pu mettre en évidence que le dérivé 8. Le dérivé 14 n'a pu être visualisé. Le dérivé 8 obtenu permet la délocalisation d'éléments régulateurs issus d'*IGH* : la 3'RR1 et E μ (Figure 72). Ceci peut donc être à l'origine de la dérégulation de potentiels oncogènes au niveau du dérivé 8.



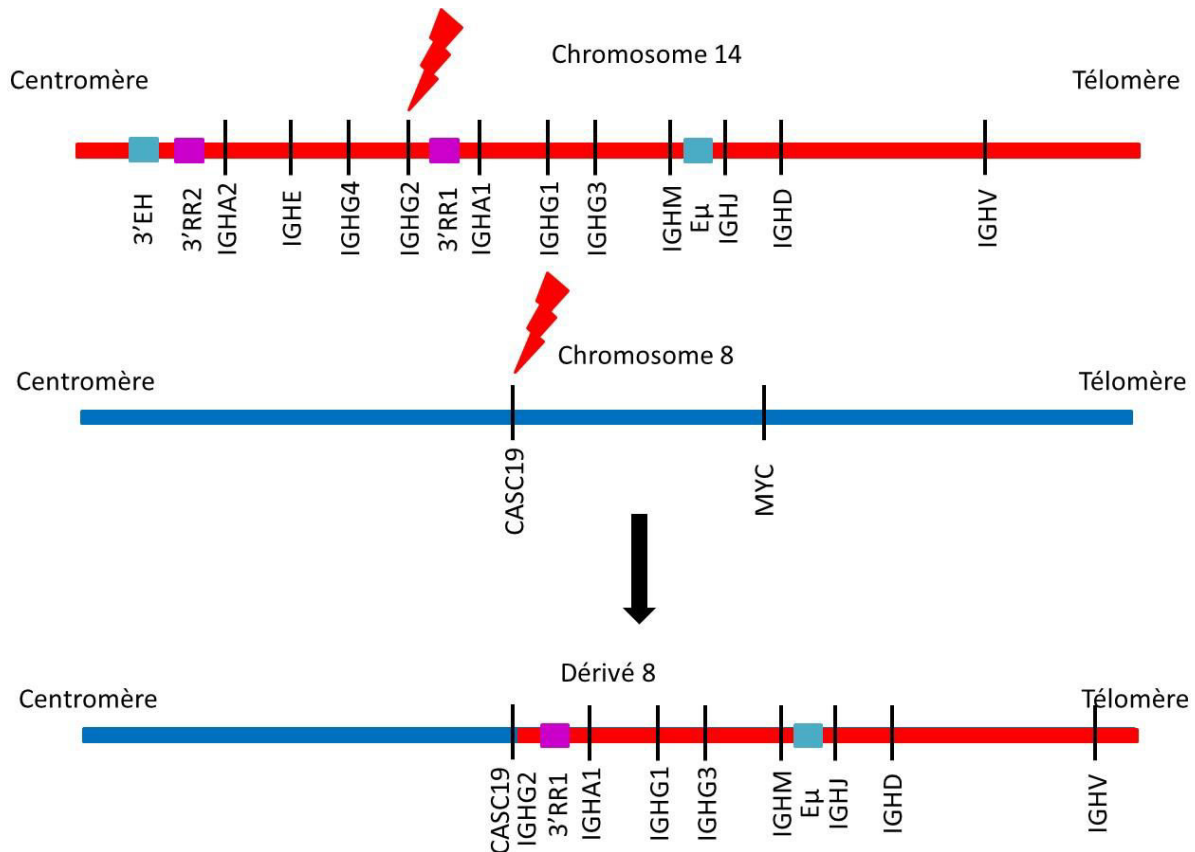


Figure 72 : Schéma représentant la translocation entre les bras long des chromosomes 8 et 14, ainsi que le dérivé 8 obtenu pour la patiente n°3. Les éléments régulateurs, enhancers et 3'RR, sont représentés en carrés bleus et violets respectivement.

Le dérivé 14, bien que non identifié, est également susceptible d'entraîner la dérégulation d'oncogènes via la 3'RR2 et 3'EH. Une analyse plus approfondie des résultats obtenus par séquençage du génome entier est susceptible de permettre d'identifier le dérivé 14. Nous allons également réaliser de la FISH à façon en ciblant *CASC19* afin de visualiser le réarrangement et éventuellement de voir si le dérivé 14 contient une partie de ce gène. Enfin, une analyse par RNA-Seq est prévue afin d'étudier les éventuelles dérégulations d'expression génique au niveau de la région 8q24.

Le mécanisme de survenue de la translocation est très probablement médié par AID au cours de la recombinaison de classe. En effet, le point de cassure au niveau du chromosome 14 touche la région switch IGHG2.

Concernant *CASC19*, il s'agit d'un lncRNA. Le point de cassure est situé au niveau de l'intron 1 (le plus long), le gène étant constitué de 3 exons. Une hyperexpression de ce gène a été décrite dans le cancer colorectal^{624,625}. Des essais *in vitro* montrent que son extinction entraîne une



inhibition de la migration cellulaire⁶²⁴. De plus, son expression est associée à la taille de la tumeur, la présence de métastases ganglionnaires et à distance, suggérant un rôle de CASC19 dans la survenue de métastases⁶²⁴.



B- Article 1 : Normal serum protein electrophoresis and mutated IGHV genes detect very slowly evolving chronic lymphocytic leukemia patients, accepté le 29 Mars 2018 dans Cancer medicine

1- Introduction

La leucémie lymphoïde chronique est une maladie très hétérogène du point de vue clinique et biologique. Le développement des techniques moléculaires à haut débit a permis de mettre en évidence de nombreuses mutations pouvant avoir un impact pronostique dans la LLC. Cependant, l'utilisation de toutes ces données n'est pas encore intégrée dans un score pronostique global. Actuellement, les techniques de NGS se développent et permettent de déterminer de nombreuses mutations géniques chez les patients atteints de LLC. En pratique, seule la mutation de *TP53* a un impact sur la prise en charge du patient, avec une adaptation du traitement en cas d'altération de ce gène.

Notre équipe s'est intéressée à l'impact de la présence d'un pic monoclonal d'isotype IgM dans la leucémie lymphoïde chronique⁴⁵⁷. Nous avons rapporté un pronostic péjoratif associé à la présence d'un pic IgM comparativement à une électrophorèse des protéines sériques normale, la présence d'une hypogammaglobulinémie ou d'un pic IgG.

Dans ce travail, nous nous sommes intéressés à l'impact pronostique de l'électrophorèse des protéines sériques (SPE) au diagnostic en comparant le statut normal versus anormal (que ce soit la présence d'un pic monoclonal ou d'une hypogammaglobulinémie), dans l'ère du NGS, afin de déterminer si nous retrouvons un impact pronostique statistiquement indépendant des nouveaux marqueurs moléculaires. L'autre intérêt de l'électrophorèse des protéines sériques est sa facilité de réalisation et son moindre coût. De plus, cet examen est généralement réalisé au diagnostic d'une LLC, principalement pour détecter une hypogammaglobulinémie pouvant nécessiter une prise en charge spécifique. Nous avons ainsi sélectionné 112 patients pour lesquels nous disposions de matériel biologique et d'une électrophorèse des protéines sériques réalisée au diagnostic, c'est-à-dire au cours de l'année du diagnostic et avant tout traitement de la LLC. Nous avons étudié les principales anomalies moléculaires décrites, à savoir la cytogénétique (*del(17p)*, *del(11q)*, trisomie 12 et *del(13q)*), les mutations (*TP53*, *SF3B1*, *NOTCH1* et *BIRC3*), ainsi que le statut mutationnel IGHV. Enfin, nous avons récupéré les données cliniques, en particulier concernant le traitement.



ORIGINAL RESEARCH

Normal serum protein electrophoresis and mutated IGHV genes detect very slowly evolving chronic lymphocytic leukemia patients

Jasmine Chauzeix^{1,2}, Marie-Pierre Laforêt¹, Mélanie Deveza¹, Liam Crowther², Elodie Marcellaud², Paco Derouault³, Anne-Sophie Lia³, François Boyer², Nicolas Barges¹, Guillaume Olombel^{2,4}, Arnaud Jaccard^{2,5}, Jean Feuillard^{1,2}, Nathalie Gachard^{1,2} & David Rizzo^{1,2}

¹Hématologie Biologique, Centre Hospitalier Universitaire de Limoges, Limoges, France

²Faculté de Médecine de Limoges, UMR CNRS 7276 CRIBL, Limoges, France

³Faculté de Médecine de Limoges, EA6309, Limoges, France

⁴Immunologie, Centre Hospitalier Universitaire de Limoges, Limoges, France

⁵Hématologie Clinique et Thérapie Cellulaire, Centre Hospitalier Universitaire de Limoges, Limoges, France

Keywords

Chronic lymphocytic leukemia, IGHV, prognosis, serum protein electrophoresis

Correspondence

David Rizzo, Centre de biologie et de recherche en santé, 2 avenue Martin Luther King, 87042 Limoges, France.
Tel: +33(0)5 55 05 61 80;
Fax: +33(0)5 55 05 61 85;
E-mail: david.rizzo@unilim.fr

Funding Information

This work was supported by "Ligue Nationale Contre le Cancer, équipe labellisée Ligue" and "Comité Orientation Recherche Cancer en Limousin" (CORC).

Received: 31 October 2017; Revised: 29 March 2018; Accepted: 29 March 2018

Cancer Medicine 2018; **7**(6):2621–2628

doi: 10.1002/cam4.1510

Abstract

More than 35 years after the Binet classification, there is still a need for simple prognostic markers in chronic lymphocytic leukemia (CLL). Here, we studied the treatment-free survival (TFS) impact of normal serum protein electrophoresis (SPE) at diagnosis. One hundred twelve patients with CLL were analyzed. The main prognostic factors (Binet stage; lymphocytosis; IGHV mutation status; *TP53*, *SP3B1*, *NOTCH1*, and *BIRC3* mutations; and cytogenetic abnormalities) were studied. The frequencies of IGHV mutation status, cytogenetic abnormalities, and *TP53*, *SP3B1*, *NOTCH1*, and *BIRC3* mutations were not significantly different between normal and abnormal SPE. Normal SPE was associated with Binet stage A, nonprogressive disease for 6 months, lymphocytosis below 30 G/L, and the absence of the IGHV3-21 gene rearrangement which is associated with poor prognosis. The TFS of patients with normal SPE was significantly longer than that of patients with abnormal SPE (log-rank test: $P = 0.0015$, with 51% untreated patients at 5.6 years and a perfect plateau afterward vs. a median TFS at 2.64 years for abnormal SPE with no plateau). Multivariate analysis using two different Cox models and bootstrapping showed that normal SPE was an independent good prognostic marker for either Binet stage, lymphocytosis, or IGHV mutation status. TFS was further increased when both normal SPE and mutated IGHV were present (log-rank test: $P = 0.008$, median not reached, plateau at 5.6 years and 66% untreated patients). A comparison with other prognostic markers suggested that normal SPE could reflect slowly advancing CLL disease. Altogether, our results show that a combination of normal SPE and mutated IGHV genes defines a subgroup of patients with CLL who evolve very slowly and who might never need treatment.

Introduction

Chronic lymphocytic leukemia (CLL), the most frequent indolent B-cell cancer with blood passage of tumor cells in the elderly in western countries, is characterized by a lymphocytosis exceeding five G/L comprising small circulating monomorphic round B lymphocytes, with constant

infiltration of bone marrow and secondary lymphoid organs [1, 2]. Diagnosis requires flow cytometry immunophenotyping of cells, usually based on the Matutes–Moreau score, which combines the expression of CD5 and CD23, weak surface membrane immunoglobulin levels and weak or absent FMC7, CD22, or CD79b expression. CLL is typically characterized by a Matutes score of 4 or 5 [3,

© 2018 The Authors. *Cancer Medicine* published by John Wiley & Sons Ltd.

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.

2621



4]. Disease progression is heterogeneous with overall survival (OS) ranging from a few years to decades.

Because some patients with CLL may have a life expectancy as long as healthy subjects of the same age without treatment while others will rapidly evolve, it has been recognized for a long time that finding reliable prognostic factors is necessary. The first and most universal clinical and biological predictive tools are the Binet and Rai staging systems that predict the clinical course [5, 6]. For example, patients with Binet Stage A CLL exhibit good treatment-free survival (TFS), with a life expectancy close to that of healthy subjects of the same age. However, it is well known that the progression of these patients is rather heterogeneous [7].

Almost twenty years after the publication of Binet and Rai stages, Damle et al. [8], and Hamblin et al. [9], showed that a greater than two percent mutation rate of the variable rearranged region of the immunoglobulin heavy chain gene (IGHV) predicts good overall survival of patients with CLL. In the early 2000s, various cytogenetic prognostic markers were identified, including isolated del(13q), which predicts very good overall survival, and del(17p), which is reliably associated with resistance to fludarabine. High-throughput sequencing of the CLL genetic landscape revealed several new gene mutations associated with a poor prognosis. For example, *NOTCH1* is associated with Richter syndrome, and the *SF3B1* mutation also seems to predict fludarabine resistance [10]. Most of these genetic markers are in fact predictive of more rapid disease progression, treatment resistance, or transformation, and very few of them identify patients who will never need to be treated. Mutated IGHV and the deletion of the long arm of chromosome 13 (del(13q)) may be the only markers that predict prolonged TFS [11–14].

To assess the cytogenetic and mutational status of CLL, each case requires high levels of technical and medical expertise, and the process is time consuming and costly [15]. To circumvent these difficulties, various other biological parameters, such as serum soluble CD23, CD38, and ZAP70 expressions, have been reported as prognostic markers [16]. However, as exemplified for antibodies detecting ZAP70, published reagents and/or methods for the detection and quantification of these markers are rather diverse [17]. To date, these markers are rarely used in the routine staging and follow-up of patients with CLL.

Universal biological analyses prescribed for the initial staging of patients with CLL include serum protein electrophoresis (SPE). SPE detects either hypogammaglobulinemia, which is supposed to predict infection risk [18] and is likely to be related to tumor burden [19], or immunoglobulin peaks. The prevalence rates of hypogammaglobulinemia and monoclonal paraprotein are 25% [20] and 11% [21], respectively. The clinical significance of

abnormal serum protein electrophoresis has been previously studied, showing a worse prognosis in the presence of serum paraprotein [19, 22]. The impact of hypogammaglobulinemia is controversial regarding either infection risk, TFS, or OS [20, 23, 24]. As most publications only report on the significance of hypogammaglobulinemia or the monoclonal component alone, the prognostic value of a normal gamma globulin serum profile (normal SPE) per se has not been reported.

Here, we studied the correlation of normal SPE with TFS and its relationship with other variables (Binet Stage, biological parameters, and genetic characteristics). We found that normal SPE is an independent good prognostic variable and that the combination of normal SPE and mutated IGHV genes predicts patients with excellent TFS who are likely to have very slowly progressing CLL and who will never need treatment after diagnosis.

Materials and Methods

Patients

We analyzed 112 samples from patients with typical CLL, diagnosed between 2004 and 2016, and 93% were diagnosed between 2004 and 2012. By flow cytometry, their Matutes score was 4 or 5 in all cases. Inclusion criteria were based on the availability of biologic samples and concomitant serum protein electrophoresis with immunofixation electrophoresis when a peak was detected. Normal concentrations of gamma globulin ranged from 6 g/L to 11 g/L. The threshold for peak detection was 1 g/L of monoclonal immunoglobulin. Patients were classified according to the Binet Stage, nonprogressive/progressive disease (known stable disease for 6 months or not), the Rossi score, and their IGHV mutation status.

DNA extraction

Genomic DNA was extracted from peripheral blood mononuclear cells using the QIAamp DNA Blood Mini Kit (Qiagen, Venlo, the Netherlands) according to the manufacturer's instructions.

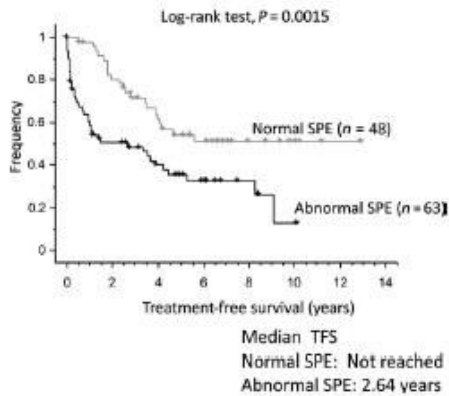
Sequencing of *SF3B1*, *NOTCH1*, *TP53*, and *BIRC3*

Patients were first screened for mutations in *SF3B1* exons 14, 15, 16, and 18 (NM_012433), *NOTCH1* exon 34 (NM_017617), and *BIRC3* exons 6 to 9 (NM_001165) by Sanger sequencing (BigDye Terminator cycle kit (Thermo Fischer Scientific, Waltham, MA) and ABI PRISM 3130xl Genetic Analyzer) after amplification of 50 ng genomic DNA. Sequences were analyzed with the Mutation Surveyor



Table 1. Patient characteristics at diagnosis (mean \pm standard deviation is given for age, lymphocytosis, hemoglobin, and platelet counts).

	Total (n = 112)	Normal SPE (n = 49)	Abnormal SPE (n = 63)	P
Sex	Men: 63.5% (n = 66) Women: 36.5% (n = 38)	Men: 63.3% (n = 31) Women: 36.7% (n = 18)	Men: 58.7% (n = 37) Women: 41.3% (n = 26)	NS
Age (years)	66.6 \pm 10.5	65.7 \pm 10.5	67.3 \pm 10.5	NS
Binet				
Stage A	81.3% (n = 91)	95.9% (n = 47)	69.8% (n = 44)	χ^2 test: P = 0.001
Stage B	15.2% (n = 17)	2.0% (n = 1)	25.4% (n = 16)	
Stage C	3.6% (n = 4)	2.0% (n = 1)	4.8% (n = 3)	
Nonprogressive disease for 6 months after diagnosis	79.5% (n = 89)	98.0% (n = 48)	65.1% (n = 41)	χ^2 test: P < 0.0001
IGHV gene mutation (M-CLL)	59.5% (n = 66)	61.2% (n = 30)	58.1% (n = 36)	NS
Lymphocytosis (G/L)	32.8 \pm 63.6	18.2 \pm 13.3	44.1 \pm 82.5	NS
Lymphocytosis >30 G/L*	25.9% (n = 29)	14.3% (n = 7)	34.9% (n = 22)	t test: P = 0.017
Hemoglobin (g/dL)	13.7 \pm 1.8	14.1 \pm 1.5	13.5 \pm 2.0	NS
Platelets (G/L)	212.2 \pm 87.7	228.1 \pm 89.1	199.7 \pm 184.5	NS
Cytogenetic				
del(17p)	0.9% (n = 1)	2.1% (n = 1)	0.0% (n = 0)	NS
del(11q)	11.7% (n = 13)	8.3% (n = 4)	14.3% (n = 9)	
Trisomy 12	15.5% (n = 17)	10.4% (n = 5)	19.4% (n = 12)	
Isolated del(13q)	28.8% (n = 32)	32.7% (n = 16)	25.8% (n = 16)	
Mutations				
NOTCH1	8.9% (n = 10)	12.2% (n = 6)	6.3% (n = 4)	NS
SF3B1	11.6% (n = 13)	6.1% (n = 3)	15.9% (n = 10)	
TP53	7.1% (n = 8)	4.1% (n = 2)	9.5% (n = 6)	
BIRC3	2.7% (n = 3)	2.0% (n = 1)	3.2% (n = 2)	

**Figure 1.** Kaplan-Meier curves for treatment-free survival according to SPE status.

DNA Analysis Software V4.0 (SoftGenetics, State College, PA). Patients not screened by Sanger sequencing were assayed by high-throughput sequencing (HTS) (Proton, Ion Torrent), using a panel that targeted the same regions. The *TP53* gene was analyzed by HTS in all cases. The panels were designed on the AmpliSeq designer platform (www.ampliseq.com). Libraries were constructed using the

Ion AmpliSeq Library kit 2.0 (Thermo Fischer Scientific) according to the manufacturer's instructions. Variants were filtered to retain exonic mutations and to obtain a prediction score of pathogenesis by SIFT and/or CADD positive or unknown on both scores. Known mutations were identified by querying the COSMIC, 1000 genomes, and dbSNP databases.

Immunoglobulin gene sequence analysis

Amplification of V, D, and J rearranged genes was performed using the Biomed-2 strategy with FR1 and FR2 primers as previously described. Sequence analysis of VDJ segments was performed as previously described [25].

Cytogenetics

Conventional cytogenetic and fluorescence in situ hybridization were performed according to conventional procedures. Quantitative multiplex PCR of short fluorescent fragment (QMPSF) was performed as described elsewhere [26].

Flow cytometry

Multiparametric flow cytometry analysis was performed on heparinized blood samples as described elsewhere [27].

Statistics

The chi-square test was used to evaluate the difference between categorical covariates for the SPE subgroups. A lymphocytosis threshold of 30 G/L was used as determined previously [28]. The effects of demographics (age and sex), Binet stage, lymphocytosis, genetics (IGHV mutation status, *SF3B1*, and *NOTCH1* mutations), cytogenetics (trisomy 12, del(11q), del(13q), and del(17p)), and serum parameter (SPE subgroups) covariates on treatment-free survival (TFS) were examined using the Cox proportional hazard model [29]. Briefly, each covariate was first tested in univariate analysis. All significant covariates with a *P*-value below 0.20 after univariate analysis were included simultaneously in the multivariate model. The significance of variables in the final model was tested by a backward stepwise process using the likelihood ratio to evaluate the effect of omitting variables. The stability of the final model was validated by performing 1000 bootstrap samples.

Results

To study the significance of normal serum protein electrophoresis in CLL, a series of 112 patients was selected. Their clinical and biological characteristics are shown in Table 1. Among them, 49 (44%) had a normal SPE in the year of diagnosis, and 63 (56%) had an abnormal SPE (hypogammaglobulinemia: 24%; gamma globulin peak: 32%). These two groups of patients with CLL were comparable in terms of age and sex, as well as hemoglobin concentrations and platelet counts.

Regarding the Binet stages of the patients, 39 of 41 patients (96%) with normal SPE were Binet stage A, while 19 of the 63 (30%) patients with abnormal SPE were Binet stage B or C (Table 1, *P* = 0.001, chi-square test). Nonprogressive CLL (known stable disease for 6 months) was found in 98% of patients with normal SPE (48/49), while 65% (41/63) of the patients with abnormal SPE had stable disease (*P* = 2×10^{-5} , chi-square test). The number of patients with lymphocytosis above 30 G/L at diagnosis was significantly lower in patients with normal SPE (14.3% in normal SPE vs. 34.9% in abnormal SPE, Student's *t* test: *P* = 0.017) (Table 1).

With a threshold of a 2% mutation rate, the number of patients with mutated (M-CLL) or unmutated (UM-CLL) IGHV genes was comparable between normal and abnormal SPE (Table 1). Analysis of the IGHV gene repertoire was performed for IGHV gene rearrangements with a frequency over 5% among all patients with CLL of this series. Patients with a normal SPE exhibited significantly increased usage of IGHV1-2. In contrast, the rearrangement of the IGHV3-21 segment, associated with a poor

Table 2. Cox univariate analysis of TFS.

	HR	LCI	UCI	P-value
Binet stage				
Binet stage A (n = 91)	1.00			
Binet stage B (n = 17)	21.44	10.1	45.6	<0.0001
Binet stage C (n = 4)	42.97	11.2	164.6	<0.0001
Lymphocytosis				
<30 G/L (n = 83)	1.00			
≥30 G/L (n = 29)	4.12	2.44	6.96	<0.0001
IGHV mutation status				
M-CLL (n = 66)	1.00			
UM-CLL (n = 45)	2.48	1.49	4.13	0.0005
Hemoglobin				
>10 g/dL (n = 108)	1.00			
<10 g/dL (n = 3)	14.37	3.14	65.91	0.0006
SPE group				
Normal SPE (n = 49)	1.00			
Abnormal SPE (n = 63)	2.41	1.42	4.10	0.001
Rossi's score				
Very Low or Low (n = 69)	1.00			
Intermediate or high (n = 41)	2.08	1.25	3.45	0.005
SF3B1 status				
SF3B1 wild type (n = 99)	1.00			
SF3B1 mutation (n = 13)	2.28	1.18	4.42	0.014
Cytogenetics				
Normal FISH or QMPSF (n = 47)	1.00			
Isolated del(13)(q14) (n = 32)	0.45	0.24	0.83	0.010
Del(11q) (n = 13)	2.15	1.05	4.40	0.036
Trisomy 12 (n = 17)	1.62	0.86	3.07	0.14
Del(17p) (n = 1) ¹	2.73	0.37	19.96	0.32
Platelet count				
>100 G/L (n = 105)	1.00			
<100 G/L (n = 6)	2.36	0.85	6.56	0.099

HR, hazard ratio; LCI, lower confidence interval; UCI, upper confidence interval; *P*-value, probability that the hazard ratio = 1 (null hypothesis).

¹Excluded criteria because of low numbers.

prognosis, was overrepresented in patients with abnormal SPE and was completely absent in patients with normal SPE (Fig. S1).

The *BIRC3* and *NOTCH1* mutation rates were low and comparable between patients with normal and abnormal SPE (Table 1). Although not significant, the *SF3B1* mutation rate was lower in patients with normal SPE (6% vs. 16% for abnormal SPE, log-rank test: *P* = 0.11). There were no differences for del(11q), trisomy 12, or isolated del(13q) (Table 1). The number of patients with del(17p) was too low to draw any conclusion.

We then studied TFS according to SPE. Patients with normal SPE had a significantly increased TFS, with a median that was not statistically reached. The plateau

Table 3. 5 statistically significant variables after Cox multivariate analysis of TFS (the two models were designed to include all nonoverlapping variables with a P-value <0.20 after univariate analysis).

	1st Model					2nd Model						
	Multivariate analysis		Internal bootstrapping validation (mean)			Multivariate analysis		Internal bootstrapping validation (mean)				
	HR	LCI	UCI	P-value	HR	LCI	UCI	% Bootstrap selection	HR	LCI	UCI	% Bootstrap selection
Binet stage	1.00				1.00				1.00			
Binet stage A (n = 91)	1.00				1.00			84.5	1.00			
Binet stage B (n = 17)	5.39	26.35	<0.0001		11.92	5.39	26.36		3.40	1.98	5.85	97
Binet stage C (n = 4)	72.34	313.49	<0.0001		72.34	16.69	313.28		2.25	1.30	3.88	78
Lymphocytosis	1.00				1.00				1.00			
<30 G/L (n = 83)	1.00				1.00			83.8	1.00			
>30 G/L (n = 19)	2.53	1.43	0.002		2.53	1.43	4.50		2.47	1.47	4.17	85.5
SPE group	1.00				1.00				1.00			
Normal SPE (n = 49)	1.00				1.00			70	1.00			
Abnormal SPE (n = 63)	2.52	1.43	0.001		2.52	1.43	4.43		2.47	1.47	4.17	85.5
IGHV mutation status	1.00				1.00				1.00			
M-CLL (n = 66)	1.00				1.00			67.5	1.00			
UM-CLL (n = 45)	1.99	1.15	0.01		1.99	1.15	3.45		1.00			
Hemoglobin	1.00				1.00				1.00			
< 10 g/dL (n = 108)	1.00				1.00				1.00			
> 10 g/dL (n = 3)	17.10	3.33	87.72	0.0007	17.10	3.33	87.72		17.10	3.33	87.68	53

HR, hazard ratio; LCI, lower confidence interval; UCI, upper confidence interval; P-value, probability that hazard ratio = 1 (null hypothesis).

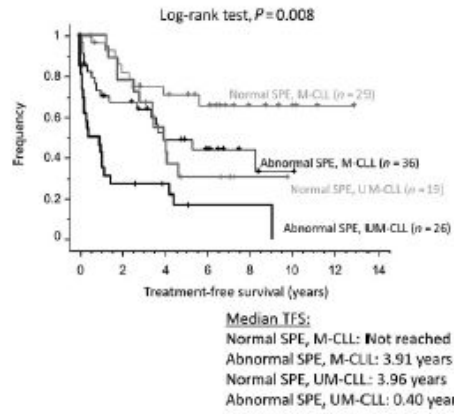


Figure 2. Kaplan-Meier curves for treatment-free survival according to SPE and IGHV mutation status.

began at 5.6 years and was at 51% of cumulated TFS. Patients with abnormal SPE had a median TFS of 2.64 years with no plateau and had a comparable TFS to patients with either hypogammaglobulinemia or an immunoglobulin peak (Figs. 1 and S2). In comparison, the median TFS of patients with mutated IGHV was 8.2 years, and the plateau began at 8.4 years, with 48% of patients being treatment-free (Fig. S3).

In addition to SPE status, the other significant variables after univariate analysis were hemoglobin levels, lymphocytosis, IGHV mutation status, *SF3B1* mutation, isolated del(13q), and del(11q). Criteria that integrated various variables, such as Binet stage or Rossi's score, were also highly significant (Table 2). Compared to normal SPE, the hazard ratio for treatment was 2.41 for abnormal SPE, a value that was compared to that of IGHV mutation status and of Rossi's score. For multivariate analysis, two Cox models were designed. The first model included the Binet stage, Rossi's score, IGHV mutation status, lymphocytosis, and SPE status, and the second model separated all biological criteria. Bootstrapping showed the stability of both models. Normal SPE was an independent variable in both. Other independent variables included the Binet stage, Rossi's score, and lymphocytosis for the first model, and IGHV mutation status and lymphocytosis for the second one (Table 3).

We looked at the relationship between SPE and either the IGHV mutation status or Rossi's score. Patients with normal SPE and mutated IGHV had an excellent prognosis (median TFS not reached after 13 years of follow-up, with a plateau at 5.6 years and 66% of accumulated TFS), while patients with normal SPE and unmutated IGHV genes or abnormal SPE and mutated IGHV genes

had comparable and shorter TFS (median TFS at 3.96 years and 3.91 years, respectively, Fig. 2). The group with abnormal SPE and unmutated IGHV genes had a very short TFS (median TFS = 0.40 years). If restricted to the 84 patients with both stage A and stable CLL, the TFS curves of patients with normal and abnormal SPE were parallel for the first 2.6 years. However, a plateau was reached at 5.6 years only for patients with normal SPE and mutated IGHV genes (Fig. S4A). Indeed, among patients with mutated IGHV genes, the TFS curves of those with normal or abnormal SPE were clearly separated if the starting time point was translated from 0 to 2.6 years (Fig. S4B).

Normal SPE had no correlation with the TFS of patients with isolated del(13q) (very low-risk Rossi's score, Fig. S5A) but was correlated with the significantly improved TFS of all other patients of Rossi's model (Fig. S5B and C). Moreover, even though their numbers were low, patients with normal SPE, mutated IGHV genes, and isolated del(13q) together had much longer TFS than the others (log-rank test: $P = 0.0295$, Fig. S6).

Altogether, our results suggest that the combination of normal SPE with mutated IGHV genes identifies patients with a very slowly evolving CLL, and even though isolated del(13q) was not selected after multivariate analysis, its presence is likely to be an additional favorable parameter.

Discussion

In this study, we showed that normal SPE is an independent good prognostic marker *per se* that predicts longer TFS in CLL, but most importantly, the combination of normal SPE and mutated IGHV genes defines a subgroup of patients with excellent TFS.

CLL has a highly variable natural course with overall survival ranging from a few years to several decades. As very well described by Binet and Rai [5, 6], the difficulty for the physician is to correctly predict disease evolution as it has a strong impact on patient medical care. It is remarkable that their scoring systems, based on a few simple clinical and biological parameters defined more than 30 years ago, are still valid and have not been supplanted by any of the "new" markers published after these princeps papers. Facing this plethora of markers, the current tendency is to search for simplicity [30–32]. In that context, the Binet and Rai scoring systems are still the keystones of CLL staging at diagnosis. Additionally, only IGHV mutation status and del(17p)/TP53 mutations are currently recommended to predict CLL progression and resistance to therapeutics [30, 31, 33].

Despite this apparent simplicity, within each group and mainly within good prognosis groups (Binet stage A or

Rai stage 0–I), patient progression is rather heterogeneous. Most of the other markers proposed in CLL staging are associated with poor prognosis when present. In other words, the absence of the marker is not inevitably associated with the absence of disease progression. IGHV mutation status and chromosome 13 abnormalities are thus particular as it is the presence of somatic mutations in the variable regions of the immunoglobulin heavy chain genes or the presence of isolated del(13q) that predicts favorable evolution [13]. Our results show that combining SPE results and IGHV mutation status improves the power of predicting a favorable disease course as the presence of normal SPE plus mutated IGHV genes is associated with very long TFS. Of note, even if the number of patients in this series was too low to form a firm conclusion, the presence of these two criteria combined with the presence of isolated del(13q) tended to define an even better prognosis group.

Hamblin et al. [9], and Opezzo et al. [34], have proposed that CLL with mutated IGHV genes could be a different disease with a more indolent course. Normal SPE appeared to be an independent parameter to predict longer TFS and was associated with Binet stage A. For mutated IGHV genes, the presence of normal SPE could also reflect a more indolent natural history. Supporting this hypothesis is the fact that various poor prognostic markers, such as the short-term evolution of the disease, lymphocytosis over 30 G/L, or the IGHV3-21 gene rearrangement, were almost always absent in patients with normal SPE. Based on their excellent TFS, patients with CLL harboring both normal SPE and mutated IGHV status might never require any specific treatment.

Acknowledgments

This work was supported by "Ligue Nationale Contre le Cancer, équipe labellisée Ligue" and "Comité Orientation Recherche Cancer en Limousin" (CORC). We thank Dr Jean-Baptiste Woillard (Pharmacology Department of Limoges University Hospital, France) for having supervised the statistical analysis, and Dr Jeanne Cook-Moreau (UMR CNRS 7276, France) for careful English editing.

Conflict of Interest

None declared.

References

1. Swerdlow, S. H., and International Agency for Research on Cancer. 2008. WHO classification of tumours of haematopoietic and lymphoid tissues. International Agency for Research on Cancer, Lyon.



2. Tamaru, J.-I. 2017. 2016 revision of the WHO classification of lymphoid neoplasms. *Rinsho Ketsueki* 58:2188–2193.
3. Matutes, E., K. Owusu-Ankomah, R. Morilla, J. Garcia Marco, A. Houlihan, T. H. Que, et al. 1994. The immunological profile of B-cell disorders and proposal of a scoring system for the diagnosis of CLL. *Leukemia* 8:1640–1645.
4. Moreau, E. J., E. Matutes, R. P. A'hern, A. M. Morilla, R. M. Morilla, K. A. Owusu-Ankomah, et al. 1997. Improvement of the chronic lymphocytic leukemia scoring system with the monoclonal antibody SN8 (CD79b). *Am. J. Clin. Pathol.* 108:378–382.
5. Binet, J. L., A. Auquier, G. Dighiero, C. Chastang, H. Pignat, J. Goasguen, et al. 1981. A new prognostic classification of chronic lymphocytic leukemia derived from a multivariate survival analysis. *Cancer* 48:198–206.
6. Rai, K. R., A. Sawitsky, E. P. Cronkite, A. D. Chanana, R. N. Levy, and B. S. Pasternack. 1975. Clinical staging of chronic lymphocytic leukemia. *Blood* 46:219–234.
7. Letestu, R., V. Lévy, V. Eclache, F. Baran-Marszak, D. Vaur, D. Naguib, et al. 2010. Prognosis of Binet stage A chronic lymphocytic leukemia patients: the strength of routine parameters. *Blood* 116:4588–4590.
8. Damle, R. N., T. Wasil, F. Fais, F. Ghiotto, A. Valetto, S. L. Allen, et al. 1999. Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. *Blood* 94:1840–1847.
9. Hamblin, T. J., Z. Davis, A. Gardiner, D. G. Oscier, and F. K. Stevenson. 1999. Unmutated Ig V(H) genes are associated with a more aggressive form of chronic lymphocytic leukemia. *Blood* 94:1848–1854.
10. Rossi, D., A. Bruscajgin, V. Spina, S. Rasi, H. Khiabani, M. Messina, et al. 2011. Mutations of the SF3B1 splicing factor in chronic lymphocytic leukemia: association with progression and fludarabine-refractoriness. *Blood* 118:6904–6908.
11. Rigolin, G. M., I. del Giudice, L. Formigaro, E. Saccenti, S. Martinelli, M. Cavallari, et al. 2015. Chromosome aberrations detected by conventional karyotyping using novel mitogens in chronic lymphocytic leukemia: clinical and biologic correlations. *Genes Chromosom. Cancer* 54:818–826.
12. Rossi, D., S. Rasi, V. Spina, A. Bruscajgin, S. Monti, C. Ciardullo, et al. 2013. Integrated mutational and cytogenetic analysis identifies new prognostic subgroups in chronic lymphocytic leukemia. *Blood* 121:1403–1412.
13. Gladstone, D. E., L. Swinnen, Y. Kasamon, A. Blackford, C. D. Gocke, C. A. Griffin, et al. 2011. Importance of immunoglobulin heavy chain variable region mutational status in del(13q) chronic lymphocytic leukemia. *Leuk. Lymphoma* 52:1873–1881.
14. Sandoval-Sus, J. D., J. C. Chavez, S. Dalia, S. M. H. Naqvi, C. Talati, L. Nodzon, et al. 2018. Association between immunoglobulin heavy-chain variable region mutational status and isolated favorable baseline genomic aberrations in chronic lymphocytic leukemia. *Leuk. Lymphoma* 59:59–68.
15. Hallek, M. 2017. Chronic lymphocytic leukemia: 2017 update on diagnosis, risk stratification, and treatment. *Am. J. Hematol.* 92:946–965.
16. Rosenquist, R., D. Cortese, S. Bhoi, L. Mansouri, and R. Gunnarsson. 2013. Prognostic markers and their clinical applicability in chronic lymphocytic leukemia: where do we stand? *Leuk. Lymphoma* 54:2351–2364.
17. Gachard, N., A. Salviat, C. Boutet, C. Arnoulet, F. Durrieu, B. Lenormand, et al. 2008. Multicenter study of ZAP-70 expression in patients with B-cell chronic lymphocytic leukemia using an optimized flow cytometry method. *Haematologica* 93:215–223.
18. Forconi, F., and P. Moss. 2015. Perturbation of the normal immune system in patients with CLL. *Blood* 126:573–581.
19. Rizzo, D., J. Chauzeix, F. Trimoreau, J. B. Woillard, F. Genevieve, A. Bouvier, et al. 2015. IgM peak independently predicts treatment-free survival in chronic lymphocytic leukemia and correlates with accumulation of adverse oncogenetic events. *Leukemia* 29:337–345.
20. Parikh, S. A., J. F. Leis, K. G. Chaffee, T. G. Call, C. A. Hanson, W. Ding, et al. 2015. Hypogammaglobulinemia in newly diagnosed chronic lymphocytic leukemia: natural history, clinical correlates, and outcomes. *Cancer* 121:2883–2891.
21. Martin, W., R. Abraham, T. Shanafelt, R. J. Clark, N. Bone, S. M. Geyer, et al. 2007. Serum-free light chain—a new biomarker for patients with B-cell non-Hodgkin lymphoma and chronic lymphocytic leukemia. *Transl. Res.* 149:231–235.
22. Xu, W., Y. H. Wang, L. Fan, C. Fang, D. X. Zhu, D. M. Wang, et al. 2011. Prognostic significance of serum immunoglobulin paraprotein in patients with chronic lymphocytic leukemia. *Leuk. Res.* 35:1060–1065.
23. Andersen, M. A., F. J. Vojdeman, M. K. Andersen, P. D. N. Brown, C. H. Geisler, O. Weis Bjerrum, et al. 2016. Hypogammaglobulinemia in newly diagnosed chronic lymphocytic leukemia is a predictor of early death. *Leuk. Lymphoma* 57:1592–1599.
24. Matutes, E., and A. Polliack. 2016. A matter of life or early death in CLL: do serum immunoglobulin levels really count? *Leuk. Lymphoma* 57:1501–1502.
25. Gachard, N., M. Parrens, I. Soubeyran, B. Petit, A. Marfak, D. Rizzo, et al. 2013. IGHV gene features and MYD88 L265P mutation separate the three marginal zone lymphoma entities and Waldenström macroglobulinemia/lymphoplasmacytic lymphomas. *Leukemia* 27:183–189.



26. Bastard, C., G. Raux, C. Fruchart, F. Parmentier, D. Vaur, D. Penther, et al. 2007. Comparison of a quantitative PCR method with FISH for the assessment of the four aneuploidies commonly evaluated in CLL patients. *Leukemia* 21:1460–1463.
27. Rizzo, D., A. Lotay, N. Gachard, I. Marfak, J. L. Faucher, F. Trimoreau, et al. 2013. Very low levels of surface CD45 reflect CLL cell fragility, are inversely correlated with trisomy 12 and are associated with increased treatment-free survival. *Am. J. Hematol.* 88:747–753.
28. Del Giudice, I., F. R. Mauro, M. S. De Propris, S. Santangelo, M. Marinelli, N. Peragine, et al. 2011. White blood cell count at diagnosis and immunoglobulin variable region gene mutations are independent predictors of treatment-free survival in young patients with stage A chronic lymphocytic leukemia. *Haematologica* 96:626–630.
29. Coordonné par Andrew Kramar et Simone Mathoulin-Pélissier. Méthodes biostatistiques appliquées à la recherche clinique en cancérologie. Available at http://www.jle.com/fr/ouvrages/e-docs/methodes_biostatistiques_appliquees_a_la_recherche_clinique_en_cancerologie_290431/ouvrage.phtml (accessed 17 October 2017).
30. Rosenquist, R., A. Rosenwald, M. Q. Du, G. Gaidano, P. Groenen, A. Wotherspoon, et al. 2016. Clinical impact of recurrently mutated genes on lymphoma diagnostics: state-of-the-art and beyond. *Haematologica* 101:1002–1009.
31. International CLL-IPI working group. 2016. An international prognostic index for patients with chronic lymphocytic leukaemia (CLL-IPI): a meta-analysis of individual patient data. *Lancet Oncol.* 17:779–790.
32. Gocke, C. D., and D. E. Gladstone. 2018. The absolute percent deviation of IGHV mutation rather than a 98% cut-off predicts survival of chronic lymphocytic leukaemia patients treated with fludarabine, cyclophosphamide and rituximab. *Br. J. Haematol.* 180:7–8.
33. Delgado, J., M. Doubek, T. Baumann, J. Kotaskova, S. Molica, P. Mozas, et al. 2017. Chronic lymphocytic leukemia: a prognostic model comprising only two biomarkers (IGHV mutational status and FISH cytogenetics) separates patients with different outcome and simplifies the CLL-IPI. *Am. J. Hematol.* 92:375–380.
34. Oppezio, P., C. Magnac, S. Bianchi, F. Vuillier, A. Tiscornia, G. Dumas, et al. 2002. Do CLL B cells correspond to naive or memory B-lymphocytes? Evidence for an active Ig switch unrelated to phenotype expression and Ig mutational pattern in B-CLL cells. *Leukemia* 16:2438–2446.

Supporting Information

Additional supporting information may be found in the online version of this article:

Figure S1. IGHV repertoire-Only IGHV gene rearrangements with a frequency over five percent of the whole series are shown (Chi² test: $P = 0.021$ for IGHV1-2 and $P = 0.0095$ for IGHV3-21).

Figure S2. Kaplan Meyer curves for treatment-free survival for normal SPE, hypogammaglobulinemia and monoclonal peak.

Figure S3. Kaplan Meyer curves for treatment-free survival according to IGHV mutation status.

Figure S4. Kaplan Meyer curves for treatment-free survival (A) restricted to patients with non-progressive CLL for 6 months and Binet stage A, according to SPE and IGHV mutation status. (B) Restricted to patients with non-progressive CLL for 6 months, Binet stage A and mutated IGHV status, according to SPE, starting at 2.6 years).

Figure S5. Kaplan Meyer curves for treatment-free survival according to SPE status and Rossi's score (A: very good Rossi's score. B: good Rossi's score. C: Intermediate Rossi's score).

Figure S6. Kaplan Meyer curves for treatment-free survival for CLL patients with mutated IGHV genes, normal SPE and isolated del(13q) versus other patients of this whole series.



3- Discussion

Nous avons étudié 112 patients atteints de LLC dont 49 avec une SPE normale et 63 avec une SPE anormale. Le groupe avec SPE normale regroupait majoritairement des patients en stade Binet A (96 %), tandis qu'un stade Binet plus avancé était plus fréquent en cas de SPE anormale (30 %). Nous avons regardé les patients considérés comme stables, c'est-à-dire n'ayant pas nécessité de traitement dans les 6 mois après le diagnostic : la proportion de patients stables était significativement plus importante en cas de SPE normale (98 % versus 65 %).

Nos résultats montrent que la présence d'une électrophorèse des protéines sériques normale au diagnostic permet de distinguer les patients ayant une TFS significativement supérieure (TFS médiane non atteinte à 13 ans et présence d'un plateau à partir de 5,6 ans et 51 % de TFS cumulative versus une TFS médiane de 2,64 ans en cas de SPE anormale). L'utilisation conjointe du statut mutationnel IGHV permet la détection de patients atteints de LLC ayant une TFS particulièrement longue (médiane de survie non atteinte à 13 ans, avec un plateau à partir de 5,6 ans et 66 % de TFS cumulative, versus des TFS médianes de 3,96 et 3,91 ans respectivement pour les groupes avec SPE normale et statut IGHV non muté d'une part, et SPE anormale et statut IGHV muté d'autre part). Dans le groupe SPE normale et statut IGHV muté, certains marqueurs de mauvais pronostic sont significativement sous-représentés. C'est le cas du réarrangement IGHV3-21 (absent) et de la lymphocytose au diagnostic avec un seuil à 30 G/L.

L'utilisation de la SPE est intéressante car il s'agit d'un examen généralement réalisé au diagnostic de la LLC, de faible coût et facilement réalisable. Malgré l'apparition de nombreux marqueurs de mauvais pronostic et y compris dans l'ère du séquençage haut débit, une électrophorèse des protéines sériques normale (c'est-à-dire immunoglobulines polyclonales sans hypogammaglobulinémie) au diagnostic demeure un marqueur indépendant de bon pronostic, permettant d'identifier des patients avec une LLC évoluant lentement et qui ne nécessiteront probablement jamais de traitement spécifique.



C- Article 2 : « Updates to « COV’COP » to easily detect CNVs in inherited and cancer diseases », manuscrit soumis dans Bioinformatics le 11/06/2018 et refusé le 18/06/2018, en attente de resoumission dans un autre journal

1- Introduction

Le logiciel Cov’Cop a été mis au point par les Drs Paco Derouault et Anne-Sophie Lia (EA6308) afin de mettre en évidence les CNV dans des pathologies héréditaires⁶⁰⁶. Ce logiciel répond à un besoin pour l’interprétation des CNV à partir de données issues de NGS avec bibliothèques préparées à partir de panels d’amplicons. En effet, de nombreux outils existent pour la détection de CNV par NGS mais sont pour la plupart développés pour des bibliothèques préparées par capture et séquencées par technologie Illumina. Les contraintes et stratégies de normalisation sont différentes de celles nécessaires pour notre stratégie NGS (amplicons et technologie Ion Torrent).

Nous avons voulu adapter l’utilisation de Cov’Cop à la détection des CNV dans des pathologies somatiques, en particulier dans la LLC pour laquelle nous devons répondre sur la présence ou non des principales anomalies (del(17p), del(11q), trisomie 12, del(13q)). Nous avons donc collaboré avec l’équipe EA6308 pour mettre au point un panel de NGS pouvant permettre la détection des CNV avec le logiciel Cov’Cop dans le cadre de pathologies lymphoïdes, ainsi qu’un second panel réduit pour la recherche ciblée de mutation et de délétion de *TP53* et d’*ATM*.

Nous avons contribué à l’amélioration du logiciel Cov’Cop avec l’intégration d’une interface graphique et d’aides visuelles à la détection automatique des CNV. Nous avons participé à son adaptation prenant en compte les spécificités relatives aux anomalies somatiques. En effet, les anomalies rencontrées peuvent être sous-clonales avec un pourcentage réduit de cellules porteuses. De plus, certains CNV tels que la del(13q) sont fréquents et peuvent induire un biais dans le calcul des ratios normalisés pour des puces comprenant un nombre limité de patients (7 à 8 échantillons par puce pour le panel lymphoïde étendu). Cette difficulté a nécessité l’application d’une double normalisation (élimination automatique des régions avec CNV dans le processus de calcul des ratios normalisés) et l’utilisation optionnelle d’un échantillon de référence validé pour ne présenter aucune anomalie de nombre (ADN témoin). Enfin, nous avons tenté de déterminer un seuil de sensibilité. Dans ce travail, certains gènes et régions chromosomiques ont été anonymisés pour protéger des données non encore publiées.



*Data and text mining***Updates to “COV’COP” to easily detect CNVs in inherited and cancer diseases**

Paco Derouault¹, Jasmine Chauzeix^{2,3}, David Rizzo^{2,3}, Corinne Magdelaine^{1,4}, Sylvie Bourthomieu^{1,5}, Karine Durand^{6,7}, H el ene Dzugan^{1,4}, Franck Sturtz^{1,4}, St ephane M erillou⁸, Anne-Sophie Lia^{1,4*}

1 Univ. Limoges, MMNP, EA 6309, F-87000 Limoges, France; 2 CHU Limoges, Service H ematologie Biologique, F-87000 Limoges, France; 3 Univ. Limoges, UMR CNRS 7276 CRIBL, F-87000, Limoges, France; 4 CHU Limoges, Service Biochimie et G en etique Mol culaire, F-87000 Limoges, France; 5 CHU Limoges, Service de Cytog en etique, F-87000 Limoges, France; 6 CHU Limoges, Service Anatomopathologie, F-87000 Limoges, France; 7 Univ. Limoges, EA CAPTur, F-87000 Limoges, France; 8 Univ. Limoges, UMR-7252X-LIM, F-87000, France

*To whom correspondence should be addressed.

Associate Editor: XXXXXXXX

Received on XXXXX; revised on XXXXX; accepted on XXXXX

Abstract

Cov’Cop 2 is a new major version of the Cov’Cop software for CNVs detection among amplicons sequencing data. New original features are provided in order to maximize CNVs detection. All the first steps of correction and normalization are identical in Cov’Cop 1 and 2. However, in this new version, three major new implementations are present. Firstly, we added a new CNV calling algorithm, based on the z-score, that automatically defines the deletion and duplication thresholds. Secondly, we optimized Cov’Cop so that it can detect deletion or duplication events from cancer data with a low ratio of cancer cells among healthy cells, especially by using control samples. Finally, we improved the user-friendly graphics interface by developing a new interactive visualization, strongly maximizing results interpretation.

Availability: <https://git.unilim.fr/merilp02/CovCop2/>

Contact: anne-sophie.lia@unilim.fr

Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 Introduction

In order to help molecular geneticists to quickly identify CNVs responsible for inherited diseases among Amplicons Sequencing data generated by NGS, we designed a user-friendly tool “Cov’Cop” (Derouault *et al.* 2017). Using the raw-read-count (RRC) value of each amplicon (provided as a coverage file for each run by the sequencer), it simultaneously analyzes all the patients using a two-stage algorithm containing correction and normalization levels to give a score for each amplicon. Theoretical score of 1 is the normal case while low (<0.5) or high (>1.5) values respectively reveal deletions or duplications.

Here, we propose a new major update to Cov’Cop with a new CNV calling algorithm, based on the z-score, that automatically defines the deletion and duplication thresholds. Also, we optimized Cov’Cop so that

it can detect deletion or duplication events from cancer data with a low ratio of cancer cells among healthy cells, especially by using control samples (data sets described in supplementary data 1). Finally, we improved the user-friendly graphics interface by developing a new interactive visualization, greatly maximizing results interpretation.

2 New features**2.1 CNV detection algorithm****2.1.1 Z-score-based CNV detection : “Z-detection”**

Cov’Cop 1 considered an amplicon deleted or duplicated according to thresholds defined by the user. A new CNV detection algorithm, based



Derouault et al.

on the z-score, has been implemented in Cov'Cop 2. For each amplicon, the z-score is calculated from all the values of a patient. The z-score following a standard normal distribution, we fixed for both deletion and duplication events a threshold corresponding to a significance level of 0.01 with one-tailed test. Thus, a negative z-score with a p-value < 0.01 indicates a deleted amplicon, while a positive z-score with a p-value < 0.01 states a duplicated amplicon. This algorithm automatically determines the best deletion and duplication thresholds based on the variability of a patient's data. We estimated the sensitivity of this new algorithm to be around 91% and the specificity around 96% (see supplementary data 2).

2.1.2 Two-stages ratio to optimize CNVs detection

The last normalization step of Cov'Cop 1 consisted in a ratio of standardized patient values that gives a theoretical value equal to 1 for a gene present in two copies, 0.5 for a deletion event and 1.5 for a duplication. In Cov'Cop 2, we optimized this detection by adding a new step: once the first ratio is calculated and first CNV detection is performed, a second ratio is calculated excluding the amplicons located inside the firstly detected CNVs, and the final CNV detection is realized. This approach is used to improve standardization in regions where a same CNV event is present in many patients (see supplementary data 3).

2.1.3 Merging CNVs

To reduce the impact of false negative amplicons on CNV detection, we added a "merge" option in Cov'Cop 2. If two CNV areas, located on the same chromosome, are disjointed by only one amplicon with a z-score duplicated or deleted at a significance level of 0.05, Cov'Cop 2 will then merge the two CNV areas to point out easily this global CNV (see supplementary data 4).

2.1.4 Reference amplicons selection or exclusion

For the normalization step, Cov'Cop 1 and 2 select a set of amplicons, which are the most stable ones among the patients of a run. These amplicons are used to normalize the values of the other amplicons. In Cov'Cop 2, the user can indicate specific amplicons to use for this normalization step (See supplementary data 5). Inversely, Cov'Cop 2 gives also the opportunity to exclude manually some amplicon data for the last ratio step of normalizations (See supplementary data 5).

2.2 CNV calling from cancer data

2.2.1 Control samples

In Cov'Cop 1, one patient was compared to all the other ones for CNVs calling and users did not have the possibility to choose any control samples as reference. Cov'Cop 2 allows it, if the user wishes to. In such a case, in the last ratio normalization step, for each amplicon, the median of controls samples is calculated. Then for each patient, the amplicon values are divided by the median calculated for the controls. Practically, Cov'Cop 2 can be run with only one control sample but more control samples will improve the result.

2.2.2 Cancer cell proportion

A cancer tissue sample may contain both healthy cells and cancer cells at the same time. This proportion may impact on the detection of CNVs. Simulated and real data demonstrate that Cov'Cop 2 is able to detect

cancer CNVs in samples containing as low as 19% of cancer cells (see supplementary data 7).

2.3 2D interactive visualization

Upgrade in visual presentation of the results is a strong new feature of Cov'Cop 2. We added an interactive 2D visualization for each patient (Figure 1). The amplicons are represented by dots on the top of their chromosomal positions on the x-axis, and the y-axis corresponds to their normalized values. Users can interactively zoom in on specific regions and navigate between data. Several information described below have also been added on the graph.

2.3.1 Local regression curve

To optimize visual CNVs detection, we introduced the possibility to display regression curves on the presented chart. We chose to implement the Loess local regression algorithm (Cleveland, 1979) to visualize easily a sudden change (See supplementary data 6.1).

2.3.2 CUSUM charts

For some data such as cancer or mosaic ones, a sample may contain at the same time "normal" and deleted/duplicated cells. The deletion/duplication detection efficiency depends then on the proportion of deleted/duplicated cells compared to normal cells and the normalized values can be close to 1. CNVs will be then very difficult to detect. Consequently to be able to observe a slight increase or decrease in values, we added the visual method named CUMulative SUMmary control chart (CUSUM) (Page, 1954). (see supplementary data 6.2).

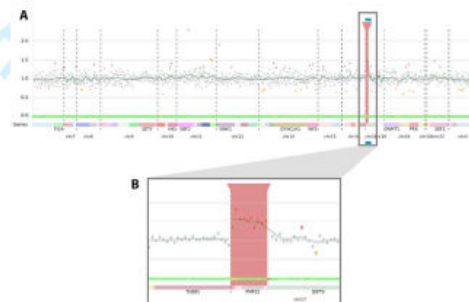


Figure 1: Visualization of Cov'Cop 2. A – General view. Each dot corresponds to an amplicon. The grey color indicates a "normal" value while a red or orange dot states a duplicated and deleted amplicon respectively. The red rectangle corresponds to the duplication of a region. B – Zoom on the duplicated region (PMP22).

3 Conclusion

Cov'Cop 2 offers a wide range of new features to interpret data from amplicon sequencing to detect CNVs. This tool works on data generated from Ion Designer (Life Technologies, CA, USA) but also from the Illumina DesignStudio (Illumina Inc., San Diego, CA, USA). The user-friendly interface associated with our new 2D visualization facilitates data exploration and manipulation allowing to process complex analyses such as those from cancer data. Cov'Cop 2 also offers to export the results in VCF format (Danecek et al, 2011) or graphical output for



Article short title1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

publications. It can also now be used in command line mode to be integrated in various pipelines (see supplementary data 8).

Acknowledgements

This work has been supported by a grant from CHU de Limoges APREL 2013 and 2015, sponsored by ARS Limousin.
Conflict of Interest: none to declared.

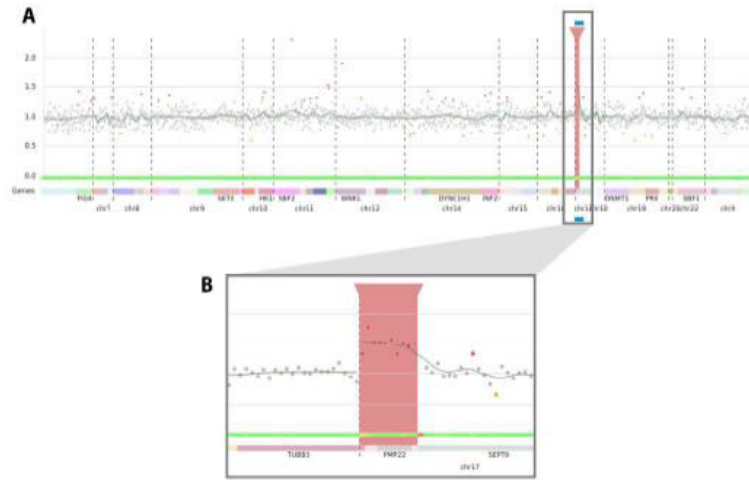
References

- Cleveland, W.S. (1979) Robust Locally Weighted Regression and Smoothing Scatterplots. *Journal of the American Statistical Association*, 74(368), 829-36.
- Danecek, P. et al. (2011) The variant call format and VCFtools. *Bioinformatics*, 27(15), 2156-58.
- Derouault, P. et al. (2017) "COV-COP" Allows to Detect CNVs Responsible for Inherited Diseases among Amplicons Sequencing Data. *Bioinformatics*, 33(10), 1586-88.
- Page, E.S. (1954) Continuous Inspection Schemes. *Biometrika*, 41(1/2), 100-115.

For Peer Review



1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



440x288mm (300 x 300 DPI)

Review



3- Discussion

Dans ce travail, nous avons activement contribué à l'amélioration du logiciel Cov'Cop pour la détection des CNV en génétique somatique des tumeurs. Par rapport aux maladies héréditaires, et à l'exception des cas de mosaïque, la génétique somatique des tumeurs est directement impactée par le degré d'infiltration de l'échantillon biologique par les cellules cancéreuses. De plus, au sein même d'une tumeur, certaines anomalies moléculaires, dont les CNV, peuvent être sous-clonales. Ainsi, la question de la sensibilité de la méthode est capitale même si la pureté de l'échantillon tumoral est élevée.

Les seuils arbitraires du ratio normalisé appliqués pour la détection automatique d'une anomalie (c'est-à-dire $\text{ratio} < 0,8$ ou $> 1,2$) étaient déjà modifiables dans la première version publiée du logiciel. Dans cette version 2, des options basées sur des tests statistiques (et donc sans *a priori*) ont été intégrées pour faciliter et standardiser la détection de ces anomalies et leur interprétation.

Une première amélioration a consisté en l'intégration d'une visualisation graphique des résultats. En effet, dans la première version publiée, le logiciel générait à partir des couvertures obtenues, un ratio normalisé par amplicon. Selon le seuil de détection appliqué, les amplicons sortant des bornes ($< 0,8$ ou $> 1,2$) étaient mis en évidence par surlignage en orange pour une délétion ou rouge pour une amplification (Figure 73 A). Il fallait donc faire défiler les résultats de tous les amplicons (ce qui, dans le cas de notre premier panel, correspond à une table de plus de 1200 lignes et autant de colonnes que d'échantillons) afin de visualiser les anomalies. Dans cette nouvelle version, un graphique est construit automatiquement pour chaque échantillon, représentant en ordonnée le ratio normalisé, et en abscisse les amplicons rangés par chromosome, séparés par des traits pointillés verticaux (Figure 73 B). Les gènes sont mis en évidence par un système de couleur. Il est également possible de sélectionner sur le graphique une zone d'intérêt (gène ou amplicon), permettant de voir directement le ratio normalisé en rapport. Le rendu de résultat consiste alors en un examen visuel rapide de la sortie graphique.



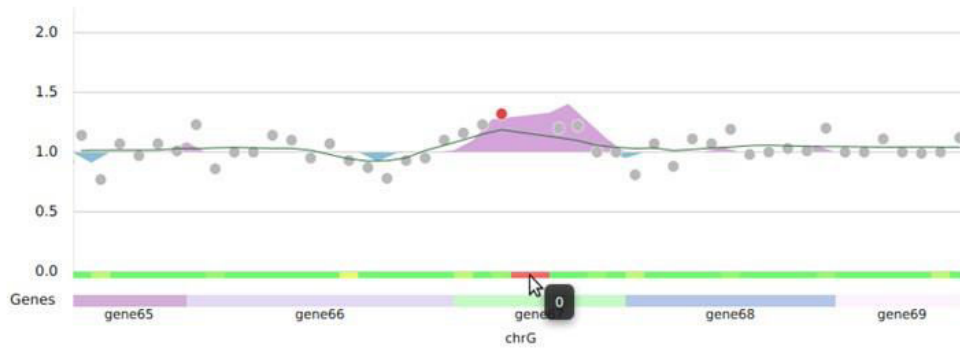


Figure 74 : Visualisation graphique du logiciel Cov'Cop2. Exemple de détection d'amplicons mal amplifiés.

Un système de couleur en bas du graphique permet de visualiser la stabilité des amplicons entre les différents patients de l'analyse (vert pour une bonne stabilité et rouge pour une mauvaise stabilité). Ici, deux amplicons sont éliminés du graphique en raison d'un score à 0, et indiqués en rouge sous le graphique.

Afin de faciliter la visualisation de certains CNV n'affectant qu'une faible proportion de cellules, une courbe de régression Loess a également été intégrée (Figure 75).

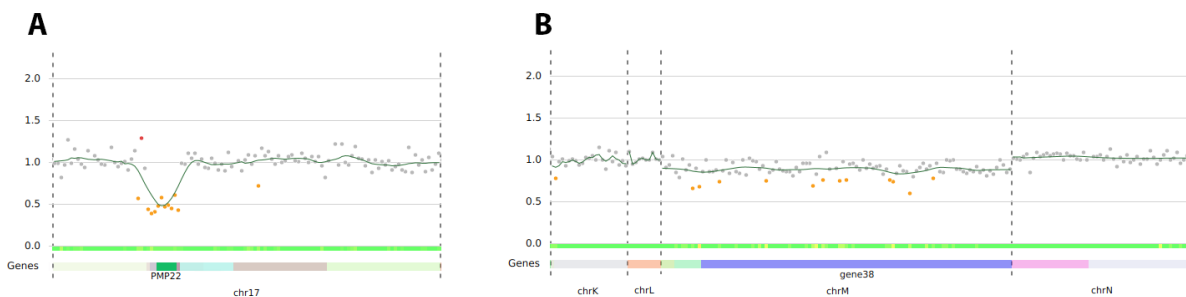


Figure 75 : Exemple de visualisation graphique fournie par le logiciel Cov'Cop2 avec intégration d'une courbe Loess en vert.

A- Exemple de visualisation de la déletion du gène *PMP22* : un pic inverse apparaît sur la courbe.

B- Exemple de déletion d'un chromosome entier dans une faible proportion de cellules (échantillon issu d'une LLC). La courbe Loess est sous la valeur de 1 pour toute la région concernée et facilite la visualisation de cette déletion.



De plus, il est possible d'appliquer l'option « CUSUM (pour Cumulative Summary control) charts » pour faciliter la visualisation de CNV sous-clonaux. Il s'agit d'un algorithme qui calcule pour chaque chromosome la somme cumulative des déviations par rapport à la moyenne pour chaque patient. Les anomalies alors détectées sont surlignées en bleu pour une délétion et en violet pour une duplication (Figure 76).

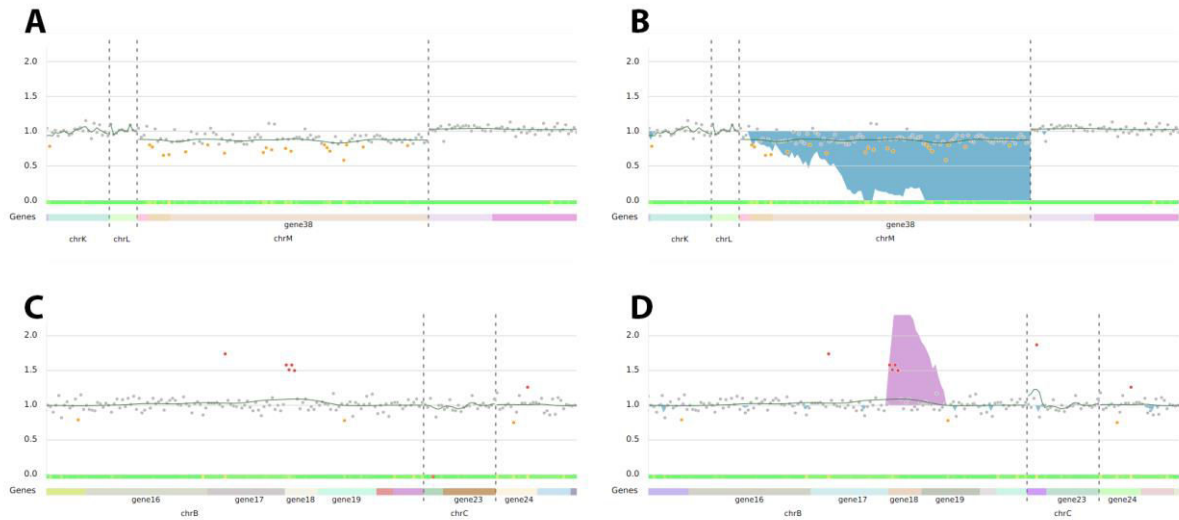


Figure 76 : Visualisation des graphiques obtenus avec le logiciel Cov'Cop2 en appliquant l'option « CUSUM chart ».

A- Délétion du chromosome M, avec un ratio proche de 1, non mise en évidence par le logiciel par visualisation standard.

B- L'option "CUSUM chart" permet d'indiquer la délétion du chromosome M en bleu.

C- Duplication d'un faible nombre d'amplicons dans le gène 18, non mise en évidence par le logiciel par visualisation standard.

D- L'option "CUSUM chart" permet de repérer la duplication du gène 18 en violet.

Une dernière option graphique permet de mettre en évidence la valeur moyenne des ratios normalisés pour chaque gène. Celle-ci est alors représentée par une ligne bleue et peut aider à l'interprétation des CNV (Figure 77).



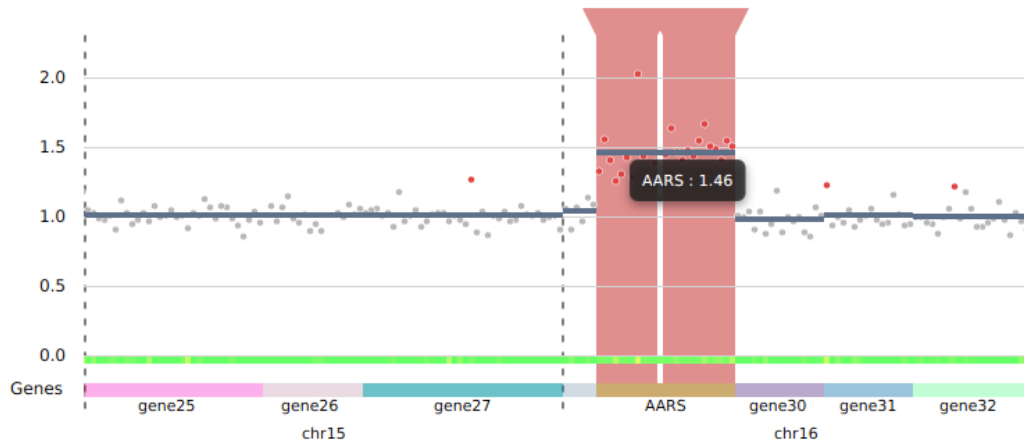


Figure 77 : Représentation graphique avec le logiciel Cov'Cop2 d'une duplication du gène AARS avec l'option de visualisation de la valeur moyenne des ratios normalisés par gène. Dans ce cas, la valeur moyenne pour le gène AARS est de 1,46.

Par ailleurs, la méthode de détection des CNV a été revue dans cette version de Cov'Cop. Initialement, le seuil de détection d'une anomalie était déterminé par l'utilisateur (*cf.* seuils arbitraires). Cependant, selon les amplicons, l'application d'un seuil fixe n'est pas optimale. Un z-score est maintenant calculé pour chaque amplicon ($z\text{-score} = (\text{Nombre de séquences lues normalisé} - \text{moyenne}) / \text{écart type}$). Celui-ci suit une loi normale. Le seuil de significativité a été fixé à une $p\text{-value}=0,01$. Toute valeur sortant de l'intervalle de référence est considérée comme anormale. Il reste toutefois possible d'appliquer un seuil donné arbitraire comme dans la version initiale de Cov'Cop. En effet, cette méthode augmente la sensibilité et diminue la spécificité de la détection d'une anomalie.

D'autre part, certaines anomalies peuvent être fréquentes dans une même puce et entraîner un biais dans le calcul du ratio normalisé. En effet, dans le cas de notre panel permettant le passage simultané de 7 à 8 échantillons par puces, il est arrivé que la moitié des patients soient porteurs d'une $del(13q)$. Dans ce cas, le logiciel rendait parfois une duplication de cette région chez les patients non porteurs de l'anomalie et la sensibilité de détection de la délétion était diminuée (Figure 78). Pour améliorer cela, une nouvelle étape de normalisation a été ajoutée (double normalisation). Celle-ci exclut du calcul les amplicons détectés comme anormaux à l'issue de la première étape de normalisation.



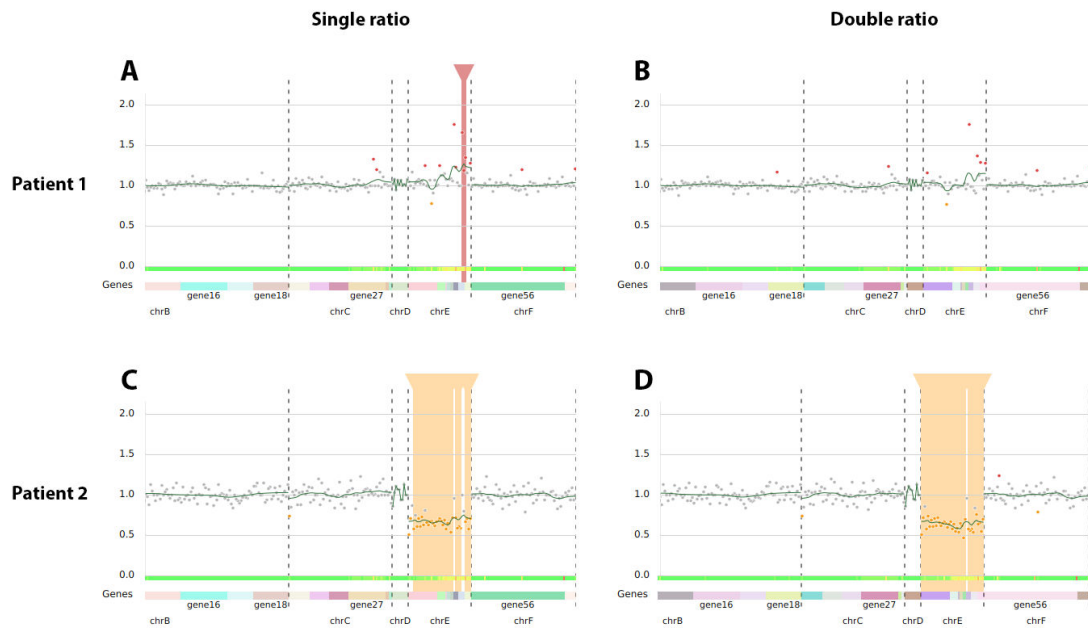


Figure 78 : Visualisation des graphiques obtenus avec CovCop avec en abscisse les régions chromosomiques et en ordonnée les ratios normalisés obtenus.

Cas d'une puce pour laquelle la moitié des patients sont porteurs de la même anomalie (délétion du bras long du chromosome 13, noté E sur la figure). Les régions en jaunes sont détectées comme délétées par le logiciel.

A- Graphique obtenu sans la normalisation en deux étapes pour le patient 1. Le logiciel met en évidence une fausse duplication au niveau du chromosome 13.

B- L'application du ratio en deux étapes améliore la stabilité de la région observée avec l'absence de détection d'anomalie au niveau du chromosome 13 pour le patient 1.

C- Graphique obtenu sans la normalisation en 2 étapes chez le patient 2 : 6 amplicons (sur 39) ne sont pas détectés comme délétés au niveau du chromosome 13. Le logiciel interprète la région comme porteuse de 3 CNV.

D- L'application du ratio en deux étapes permet d'améliorer la détection des amplicons délétés (seuls 3 amplicons sont considérés comme normaux), et un seul amplicon considéré comme normal coupe la région d'intérêt.

De plus, dans certains cas, le ratio normalisé d'un amplicon isolé au sein d'une région délétée peut être non significatif, donnant comme résultat brut issu du logiciel deux délétions chromosomiques proches. Une option permet maintenant de fusionner ces régions si les amplicons non considérés comme délétés ont un z-score significatif au seuil de 0,05 (contre 0,01 initialement) (Figure 79).



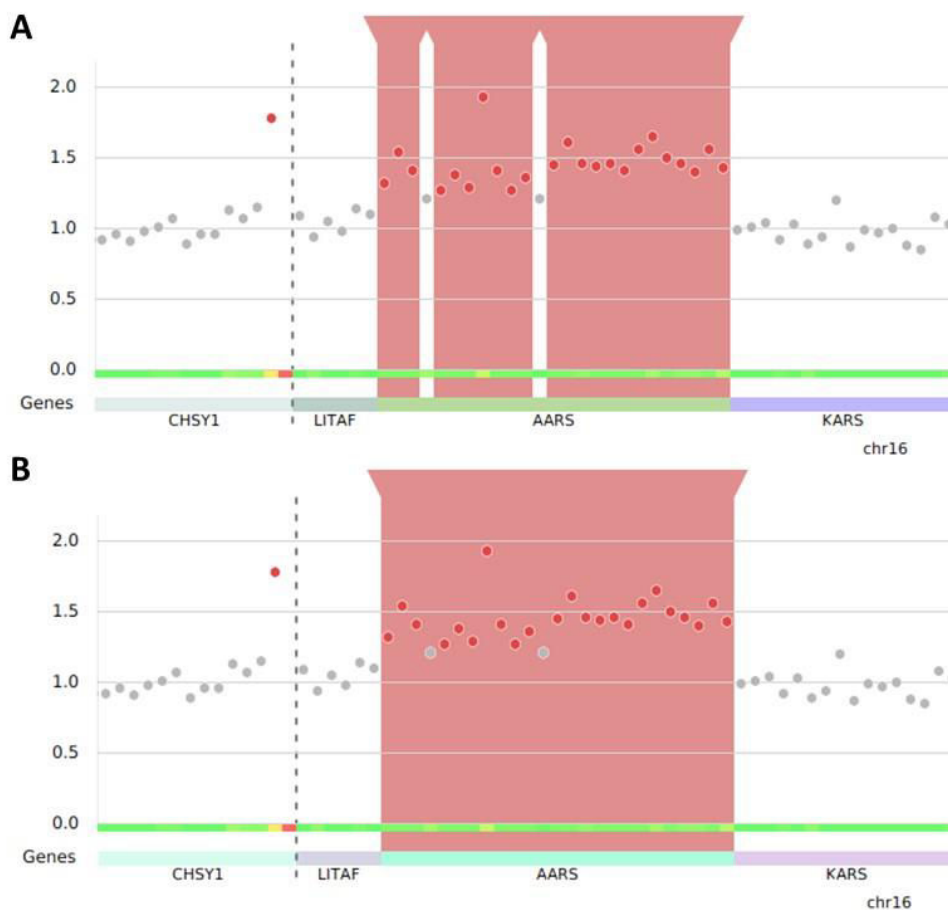


Figure 79 : Visualisation graphique à partir du logiciel Cov'Cop2 de la duplication du gène AARS.
A- Représentation graphique sans l'option permettant de fusionner les amplicons ayant un z-score significatif avec un p compris entre 0,01 et 0,05. Le logiciel détecte trois CNV différents.
B- Représentation graphique avec l'option permettant de fusionner les amplicons ayant un z-score significatif avec un p compris entre 0,01 et 0,05. La duplication détectée inclut le gène AARS en entier. Les amplicons fusionnés par cette méthode sont distinguables par coloration en gris.

Par ailleurs, l'utilisateur peut maintenant sélectionner manuellement des amplicons de référence pour l'analyse, ou à l'inverse exclure certains amplicons.

En application à la problématique des pathologies somatiques, il est possible au moment de lancer l'analyse de déterminer le contexte somatique versus germlinal de détection des CNV. Dans un contexte somatique, la détermination d'un ou plusieurs échantillons contrôles sains, servant alors de référence, est une option intégrée au lancement de l'analyse. L'application d'un seul contrôle sain par puce est possible. L'exemple d'une puce comportant 8 patients dont 5 présentent une del(13q) est présenté en Figure 80 : l'assignation d'un contrôle sain permet de stabiliser la région 13q en évitant la détection de fausse duplication chez les sujets normaux, et en augmentant la sensibilité de détection de cette anomalie.



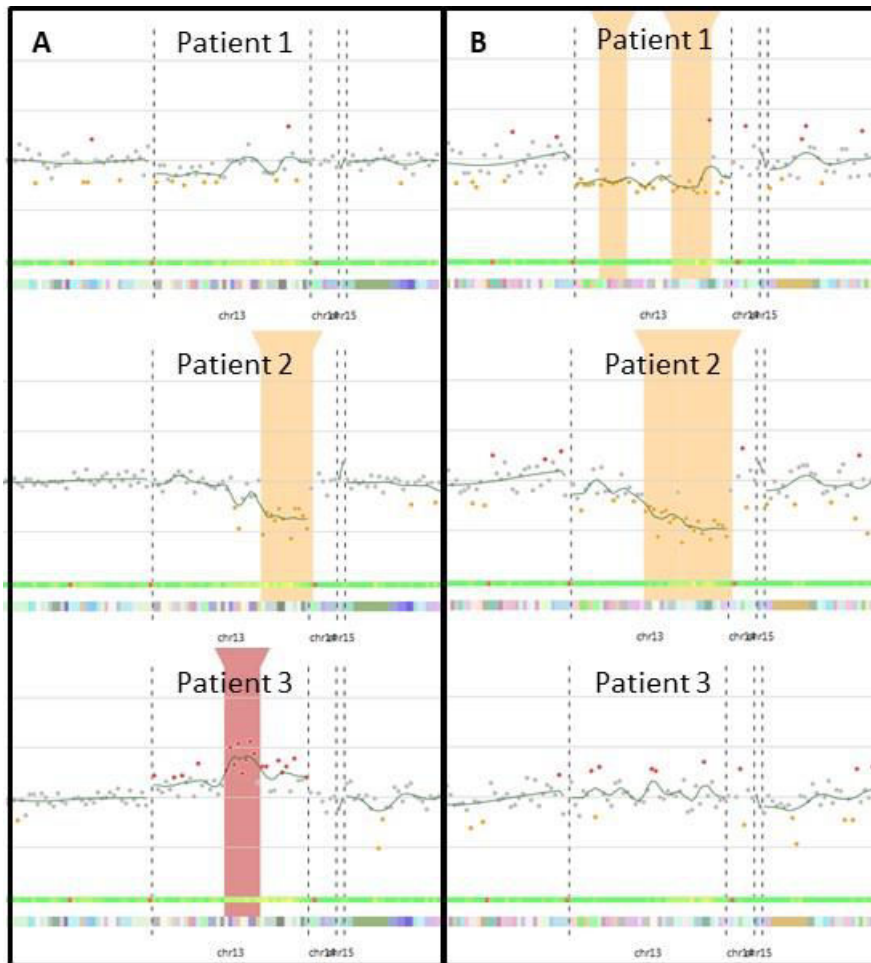


Figure 80 : Visualisation graphique à partir du logiciel Cov'Cop2 du chromosome 13 chez 3 patients. Cas d'une puce comprenant 8 patients dont 5 sont porteurs d'une del(13q).

A- Avant assignation d'un contrôle sain, la del(13q) n'est pas identifiée chez le patient 1, elle est détectée chez le patient 2. Une fausse duplication est visualisée chez le patient 3 ne présentant pas d'anomalie au niveau du locus 13q.

B- Après assignation d'un contrôle sain, la del(13q) est détectée le patient 1. Le locus est stabilisé avec l'absence de fausse duplication identifiée chez le patient 3. La del(13q) du patient 2 est mieux visible avec plus d'amplicons significativement touchés.

Enfin, nous avons tenté de déterminer un seuil d'infiltration minimum par des cellules porteuses d'un CNV. Pour cela, une matrice de simulation a été générée pour la détection d'une délétion complète d'un gène couvert par 80 amplicons (sur un panel de 2394). Celle-ci a été construite à partir des données de couverture obtenues avec une puce contenant 16 patients. Le nombre de lectures en cas de délétion a été calculé de la façon suivante :

$$SRC_i = RRC_i \times (1 - \text{Proportion de cellules cancéreuses}) + \frac{RRC_i}{2} \times \text{Proportion de cellules cancéreuses}$$



avec SRC_i, la valeur simulée de l'amplicon *i*, RRC_i le nombre de lectures de l'amplicon *i* et la proportion de cellules cancéreuses correspondant à la quantité relative de cellules porteuses de l'anomalie (valeur comprise entre 0 et 1, simulé par étapes de 0,1). Ce CNV était alors détectable grâce à l'option « CUSUM charts » pour une infiltration tumorale fixée à 15 % des cellules cancéreuses et apparaît clairement lorsque l'infiltration atteint 20 % (Figure 81). L'utilisation isolée du z-score ne permet la détection de la délétion que lorsque 40 % des cellules sont porteuses de l'anomalie. Enfin, la délétion quasi-complète du gène (concernant 67 amplicons sur 80) détectée par le z-score n'est visible qu'à 60 %.

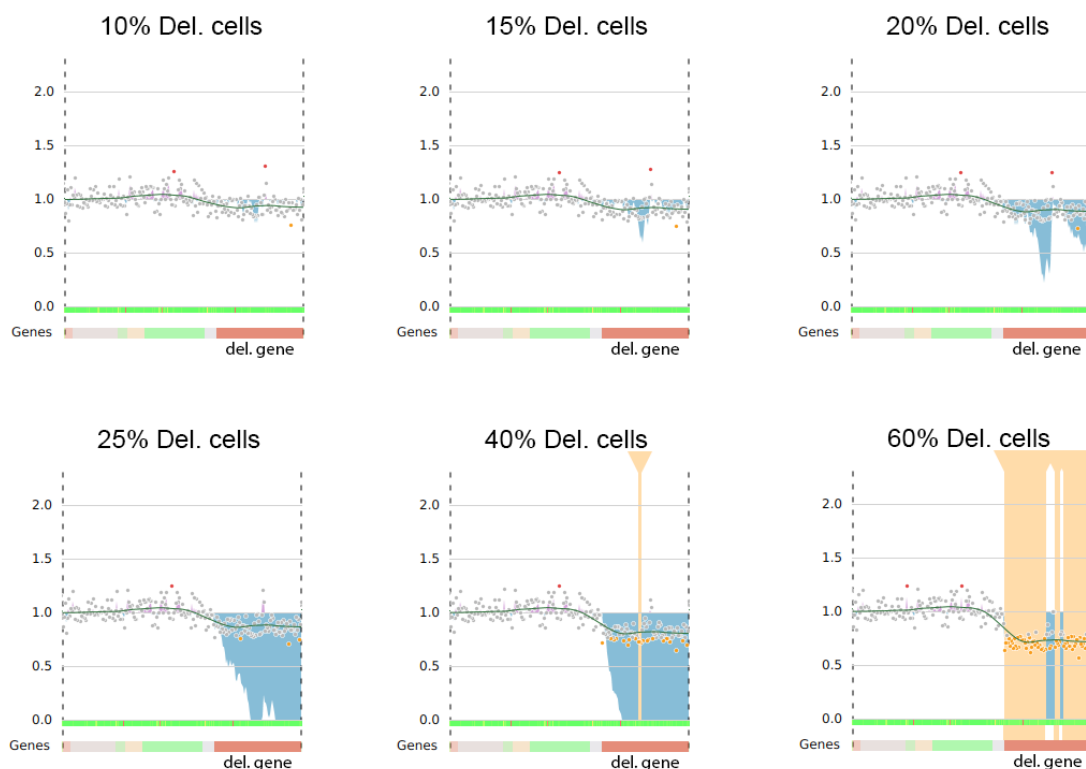


Figure 81 : Visualisation graphique avec le logiciel Cov'Cop2 de la simulation de la délétion d'un gène (visualisé par un trait rouge-orangé sous les graphiques), avec différentes proportions de cellules contenant la délétion.

L'option « CUSUM charts » (surlignage en bleu de la courbe) apparaît dès 15 à 20 % de cellules contenant l'anomalie. La détection grâce au z-score apparaît à partir d'une infiltration à 40 % de cellules cancéreuses.

Ces données ont été confirmées expérimentalement à partir de nos résultats obtenus sur notre série de cas de LLC. Nous avons choisi d'illustrer la sensibilité de la méthode en comparant les profils de 5 patients présentant un clone avec délétion 11q (correspondant à



une délétion d'*ATM*). La proportion du clone avait été précédemment évaluée par caryotype et/ou FISH. La délétion n'est pas détectée lorsque l'anomalie n'est présente que dans 17,5 % des cellules. L'option CUSUM charts permet la détection pour une proportion de 19,5 % (Figure 82).

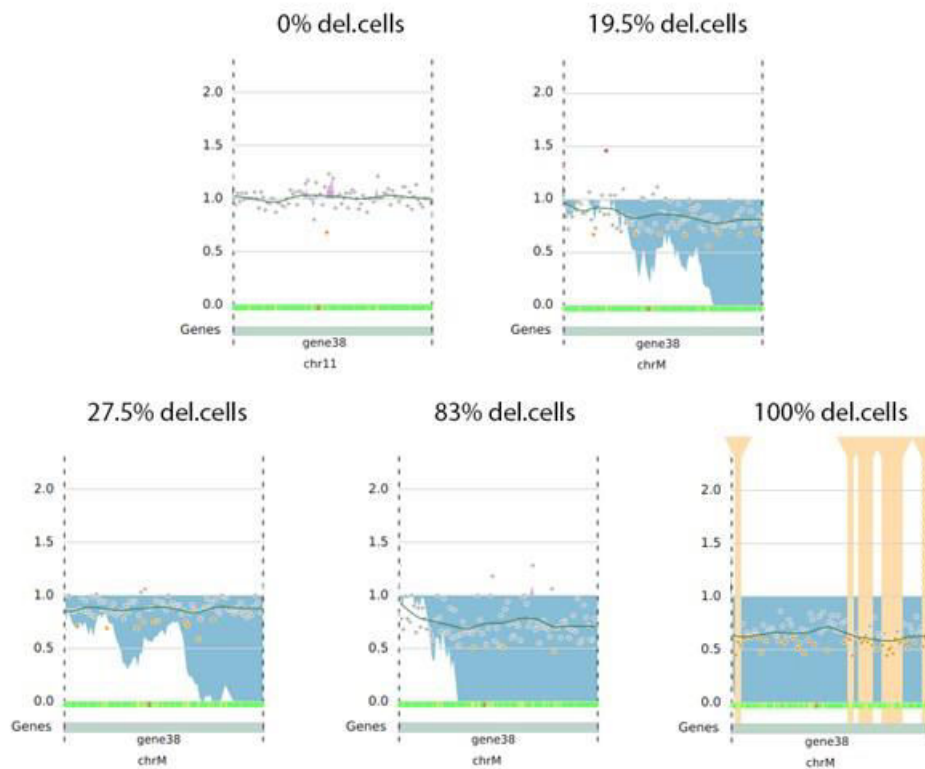


Figure 82 : Visualisation graphique avec le logiciel Cov'Cop2 de la délétion du gène 38 (soit *ATM*) chez différents patients présentant une infiltration clonale variant entre 0 et 100 % (estimation par caryotype et/ou FISH).

L'option « CUSUM charts » permet la détection de l'anomalie pour une proportion de cellules cancéreuses estimée à 19,5 % dans cet exemple.

Dans cet exemple, le seuil de sensibilité de la technique a donc été démontré à 19,5 % pour cette cible. La méthode a été évaluée à plus large échelle sur la série de 82 patients de l'article 3 ainsi qu'au cours de l'étude prospective en routine diagnostique sur les 12 derniers mois, soit environ 64 autres patients. Les résultats montrent que le seuil d'infiltration minimum détectable est dépendant de la qualité des bibliothèques, de l'efficacité d'amplification des amplicons (profondeur globale), du nombre d'amplicons ciblant la région d'intérêt et du nombre d'échantillons. Cette validation de l'approche Cov'Cop dans une série réelle de patients est, entre autre, discutée de façon plus approfondie dans l'article 3.



Au final, cette nouvelle version du logiciel fournit une visualisation rapide des résultats obtenus par patients. Des options d'aide à l'interprétation, en particulier pour les anomalies sous-clonales, ont été intégrées. Enfin, l'utilisation d'un contrôle sain dans la recherche d'anomalies dans un contexte somatique permet de corriger les biais induits par des anomalies fréquentes au sein d'une même puce. Ce logiciel est un outil performant pour la détection des CNV à partir de données de NGS basé sur des librairies par amplicons.



D- Evaluation du logiciel CovCop dans la LLC

1- Introduction

Il existe un grand nombre de facteurs pronostiques dans la leucémie lymphoïde chronique. Il est actuellement indispensable d'étudier avant toute mise en place thérapeutique le statut de *TP53* sur le plan cytogénétique (délétion ou non) et mutationnel. Ceci nécessite la réalisation de plusieurs examens, à savoir un caryotype, de la QMPSF +/- de la FISH en cas d'ininterprétabilité de la QMPSF pour certains loci et du NGS pour la détermination des mutations géniques, en particulier de *TP53*. Afin de simplifier le nombre d'analyses à réaliser et d'optimiser le coût et le délai de rendu des résultats, nous avons voulu tenter de remplacer la QMPSF par le NGS pour la détermination des principales anomalies cytogénétiques de la LLC, à savoir la del(17p), la del(11q), la trisomie 12 et la del(13q).

Pour cela, nous avons collaboré avec l'EA6308 pour mettre au point notre panel de NGS pour la détection de ces anomalies grâce au logiciel Cov'Cop. De plus, nous avons intégré une nouvelle méthode de détection des CNV basée sur l'analyse de la VAF des SNP obtenue par NGS. En effet, en théorie, un SNP hétérozygote est sensé avoir une VAF proche de 50 %. Une déviation de cette valeur peut indiquer une délétion ou une duplication d'un des deux allèles. Nous avons donc mis au point un script utilisable avec le logiciel R permettant la sélection des SNP détectés chez les patients et la construction d'un graphique automatique pour visualisation des résultats.

Nous avons testé ces deux approches (Cov'Cop2 et VAF des SNP) sur une cohorte de 82 patients atteints de LLC pour lesquels nous disposons des données concernant ces anomalies par cytogénétique et/ou QMPSF. Nous avons déterminé les performances de cette stratégie en comparaison aux techniques actuellement utilisées, à savoir le caryotype et la FISH et/ou la QMPSF. Enfin, nous avons mis en évidence des anomalies supplémentaires par NGS, non identifiables par cytogénétique (CNV de petite taille, UPD).

2- Résultats

Nous avons participé activement à l'amélioration du logiciel Cov'Cop2 pour la détection des CNV (voir article 2, page 203). En particulier, une représentation graphique des résultats générée automatiquement a été intégrée dans cette nouvelle version. Un exemple de visualisation de trisomie 12 est présenté dans la Figure 83-A. Afin d'améliorer les performances analytiques de la détection des CNV par NGS, nous avons également mis au point un algorithme (script R) pour la détection de SNP identifiés par NGS et permettant également la génération automatique d'un graphique représentant les résultats de la VAF en



ordonnée pour chaque SNP, ces-derniers étant regroupés par amplicons, dans l'ordre chromosomique en abscisse. Un exemple de trisomie 12 est présenté dans la Figure 83-B.

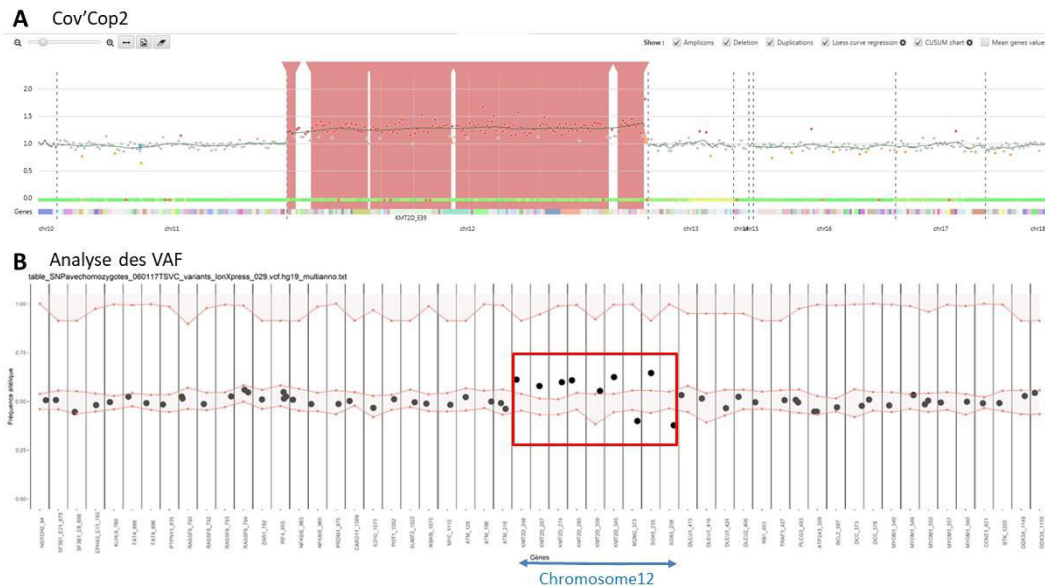


Figure 83 : Visualisation d'une trisomie 12 par le logiciel Cov'Cop2 et l'analyse des VAF.

A- La trisomie 12 est identifiée par surlignage en rouge. La courbe Loess en vert montre un ratio normalisé augmenté pour l'ensemble des amplicons localisés sur ce chromosome, bien que certains ne soient pas significatifs au seuil choisi (1,15).

B- Le graphique construit sous R représente en abscisse les différents amplicons pour lesquels un SNP a été identifié, rangés du chromosome 1 au chromosome X, et en ordonnée la VAF obtenue par NGS. Les points noirs signifient les SNP identifiés par le script. Les régions gris-rouge au milieu et au-dessus du graphique indiquent les zones de valeurs non significativement différentes de la normale pour des SNP homozygotes (en haut du graphique) ou hétérozygotes (au milieu). Le rectangle rouge encadre les SNP pour lesquels la VAF est significativement anormale. Ceux-ci sont situés au niveau du chromosome 12.

Nous avons ainsi analysé les données de 82 patients atteints de LLC passés en NGS sur notre panel de 70 gènes. Nous avons comparé les résultats obtenus pour les 4 anomalies cytogénétiques principales de la LLC (c'est-à-dire la del(13q), la trisomie 12, la del(11q) et la del(17p)) entre le NGS seul (approches Cov'Cop2 et/ou VAF) et les techniques conventionnelles (cytogénétique et/ou QMPSF). Les résultats sont présentés Figure 84 pour chacune des 4 anomalies cytogénétiques.



A- NGS : CovCop + VAF

1		del(13q)		2		Trisomie 12	
		NGS				NGS	
		positive	negative			positive	negative
Cytogénétique +/-QMPFSF	positive	38	6	Cytogénétique +/-QMPFSF	positive	9	0
	negative	3	35		negative	0	73
3		del(11q)		4		del(17p)	
		NGS				NGS	
		positive	negative			positive	negative
Cytogénétique +/-QMPFSF	positive	12	5	Cytogénétique +/-QMPFSF	positive	5	4
	negative	3	62		negative	3	70

B- NGS : CovCop seul

1		del(13q)		2		Trisomie 12	
		NGS				NGS	
		positive	negative			positive	negative
Cytogénétique +/-QMPFSF	positive	35	9	Cytogénétique +/-QMPFSF	positive	9	0
	negative	3	35		negative	0	73
3		del(11q)		4		del(17p)	
		NGS				NGS	
		positive	negative			positive	negative
Cytogénétique +/-QMPFSF	positive	12	5	Cytogénétique +/-QMPFSF	positive	4	5
	negative	0	65		negative	0	73

Figure 84 : Tables de contingence entre l'approche par NGS (Cov'Cop2 et/ou VAF) et techniques conventionnelles.

A- Résultats des CNV sur la base des données NGS avec les 2 approches (Cov'Cop2 et VAF)

B- Résultats des CNV obtenus par NGS avec le logiciel Cov'Cop2 uniquement

Résultats donnés pour chacune des 4 principales anomalies cytogénétiques de la LLC : la del(13q) (1), la trisomie 12 (2), la del(11q) (3) et la del(17p) (4).

Pour la del(13q), l'approche par NGS entraîne 6 faux négatifs et 9 avec l'approche Cov'Cop seule (Figure 84-1). Les 3 cas non détectés par Cov'Cop2 sont issus de puces enrichies en patients porteurs de del(13q) (n=4 sur les 2 puces concernées) et en l'absence de contrôle sain, pouvant expliquer un manque de sensibilité de l'approche Cov'Cop2. La VAF a alors permis de détecter ces patients. L'approche VAF peut toutefois être en défaut en l'absence de SNP sur la région d'intérêt explorée. Dans le cas de la del(13q), un seul patient parmi les 82 étudiés était non contributif par absence de SNP sur cette région. Enfin, deux cas sont détectés comme positifs pour la del(13q) (délétion de type 1, c'est-à-dire *RB1* conservé), avec le logiciel Cov'Cop2 et l'approche VAF, alors que le caryotype est normal. Pour l'un de ces deux cas, la FISH a été réalisée et est négative, tandis que la QMPFSF est également positive (réalisée sur un prélèvement antérieur). Il s'agit donc d'une délétion de petite taille, non visible par FISH. Dans certains de ces cas, nous disposons des résultats par caryotype et/ou FISH, nous permettant une quantification relative de la présence de l'anomalie parmi les cellules de



l'échantillon. Les résultats montrent un proportion de cellules porteuses de la del(13q) inférieure à 30 % (Figure 85-A), la fréquence la plus élevée étant à 24 %.

La concordance entre les différentes techniques utilisées est parfaite pour la trisomie 12 (Figure 84-2). Les cas de notre série présentent un proportion de cellules porteuses de l'anomalie supérieure ou égale à 25 % (Figure 85-B).

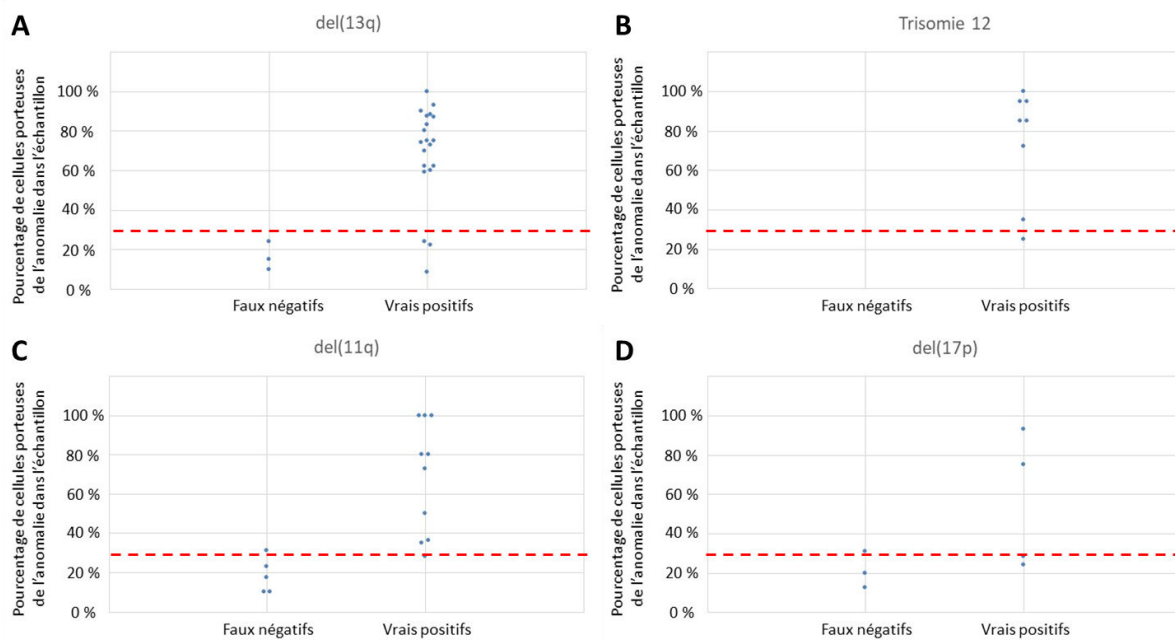


Figure 85 : Graphiques représentant par anomalie cytogénétique les faux négatifs et vrais positifs par technique NGS par rapport à la technique conventionnelle (cytogénétique +/- QMPSF) selon le pourcentage de cellules porteuses de l'anomalie dans l'échantillon.

Les graphiques sont construits pour chaque anomalie cytogénétique principale de la LLC : la del(13q) (A), la trisomie 12 (B), la del(11q) (C) et la del(17p) (D).

Le trait pointillé rouge indique le seuil de 30 % de cellules porteuses de l'anomalie.

Concernant la del(11q), l'approche par NGS entraîne 5 faux négatifs (Figure 84-3). Parmi ceux-ci, la fréquence de l'anomalie la plus élevée est de 31 % (Figure 85-C). La fréquence de cellules tumorales la plus faible permettant de détecter une del(11q) dans notre série est de 28 % par NGS. La technique par NGS détecte également trois cas de del(11q) qui n'avaient pas été détectés initialement par cytogénétique. Dans un premier cas, la détection par l'approche VAF est un faux positif. En effet, le SNP identifié est à la limite du seuil de détection par rapport aux valeurs d'un SNP homozygote (Figure 86).



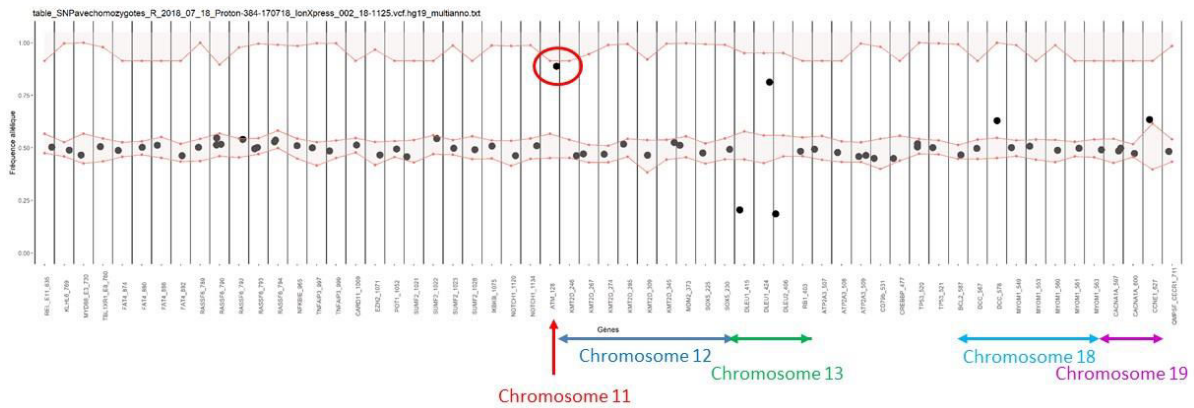


Figure 86 : Visualisation graphique des VAF des SNP détectés par NGS : cas d'un faux positif pour la del(11q).

Un seul SNP est retrouvé au niveau d'*ATM* (entouré en rouge). Sa VAF est proche du seuil de normalité d'un SNP homozygote. On observe par ailleurs une del(13q), une anomalie au niveau d'un SNP sur le chromosome 18 et un SNP dont la VAF est à la limite du seuil de normalité pour le chromosome 19. Le caryotype ne montre pas d'anomalie des chromosomes 18 et 19 pour ce patient.

Pour les deux autres cas, la QMPSF avait été réalisée et était négative. Nous avons fait une FISH du locus *ATM* pour l'un des deux patients qui ne détecte pas d'anomalie. Ces cas ne sont détectés qu'avec les VAF. Il s'agit probablement de cas d'UPD, c'est-à-dire que le locus est bien présent en deux exemplaires, mais qu'il existe pour une proportion des cellules une perte d'hétérozygotie par perte d'un allèle et duplication du second (perte d'hétérozygotie à nombre de copies neutre). Ces patients sont par ailleurs porteurs de mutations d'*ATM* (2 mutations avec charge allélique à environ 39 % pour l'un et 1 mutation avec charge allélique à 10 % pour l'autre).

Enfin, pour la del(17p), l'approche par NGS entraîne 4 faux négatifs pour des cas ayant une charge allélique inférieure ou égale à 31 % (Figure 84-4 et Figure 85-D). Un cas n'est détecté qu'avec l'approche VAF. La del(17p) est détectée dans certains cas jusqu'à des proportions de cellules porteuses de l'anomalie de 10 % pour deux cas de notre série. Enfin, trois cas sont détectés comme anormaux par NGS, alors que le caryotype, la FISH et la QMPSF sont normaux pour le locus *TP53*. L'anomalie est détectée par l'approche VAF. Il s'agit probablement de cas d'UPD. Un exemple est présenté Figure 87. Parmi ces trois patients, deux sont mutés pour *TP53* : l'un présente 3 mutations différentes de l'exon 8 avec des charges alléliques à 31 % pour l'une et 3 % pour les 2 autres. L'autre présente une mutation de l'exon 5 avec une charge allélique à 88 %. Dans ce dernier cas, l'UPD entraîne une duplication de l'allèle muté.



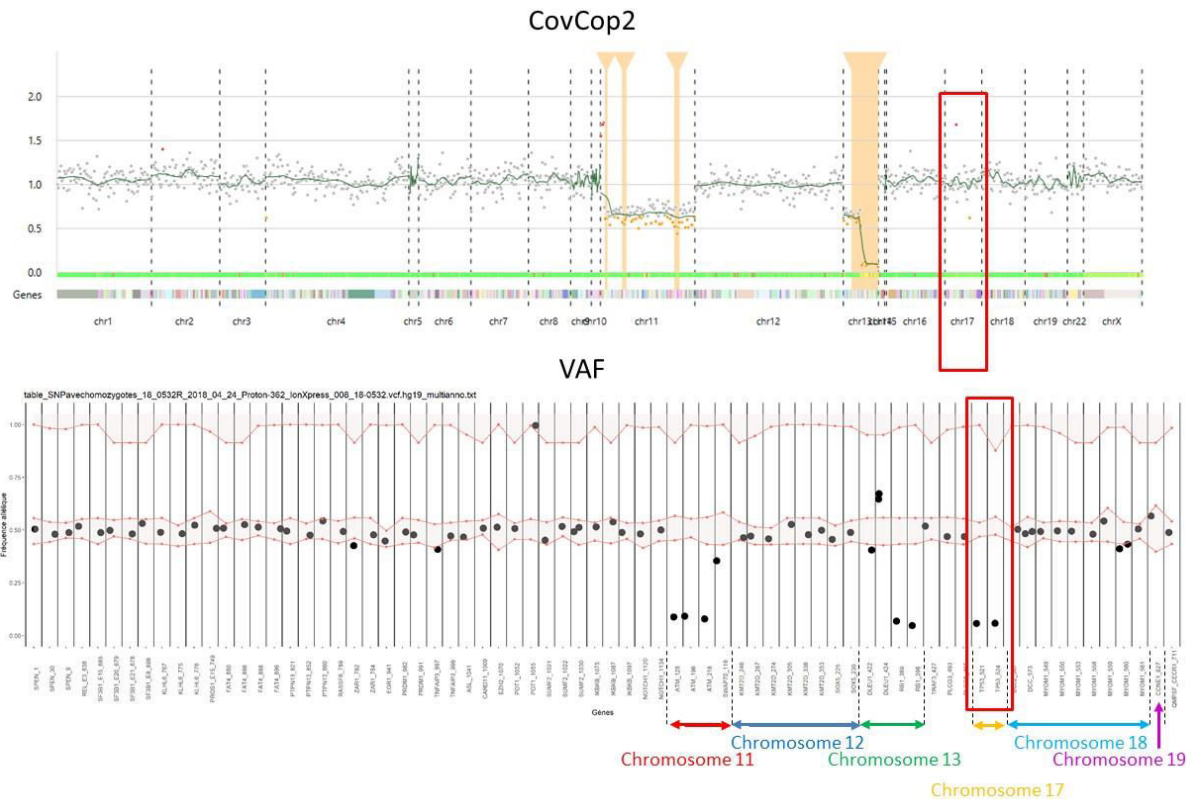


Figure 87 : Visualisation des graphiques des résultats avec le logiciel CovCop2 et les VAF : exemple d'un cas de disomie uniparentale touchant *TP53*.
Le rectangle rouge entoure la région couvrant le chromosome 17. Aucune anomalie n'est visible avec le logiciel Cov'Cop2 (en haut). Les VAF sont anormales (en bas). On observe également une del(11q) et une del(13q) avec les deux approches. La del(13q) emporte *RB1* et *DLEU2*. Elle est homozygote pour *DLEU2* et hétérozygote pour *RB1*.

Au total, pour ces 4 anomalies, le seuil de sensibilité de la technique de détection des CNV par NGS est estimée autour de 30 %. L'approche VAF permet de détecter quelques anomalies non détectées par Cov'Cop2 (un cas de del(17p) notamment). Cette approche permet par ailleurs de détecter d'autres anomalies non identifiables par cytogénétique ou QMPSF. C'est le cas d'UPD identifiées pour *TP53*, mais aussi pour d'autres gènes tels que *EGR1* ou *EZH2*, bien que leur impact pronostique ne soit pas connu actuellement. De plus, notre panel nous a permis d'identifier d'autres anomalies cytogénétiques d'intérêt pronostique (amplification de *MYC*, délétion 6q...) ou d'impact non connu (délétion hétérozygote de *POT1* avec mutation de l'autre allèle chez un patient ayant un caryotype rendu normal).

Nous avons ensuite étudié l'impact de l'évaluation par approche NGS seule (Cov'Cop2 et VAF) sur le pronostic des patients. Pour cela, nous avons classé les différents cas selon le score de Rossi à partir des données obtenues par techniques conventionnelles d'une part, et par NGS seul d'autre part (Figure 88). La présence de faux négatifs pour la del(13q), la del(17p) et la



del(11q) par NGS entraîne des reclassements pour 11 patients, dont 6 cas étant reclassés vers des catégories de risque inférieur. Ceci est principalement lié à la mauvaise détection de la del(13q), faisant passer 3 patients du groupe de très bon risque vers le groupe de bon risque. Deux cas basculent vers le groupe de risque intermédiaire, l'un étant un faux positif pour la del(11q) par l'approche VAF, et l'autre ayant une probable UPD au niveau d'ATM. La détection d'une probable UPD au niveau de TP53 par la méthode VAF fait passer un cas du groupe de très bon risque vers le groupe de haut risque. Deux del(17p) ne sont pas détectées par NGS et ce qui entraîne le reclassement de ces cas vers le groupe de risque très faible. Enfin, trois del(11q) sont non identifiées par NGS et sont à l'origine du reclassement vers le groupe de risque faible (2 cas) ou de très faible (1 cas). Au total, on observe principalement une augmentation du nombre de patients dans le groupe de risque faible par non détection d'anomalies, et en particulier de la del(13q).

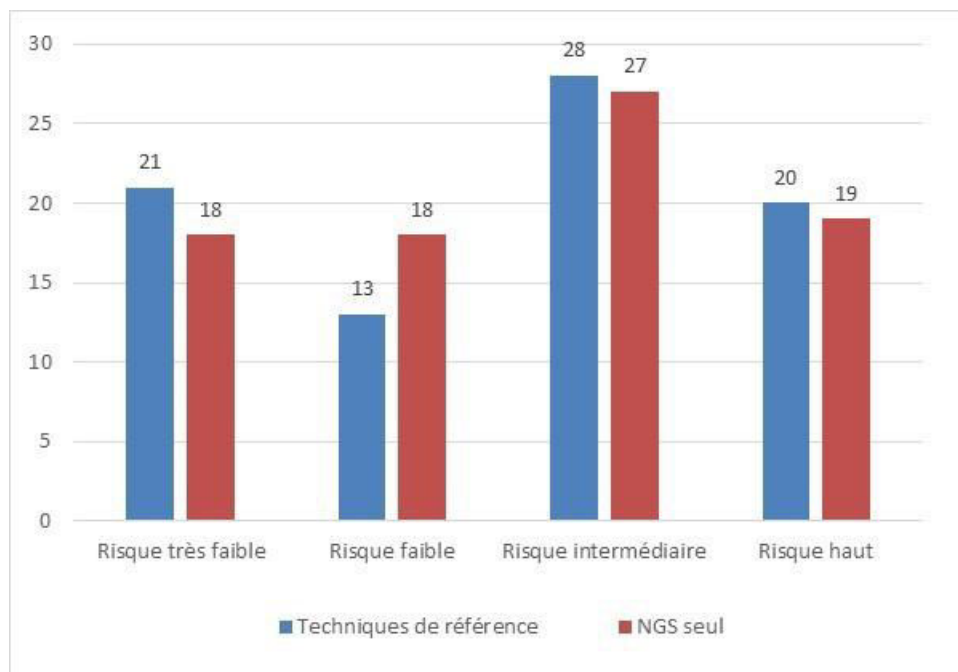


Figure 88 : Classement des patients selon le score de Rossi à partir de la cytogénétique (en bleu) ou du NGS seul (en rouge) pour la caractérisation de la del(13q), la trisomie 12, la del(11q) et la del(17p).

A partir de ces données, nous avons regardé l'impact de ces reclassements sur la TFS en fonction du score de Rossi. Nous avons alors sélectionné les patients pour lesquels l'échantillon analysé était prélevé avant tout traitement pour la LLC (soit n = 50). Dans un premier temps nous avons construit des courbes de survie (analyse de Kaplan Meyer pour la TFS) selon les 4 catégories de risque. Le groupe de haut risque comprend 7 patients, quelles que soient les techniques utilisées pour la détermination des CNV. Les résultats obtenus pour



cette catégorie sont aberrants et ne sont pas interprétables compte-tenu du faible nombre de patients. Nous les avons donc exclus de l'analyse en groupes séparés (Figure 89-A). Nous avons ensuite regroupé les catégories de risque faible et très faible d'une part, et de risque intermédiaire et haut d'autre part (Figure 89-B). Les résultats sont non significatifs après regroupement quelle que soit la méthode utilisée pour la détection des CNV, bien que les résultats soient meilleurs (p-value plus faible) en cas de classement avec les techniques conventionnelles.

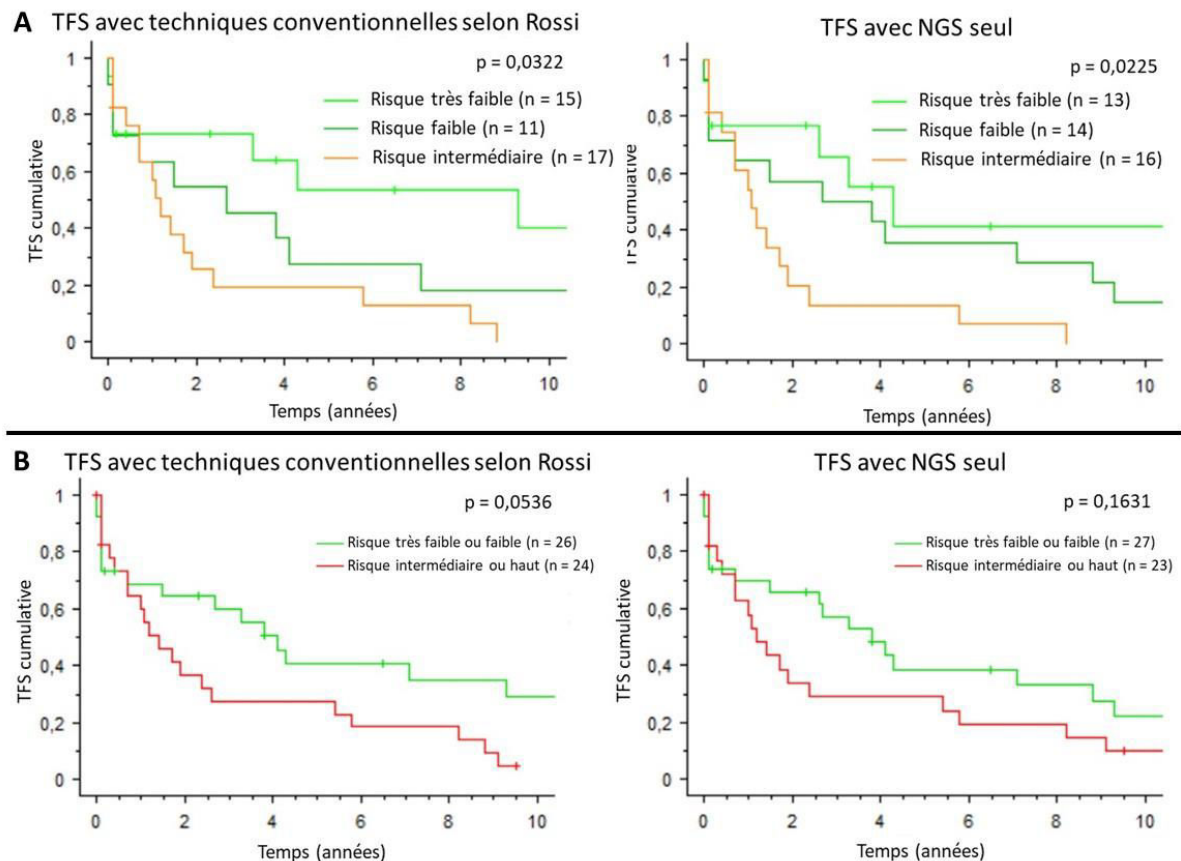


Figure 89 : Courbes de survie (TFS) selon le score de Rossi avec classement pour les anomalies cytogénétiques tenant compte des techniques conventionnelles ou du NGS seul.

A- Classement en 4 groupes pronostiques selon le score de Rossi, en tenant compte des résultats des techniques conventionnelles (cytogénétique +/- QMPSF) (à gauche) ou du NGS seul (à droite). Le groupe de risque haut ne comporte que 7 patients et n'est pas contributif. Il a donc été retiré de l'analyse.

B- Classement en 2 groupes pronostiques selon le score de Rossi, avec regroupement des groupes de risque faible et très faible d'une part, et de risque intermédiaire et haut d'autre part. Les anomalies cytogénétiques sont déterminées selon les résultats des techniques conventionnelles (cytogénétique +/- QMPSF) (à gauche) ou du NGS seul (à droite).



3- Courte discussion

Dans ce travail, nous avons étudié les performances analytiques du NGS pour la détection des CNV. Nous avons développé un nouvel outil de détection des CNV par NGS basé sur la VAF des SNP. Nous avons combiné l'utilisation cette approche avec le logiciel Cov'Cop2 pour détecter les CNV dans une cohorte de 82 patients atteints de LLC, et nous avons comparé ces résultats à ceux obtenus par technique de référence, à savoir la cytogénétique (caryotype + FISH).

L'approche par NGS entraîne de faux négatifs pour la del(17p), la del(11q) et la del(13q). Le seuil de sensibilité est estimé à 30 %. Ce seuil est variable selon les anomalies et selon les puces analysées. Par exemple, pour la del(13q), certaines anomalies vont parfois être détectées à des pourcentages de cellules porteuses de l'anomalie de 8,5 %, alors que dans certains cas un pourcentage à 24 % n'est pas suffisant. Le seuil de détection dépend du nombre d'amplicons au niveau de la région d'intérêt et de leur efficacité d'amplification. De plus, malgré l'apport des nouvelles options d'aide à la détection des CNV du logiciel Cov'Cop2, et en particulier l'assignation d'un contrôle sain, certaines anomalies fréquentes telles que la del(13q) peuvent être ininterprétables si la puce contient trop de patients porteurs de l'anomalie. L'analyse est également tributaire de la qualité des librairies, qui, lorsqu'elle est mauvaise, entraîne une grande dispersion des ratios normalisés autour de 1, et rendant ininterprétable la détection d'anomalies par le logiciel Cov'Cop2 (Figure 90).

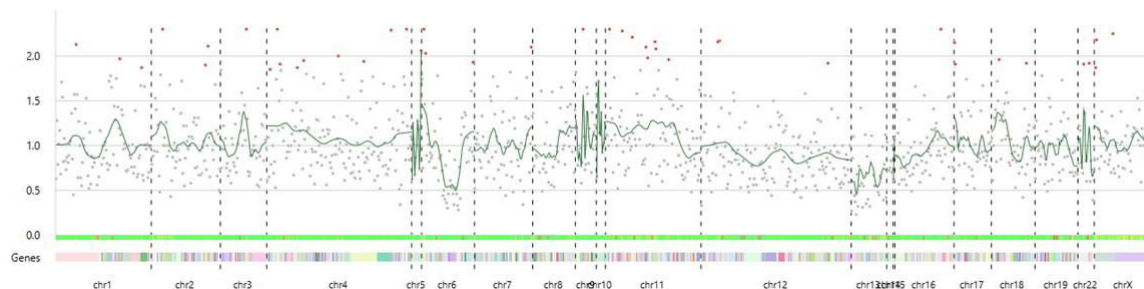


Figure 90 : Visualisation avec le logiciel Cov'Cop2 d'un cas ininterprétable avec grande dispersion des ratios normalisés.

Ce manque de sensibilité peut poser problème pour la détection de la del(17p). En effet, les recommandations actuelles requièrent un seuil de détection à 10 % pour les anomalies de *TP53*³⁵³. Notre ancienne méthodologie reposait sur un rendu des 4 anomalies cytogénétiques principales basé sur la QMPFS. Cette dernière technique a un seuil de sensibilité de 25 %, ce qui est également insuffisant actuellement⁶²⁶. L'approche par NGS est donc valide pour remplacer la QMPFS, mais n'est pas suffisante pour exclure une del(17p) au seuil de 10 %, ceci



étant décisionnel pour la mise en place thérapeutique. La FISH reste alors la technique de référence avec un seuil de sensibilité estimé à 8 % dans notre laboratoire.

Le NGS permet d'identifier les deux types de del(13q), c'est-à-dire entraînant la perte de *RB1* (type 2) ou non (type 1)³⁷¹. Ceci peut être intéressant dans la mesure où le pronostic n'est alors pas le même^{370,371}. Il permet également de détecter des délétions de petite taille, parfois non visibles par FISH.

Par ailleurs, l'analyse par NGS détecte dans certains cas une anomalie avec l'approche VAF uniquement, le logiciel Cov'Cop rendant un résultat normal. Il peut alors s'agir d'UPD, sans délétion de *TP53*, et dont la valeur pronostique n'est pas connue. Ces anomalies sont classiquement détectées par technique de SNP-array. Elles ont déjà été identifiées dans la LLC⁶²⁷⁻⁶²⁹. Il a été décrit des cas d'UPD touchant un allèle muté pour *TP53*, ce que entraînerait alors sa duplication et un pronostic semblable à la double inactivation du gène (double mutation ou mutation et délétion)⁶²⁸. Ce type d'anomalie serait plus fréquent en cas de LLC de stade avancé⁶²⁸. Dans notre série, un patient présente une UPD au niveau de *TP53* détectée par VAF avec duplication de l'allèle muté et une charge allélique de la mutation à 88 %. Il s'agit d'une LLC ré-évolutive 4 ans après traitement par FCR. Des UPD touchant le chromosome 13 en présence de del(13q), en particulier homozygotes, sont décrites dans plusieurs articles^{627,629,630}. Leur impact n'est actuellement pas connu. Certains auteurs rapportent des cas impliquant *ATM*⁶²⁹. Des UPD d'autres loci sont également décrites. Par exemple, une équipe rapporte un cas où *BCL2* est touché, cette anomalie étant acquise au cours de l'évolution⁶³¹. Par ailleurs, une étude retrouve une association entre l'augmentation du nombre d'UPD et des télomères raccourcis⁶³². Un raccourcissement des télomères est susceptible d'entraîner une instabilité chromosomique globale, responsable d'erreurs génétiques par défaut de réparation de l'ADN et aboutissant à des UPD.

Notre technique de détection des CNV par NGS permet la détection d'UPD, ce qui semble être une donnée intéressante. La présence de ces anomalies pourrait être étudiée à l'avenir comme marqueur pronostique, en tant que marqueur global, quel que soit le locus touché, ou en fonction du locus impliqué.

En pratique, actuellement, seul le résultat de la del(17p) a un impact sur la prise en charge thérapeutique des patients. Le NGS est réalisé systématiquement pour la recherche de mutation de *TP53* avant toute mise en place de traitement. La détection d'une mutation de ce gène est alors suffisante pour la prise de décision thérapeutique. En revanche, pour les cas non mutés et en l'absence de détection de del(17p) par NGS ou la détection avec l'approche VAF seule (pouvant indiquer une UPD), la cytogénétique reste nécessaire pour le rendu de résultat.

En conclusion, notre méthode de détection des CNV par NGS permet de remplacer la QMPSF, et de ne réaliser qu'un nombre limité de FISH (cas *TP53* non muté et pour lesquels la del(17p) n'est pas mise en évidence par Cov'Cop2). Elle permet en plus de détecter des CNV de petite taille ainsi que des anomalies potentiellement intéressantes telles que les UPD.



Chapitre 6 : Discussion générale et perspectives

Dans ce travail, nous avons tenté de caractériser l'impact des anomalies moléculaires dans l'histoire naturelle de la LLC. Un aspect fondamental de cette thèse concerne la recherche de nouveaux mécanismes oncogéniques via l'étude de trois remaniements impliquant les gènes des immunoglobulines dans la LLC/lymphome lymphocytaire. Une partie plus appliquée aborde le rôle pronostique des immunoglobulines par l'étude de l'électrophorèse des protéines sériques dans la LLC. Pour cela, nous avons développé et amélioré des outils permettant la caractérisation des anomalies de nombre par NGS.

A- Remaniements atypiques impliquant les gènes des immunoglobulines dans la LLC/lymphome lymphocytaire

Nous nous sommes concentrés sur trois cas de LLC/lymphome lymphocytaire présentant en cytogénétique des translocations originales impliquant les gènes des immunoglobulines et un locus inconnu. Notre étude a permis d'identifier pour chaque dossier les points de cassures et les gènes partenaires. Ceci nous permet de formuler des hypothèses concernant le mécanisme de survenue de ces translocations, ainsi que sur leur potentiel impact dans la pathogénèse.

1- Mécanisme de survenue des remaniements impliquant les gènes des immunoglobulines

a- Mécanismes au niveau du locus des immunoglobulines

Chacun des trois remaniements identifiés possède des caractéristiques propres et semblent résulter de mécanismes de survenue distincts.

Chez le patient 1, un remaniement entre les chromosomes 14 et 17, impliquant *IGH* avait été mis en évidence. Les points de cassures identifiés dans ce travail impliquent *IGHJ4* et *IGHD4-17*, et sont situés au niveau de RSS. Le mécanisme de survenue du remaniement est très probablement lié à l'activité des enzymes RAG au cours du réarrangement VDJ au niveau du locus *IGH*. Un autre argument en faveur de ce mécanisme est l'insertion de nucléotides pouvant mimer des N-nucléotides au niveau des jonctions. On note dans le cas de la jonction



IGHJ4-chromosome 17 un possible P-nucléotide : on retrouve à ce niveau l'insertion de nucléotides au hasard (chromosome 14-TTTACGATC-chromosome 17), et la première thymine insérée peut correspondre à un P-nucléotide dans la mesure où le nucléotide adjacent tronqué est une adénine.

Dans le cas du patient 2, une translocation entre les chromosomes 17 et 22, impliquant *IGL* avait été identifiée. Le point de cassure est situé au niveau des régions variables d'*IGL*, et touche un gène non fonctionnel (*IGLV3-6*), à distance du RSS. Ces éléments sont des arguments en défaveur d'un mécanisme médié par RAG. En particulier, les points de cassure dans les translocations supposées liées au mécanisme de recombinaison VDJ sont situés directement au niveau du signal. En revanche, un mécanisme lié à AID au cours du processus d'hypermutation somatique est potentiellement en cause. Les arguments en faveur sont la localisation du point de cassure au niveau des régions variables d'*IGL*, l'identification d'un motif CGCG au niveau du point de cassure sur le chromosome 17 et la présence d'une mutation ponctuelle, supposée somatique à proximité du point de cassure au niveau d'*IGL* (G>A dans 1 des 4 séquences lues). Etant donnée la faible couverture au niveau de la base concernée par cette dernière mutation, il pourrait s'agir d'une erreur de séquençage. Cependant, elle est située au niveau d'un motif WGCW suggérant une potentielle cible d'AID et confortant la présence de cette mutation. Le séquençage par Sanger permettra de vérifier la présence celle-ci.

Chez le troisième patient, une translocation impliquant *IGH* et le locus 8q24 avait été identifiée. Le point de cassure au niveau du chromosome 14 se situe dans la région switch d'*IGG2*, au niveau d'une cytidine adjacente à un motif WGCW. Ceci suggère un mécanisme de survenue de la translocation médié par AID au cours de la recombinaison de classe isotypique.

Au total, les mécanismes de survenue de ces trois translocations au niveau du locus des immunoglobulines semblent liés aux enzymes RAG et AID.

b- Mécanismes au niveau du gène partenaire

Chez le patient 1, il n'a pas été mis en évidence de séquence type RSS cryptique pouvant évoquer une implication de RAG ou de séquence WGCW ou CG, cibles d'AID ou encore de motif RGYW, hotspot d'AID pour l'hypermutation somatique, ne permettant pas de suggérer un mécanisme particulier de survenue de la cassure double brin au niveau du chromosome 17. Une implication de RAG dans la survenue de la cassure double brin au niveau du locus d'immunoglobuline étant suspecté, une analyse par ChiP-Seq de l'interaction de ces enzymes avec la région au niveau du chromosome 17 serait intéressante. En effet, il a été montré une interaction possible de RAG avec certaines régions chromatiniennes riches en H3K4me3 (associée à une activation de la transcription) ou directe avec les histones⁶¹. Il est possible qu'une erreur de recombinaison liée à l'activité de ces enzymes soit en cause. Une autre



hypothèse est qu'il s'agisse de cassures double brin générées de façon non spécifique par d'autres mécanismes (radiations ionisantes, radicaux libres oxygénés...).

Le mécanisme responsable de la cassure double brin au niveau du gène peut être lié à l'action d'AID dans deux cas. Pour la patiente 2, un motif CGCG est retrouvé au niveau du point de cassure sur le chromosome 17, ce qui suggère une implication d'AID au cours de l'hypermutation somatique. Dans le cas du patient 3, le point de cassure au niveau du chromosome 8 est situé à proximité d'un dinucléotide CG, pouvant être un argument en faveur d'un rôle d'AID. Il s'agit probablement dans ces deux cas d'un mécanisme type AID-AID tel que décrit par Lieber *et al.*⁴⁷. Cela signifie que l'enzyme AID est responsable de la survenue de cassure double brin au niveau du gène des immunoglobulines et du gène partenaire.

c- Mécanisme de remaniement complexe dans le cas n°1

Le remaniement observé chez le patient 1 est complexe. Nous pensions initialement qu'il s'agissait d'une translocation. Nos résultats montrent que c'est une insertion complexe d'une portion d'*IGH* dans le chromosome 17. Nous ne connaissons pas exactement le fragment délocalisé. Le mécanisme de survenue semble faire intervenir plus de 3 cassures double brin (au moins au niveau d'*IGHJ4*, *IGHD4-17* sur le chromosome 14 et *B3GNTL1* sur le chromosome 17). Nous pouvons évoquer un processus en plusieurs étapes dans l'apparition de ce remaniement avec par exemple une première translocation au cours du réarrangement VDJ et un ou des remaniements ultérieurs aboutissant à cette insertion complexe. La probabilité de survenue de plusieurs cassures double brin apparaissant de façon successive dans deux régions chromosomiques restreintes semble faible. Un autre mode de survenue proposé est l'apparition de cassures double brin au niveau du chromosome 14 ayant entraîné la formation d'un fragment circularisé (potentiellement au cours du réarrangement VDJ) avec insertion de celui-ci dans le chromosome 17. Enfin un phénomène de chromothripsis est également possible. Afin d'avancer plus dans le mécanisme de cette translocation, il est nécessaire de connaître plus précisément les régions d'*IGH* insérées dans le chromosome 17. Pour cela, nous allons utiliser les données obtenues par séquençage du génome entier et générer de nouveaux alignements (*de novo* et alignements contre référence avec des paramètres de stringence diminués) afin de tenter de visualiser des jonctions chromosome 14-chromosome 14 aberrantes. L'alignement *de novo* risque d'être difficile à analyser étant donnée la structure du gène *IGH* avec de nombreuses régions de séquence très proche. Les éventuels résultats d'intérêt mis en évidence seront confirmés ensuite par PCR standard.



2- Mécanisme de dérégulation oncogénique

Une autre question importante est l'impact de ces translocations sur la lymphomagenèse. Les gènes des immunoglobulines contiennent des éléments régulateurs tels que l'enhancer E μ ou les 3'RR. Leur rapprochement d'un oncogène par translocation peut entraîner la surexpression de celui-ci.

Nous avons vu que dans les cas 2 et 3, des éléments régulateurs sont potentiellement impliqués (enhancer 3'E λ et 3'RR1 respectivement). Les translocations sont donc susceptibles d'entraîner la surexpression d'oncogènes au niveau des loci partenaires. Dans les deux cas, les gènes en cause ne sont pas connus. Il s'agit d'*AC139099.4* en 17q25 et *CASC19* en 8q24. Tous deux sont des lncRNA. Le premier est mal connu. Il n'a pas été décrit comme étant impliqué dans des remaniements dans la littérature jusqu'à présent. Sa dérégulation n'est pas rapportée dans les cancers. L'impact potentiel d'une dérégulation de celui-ci est donc difficile à anticiper. Le deuxième est peu décrit. Quelques études rapportent une surexpression dans le cancer colorectal^{624,625,633}. Son impact dans la lymphomagenèse n'est pas connu. Le point de cassure se situe au niveau de l'intron 1. La translocation est alors susceptible soit d'entraîner la surexpression de ce lncRNA, soit d'interrompre son expression. Suivant son mode d'action et ses cibles, un rôle dans la lymphomagenèse est possible. Il est également probable que cette translocation entraîne la surexpression d'autres gènes à proximité (gènes situés en amont de *CASC19* sur le dérivé 8 identifié). Un aspect intéressant est la localisation du point de cassure en 8q24. Ce locus correspond à une région de 2 Mb à proximité du gène *MYC*, et est classiquement nommé le « désert de gènes » en raison du faible nombre de gènes fonctionnels annotés dans cette région⁶³⁴. Il est connu comme étant parmi les loci les plus fréquemment touchés par des anomalies du nombre de copies dans les cancers⁶³⁵. Il contient de nombreux gènes associés à une susceptibilité au cancer tels que *PCAT1* (pour Prostate Cancer Associated Transcript 1) ou *CCAT1* (pour Colon Cancer Associated Transcript 1). Un autre lncRNA localisé dans cette région et situé en aval de *MYC* est impliqué dans des translocations impliquant les gènes des immunoglobulines dans le lymphome de Burkitt. Il s'agit de *PVT1*. Son implication et ses effets dans les cancers sont actuellement intensivement étudiés avec déjà 71 publications répertoriées dans Pubmed parues en 2018 sur le sujet. L'expression de ce lncRNA est corrélée à celle de la protéine *MYC* dans des tumeurs humaines⁶³⁶. De plus, des essais *in vitro* d'inhibition de l'expression de *PVT1* par CRISPR-Cas9 (CRISPR pour Clustered Regularly Interspaced Short Palindromic Repeats et Cas9 pour CRISPR-associated protein 9) dans une lignée cellulaire de cancer colorectal dépendant de *MYC* ont été réalisés : ils montrent une diminution du potentiel oncogénique de *MYC* en l'absence de *PVT1*⁶³⁶. D'autres essais ont été réalisés par ARN-interférence de *PVT1* sur des lignées cellulaires issues d'hémopathies (myélome, leucémie myéloïde chronique et leucémie aiguë myéloblastique) : les résultats montrent une dégradation de *MYC*, une diminution de la prolifération cellulaire et une augmentation de l'apoptose⁶³⁷. Ce lncRNA est susceptible de devenir une cible thérapeutique dans les cancers^{636,637}. L'identification d'une translocation impliquant *CASC19*, un lncRNA situé dans le locus 8q24, est donc particulièrement intéressante. Un mécanisme d'action semblable à celui de *PVT1* est probable. Une analyse par



RNA-Seq de l'échantillon de ce patient nous permettra d'identifier d'éventuelles surexpressions de ce lncRNA, de *MYC*, mais aussi des autres gènes non codants à proximités tels que *PVT1*. A partir de ces résultats, nous pourrions imaginer des essais permettant de moduler l'expression de *CASC19* pour voir son impact sur l'expression de *MYC* et sur l'apoptose par exemple.

Dans le premier cas, nous ne connaissons pas les éléments régulateurs du gène des immunoglobulines qui sont délocalisés au cours du remaniement. Il est alors plus difficile de formuler des hypothèses sur son effet. Le gène impliqué sur le chromosome 17 est *B3GNTL1*. Il se peut que ce remaniement entraîne une modification des voies de glycosylation dans la cellule. Un aspect intéressant est la proximité du point de cassure identifié dans ce cas avec celui mis en évidence chez la patiente 2, en 17q25. Le gène touché dans le cas 1 (*B3GNTL1*) est situé seulement 200 kb en amont du lncRNA *AC139099.4* impliqué dans le cas 2. Il est possible que, de même que pour le locus 8q24, cette région possède un oncogène particulier avec ARN non codants régulateurs à proximité. On peut imaginer une possible dérégulation de *B3GNTL1* par modification de l'expression d'*AC139099.4*, ce-dernier pouvant avoir un effet stabilisant ou amplificateur de *B3GNTL1*, ou bien à l'inverse inhibiteur. Les lncRNA pouvant avoir de nombreux effets, il se peut qu'une régulation génique d'un oncogène situé à distance soit également en cause. L'analyse par RNA-Seq de l'expression des gènes situés dans ce locus 17q25, en particulier, *B3GNTL1*, *METRNL* et *AC139099.4*, mais aussi de l'expression de l'ensemble du transcriptome nous permettra d'avancer de nouvelles hypothèses sur le mécanisme de dérégulation et le potentiel pouvoir oncogénique de ces remaniements. Enfin, dans l'hypothèse où aucun élément régulateur n'est délocalisé sur le chromosome 17 au cours du remaniement, un autre mode de dérégulation peut être envisagé tel que la disruption d'un gène avec pour conséquence la production d'une protéine tronquée de fonction modifiée ou nulle. Un exemple dans la littérature est celui d'une translocation entre *IGH* et un gène codant une sulfotransférase dans la LLC : ceci entraîne la synthèse de 2 protéines tronquées, aboutissant à une anomalie de distribution cellulaire de celles-ci. La conséquence est une dérégulation d'une voie de signalisation dépendante de cette sulfotransférase et spécifique de la lignée hématopoïétique³⁸².

3- Modèles de lymphomagénèse

L'apparition des remaniements étudiés semble en relation avec les mécanismes de recombinaison VDJ (cas n°1, au moins en partie), d'hypermutation somatique (cas n°2) et de recombinaison de classe isotypique (cas n°3). Weigert et Weinstock proposent trois modèles de lymphomagénèse (Figure 30)⁷⁶. Le premier correspond à l'acquisition de l'anomalie responsable de la pathologie à un stade unique de différenciation lymphocytaire. Le modèle 2 propose deux étapes, la première survenant aux stades précoces de différenciation, conférant un avantage prolifératif, puis la deuxième à des stades plus tardifs. Enfin, dans le troisième modèle, plusieurs anomalies sont acquises successivement au cours de la



différenciation lymphoïde, la première survenant au stade de progéniteur non engagé dans une lignée. Certaines cellules peuvent conserver des capacités d'auto-renouvellement et constituer des réservoirs pour diversification clonale.

Le cas 1 appartiendrait au modèle 2 dans lequel le remaniement est susceptible d'apparaître au cours du remaniement VDJ au niveau des stades pro-B et pré-B médullaires, puis acquisition de nouvelles anomalies acquises plus tardivement. Ceci est compatible avec la mise en évidence d'une anomalie sous-clonale au caryotype (la t(13;14)). Ainsi, le remaniement d'intérêt serait un événement primaire dans la lymphomagénèse de ce cas, potentiellement associé à l'expansion clonale ou une résistance à l'apoptose. Ceci est cohérent avec la clinique observée chez ce patient, plutôt indolente, sans nécessité de mise en place thérapeutique.

Le cas n°2 serait quant à lui à relier au modèle 1, avec acquisition de l'anomalie au stade centre germinatif. Il s'agirait là encore d'un événement primaire. Cette anomalie est probablement associée à l'expansion clonale et la prolifération. Après deux lignes de traitement successives, la patiente est actuellement en abstention thérapeutique depuis plus de deux ans.

Enfin, le cas n°3 est supposé apparaître au cours de la recombinaison de classe, au stade centre germinatif. Il pourrait être apparenté au modèle 3 dans la mesure où plusieurs clones différents sont identifiés au diagnostic, tous ne présentant pas la t(8;14). On peut imaginer l'apparition d'anomalies à des stades plus précoces, et le remaniement impliquant le gène des immunoglobulines serait un événement secondaire acquis au stade centre germinatif. Il est toutefois intéressant de noter que le clone porteur de ce remaniement semble sélectionné au cours du temps, après chimiothérapie. Il s'agirait donc d'une anomalie associée à la prolifération, potentiellement impliquée dans l'agressivité de la LLC chez cette patiente.

4- Perspectives

Ce travail a permis d'identifier les points de cassures pour 3 remaniements atypiques impliquant les gènes des immunoglobulines. Il ouvre de grandes perspectives dans la découverte de nouveaux mécanismes de dérégulation génique, en particulier avec l'implication de lncRNA. Ces derniers, initialement peu explorés à cause de leur caractère non codant, sont particulièrement étudiés à l'heure actuelle grâce au développement des nouvelles techniques de séquençage haut débit, et semblent prometteurs dans la compréhension des mécanismes de régulation génique, en particulier dans les cancers et la lymphomagénèse. Cependant, il reste encore de nombreux aspects à éclaircir. En premier lieu, l'étude d'expression génique par RNA-Seq nous permettra de formuler des hypothèses plus avancées sur le mode de dérégulation produit par ces remaniements, et d'envisager de nouvelles expériences à mener pour expliquer leur rôle dans la lymphomagénèse. Il est possible que, comme PVT1, de nouvelles cibles thérapeutiques puissent être identifiées.



Par ailleurs, le remaniement du premier patient n'a pu être détecté initialement que par FISH. Il s'agit d'une anomalie cryptique au caryotype. Il est donc possible que ce type d'anomalie soit plus fréquent dans la mesure où l'analyse par FISH des loci des immunoglobulines n'est pas systématiquement réalisée. Il serait intéressant de dépister de nouveaux cas de lymphomes afin d'identifier des remaniements impliquant les loci 17q25 ou 8q24 en amont de *MYC*, ou encore de nouveaux loci non précédemment identifiés.

B- Impact pronostique de l'électrophorèse des protéines sériques dans la LLC

Avec l'avènement du NGS, de nombreuses mutations géniques ont été mises en évidence dans la LLC à des fréquences plus ou moins importantes. Les anomalies récurrentes telles que celles affectant *TP53*, *NOTCH1*, *SF3B1* ou *BIRC3* ont été étudiées dans des séries de patients et leur impact pronostique péjoratif a pu être mis en évidence^{360,361,422,445,446,448}. L'apparition d'un nombre important de nouveaux marqueurs pronostiques moléculaires dans la LLC rend de plus en plus complexe l'analyse et un score intégrant ces différentes variables reste à établir. Récemment, un nouveau score a été proposé : le CLL-IPI. Celui-ci ne prend en compte sur le plan moléculaire que le statut mutationnel *IGHV* et de *TP53*⁴⁹⁷. L'intégration de l'ensemble de ces nouveaux marqueurs semble difficile sur le plan pratique et seuls ceux ayant un impact suffisamment fort sont donc utilisés. En pratique, aucun de ces marqueurs pronostiques moléculaires ne constitue en lui-même une indication de traitement.

Nous avons étudié dans ce travail un marqueur sérique habituellement analysé au diagnostic et au cours du suivi dans la LLC : il s'agit de l'électrophorèse des protéines sériques. L'impact pronostique d'une électrophorèse des protéines sérique normale n'a pas été étudié jusqu'ici dans la littérature, l'hypogammaglobulinémie et/ou la présence d'un composant monoclonal, évalués par différentes techniques, ayant été débattus comme marqueurs pronostiques^{327,457,504-509}. Nous avons donc analysé l'impact d'une électrophorèse des protéines sériques normale (c'est-à-dire immunoglobulines polyclonales sans hypogammaglobulinémie) versus anormale au diagnostic dans la LLC. L'intérêt de ce marqueur est sa simplicité de mesure et son moindre coût. Nous avons pu montrer qu'une électrophorèse des protéines sérique normale permet d'identifier des patients avec un bon pronostic. Elle est associée au stade Binet A et à une lymphocytose au diagnostic inférieure à 30 G/L. La combinaison de ce marqueur avec le statut *IGHV* muté permet d'identifier une population particulière de patients ayant une survie sans traitement significativement plus longue que les autres cas de LLC, et qui ne nécessiteront probablement jamais de traitement spécifique de cette maladie.

L'électrophorèse des protéines sériques anormale serait en fait un marqueur d'évolutivité. L'hypogammaglobulinémie peut survenir d'emblée ou apparaître lors du suivi. Elle pourrait



être attribuable à un déficit fonctionnel des lymphocytes B CD5- n'appartenant pas au clone de la LLC⁶³⁸. Les sous-types IgG3 et IgG4 sont les plus touchés. La présence d'un pic monoclonal peut quant à elle être attribuable à une sécrétion par le clone de LLC. Cependant, il existe des cas pour lesquels une discordance est observée concernant l'isotype de la chaîne légère (Kappa ou Lambda) exprimée par le clone de LLC, identifiée par cytométrie en flux et celle de l'immunoglobuline monoclonale retrouvée au niveau sérique. Ceci serait observable dans 34 % des cas⁶³⁹. Elle pourrait alors être sécrétée par des lymphocytes B n'appartenant pas au clone et être un reflet de l'activité du microenvironnement.

En conclusion, l'électrophorèse des protéines sériques fournit des informations intéressantes non seulement par la détection des patients présentant une hypogammaglobulinémie et donc à risque de développer des infections, mais aussi sur le plan pronostique, permettant d'identifier un groupe de patients avec une évolution favorable lorsqu'elle est normale au diagnostic, en particulier en association à un statut IGHV muté. Ceci reste valable dans l'ère du séquençage haut débit.

C- Utilisation du séquençage haut débit pour la détection d'anomalies de nombre

Nous avons mis au point un panel de séquençage haut débit par amplicons permettant l'identification des principales mutations récurrentes décrites dans les lymphomes. Nous avons développé ce panel afin de pouvoir également détecter les principales anomalies de nombre de la LLC : la del(13q), la trisomie 12, la del(11q) et la del(17p). Nous avons collaboré avec l'EA6308 afin d'améliorer le logiciel Cov'Cop développé initialement pour l'identification de CNV dans les maladies héréditaires⁶⁰⁶. Diverses options ont été intégrées afin de faciliter la détection des CNV, en particulier dans les cas où l'anomalie peut être présente dans une minorité de cellules dans l'échantillon analysé (mosaïques de certaines pathologies constitutionnelles ou cas des tumeurs avec infiltration partielle et/ou anomalies sous-clonales). Une amélioration majeure est la construction automatique d'un graphique représentant les ratios normalisés de chaque amplicon dans l'ordre chromosomique permettant l'analyse rapide des résultats. De plus, il est maintenant possible d'assigner un ou plusieurs contrôles sains dans l'analyse, ce qui permet d'améliorer la stabilité de certaines régions pouvant être fréquemment anormales, comme c'est le cas avec la del(13q) dans la LLC.

L'analyse de nos résultats obtenus par NGS concernant les mutations nous a permis d'identifier une bonne stabilité de la VAF d'un certain nombre de SNP avec des valeurs très proches de 50 % en cas d'hétérozygotie. Nous avons donc voulu utiliser ces résultats pour détecter la présence de CNV. Pour cela, nous avons développé un script utilisable avec le logiciel R, permettant de sélectionner automatiquement les SNP d'un échantillon et de



représenter leur VAF graphiquement. Nous avons évalué les performances analytiques du NGS en combinant l'utilisation du logiciel Cov'Cop2 et l'analyse des VAF de SNP dans une série de 82 patients atteints de LLC, en comparaison à la technique de référence (cytogénétique et/ou QMPSF). Nos résultats nous permettent d'estimer un seuil de sensibilité à 30 %. Certaines anomalies peuvent être détectées avec un pourcentage de cellules porteuses de l'anomalie dans l'échantillon inférieur à ce seuil. La qualité de la librairie est primordiale pour une bonne interprétation des résultats. En effet, en cas de mauvaise qualité, les ratios obtenus sont très dispersés et les anomalies sont alors difficilement interprétables.

L'approche VAF peut permettre de détecter des anomalies initialement non identifiées avec le logiciel Cov'Cop2 et d'améliorer la sensibilité de l'analyse. Cependant, elle détecte également de probables UPD dont la significations pronostique au niveau de *TP53* et *ATM* reste discutable en l'absence de mutation de ces gènes. Nous avons donc décidé de vérifier par FISH ces CNV lorsqu'ils ne sont identifiés que par VAF, et non par Cov'Cop2.

Nous avons classé les patients en groupes pronostiques selon le score de Rossi à partir des données du NGS seul ou des résultats des techniques conventionnelles (cytogénétiques et/ou QMPSF)³⁶¹. Malgré un manque de sensibilité de notre technique pour la détection de certains CNV (del(13q), del(11q) et del(17p)), nous montrons que l'impact sur la survie sans traitement est peu marqué pour les patients de notre série.

Dans ce travail, nous avons privilégié une bonne spécificité, le seuil de sensibilité de 30 % étant insuffisant pour exclure une délétion de *TP53* selon les recommandations actuelles³⁵³. Ainsi, notre approche de détection des CNV par NGS permet de remplacer l'analyse par QMPSF dont la sensibilité à 25 % est également insuffisante dans ce contexte⁶²⁶. Elle ne permet pas en revanche de s'affranchir totalement de l'analyse par FISH, dont le seuil est estimé à 8 % dans notre laboratoire pour le locus *TP53*, demeurant la technique de référence. Notre approche par NGS permet de diminuer le nombre d'analyses par FISH, en réservant son utilisation à la détection de la del(17p) lorsque celle-ci n'est pas détectable par NGS et en l'absence de mutation de *TP53*.

La détection d'UPD par cette technique est intéressante. La présence d'UPD peut permettre de dupliquer un allèle muté et entraîner la double inactivation d'un gène, comme par exemple pour *TP53*⁶²⁸. Un cas de notre série est concerné. Il s'agit donc d'une donnée intéressante. En revanche la valeur pronostique de ce type d'anomalie en l'absence de mutation est inconnu. Certains auteurs ont rapporté un lien entre le nombre d'UPD et un raccourcissement des télomères responsable d'une instabilité chromosomique globale⁶³². Ainsi, la détermination de la présence d'UPD pourrait être un reflet de l'instabilité chromosomique dans la LLC, à l'origine d'accumulation d'erreurs génétiques par défaut de réparation de l'ADN. Il serait intéressant de regarder l'impact pronostique des UPD dans la LLC, soit de façon ciblée (par exemple UPD de *TP53* avec mutation concomitante ou non) ou de façon globale : présence ou non d'UPD quelque soit le locus touché. Cependant, notre approche par NGS ne permet de détecter que certaines UPD ciblées par notre panel de 70 gènes et l'étude de l'impact des UPD dans la LLC en tant qu'altération globale devrait être analysée par SNP-array.



Conclusion

La question posée était celle du poids des anomalies moléculaires et en particulier des gènes des immunoglobulines dans l'histoire naturelle de la LLC, tant pour le développement de la maladie que pour son pronostic.

Au plan mécanistique, ce travail de thèse a permis de déterminer les points de cassure de trois remaniements des loci des immunoglobulines atypiques dans la LLC/lymphome lymphocytaire. Il met en évidence dans deux cas l'implication de lncRNA (cas 2 et 3). Il ouvre donc la voie pour l'étude de ceux-ci et de leurs gènes cibles dans la LLC. Deux cas ont également des points de cassure dans une région chromosomique proche (espacés de seulement 200 kb en 17q25 pour les cas 1 et 2) et il pourrait s'agir d'une région importante dans la lymphomagenèse, de même que pour la région 8q24 contenant *MYC* et de nombreux gènes non codants jusqu'ici peu explorés. Enfin, un de nos remaniements est cryptique au caryotype et démontre l'importance d'une analyse systématique par FISH pour les loci des immunoglobulines pour la détection de nouveaux remaniements atypiques apparentés ou non dans les lymphomes B.

D'un point de vue pronostique, nous avons démontré qu'un marqueur simple et peu coûteux reste, dans l'ère du NGS, un marqueur pronostique fort dans la LLC. Il s'agit de l'électrophorèse des protéines sériques qui, lorsqu'elle est normale (c'est-à-dire immunoglobulines polyclonales sans hypogammaglobulinémie) au diagnostic, permet d'identifier des patients avec une excellente survie sans traitement. En particulier, la combinaison de ce marqueur avec un statut IGHV muté permet d'identifier des cas de LLC qui n'auront probablement jamais besoin d'un traitement spécifique de leur hémopathie.

De façon connexe, nous avons mis au point un outil de détection des CNV par NGS performant. Malgré un manque de sensibilité, il permet de simplifier l'analyse génétique des cas de LLC au laboratoire. Il est capable également de détecter des UPD dont la signification pronostique n'est actuellement pas établie. Ceci pourrait être un marqueur d'instabilité génétique et avoir un impact pronostique qu'il serait intéressant d'étudier.



Références bibliographiques

1. Swerdlow, S. H. WHO classification of tumours of haematopoietic and lymphoid tissues: ... reflects the views of a working group that convened for an Editorial and Consensus Conference at the International Agency for Research on Cancer (IARC), Lyon, October 25 - 27, 2007. (Internat. Agency for Research on Cancer, 2008).
2. Armitage, J. O., Mauch, P. M., Harris, N. L., Coiffier, B. & Dalla-Favera, R. Non-Hodgkin lymphomas. (Wolters Kluwer Health, 2013).
3. Schroeder, H. W. & Cavacini, L. Structure and function of immunoglobulins. *J. Allergy Clin. Immunol.* **125**, S41-52 (2010).
4. Owen, J., Punt, J. & Stranford, S. Immunologie : Le cours de Janis Kuby avec questions de révision. (Dunod).
5. Schatz, D. G. & Ji, Y. Recombination centres and the orchestration of V(D)J recombination. *Nat. Rev. Immunol.* **11**, 251–263 (2011).
6. Dudley, D. D., Chaudhuri, J., Bassing, C. H. & Alt, F. W. Mechanism and control of V(D)J recombination versus class switch recombination: similarities and differences. *Adv. Immunol.* **86**, 43–112 (2005).
7. Jung, D., Giallourakis, C., Mostoslavsky, R. & Alt, F. W. Mechanism and control of V(D)J recombination at the immunoglobulin heavy chain locus. *Annu. Rev. Immunol.* **24**, 541–570 (2006).
8. Janeway, C. A., Murphy, K., Travers, P., Walport, M. & Masson, P. L. Immunobiologie. (De Boeck, 2009).
9. Meier, J. T. & Lewis, S. M. P nucleotides in V(D)J recombination: a fine-structure analysis. *Mol Cell Biol* **13**, 1078–1092 (1993).
10. Parham, P. & Masson, P. L. Le système immunitaire. (De Boeck Université, 2003).
11. Goossens, T., Klein, U. & Küppers, R. Frequent occurrence of deletions and duplications during somatic hypermutation: implications for oncogene translocations and heavy chain disease. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 2463–2468 (1998).
12. Bross, L. et al. DNA double-strand breaks in immunoglobulin genes undergoing somatic hypermutation. *Immunity* **13**, 589–597 (2000).



13. Pasqualucci, L. et al. BCL-6 mutations in normal germinal center B cells: evidence of somatic hypermutation acting outside Ig loci. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 11816–11821 (1998).
14. Shen, H. M., Peters, A., Baron, B., Zhu, X. & Storb, U. Mutation of BCL-6 gene in normal B cells by the process of somatic hypermutation of Ig genes. *Science* **280**, 1750–1752 (1998).
15. Müschen, M. et al. Somatic mutation of the CD95 gene in human B cells as a side-effect of the germinal center reaction. *J. Exp. Med.* **192**, 1833–1840 (2000).
16. Gordon, M. S., Kanegai, C. M., Doerr, J. R. & Wall, R. Somatic hypermutation of the B cell receptor genes B29 (Igbeta, CD79b) and mb1 (Igalpha, CD79a). *Proc. Natl. Acad. Sci. U.S.A.* **100**, 4126–4131 (2003).
17. Honjo, T., Kinoshita, K. & Muramatsu, M. Molecular mechanism of class switch recombination: linkage with somatic hypermutation. *Annu. Rev. Immunol.* **20**, 165–196 (2002).
18. Revy, P. et al. Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). *Cell* **102**, 565–575 (2000).
19. Okazaki, I., Kinoshita, K., Muramatsu, M., Yoshikawa, K. & Honjo, T. The AID enzyme induces class switch recombination in fibroblasts. *Nature* **416**, 340–345 (2002).
20. Yoshikawa, K. et al. AID enzyme-induced hypermutation in an actively transcribed gene in fibroblasts. *Science* **296**, 2033–2036 (2002).
21. Hwang, J. K., Alt, F. W. & Yeap, L.-S. Related Mechanisms of Antibody Somatic Hypermutation and Class Switch Recombination. *Microbiol Spectr* **3**, MDNA3-0037-2014 (2015).
22. Lefranc, M. P. Nomenclature of the human immunoglobulin heavy (IGH) genes. *Exp. Clin. Immunogenet.* **18**, 100–116 (2001).
23. van Dongen, J. J. M. et al. Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: report of the BIOMED-2 Concerted Action BMH4-CT98-3936. *Leukemia* **17**, 2257–2317 (2003).
24. Pinaud, E. et al. The IgH locus 3' regulatory region: pulling the strings from behind. *Adv. Immunol.* **110**, 27–70 (2011).
25. Immunoglobulin heavy (IGH) orphans Gene Family | HUGO Gene Nomenclature Committee. Available at: <https://www.genenames.org/cgi-bin/genefamilies/set/350>. (Accessed: 19th November 2017)



26. Lefranc, M. P. Nomenclature of the human immunoglobulin kappa (IGK) genes. *Exp. Clin. Immunogenet.* **18**, 161–174 (2001).
27. Immunoglobulin kappa (IGK) orphans Gene Family | HUGO Gene Nomenclature Committee. Available at: <https://www.genenames.org/cgi-bin/genefamilies/set/352>. (Accessed: 19th November 2017)
28. Lefranc, M. P. Nomenclature of the human immunoglobulin lambda (IGL) genes. *Exp. Clin. Immunogenet.* **18**, 242–254 (2001).
29. Immunoglobulin lambda (IGL) orphans Gene Family | HUGO Gene Nomenclature Committee. Available at: <https://www.genenames.org/cgi-bin/genefamilies/set/354>. (Accessed: 19th November 2017)
30. Seda, V. & Mraz, M. B-cell receptor signalling and its crosstalk with other pathways in normal and malignant cells. *Eur. J. Haematol.* **94**, 193–205 (2015).
31. Young, R. M. & Staudt, L. M. Targeting pathological B cell receptor signalling in lymphoid malignancies. *Nat Rev Drug Discov* **12**, 229–243 (2013).
32. Pieper, K., Grimbacher, B. & Eibel, H. B-cell biology and development. *J. Allergy Clin. Immunol.* **131**, 959–971 (2013).
33. Meffre, E., Casellas, R. & Nussenzweig, M. C. Antibody regulation of B cell development. *Nat. Immunol.* **1**, 379–385 (2000).
34. Ye, M. & Graf, T. Early decisions in lymphoid development. *Current Opinion in Immunology* **19**, 123–128 (2007).
35. LeBien, T. W. Fates of human B-cell precursors. *Blood* **96**, 9–23 (2000).
36. Hieter, P. A., Korsmeyer, S. J., Waldmann, T. A. & Leder, P. Human immunoglobulin kappa light-chain genes are deleted or rearranged in lambda-producing B cells. *Nature* **290**, 368–372 (1981).
37. Korsmeyer, S. J. et al. Normal human B cells display ordered light chain gene rearrangements and deletions. *J. Exp. Med.* **156**, 975–985 (1982).
38. Zhang, Z., Burrows, P. D. & Cooper, M. D. The molecular basis and biological significance of VH replacement. *Immunol. Rev.* **197**, 231–242 (2004).
39. Luning Prak, E. T., Monestier, M. & Eisenberg, R. A. B cell receptor editing in tolerance and autoimmunity. *Ann. N. Y. Acad. Sci.* **1217**, 96–121 (2011).
40. Nemazee, D. Receptor editing in B cells. *Adv. Immunol.* **74**, 89–126 (2000).



41. Küppers, R. Mechanisms of B-cell lymphoma pathogenesis. *Nat. Rev. Cancer* **5**, 251–262 (2005).
42. Allen, C. D. C., Okada, T., Tang, H. L. & Cyster, J. G. Imaging of germinal center selection events during affinity maturation. *Science* **315**, 528–531 (2007).
43. Schwickert, T. A. et al. In vivo imaging of germinal centres reveals a dynamic open structure. *Nature* **446**, 83–87 (2007).
44. Shapiro-Shelef, M. & Calame, K. Regulation of plasma-cell development. *Nat. Rev. Immunol.* **5**, 230–242 (2005).
45. Good-Jacobson, K. L. & Tarlinton, D. M. Multiple routes to B-cell memory. *Int. Immunol.* **24**, 403–408 (2012).
46. Swerdlow, S. H. et al. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood* **127**, 2375–2390 (2016).
47. Lieber, M. R. Mechanisms of human lymphoid chromosomal translocations. *Nat. Rev. Cancer* **16**, 387–398 (2016).
48. Küppers, R. & Dalla-Favera, R. Mechanisms of chromosomal translocations in B cell lymphomas. *Oncogene* **20**, 5580–5594 (2001).
49. Limpens, J. et al. Bcl-2/JH rearrangements in benign lymphoid tissues with follicular hyperplasia. *Oncogene* **6**, 2271–2276 (1991).
50. Summers, K. E. et al. Frequency of the Bcl-2/IgH rearrangement in normal individuals: implications for the monitoring of disease in patients with follicular lymphoma. *J. Clin. Oncol.* **19**, 420–424 (2001).
51. Lieber, M. R., Yu, K. & Raghavan, S. C. Roles of nonhomologous DNA end joining, V(D)J recombination, and class switch recombination in chromosomal translocations. *DNA Repair (Amst.)* **5**, 1234–1245 (2006).
52. Raghavan, S. C., Swanson, P. C., Wu, X., Hsieh, C.-L. & Lieber, M. R. A non-B-DNA structure at the Bcl-2 major breakpoint region is cleaved by the RAG complex. *Nature* **428**, 88–93 (2004).
53. Roth, D. B. & Craig, N. L. VDJ recombination: a transposase goes to work. *Cell* **94**, 411–414 (1998).
54. Hiom, K., Melek, M. & Gellert, M. DNA transposition by the RAG1 and RAG2 proteins: a possible source of oncogenic translocations. *Cell* **94**, 463–470 (1998).



55. Jäger, U. et al. Follicular lymphomas' BCL-2/IgH junctions contain templated nucleotide insertions: novel insights into the mechanism of t(14;18) translocation. *Blood* **95**, 3520–3529 (2000).
56. Welzel, N. et al. Templated nucleotide addition and immunoglobulin JH-gene utilization in t(11;14) junctions: implications for the mechanism of translocation and the origin of mantle cell lymphoma. *Cancer Res.* **61**, 1629–1636 (2001).
57. Blum, K. A., Lozanski, G. & Byrd, J. C. Adult Burkitt leukemia and lymphoma. *Blood* **104**, 3009–3020 (2004).
58. Wyatt, R. T., Rudders, R. A., Zelenetz, A., Delellis, R. A. & Krontiris, T. G. BCL2 oncogene translocation is mediated by a chi-like consensus. *J. Exp. Med.* **175**, 1575–1588 (1992).
59. Jaeger, U. et al. Molecular mechanism of the t(14;18)--a model for lymphoid-specific chromosomal translocations. *Leuk. Lymphoma* **14**, 197–202 (1994).
60. Agrawal, A., Eastman, Q. M. & Schatz, D. G. Transposition mediated by RAG1 and RAG2 and its implications for the evolution of the immune system. *Nature* **394**, 744–751 (1998).
61. Maman, Y., Teng, G., Seth, R., Kleinstein, S. H. & Schatz, D. G. RAG1 targeting in the genome is dominated by chromatin interactions mediated by the non-core regions of RAG1 and RAG2. *Nucleic Acids Res* **44**, 9624–9637 (2016).
62. Chesi, M. et al. The t(4;14) translocation in myeloma dysregulates both FGFR3 and a novel gene, MMSET, resulting in IgH/MMSET hybrid transcripts. *Blood* **92**, 3025–3034 (1998).
63. Gabrea, A., Bergsagel, P. L., Chesi, M., Shou, Y. & Kuehl, W. M. Insertion of excised IgH switch sequences causes overexpression of cyclin D1 in a myeloma tumor cell. *Mol. Cell* **3**, 119–123 (1999).
64. Meng, F.-L. et al. Convergent transcription at intragenic super-enhancers targets AID-initiated genomic instability. *Cell* **159**, 1538–1548 (2014).
65. Pefanis, E. et al. Noncoding RNA transcription targets AID to divergently transcribed loci in B cells. *Nature* **514**, 389–393 (2014).
66. Papavasiliou, F. N. & Schatz, D. G. Cell-cycle-regulated DNA double-stranded breaks in somatic hypermutation of immunoglobulin genes. *Nature* **408**, 216–221 (2000).
67. Cario, G., Stadt, U. Z., Reiter, A., Welte, K. & Sykora, K. W. Variant translocations in sporadic Burkitt's lymphoma detected in fresh tumour material: analysis of three cases. *Br. J. Haematol.* **110**, 537–546 (2000).



68. Ohno, H. Pathogenetic and clinical implications of non-immunoglobulin ; BCL6 translocations in B-cell non-Hodgkin's lymphoma. *J Clin Exp Hematop* **46**, 43–53 (2006).
69. Yoshida, S. et al. Identification of heterologous translocation partner genes fused to the BCL6 gene in diffuse large B-cell lymphomas: 5'-RACE and LA - PCR analyses of biopsy samples. *Oncogene* **18**, 7994–7999 (1999).
70. Akasaka, H. et al. Molecular anatomy of BCL6 translocations revealed by long-distance polymerase chain reaction-based assays. *Cancer Res.* **60**, 2335–2341 (2000).
71. Twa, D. D. W. et al. Genomic rearrangements involving programmed death ligands are recurrent in primary mediastinal large B-cell lymphoma. *Blood* **123**, 2062–2065 (2014).
72. Chong, L. C. et al. Comprehensive characterization of programmed death ligand structural rearrangements in B-cell non-Hodgkin lymphomas. *Blood* **128**, 1206–1213 (2016).
73. Blank, C., Gajewski, T. F. & Mackensen, A. Interaction of PD-L1 on tumor cells with PD-1 on tumor-specific T cells as a mechanism of immune evasion: implications for tumor immunotherapy. *Cancer Immunol. Immunother.* **54**, 307–314 (2005).
74. Pasqualucci, L. et al. Hypermutation of multiple proto-oncogenes in B-cell diffuse large-cell lymphomas. *Nature* **412**, 341–346 (2001).
75. Compagno, M. et al. Mutations of multiple genes cause deregulation of NF-kappaB in diffuse large B-cell lymphoma. *Nature* **459**, 717–721 (2009).
76. Weigert, O. & Weinstock, D. M. The evolving contribution of hematopoietic progenitor cells to lymphomagenesis. *Blood* **120**, 2553–2561 (2012).
77. Basso, K. & Dalla-Favera, R. Germinal centres and B cell lymphomagenesis. *Nature Reviews Immunology* **15**, 172–184 (2015).
78. Bouzelfen, A. et al. HACE1 is a putative tumor suppressor gene in B-cell lymphomagenesis and is down-regulated by both deletion and epigenetic alterations. *Leuk. Res.* **45**, 90–100 (2016).
79. Jarosova, M. et al. Chromosome 6q deletion correlates with poor prognosis and low relative expression of FOXO3 in chronic lymphocytic leukemia patients. *Am. J. Hematol.* **92**, E604–E607 (2017).
80. Krajnović, M., Radojković, M., Davidović, R., Dimitrijević, B. & Krtolica, K. Prognostic significance of epigenetic inactivation of p16, p15, MGMT and DAPK genes in follicular lymphoma. *Med. Oncol.* **30**, 441 (2013).



81. Laherty, C. D., Hu, H. M., Opipari, A. W., Wang, F. & Dixit, V. M. The Epstein-Barr virus LMP1 gene product induces A20 zinc finger protein expression by activating nuclear factor kappa B. *J. Biol. Chem.* **267**, 24157–24160 (1992).
82. Luftig, M. et al. Epstein-Barr virus latent infection membrane protein 1 TRAF-binding site induces NIK/IKK alpha-dependent noncanonical NF-kappaB activation. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 141–146 (2004).
83. Guasparri, I., Keller, S. A. & Cesarman, E. KSHV vFLIP is essential for the survival of infected lymphoma cells. *J. Exp. Med.* **199**, 993–1003 (2004).
84. Carbone, A., Ghoghini, A. & Dotti, G. EBV-associated lymphoproliferative disorders: classification and treatment. *Oncologist* **13**, 577–585 (2008).
85. van Baarle, D. et al. Dysfunctional Epstein-Barr virus (EBV)-specific CD8(+) T lymphocytes and increased EBV load in HIV-1 infected individuals progressing to AIDS-related non-Hodgkin lymphoma. *Blood* **98**, 146–155 (2001).
86. Küppers, R., Engert, A. & Hansmann, M.-L. Hodgkin lymphoma. *J. Clin. Invest.* **122**, 3439–3447 (2012).
87. Wang, L. W., Jiang, S. & Gewurz, B. E. Epstein-Barr Virus LMP1-Mediated Oncogenicity. *J. Virol.* **91**, (2017).
88. Caldwell, R. G., Wilson, J. B., Anderson, S. J. & Longnecker, R. Epstein-Barr virus LMP2A drives B cell development and survival in the absence of normal B cell receptor signals. *Immunity* **9**, 405–411 (1998).
89. Ritz, O. et al. Downregulation of internal enhancer activity contributes to abnormally low immunoglobulin expression in the MedB-1 mediastinal B-cell lymphoma cell line. *J. Pathol.* **205**, 336–348 (2005).
90. Arguello, M. et al. Disruption of the B-cell specific transcriptional program in HHV-8 associated primary effusion lymphoma cell lines. *Oncogene* **22**, 964–973 (2003).
91. Klein, U. et al. Gene expression profile analysis of AIDS-related primary effusion lymphoma (PEL) suggests a plasmablastic derivation and identifies PEL-specific transcripts. *Blood* **101**, 4115–4121 (2003).
92. Borche, L., Lim, A., Binet, J. L. & Dighiero, G. Evidence that chronic lymphocytic leukemia B lymphocytes are frequently committed to production of natural autoantibodies. *Blood* **76**, 562–569 (1990).
93. Stoeber, Z. M. et al. Production of autoantibodies by CD5-expressing B lymphocytes from patients with chronic lymphocytic leukemia. *J. Exp. Med.* **169**, 255–268 (1989).



94. Mann, D. L. et al. HTLV-I--associated B-cell CLL: indirect role for retrovirus in leukemogenesis. *Science* **236**, 1103–1106 (1987).
95. Fais, F. et al. Chronic lymphocytic leukemia B cells express restricted sets of mutated and unmutated antigen receptors. *J. Clin. Invest.* **102**, 1515–1525 (1998).
96. Messmer, B. T. et al. Multiple distinct sets of stereotyped antigen receptors indicate a role for antigen in promoting chronic lymphocytic leukemia. *J. Exp. Med.* **200**, 519–525 (2004).
97. Damle, R. N. et al. B-cell chronic lymphocytic leukemia cells express a surface membrane phenotype of activated, antigen-experienced B lymphocytes. *Blood* **99**, 4087–4093 (2002).
98. Thompsett, A. R., Ellison, D. W., Stevenson, F. K. & Zhu, D. V(H) gene sequences from primary central nervous system lymphomas indicate derivation from highly mutated germinal center B cells with ongoing mutational activity. *Blood* **94**, 1738–1746 (1999).
99. Dighiero, G. et al. Autoantibody activity of immunoglobulins isolated from B-cell follicular lymphomas. *Blood* **78**, 581–585 (1991).
100. Bahler, D. W. & Levy, R. Clonal evolution of a follicular lymphoma: evidence for antigen selection. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 6770–6774 (1992).
101. Zhu, D. et al. Acquisition of potential N-glycosylation sites in the immunoglobulin variable region by somatic mutation is a distinctive feature of follicular lymphoma. *Blood* **99**, 2562–2568 (2002).
102. Quinn, E. R. et al. The B-cell receptor of a hepatitis C virus (HCV)-associated non-Hodgkin lymphoma binds the viral E2 envelope protein, implicating HCV in lymphomagenesis. *Blood* **98**, 3745–3749 (2001).
103. Hermine, O. et al. Regression of splenic lymphoma with villous lymphocytes after treatment of hepatitis C virus infection. *N. Engl. J. Med.* **347**, 89–94 (2002).
104. Hussell, T., Isaacson, P. G., Crabtree, J. E. & Spencer, J. Helicobacter pylori-specific tumour-infiltrating T cells provide contact dependent help for the growth of malignant B cells in low-grade gastric lymphoma of mucosa-associated lymphoid tissue. *J. Pathol.* **178**, 122–127 (1996).
105. Hussell, T., Isaacson, P. G., Crabtree, J. E., Dogan, A. & Spencer, J. Immunoglobulin specificity of low grade B cell gastrointestinal lymphoma of mucosa-associated lymphoid tissue (MALT) type. *Am. J. Pathol.* **142**, 285–292 (1993).
106. Dogan, A. et al. Follicular lymphomas contain a clonally linked but phenotypically distinct neoplastic B-cell population in the interfollicular zone. *Blood* **91**, 4708–4714 (1998).



107. Umetsu, D. T., Esserman, L., Donlon, T. A., DeKruyff, R. H. & Levy, R. Induction of proliferation of human follicular (B type) lymphoma cells by cognate interaction with CD4+ T cell clones. *J. Immunol.* **144**, 2550–2557 (1990).
108. Johnson, P. W. et al. Isolated follicular lymphoma cells are resistant to apoptosis and can be grown in vitro in the CD40/stromal cell system. *Blood* **82**, 1848–1857 (1993).
109. Küppers, R. Molecular biology of Hodgkin lymphoma. *Hematology Am Soc Hematol Educ Program* 491–496 (2009). doi:10.1182/asheducation-2009.1.491
110. Nobili, L., Lionetti, M. & Neri, A. Long non-coding RNAs in normal and malignant hematopoiesis. *Oncotarget* **7**, 50666–50681 (2016).
111. LNCipedia. Available at: <https://lncipedia.org/>. (Accessed: 26th September 2018)
112. Volders, P.-J. et al. An update on LNCipedia: a database for annotated human lncRNA sequences. *Nucleic Acids Res.* **43**, D174-180 (2015).
113. NONCODE. Available at: <http://www.noncode.org/>. (Accessed: 26th September 2018)
114. Winkle, M., Kluiver, J. L., Diepstra, A. & van den Berg, A. Emerging roles for long noncoding RNAs in B-cell development and malignancy. *Crit. Rev. Oncol. Hematol.* **120**, 77–85 (2017).
115. Schmidt, L. H. et al. The long noncoding MALAT-1 RNA indicates a poor prognosis in non-small cell lung cancer and induces migration and tumor growth. *J Thorac Oncol* **6**, 1984–1992 (2011).
116. Petri, A. et al. Long Noncoding RNA Expression during Human B-Cell Development. *PLoS ONE* **10**, e0138236 (2015).
117. Ranzani, V. et al. The long intergenic noncoding RNA landscape of human lymphocytes highlights the regulation of T cell differentiation by linc-MAF-4. *Nat. Immunol.* **16**, 318–325 (2015).
118. Casero, D. et al. Long non-coding RNA profiling of human lymphoid progenitor cells reveals transcriptional divergence of B cell and T cell lineages. *Nat. Immunol.* **16**, 1282–1291 (2015).
119. Nakamura, Y. et al. The GAS5 (growth arrest-specific transcript 5) gene fuses to BCL6 as a result of t(1;3)(q25;q27) in a patient with B-cell lymphoma. *Cancer Genet. Cytogenet.* **182**, 144–149 (2008).



120. Mourtada-Maarabouni, M. & Williams, G. T. Role of GAS5 noncoding RNA in mediating the effects of rapamycin and its analogues on mantle cell lymphoma cells. *Clin Lymphoma Myeloma Leuk* **14**, 468–473 (2014).
121. Does, G. M. et al. Chronic lymphocytic leukaemia and small lymphocytic lymphoma: overview of the descriptive epidemiology. *Br. J. Haematol.* **139**, 809–819 (2007).
122. Mak, V. et al. Preservation of lower incidence of chronic lymphocytic leukemia in Chinese residents in British Columbia: a 26-year survey from 1983 to 2008. *Leuk. Lymphoma* **55**, 824–827 (2014).
123. Monnereau, A. et al. Estimation nationale de l'incidence et de la mortalité par cancer en France entre 1980 et 2012: étude à partir des registres des cancers du réseau Francim. (Institut de veille sanitaire, 2013).
124. Hallek, M. Chronic lymphocytic leukemia: 2017 update on diagnosis, risk stratification, and treatment. *Am. J. Hematol.* **92**, 946–965 (2017).
125. Cancer statistics, 2015 - Siegel - 2015 - CA: A Cancer Journal for Clinicians - Wiley Online Library. Available at: <https://onlinelibrary.wiley.com/doi/abs/10.3322/caac.21254>. (Accessed: 25th July 2018)
126. Brown, J. R. Inherited predisposition to chronic lymphocytic leukemia. *Expert Rev Hematol* **1**, 51–61 (2008).
127. Mauro, F. R. et al. Clinical features and outcome of familial chronic lymphocytic leukemia. *Haematologica* **91**, 1117–1120 (2006).
128. Jønsson, V. et al. CLL family 'Pedigree 14' revisited: 1947–2004. *Leukemia* **19**, 1025–1028 (2005).
129. Ishibe, N. et al. Clinical characteristics of familial B-CLL in the National Cancer Institute Familial Registry. *Leuk. Lymphoma* **42**, 99–108 (2001).
130. Goldin, L. R., Pfeiffer, R. M., Li, X. & Hemminki, K. Familial risk of lymphoproliferative tumors in families of patients with chronic lymphocytic leukemia: results from the Swedish Family-Cancer Database. *Blood* **104**, 1850–1854 (2004).
131. Sellick, G. S. et al. A high-density SNP genome-wide linkage search of 206 families identifies susceptibility loci for chronic lymphocytic leukemia. *Blood* **110**, 3326–3333 (2007).
132. Crowther-Swanepoel, D. et al. Genetic variation in CXCR4 and risk of chronic lymphocytic leukemia. *Blood* **114**, 4843–4846 (2009).



133. Brown, J. R. Inherited susceptibility to chronic lymphocytic leukemia: evidence and prospects for the future. *Ther Adv Hematol* **4**, 298–308 (2013).
134. Kipps, T. J. et al. Chronic lymphocytic leukaemia. *Nat Rev Dis Primers* **3**, 16096 (2017).
135. Di Bernardo, M. C. et al. A genome-wide association study identifies six susceptibility loci for chronic lymphocytic leukemia. *Nat. Genet.* **40**, 1204–1210 (2008).
136. Berndt, S. I. et al. Meta-analysis of genome-wide association studies discovers multiple loci for chronic lymphocytic leukemia. *Nat Commun* **7**, 10933 (2016).
137. Richardson, D. B., Wing, S., Schroeder, J., Schmitz-Feuerhake, I. & Hoffmann, W. Ionizing radiation and chronic lymphocytic leukemia. *Environ. Health Perspect.* **113**, 1–5 (2005).
138. Radivoyevitch, T., Sachs, R. K., Gale, R. P., Smith, M. R. & Hill, B. T. Ionizing radiation exposures in treatments of solid neoplasms are not associated with subsequent increased risks of chronic lymphocytic leukemia. *Leuk. Res.* **43**, 9–12 (2016).
139. Vrijheid, M. et al. Ionizing radiation and risk of chronic lymphocytic leukemia in the 15-country study of nuclear industry workers. *Radiat. Res.* **170**, 661–665 (2008).
140. Schwartz, G. G. & Klug, M. G. Incidence rates of chronic lymphocytic leukemia in US states are associated with residential radon levels. *Future Oncol* **12**, 165–174 (2016).
141. Oancea, S. C. et al. County level incidence rates of chronic lymphocytic leukemia are associated with residential radon levels. *Future Oncol* **13**, 1873–1881 (2017).
142. Kreuzer, M., Sobotzki, C., Fenske, N., Marsh, J. W. & Schnelzer, M. Leukaemia mortality and low-dose ionising radiation in the WISMUT uranium miner cohort (1946-2013). *Occup Environ Med* **74**, 252–258 (2017).
143. Portier, C. J. & Wolfe, M. S. Assessment of health effects from exposure to power-line frequency electric and magnetic fields. NIH publication **98**, 3981 (1998).
144. Ahlbom, I. C. et al. Review of the epidemiologic literature on EMF and Health. *Environ. Health Perspect.* **109 Suppl 6**, 911–933 (2001).
145. Tynes, T. & Haldorsen, T. Residential and occupational exposure to 50 Hz magnetic fields and hematological cancers in Norway. *Cancer Causes Control* **14**, 715–720 (2003).
146. Smedby, K. E. et al. Ultraviolet radiation exposure and risk of malignant lymphomas. *J. Natl. Cancer Inst.* **97**, 199–209 (2005).



147. Kricker, A. et al. Personal sun exposure and risk of non Hodgkin lymphoma: a pooled analysis from the Interlymph Consortium. *Int. J. Cancer* **122**, 144–154 (2008).
148. Burmeister, L. F., Van Lier, S. F. & Isacson, P. Leukemia and farm practices in Iowa. *Am. J. Epidemiol.* **115**, 720–728 (1982).
149. Blair, A. & White, D. W. Leukemia cell types and agricultural practices in Nebraska. *Arch. Environ. Health* **40**, 211–214 (1985).
150. Brown, L. M. et al. Pesticide exposures and other agricultural risk factors for leukemia among men in Iowa and Minnesota. *Cancer Res.* **50**, 6585–6591 (1990).
151. Zheng, T., Blair, A., Zhang, Y., Weisenburger, D. D. & Zahm, S. H. Occupation and risk of non-Hodgkin's lymphoma and chronic lymphocytic leukemia. *J. Occup. Environ. Med.* **44**, 469–474 (2002).
152. Miligi, L. et al. Non-Hodgkin's lymphoma, leukemia, and exposures in agriculture: results from the Italian multicenter case-control study. *Am. J. Ind. Med.* **44**, 627–636 (2003).
153. Cocco, P. et al. Lymphoma risk and occupational exposure to pesticides: results of the Epilymph study. *Occup Environ Med* **70**, 91–98 (2013).
154. Slager, S. L. et al. Medical history, lifestyle, family history, and occupational risk factors for chronic lymphocytic leukemia/small lymphocytic lymphoma: the InterLymph Non-Hodgkin Lymphoma Subtypes Project. *J. Natl. Cancer Inst. Monographs* **2014**, 41–51 (2014).
155. Linet, M. S. et al. Chronic lymphocytic leukaemia: an overview of aetiology in light of recent developments in classification and pathogenesis. *Br. J. Haematol.* **139**, 672–686 (2007).
156. Tsai, H.-T. et al. Dietary factors and risk of chronic lymphocytic leukemia and small lymphocytic lymphoma: a pooled analysis of two prospective studies. *Cancer Epidemiol. Biomarkers Prev.* **19**, 2680–2684 (2010).
157. Solans, M. et al. Adherence to the Western, Prudent, and Mediterranean dietary patterns and chronic lymphocytic leukemia in the MCC-Spain study. *Haematologica* (2018). doi:10.3324/haematol.2018.192526
158. Linet, M. S. et al. Chronic lymphocytic leukemia and acquired disorders affecting the immune system: a case-control study. *J. Natl. Cancer Inst.* **77**, 371–378 (1986).
159. Landgren, O. et al. Patterns of autoimmunity and subsequent chronic lymphocytic leukemia in Nordic countries. *Blood* **108**, 292–296 (2006).
160. Cerhan, J. R. et al. Blood transfusions and risk of non-Hodgkin's lymphoma subtypes and chronic lymphocytic leukemia. *Cancer Epidemiol. Biomarkers Prev.* **10**, 361–368 (2001).



161. Landgren, O., Gridley, G., Check, D., Caporaso, N. E. & Morris Brown, L. Acquired immune-related and inflammatory conditions and subsequent chronic lymphocytic leukaemia. *Br. J. Haematol.* **139**, 791–798 (2007).
162. Landgren, O. et al. Respiratory tract infections and subsequent risk of chronic lymphocytic leukemia. *Blood* **109**, 2198–2201 (2007).
163. Zhang, S. & Kipps, T. J. The pathogenesis of chronic lymphocytic leukemia. *Annu Rev Pathol* **9**, 103–118 (2014).
164. Caligaris-Cappio, F., Gobbi, M., Bofill, M. & Janossy, G. Infrequent normal B lymphocytes express features of B-chronic lymphocytic leukemia. *J. Exp. Med.* **155**, 623–628 (1982).
165. Caligaris-Cappio, F. B-chronic lymphocytic leukemia: a malignancy of anti-self B cells. *Blood* **87**, 2615–2620 (1996).
166. Dono, M. et al. CD5+ B cells with the features of subepithelial B cells found in human tonsils. *Eur. J. Immunol.* **37**, 2138–2147 (2007).
167. Chiorazzi, N. & Ferrarini, M. Cellular origin(s) of chronic lymphocytic leukemia: cautionary notes and additional considerations and possibilities. *Blood* **117**, 1781–1791 (2011).
168. Forconi, F. et al. The normal IGHV1-69-derived B-cell repertoire contains stereotypic patterns characteristic of unmutated CLL. *Blood* **115**, 71–77 (2010).
169. Klein, U. et al. Gene expression profiling of B cell chronic lymphocytic leukemia reveals a homogeneous phenotype related to memory B cells. *J. Exp. Med.* **194**, 1625–1638 (2001).
170. Rosenwald, A. et al. Relation of gene expression phenotype to immunoglobulin mutation genotype in B cell chronic lymphocytic leukemia. *J. Exp. Med.* **194**, 1639–1647 (2001).
171. Seifert, M. et al. Cellular origin and pathophysiology of chronic lymphocytic leukemia. *J. Exp. Med.* **209**, 2183–2198 (2012).
172. Widhopf, G. F. et al. Nonstochastic pairing of immunoglobulin heavy and light chains expressed by chronic lymphocytic leukemia B cells is predicated on the heavy chain CDR3. *Blood* **111**, 3137–3144 (2008).
173. Lanemo Myhrinder, A. et al. A new perspective: molecular motifs on oxidized LDL, apoptotic cells, and bacteria are targets for chronic lymphocytic leukemia antibodies. *Blood* **111**, 3838–3848 (2008).



174. Chu, C. C. et al. Chronic lymphocytic leukemia antibodies with a common stereotypic rearrangement recognize nonmuscle myosin heavy chain IIA. *Blood* **112**, 5122–5129 (2008).
175. Catera, R. et al. Chronic lymphocytic leukemia cells recognize conserved epitopes associated with apoptosis and oxidation. *Mol. Med.* **14**, 665–674 (2008).
176. Kostareli, E. et al. Antigen receptor stereotypy across B-cell lymphoproliferations: the case of IGHV4-59/IGKV3-20 receptors with rheumatoid factor activity. *Leukemia* **26**, 1127–1131 (2012).
177. Steininger, C. et al. Recombinant antibodies encoded by IGHV1-69 react with pUL32, a phosphoprotein of cytomegalovirus and B-cell superantigen. *Blood* **119**, 2293–2301 (2012).
178. Stamatopoulos, K. Antigens in CLL: themes and variations. *Blood* **115**, 3855–3856 (2010).
179. Berdasco, M. & Esteller, M. Aberrant epigenetic landscape in cancer: how cellular identity goes awry. *Dev. Cell* **19**, 698–711 (2010).
180. Lipsanen, V., Leinonen, P., Alhonen, L. & Jänne, J. Hypomethylation of ornithine decarboxylase gene and erb-A1 oncogene in human chronic lymphatic leukemia. *Blood* **72**, 2042–2044 (1988).
181. Wahlfors, J. et al. Genomic hypomethylation in human chronic lymphocytic leukemia. *Blood* **80**, 2074–2080 (1992).
182. Stach, D. et al. Capillary electrophoretic analysis of genomic DNA methylation levels. *Nucleic Acids Res.* **31**, E2 (2003).
183. Fabris, S. et al. Biological and clinical relevance of quantitative global methylation of repetitive DNA sequences in chronic lymphocytic leukemia. *Epigenetics* **6**, 188–194 (2011).
184. Kulis, M. et al. Epigenomic analysis detects widespread gene-body DNA hypomethylation in chronic lymphocytic leukemia. *Nat. Genet.* **44**, 1236–1242 (2012).
185. Oakes, C. C. et al. DNA methylation dynamics during B cell maturation underlie a continuum of disease phenotypes in chronic lymphocytic leukemia. *Nat. Genet.* **48**, 253–264 (2016).
186. Mansouri, L., Wierzbinska, J. A., Plass, C. & Rosenquist, R. Epigenetic deregulation in chronic lymphocytic leukemia: Clinical and biological impact. *Semin. Cancer Biol.* (2018). doi:10.1016/j.semcancer.2018.02.001
187. Corcoran, M. et al. ZAP-70 methylation status is associated with ZAP-70 expression status in chronic lymphocytic leukemia. *Haematologica* **90**, 1078–1088 (2005).



188. Chantepie, S. P. et al. ZAP-70 intron1 DNA methylation status: determination by pyrosequencing in B chronic lymphocytic leukemia. *Leuk. Res.* **34**, 800–808 (2010).
189. Claus, R. et al. Quantitative DNA methylation analysis identifies a single CpG dinucleotide important for ZAP-70 expression and predictive of prognosis in chronic lymphocytic leukemia. *J. Clin. Oncol.* **30**, 2483–2491 (2012).
190. Kulis, M. et al. Whole-genome fingerprint of the DNA methylome during human B cell differentiation. *Nat. Genet.* **47**, 746–756 (2015).
191. Cahill, N. et al. 450K-array analysis of chronic lymphocytic leukemia cells reveals global DNA methylation to be relatively stable over time and similar in resting and proliferative compartments. *Leukemia* **27**, 150–158 (2013).
192. Oakes, C. C. et al. Evolution of DNA methylation is linked to genetic aberrations in chronic lymphocytic leukemia. *Cancer Discov* **4**, 348–361 (2014).
193. Landau, D. A. et al. Locally disordered methylation forms the basis of intratumor methylome variation in chronic lymphocytic leukemia. *Cancer Cell* **26**, 813–825 (2014).
194. Ozdağ, H. et al. Differential expression of selected histone modifier genes in human solid cancers. *BMC Genomics* **7**, 90 (2006).
195. Papakonstantinou, N. et al. The histone methyltransferase EZH2 as a novel prosurvival factor in clinically aggressive chronic lymphocytic leukemia. *Oncotarget* **7**, 35946–35959 (2016).
196. Varambally, S. et al. Genomic loss of microRNA-101 leads to overexpression of histone methyltransferase EZH2 in cancer. *Science* **322**, 1695–1699 (2008).
197. Papakonstantinou, N. et al. Differential microRNA profiles and their functional implications in different immunogenetic subsets of chronic lymphocytic leukemia. *Mol. Med.* **19**, 115–123 (2013).
198. Iannetti, A. et al. Regulation of p53 and Rb links the alternative NF- κ B pathway to EZH2 expression and cell senescence. *PLoS Genet.* **10**, e1004642 (2014).
199. Neo, W. H., Lim, J. F., Grumont, R., Gerondakis, S. & Su, I.-H. c-Rel regulates Ezh2 expression in activated lymphocytes and malignant lymphoid cells. *J. Biol. Chem.* **289**, 31693–31707 (2014).
200. Puente, X. S. et al. Non-coding recurrent mutations in chronic lymphocytic leukaemia. *Nature* **526**, 519–524 (2015).



201. Jones, S. et al. Frequent mutations of chromatin remodeling gene ARID1A in ovarian clear cell carcinoma. *Science* **330**, 228–231 (2010).
202. Wang, K. et al. Exome sequencing identifies frequent mutation of ARID1A in molecular subtypes of gastric cancer. *Nat. Genet.* **43**, 1219–1223 (2011).
203. Guan, B., Wang, T.-L. & Shih, I.-M. ARID1A, a factor that promotes formation of SWI/SNF-mediated chromatin remodeling, is a tumor suppressor in gynecologic cancers. *Cancer Res.* **71**, 6718–6727 (2011).
204. Van Rechem, C., Boulay, G. & Leprince, D. HIC1 interacts with a specific subunit of SWI/SNF complexes, ARID1A/BAF250A. *Biochem. Biophys. Res. Commun.* **385**, 586–590 (2009).
205. Rodríguez, D. et al. Mutations in CHD2 cause defective association with active chromatin in chronic lymphocytic leukemia. *Blood* **126**, 195–202 (2015).
206. Quesada, V. et al. Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. *Nature genetics* (2011).
207. Edmunds, J. W., Mahadevan, L. C. & Clayton, A. L. Dynamic histone H3 methylation during gene induction: HYPB/Setd2 mediates all H3K36 trimethylation. *EMBO J.* **27**, 406–420 (2008).
208. Parker, H. et al. Genomic disruption of the histone methyltransferase SETD2 in chronic lymphocytic leukaemia. *Leukemia* **30**, 2179–2186 (2016).
209. Pellestor, F. Chromothripsis et péri-conceptologie. *Morphologie* **100**, 161 (2016).
210. Stephens, P. J. et al. Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell* **144**, 27–40 (2011).
211. Srivastava, S., Tsongalis, G. J. & Kaur, P. Recent advances in microRNA-mediated gene regulation in chronic lymphocytic leukemia. *Clin. Biochem.* **46**, 901–908 (2013).
212. Esquela-Kerscher, A. & Slack, F. J. Oncomirs - microRNAs with a role in cancer. *Nat. Rev. Cancer* **6**, 259–269 (2006).
213. Di Leva, G., Briskin, D. & Croce, C. M. MicroRNA in cancer: new hopes for antineoplastic chemotherapy. *Ups. J. Med. Sci.* **117**, 202–216 (2012).
214. Calin, G. A. et al. Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 15524–15529 (2002).



215. Pfeifer, D. et al. Genome-wide analysis of DNA copy number changes and LOH in CLL using high-density SNP arrays. *Blood* **109**, 1202–1210 (2007).
216. Calin, G. A. et al. A MicroRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. *N. Engl. J. Med.* **353**, 1793–1801 (2005).
217. Raveche, E. S. et al. Abnormal microRNA-16 locus with synteny to human 13q14 linked to CLL in NZB mice. *Blood* **109**, 5079–5086 (2007).
218. Cimmino, A. et al. miR-15 and miR-16 induce apoptosis by targeting BCL2. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 13944–13949 (2005).
219. Hanlon, K., Rudin, C. E. & Harries, L. W. Investigating the targets of MIR-15a and MIR-16-1 in patients with chronic lymphocytic leukemia (CLL). *PLoS ONE* **4**, e7169 (2009).
220. Fulci, V. et al. Quantitative technologies establish a novel microRNA profile of chronic lymphocytic leukemia. *Blood* **109**, 4944–4951 (2007).
221. Pekarsky, Y. et al. Tcl1 expression in chronic lymphocytic leukemia is regulated by miR-29 and miR-181. *Cancer Res.* **66**, 11590–11593 (2006).
222. Mott, J. L., Kobayashi, S., Bronk, S. F. & Gores, G. J. mir-29 regulates Mcl-1 protein expression and apoptosis. *Oncogene* **26**, 6133–6140 (2007).
223. Herling, M. et al. TCL1 shows a regulated expression pattern in chronic lymphocytic leukemia that correlates with molecular subtypes and proliferative state. *Leukemia* **20**, 280–285 (2006).
224. Gaudio, E. et al. Tcl1 interacts with Atm and enhances NF- κ B activation in hematologic malignancies. *Blood* **119**, 180–187 (2012).
225. Visone, R., Veronese, A., Balatti, V. & Croce, C. M. MiR-181b: new perspective to evaluate disease progression in chronic lymphocytic leukemia. *Oncotarget* **3**, 195–202 (2012).
226. Calin, G. A. & Croce, C. M. Genomics of chronic lymphocytic leukemia microRNAs as new players with clinical significance. *Semin. Oncol.* **33**, 167–173 (2006).
227. Pallasch, C. P. et al. miRNA deregulation by epigenetic silencing disrupts suppression of the oncogene PLAG1 in chronic lymphocytic leukemia. *Blood* **114**, 3255–3264 (2009).
228. Baer, C. et al. Extensive promoter DNA hypermethylation and hypomethylation is associated with aberrant microRNA expression in chronic lymphocytic leukemia. *Cancer Res.* **72**, 3775–3785 (2012).



229. Wong, K.-Y. et al. Epigenetic inactivation of the MIR129-2 in hematological malignancies. *J Hematol Oncol* **6**, 16 (2013).
230. Wang, L. Q. et al. Epigenetic inactivation of miR-9 family microRNAs in chronic lymphocytic leukemia--implications on constitutive activation of NFκB pathway. *Mol. Cancer* **12**, 173 (2013).
231. Calin, G. A. et al. MicroRNA profiling reveals distinct signatures in B cell chronic lymphocytic leukemias. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 11755–11760 (2004).
232. Rossi, S. et al. microRNA fingerprinting of CLL patients with chromosome 17p deletion identify a miR-21 score that stratifies early survival. *Blood* **116**, 945–952 (2010).
233. Asslaber, D. et al. microRNA-34a expression correlates with MDM2 SNP309 polymorphism and treatment-free survival in chronic lymphocytic leukemia. *Blood* **115**, 4191–4197 (2010).
234. Döhner, H. et al. Genomic aberrations and survival in chronic lymphocytic leukemia. *N. Engl. J. Med.* **343**, 1910–1916 (2000).
235. Klein, U. et al. The DLEU2/miR-15a/16-1 cluster controls B cell proliferation and its deletion leads to chronic lymphocytic leukemia. *Cancer Cell* **17**, 28–40 (2010).
236. Lerner, M. et al. DLEU2, frequently deleted in malignancy, functions as a critical host gene of the cell cycle inhibitory microRNAs miR-15a and miR-16-1. *Exp. Cell Res.* **315**, 2941–2952 (2009).
237. Garding, A. et al. Epigenetic upregulation of lncRNAs at 13q14.3 in leukemia is linked to the In Cis downregulation of a gene cluster that targets NF-κB. *PLoS Genet.* **9**, e1003373 (2013).
238. Blume, C. J. et al. p53-dependent non-coding RNA networks in chronic lymphocytic leukemia. *Leukemia* **29**, 2015–2023 (2015).
239. Naganuma, T. et al. Alternative 3'-end processing of long noncoding RNA initiates construction of nuclear paraspeckles. *EMBO J.* **31**, 4020–4034 (2012).
240. Sattari, A. et al. Upregulation of long noncoding RNA MIAT in aggressive form of chronic lymphocytic leukemias. *Oncotarget* **7**, 54174–54182 (2016).
241. Ip, J. Y. & Nakagawa, S. Long non-coding RNAs in nuclear bodies. *Dev. Growth Differ.* **54**, 44–54 (2012).



242. Wang, L. Q., Wong, K. Y., Li, Z. H. & Chim, C. S. Epigenetic silencing of tumor suppressor long non-coding RNA BM742401 in chronic lymphocytic leukemia. *Oncotarget* **7**, 82400–82410 (2016).
243. Ferreira, P. G. et al. Transcriptome characterization by RNA sequencing identifies a major molecular and clinical subdivision in chronic lymphocytic leukemia. *Genome Res.* **24**, 212–226 (2014).
244. Ronchetti, D. et al. lncRNA profiling in early-stage chronic lymphocytic leukemia identifies transcriptional fingerprints with relevance in clinical outcome. *Blood Cancer J* **6**, e468 (2016).
245. Caligaris-Cappio, F., Bertilaccio, M. T. S. & Scielzo, C. How the microenvironment wires the natural history of chronic lymphocytic leukemia. *Semin. Cancer Biol.* **24**, 43–48 (2014).
246. Burger, J. A., Ghia, P., Rosenwald, A. & Caligaris-Cappio, F. The microenvironment in mature B-cell malignancies: a target for new treatment strategies. *Blood* **114**, 3367–3375 (2009).
247. Ghia, P. et al. Chronic lymphocytic leukemia B cells are endowed with the capacity to attract CD4+, CD40L+ T cells by producing CCL22. *Eur. J. Immunol.* **32**, 1403–1413 (2002).
248. Ruan, J. et al. Magnitude of stromal hemangiogenesis correlates with histologic subtype of non-Hodgkin's lymphoma. *Clin. Cancer Res.* **12**, 5622–5631 (2006).
249. Tsukada, N., Burger, J. A., Zvaifler, N. J. & Kipps, T. J. Distinctive features of 'nurselike' cells that differentiate in the context of chronic lymphocytic leukemia. *Blood* **99**, 1030–1037 (2002).
250. Herishanu, Y. et al. The lymph node microenvironment promotes B-cell receptor signaling, NF-kappaB activation, and tumor proliferation in chronic lymphocytic leukemia. *Blood* **117**, 563–574 (2011).
251. Choi, M. Y., Kashyap, M. K. & Kumar, D. The chronic lymphocytic leukemia microenvironment: Beyond the B-cell receptor. *Best Pract Res Clin Haematol* **29**, 40–53 (2016).
252. Burger, J. A. et al. Blood-derived nurse-like cells protect chronic lymphocytic leukemia B cells from spontaneous apoptosis through stromal cell-derived factor-1. *Blood* **96**, 2655–2663 (2000).
253. Filip, A. A. et al. Circulating microenvironment of CLL: are nurse-like cells related to tumor-associated macrophages? *Blood Cells Mol. Dis.* **50**, 263–270 (2013).
254. Jia, L. et al. Extracellular HMGB1 promotes differentiation of nurse-like cells in chronic lymphocytic leukemia. *Blood* **123**, 1709–1719 (2014).



255. Cols, M. et al. Stromal endothelial cells establish a bidirectional crosstalk with chronic lymphocytic leukemia cells through the TNF-related factors BAFF, APRIL, and CD40L. *J. Immunol.* **188**, 6071–6083 (2012).
256. Seiffert, M. et al. Soluble CD14 is a novel monocyte-derived survival factor for chronic lymphocytic leukemia cells, which is induced by CLL cells in vitro and present at abnormally high levels in vivo. *Blood* **116**, 4223–4230 (2010).
257. Filip, A. A., Ciseł, B. & Wąsik-Szczepanek, E. Guilty bystanders: nurse-like cells as a model of microenvironmental support for leukemic lymphocytes. *Clin. Exp. Med.* **15**, 73–83 (2015).
258. Purroy, N. et al. Co-culture of primary CLL cells with bone marrow mesenchymal cells, CD40 ligand and CpG ODN promotes proliferation of chemoresistant CLL cells phenotypically comparable to those proliferating in vivo. *Oncotarget* **6**, 7632–7643 (2015).
259. Jitschin, R. et al. Stromal cell-mediated glycolytic switch in CLL cells involves Notch-c-Myc signaling. *Blood* **125**, 3432–3436 (2015).
260. Fukuda, T. et al. Antisera induced by infusions of autologous Ad-CD154-leukemia B cells identify ROR1 as an oncofetal antigen and receptor for Wnt5a. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 3047–3052 (2008).
261. Daneshmanesh, A. H. et al. Ror1, a cell surface receptor tyrosine kinase is expressed in chronic lymphocytic leukemia and may serve as a putative target for therapy. *Int. J. Cancer* **123**, 1190–1195 (2008).
262. Choudhury, A. et al. Silencing of ROR1 and FMOD with siRNA results in apoptosis of CLL cells. *Br. J. Haematol.* **151**, 327–335 (2010).
263. Bagnara, D. et al. A novel adoptive transfer model of chronic lymphocytic leukemia suggests a key role for T lymphocytes in the disease. *Blood* **117**, 5463–5472 (2011).
264. Riches, J. C. et al. T cells from CLL patients exhibit features of T-cell exhaustion but retain capacity for cytokine production. *Blood* **121**, 1612–1621 (2013).
265. Maki, G. et al. NK resistance of tumor cells from multiple myeloma and chronic lymphocytic leukemia patients: implication of HLA-G. *Leukemia* **22**, 998–1006 (2008).
266. Zhang, S. et al. Targeting chronic lymphocytic leukemia cells with a humanized monoclonal antibody specific for CD44. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 6127–6132 (2013).
267. Stevenson, F. K., Krysov, S., Davies, A. J., Steele, A. J. & Packham, G. B-cell receptor signaling in chronic lymphocytic leukemia. *Blood* **118**, 4313–4320 (2011).



268. Mockridge, C. I. et al. Reversible anergy of sIgM-mediated signaling in the two subsets of CLL defined by VH-gene mutational status. *Blood* **109**, 4424–4431 (2007).
269. Muzio, M. et al. Constitutive activation of distinct BCR-signaling pathways in a subset of CLL patients: a molecular signature of anergy. *Blood* **112**, 188–195 (2008).
270. Lanham, S. et al. Differential signaling via surface IgM is associated with VH gene mutational status and CD38 expression in chronic lymphocytic leukemia. *Blood* **101**, 1087–1093 (2003).
271. Chen, L. et al. ZAP-70 enhances IgM signaling independent of its kinase activity in chronic lymphocytic leukemia. *Blood* **111**, 2685–2692 (2008).
272. Gobessi, S. et al. ZAP-70 enhances B-cell-receptor signaling despite absent or inefficient tyrosine kinase activation in chronic lymphocytic leukemia and lymphoma B cells. *Blood* **109**, 2032–2039 (2007).
273. Wickremasinghe, R. G., Prentice, A. G. & Steele, A. J. Aberrantly activated anti-apoptotic signalling mechanisms in chronic lymphocytic leukaemia cells: clues to the identification of novel therapeutic targets. *Br. J. Haematol.* **153**, 545–556 (2011).
274. Zupo, S. et al. CD38 expression distinguishes two groups of B-cell chronic lymphocytic leukemias with different responses to anti-IgM antibodies and propensity to apoptosis. *Blood* **88**, 1365–1374 (1996).
275. Deaglio, S. et al. CD38 is a signaling molecule in B-cell chronic lymphocytic leukemia cells. *Blood* **102**, 2146–2155 (2003).
276. Deaglio, S. et al. CD38 and CD100 lead a network of surface receptors relaying positive signals for B-CLL growth and survival. *Blood* **105**, 3042–3050 (2005).
277. Poggi, A. et al. Engagement of CD31 delivers an activating signal that contributes to the survival of chronic lymphocytic leukaemia cells. *Br. J. Haematol.* **151**, 252–264 (2010).
278. Muzio, M. et al. Expression and function of toll like receptors in chronic lymphocytic leukaemia cells. *Br. J. Haematol.* **144**, 507–516 (2009).
279. Liang, X. et al. Toll-like receptor 9 signaling by CpG-B oligodeoxynucleotides induces an apoptotic pathway in human chronic lymphocytic leukemia B cells. *Blood* **115**, 5041–5052 (2010).
280. Fonte, E. et al. In vitro sensitivity of CLL cells to fludarabine may be modulated by the stimulation of Toll-like receptors. *Clin. Cancer Res.* **19**, 367–379 (2013).



281. Bertilaccio, M. T. S. et al. Lack of TIR8/SIGIRR triggers progression of chronic lymphocytic leukemia in mouse models. *Blood* **118**, 660–669 (2011).
282. Burger, J. A. Chemokines and chemokine receptors in chronic lymphocytic leukemia (CLL): from understanding the basics towards therapeutic targeting. *Semin. Cancer Biol.* **20**, 424–430 (2010).
283. Panayiotidis, P., Jones, D., Ganeshaguru, K., Foroni, L. & Hoffbrand, A. V. Human bone marrow stromal cells prevent apoptosis and support the survival of chronic lymphocytic leukaemia cells in vitro. *Br. J. Haematol.* **92**, 97–103 (1996).
284. Lagneaux, L., Delforge, A., Bron, D., De Bruyn, C. & Stryckmans, P. Chronic lymphocytic leukemic B cells but not normal B cells are rescued from apoptosis by contact with normal bone marrow stromal cells. *Blood* **91**, 2387–2396 (1998).
285. Calissano, C. et al. Intraclonal complexity in chronic lymphocytic leukemia: fractions enriched in recently born/divided and older/quiescent cells. *Mol. Med.* **17**, 1374–1382 (2011).
286. Richardson, S. J. et al. ZAP-70 expression is associated with enhanced ability to respond to migratory and survival signals in B-cell chronic lymphocytic leukemia (B-CLL). *Blood* **107**, 3584–3592 (2006).
287. Scielzo, C. et al. HS1 has a central role in the trafficking and homing of leukemic B cells. *Blood* **116**, 3537–3546 (2010).
288. Gomez, T. S. et al. HS1 functions as an essential actin-regulatory adaptor protein at the immune synapse. *Immunity* **24**, 741–752 (2006).
289. Dighiero, G. & Binet, J.-L. When and How to Treat Chronic Lymphocytic Leukemia. *New England Journal of Medicine* **343**, 1799–1801 (2000).
290. Giles, F. J., O'Brien, S. M. & Keating, M. J. Chronic lymphocytic leukemia in (Richter's) transformation. *Semin. Oncol.* **25**, 117–125 (1998).
291. Tamaru, J.-I. 2016 revision of the WHO classification of lymphoid neoplasms. *Rinsho Ketsueki* **58**, 2188–2193 (2017).
292. Tsimberidou, A. M. et al. Assessment of chronic lymphocytic leukemia and small lymphocytic lymphoma by absolute lymphocyte counts in 2,126 patients: 20 years of experience at the University of Texas M.D. Anderson Cancer Center. *J. Clin. Oncol.* **25**, 4648–4656 (2007).
293. Daudignon, A. et al. Increased trisomy 12 frequency and a biased IgVH 3-21 gene usage characterize small lymphocytic lymphoma. *Leuk. Res.* **34**, 580–584 (2010).



294. Martínez-Trillos, A. et al. The mutational landscape of small lymphocytic lymphoma compared to non-early stage chronic lymphocytic leukemia. *Leukemia & Lymphoma* **0**, 1–9 (2017).
295. Strati, P. & Shanafelt, T. D. Monoclonal B-cell lymphocytosis and early-stage chronic lymphocytic leukemia: diagnosis, natural history, and risk stratification. *Blood* **126**, 454–462 (2015).
296. Marti, G. E. et al. Diagnostic criteria for monoclonal B-cell lymphocytosis. *Br. J. Haematol.* **130**, 325–332 (2005).
297. Shanafelt, T. D., Ghia, P., Lanasa, M. C., Landgren, O. & Rawstron, A. C. Monoclonal B-cell lymphocytosis (MBL): biology, natural history and clinical management. *Leukemia* **24**, 512–520 (2010).
298. Fazi, C. et al. General population low-count CLL-like MBL persists over time without clinical progression, although carrying the same cytogenetic abnormalities of CLL. *Blood* **118**, 6618–6625 (2011).
299. Rawstron, A. C. et al. Monoclonal B-cell lymphocytosis and chronic lymphocytic leukemia. *N. Engl. J. Med.* **359**, 575–583 (2008).
300. Mulligan, C. S., Thomas, M. E. & Mulligan, S. P. Monoclonal B-cell lymphocytosis and chronic lymphocytic leukemia. *N. Engl. J. Med.* **359**, 2065–2066; author reply 2066 (2008).
301. Mowery, Y. M. & Lanasa, M. C. Clinical aspects of monoclonal B-cell lymphocytosis. *Cancer Control* **19**, 8–17 (2012).
302. Shanafelt, T. D. et al. Survival of patients with clinically identified monoclonal B-cell lymphocytosis (MBL) relative to the age- and sex-matched general population. *Leukemia* **26**, 373–376 (2012).
303. Moreira, J. et al. Infectious complications among individuals with clinical monoclonal B-cell lymphocytosis (MBL): a cohort study of newly diagnosed cases compared to controls. *Leukemia* **27**, 136–141 (2013).
304. Ghia, P. et al. Monoclonal CD5+ and CD5- B-lymphocyte expansions are frequent in the peripheral blood of the elderly. *Blood* **103**, 2337–2342 (2004).
305. Nieto, W. G. et al. Increased frequency (12%) of circulating chronic lymphocytic leukemia-like B-cell clones in healthy subjects using a highly sensitive multicolor flow cytometry approach. *Blood* **114**, 33–37 (2009).
306. Rossi, D., Spina, V. & Gaidano, G. Biology and treatment of Richter syndrome. *Blood* **131**, 2761–2772 (2018).



307. Mao, Z. et al. IgVH mutational status and clonality analysis of Richter's transformation: diffuse large B-cell lymphoma and Hodgkin lymphoma in association with B-cell chronic lymphocytic leukemia (B-CLL) represent 2 different pathways of disease evolution. *Am. J. Surg. Pathol.* **31**, 1605–1614 (2007).
308. Soilleux, E. J. et al. Diagnostic dilemmas of high-grade transformation (Richter's syndrome) of chronic lymphocytic leukaemia: results of the phase II National Cancer Research Institute CHOP-OR clinical trial specialist haemato-pathology central review. *Histopathology* **69**, 1066–1076 (2016).
309. Rossi, D. et al. The genetics of Richter syndrome reveals disease heterogeneity and predicts survival after transformation. *Blood* **117**, 3391–3401 (2011).
310. Fabbri, G. et al. Genetic lesions associated with chronic lymphocytic leukemia transformation to Richter syndrome. *J. Exp. Med.* **210**, 2273–2288 (2013).
311. Xiao, W. et al. Hodgkin lymphoma variant of Richter transformation: morphology, Epstein-Barr virus status, clonality, and survival analysis-with comparison to Hodgkin-like lesion. *Hum. Pathol.* **55**, 108–116 (2016).
312. Parikh, S. A. et al. Hodgkin transformation of chronic lymphocytic leukemia: Incidence, outcomes, and comparison to de novo Hodgkin lymphoma. *Am. J. Hematol.* **90**, 334–338 (2015).
313. Molica, S. Second neoplasms in chronic lymphocytic leukemia: incidence and pathogenesis with emphasis on the role of different therapies. *Leuk. Lymphoma* **46**, 49–54 (2005).
314. Beiggi, S. et al. Increased risk of second malignancies in chronic lymphocytic leukaemia patients as compared with follicular lymphoma patients: a Canadian population-based study. *Br J Cancer* **109**, 1287–1290 (2013).
315. Schöllkopf, C., Rosendahl, D., Rostgaard, K., Pipper, C. & Hjalgrim, H. Risk of second cancer after chronic lymphocytic leukemia. *Int. J. Cancer* **121**, 151–156 (2007).
316. Tsimberidou, A.-M. et al. Other malignancies in chronic lymphocytic leukemia/small lymphocytic lymphoma. *J. Clin. Oncol.* **27**, 904–910 (2009).
317. Bernués, M. et al. Genetics of lymphocytes influences the emergence of second cancer in chronic lymphocytic leukemia. *Anticancer Res.* **34**, 2311–2314 (2014).
318. Tsang, M. & Parikh, S. A. A Concise Review of Autoimmune Cytopenias in Chronic Lymphocytic Leukemia. *Curr Hematol Malig Rep* **12**, 29–38 (2017).



319. Moreno, C. et al. Autoimmune cytopenia in chronic lymphocytic leukemia: prevalence, clinical associations, and prognostic significance. *Blood* **116**, 4771–4776 (2010).
320. Zent, C. S. et al. The prognostic significance of cytopenia in chronic lymphocytic leukaemia/small lymphocytic lymphoma. *Br. J. Haematol.* **141**, 615–621 (2008).
321. Weiss, R. B., Freiman, J., Kweder, S. L., Diehl, L. F. & Byrd, J. C. Hemolytic anemia after fludarabine therapy for chronic lymphocytic leukemia. *J. Clin. Oncol.* **16**, 1885–1889 (1998).
322. Gonzalez, H. et al. Severe autoimmune hemolytic anemia in eight patients treated with fludarabine. *Hematol Cell Ther* **40**, 113–118 (1998).
323. Reda, G. et al. Low-dose alemtuzumab-associated immune thrombocytopenia in chronic lymphocytic leukemia. *Am. J. Hematol.* **87**, 936–937 (2012).
324. Dearden, C. et al. The prognostic significance of a positive direct antiglobulin test in chronic lymphocytic leukemia: a beneficial effect of the combination of fludarabine and cyclophosphamide on the incidence of hemolytic anemia. *Blood* **111**, 1820–1826 (2008).
325. Visco, C. et al. Autoimmune cytopenias in chronic lymphocytic leukemia at disease presentation in the modern treatment era: is stage C always stage C? *Leuk. Lymphoma* **55**, 1261–1265 (2014).
326. Kyasa, M. J., Parrish, R. S., Schichman, S. A. & Zent, C. S. Autoimmune cytopenia does not predict poor prognosis in chronic lymphocytic leukemia/small lymphocytic lymphoma. *Am. J. Hematol.* **74**, 1–8 (2003).
327. Parikh, S. A. et al. Hypogammaglobulinemia in newly diagnosed chronic lymphocytic leukemia: Natural history, clinical correlates, and outcomes. *Cancer* **121**, 2883–2891 (2015).
328. Chiorazzi, N. et al. T cell helper defect in patients with chronic lymphocytic leukemia. *J. Immunol.* **122**, 1087–1090 (1979).
329. Kay, N. E. Abnormal T-cell subpopulation function in CLL: excessive suppressor (T gamma) and deficient helper (T mu) activity with respect to B-cell proliferation. *Blood* **57**, 418–420 (1981).
330. Kontoyiannis, D. P. et al. Impaired bactericidal but not fungicidal activity of polymorphonuclear neutrophils in patients with chronic lymphocytic leukemia. *Leuk. Lymphoma* **54**, 1730–1733 (2013).
331. Nowakowski, G. S. et al. Percentage of smudge cells on routine blood smear predicts survival in chronic lymphocytic leukemia. *J. Clin. Oncol.* **27**, 1844–1849 (2009).



332. Johansson, P. et al. Percentage of smudge cells determined on routine blood smears is a novel prognostic factor in chronic lymphocytic leukemia. *Leuk. Res.* **34**, 892–898 (2010).
333. Gogia, A. et al. Prognostic and predictive significance of smudge cell percentage on routine blood smear in chronic lymphocytic leukemia. *Clin Lymphoma Myeloma Leuk* **14**, 514–517 (2014).
334. Matutes, E. et al. The immunological profile of B-cell disorders and proposal of a scoring system for the diagnosis of CLL. *Leukemia* **8**, 1640–1645 (1994).
335. Moreau, E. J. et al. Improvement of the chronic lymphocytic leukemia scoring system with the monoclonal antibody SN8 (CD79b). *Am. J. Clin. Pathol.* **108**, 378–382 (1997).
336. Alapat, D. et al. Diagnostic usefulness and prognostic impact of CD200 expression in lymphoid malignancies and plasma cell myeloma. *Am. J. Clin. Pathol.* **137**, 93–100 (2012).
337. Gahrton, G., Juliusson, G., Robèrt, K. H. & Friberg, K. Role of chromosomal abnormalities in chronic lymphocytic leukemia. *Blood Rev.* **1**, 183–192 (1987).
338. Han, T., Henderson, E. S., Emrich, L. J. & Sandberg, A. A. Prognostic significance of karyotypic abnormalities in B cell chronic lymphocytic leukemia: an update. *Semin. Hematol.* **24**, 257–263 (1987).
339. Juliusson, G. et al. Prognostic subgroups in B-cell chronic lymphocytic leukemia defined by specific chromosomal abnormalities. *N. Engl. J. Med.* **323**, 720–724 (1990).
340. Nguyen-Khac, F. [Cytogenetic markers in chronic lymphocytic leukemia]. *Ann. Biol. Clin. (Paris)* **68**, 273–276 (2010).
341. Decker, T. et al. Immunostimulatory CpG-oligonucleotides cause proliferation, cytokine production, and an immunogenic phenotype in chronic lymphocytic leukemia B cells. *Blood* **95**, 999–1006 (2000).
342. Decker, T. et al. Immunostimulatory CpG-oligonucleotides induce functional high affinity IL-2 receptors on B-CLL cells: costimulation with IL-2 results in a highly immunogenic phenotype. *Exp. Hematol.* **28**, 558–568 (2000).
343. Dicker, F., Schnittger, S., Haferlach, T., Kern, W. & Schoch, C. Immunostimulatory oligonucleotide-induced metaphase cytogenetics detect chromosomal aberrations in 80% of CLL patients: A study of 132 CLL cases with correlation to FISH, IgVH status, and CD38 expression. *Blood* **108**, 3152–3160 (2006).
344. DE Braekeleer, M. et al. Immunoglobulin gene translocations in chronic lymphocytic leukemia: A report of 35 patients and review of the literature. *Mol Clin Oncol* **4**, 682–694 (2016).



345. Xu-Monette, Z. Y. et al. Dysfunction of the TP53 tumor suppressor gene in lymphoid malignancies. *Blood* **119**, 3668–3683 (2012).
346. Rosenquist, R., Cortese, D., Bhoi, S., Mansouri, L. & Gunnarsson, R. Prognostic markers and their clinical applicability in chronic lymphocytic leukemia: where do we stand? *Leuk. Lymphoma* **54**, 2351–2364 (2013).
347. Gunnarsson, R. et al. Array-based genomic screening at diagnosis and during follow-up in chronic lymphocytic leukemia. *Haematologica* **96**, 1161–1169 (2011).
348. Zenz, T. et al. Monoallelic TP53 inactivation is associated with poor prognosis in chronic lymphocytic leukemia: results from a detailed genetic characterization with long-term follow-up. *Blood* **112**, 3322–3329 (2008).
349. Malcikova, J. et al. Monoallelic and biallelic inactivation of TP53 gene in chronic lymphocytic leukemia: selection, impact on survival, and response to DNA damage. *Blood* **114**, 5307–5314 (2009).
350. Tam, C. S. et al. De novo deletion 17p13.1 chronic lymphocytic leukemia shows significant clinical heterogeneity: the M. D. Anderson and Mayo Clinic experience. *Blood* **114**, 957–964 (2009).
351. Van Dyke, D. L. et al. The Dohner fluorescence in situ hybridization prognostic classification of chronic lymphocytic leukaemia (CLL): the CLL Research Consortium experience. *Br. J. Haematol.* **173**, 105–113 (2016).
352. Dicker, F. et al. The detection of TP53 mutations in chronic lymphocytic leukemia independently predicts rapid disease progression and is highly correlated with a complex aberrant karyotype. *Leukemia* **23**, 117–124 (2009).
353. Malcikova, J. et al. ERIC recommendations for TP53 mutation analysis in chronic lymphocytic leukemia—update on methodological approaches and results interpretation. *Leukemia* **32**, 1070–1080 (2018).
354. Döhner, H. et al. p53 gene deletion predicts for poor survival and non-response to therapy with purine analogs in chronic B-cell leukemias. *Blood* **85**, 1580–1589 (1995).
355. Chavez, J. C. et al. Genomic aberrations deletion 11q and deletion 17p independently predict for worse progression-free and overall survival after allogeneic hematopoietic cell transplantation for chronic lymphocytic leukemia. *Leuk. Res.* **38**, 1165–1172 (2014).
356. Rosenquist, R. et al. Clinical impact of recurrently mutated genes on lymphoma diagnostics: state-of-the-art and beyond. *Haematologica* **101**, 1002–1009 (2016).



357. Austen, B. et al. Mutation status of the residual ATM allele is an important determinant of the cellular response to chemotherapy and survival in patients with chronic lymphocytic leukemia containing an 11q deletion. *J. Clin. Oncol.* **25**, 5448–5457 (2007).
358. Jiang, Y. et al. ATM function and its relationship with ATM gene mutations in chronic lymphocytic leukemia with the recurrent deletion (11q22.3-23.2). *Blood Cancer J* **6**, e465 (2016).
359. Skowronska, A. et al. Biallelic ATM inactivation significantly reduces survival in patients treated on the United Kingdom Leukemia Research Fund Chronic Lymphocytic Leukemia 4 trial. *J. Clin. Oncol.* **30**, 4524–4532 (2012).
360. Rossi, D. et al. Disruption of BIRC3 associates with fludarabine chemorefractoriness in TP53 wild-type chronic lymphocytic leukemia. *Blood* **119**, 2854–2862 (2012).
361. Rossi, D. et al. Integrated mutational and cytogenetic analysis identifies new prognostic subgroups in chronic lymphocytic leukemia. *Blood* **121**, 1403–1412 (2013).
362. Matutes, E. et al. Trisomy 12 defines a group of CLL with atypical morphology: correlation between cytogenetic, clinical and laboratory features in 544 patients. *Br. J. Haematol.* **92**, 382–388 (1996).
363. Quijano, S. et al. Impact of trisomy 12, del(13q), del(17p), and del(11q) on the immunophenotype, DNA ploidy status, and proliferative rate of leukemic B-cells in chronic lymphocytic leukemia. *Cytometry B Clin Cytom* **74**, 139–149 (2008).
364. Sellmann, L. et al. Trisomy 19 is associated with trisomy 12 and mutated IGHV genes in B-chronic lymphocytic leukaemia. *Br. J. Haematol.* **138**, 217–220 (2007).
365. Ibbotson, R. et al. Coexistence of trisomies of chromosomes 12 and 19 in chronic lymphocytic leukemia occurs exclusively in the rare IgG-positive variant. *Leukemia* **26**, 170–172 (2012).
366. Balatti, V. et al. NOTCH1 mutations in CLL associated with trisomy 12. *Blood* **119**, 329–331 (2012).
367. Del Giudice, I. et al. NOTCH1 mutations in +12 chronic lymphocytic leukemia (CLL) confer an unfavorable prognosis, induce a distinctive transcriptional profiling and refine the intermediate prognosis of +12 CLL. *Haematologica* **97**, 437–441 (2012).
368. Cosson, A. et al. 14q deletions are associated with trisomy 12, NOTCH1 mutations and unmutated IGHV genes in chronic lymphocytic leukemia and small lymphocytic lymphoma. *Genes Chromosomes Cancer* **53**, 657–666 (2014).



369. Gunn, S. R. et al. Whole-genome scanning by array comparative genomic hybridization as a clinical tool for risk assessment in chronic lymphocytic leukemia. *J Mol Diagn* **10**, 442–451 (2008).
370. Dal Bo, M. et al. 13q14 deletion size and number of deleted cells both influence prognosis in chronic lymphocytic leukemia. *Genes Chromosomes Cancer* **50**, 633–643 (2011).
371. Ouillette, P. et al. The prognostic significance of various 13q14 deletions in chronic lymphocytic leukemia. *Clin. Cancer Res.* **17**, 6778–6790 (2011).
372. Hernández, J. A. et al. A high number of losses in 13q14 chromosome band is associated with a worse outcome and biological differences in patients with B-cell chronic lymphoid leukemia. *Haematologica* **94**, 364–371 (2009).
373. Stilgenbauer, S. et al. Clonal evolution in chronic lymphocytic leukemia: acquisition of high-risk genomic aberrations associated with unmutated VH, resistance to therapy, and short survival. *Haematologica* **92**, 1242–1245 (2007).
374. Chapiro, E. et al. Gain of the short arm of chromosome 2 (2p) is a frequent recurring chromosome aberration in untreated chronic lymphocytic leukemia (CLL) at advanced stages. *Leuk. Res.* **34**, 63–68 (2010).
375. Jarosova, M. et al. Gain of chromosome 2p in chronic lymphocytic leukemia: significant heterogeneity and a new recurrent dicentric rearrangement. *Leuk. Lymphoma* **51**, 304–313 (2010).
376. Edelmann, J. et al. High-resolution genomic profiling of chronic lymphocytic leukemia reveals new recurrent genomic alterations. *Blood* **120**, 4783–4794 (2012).
377. Nguyen-Khac, F., Borie, C., Callet-Bauchu, E., Eclache, V. & Struski, S. Cytogenetics in the management of chronic lymphocytic leukemia: an update by the Groupe francophone de cytogénétique hématologique (GFCH). *Ann. Biol. Clin. (Paris)* **74**, 561–567 (2016).
378. Baliakas, P. et al. Chromosomal translocations and karyotype complexity in chronic lymphocytic leukemia: a systematic reappraisal of classic cytogenetic data. *Am. J. Hematol.* **89**, 249–255 (2014).
379. Gunnarsson, R. et al. Large but not small copy-number alterations correlate to high-risk genomic aberrations and survival in chronic lymphocytic leukemia: a high-resolution genomic screening of newly diagnosed patients. *Leukemia* **24**, 211–215 (2010).
380. Kujawski, L. et al. Genomic complexity identifies patients with aggressive chronic lymphocytic leukemia. *Blood* **112**, 1993–2003 (2008).



381. Haferlach, C., Dicker, F., Schnittger, S., Kern, W. & Haferlach, T. Comprehensive genetic characterization of CLL: a study on 506 cases analysed with chromosome banding analysis, interphase FISH, IgV(H) status and immunophenotyping. *Leukemia* **21**, 2442–2451 (2007).
382. Schmidt, H. H. et al. Deregulation of the carbohydrate (chondroitin 4) sulfotransferase 11 (CHST11) gene in a B-cell chronic lymphocytic leukemia with a t(12;14)(q23;q32). *Oncogene* **23**, 6991–6996 (2004).
383. Aamot, H. V., Bjørnslett, M., Delabie, J. & Heim, S. t(14;22)(q32;q11) in non-Hodgkin lymphoma and myeloid leukaemia: molecular cytogenetic investigations. *Br. J. Haematol.* **130**, 845–851 (2005).
384. Véronèse, L. et al. Telomeres and chromosomal instability in chronic lymphocytic leukemia. *Leukemia* **27**, 490–493 (2013).
385. Kong, C. M., Lee, X. W. & Wang, X. Telomere shortening in human diseases. *FEBS J.* **280**, 3180–3193 (2013).
386. Artandi, S. E. & DePinho, R. A. Telomeres and telomerase in cancer. *Carcinogenesis* **31**, 9–18 (2010).
387. Bechter, O. E. et al. Telomere length and telomerase activity predict survival in patients with B cell chronic lymphocytic leukemia. *Cancer Res.* **58**, 4918–4922 (1998).
388. Roos, G. et al. Short telomeres are associated with genetic complexity, high-risk genomic aberrations, and short survival in chronic lymphocytic leukemia. *Blood* **111**, 2246–2252 (2008).
389. Rossi, D. et al. Telomere length is an independent predictor of survival, treatment requirement and Richter's syndrome transformation in chronic lymphocytic leukemia. *Leukemia* **23**, 1062–1072 (2009).
390. Mansouri, L. et al. Short telomere length is associated with NOTCH1/SF3B1/TP53 aberrations and poor outcome in newly diagnosed chronic lymphocytic leukemia patients. *Am. J. Hematol.* **88**, 647–651 (2013).
391. Ghione, P. et al. Progressive telomere shortening is part of the natural history of chronic lymphocytic leukaemia and impacts clinical outcome: evidences from long term follow-up. *Br. J. Haematol.* **181**, 693–695 (2018).
392. Guièze, R. et al. Telomere status in chronic lymphocytic leukemia with TP53 disruption. *Oncotarget* **7**, 56976–56985 (2016).



393. Rampazzo, E. et al. Telomere length and telomerase levels delineate subgroups of B-cell chronic lymphocytic leukemia with different biological characteristics and clinical outcomes. *Haematologica* **97**, 56–63 (2012).
394. Ishibe, N. et al. Telomere length and heavy-chain mutation status in familial chronic lymphocytic leukemia. *Leuk. Res.* **26**, 791–794 (2002).
395. Hultdin, M. et al. Association between telomere length and V(H) gene mutation status in chronic lymphocytic leukaemia: clinical and biological implications. *Br. J. Cancer* **88**, 593–598 (2003).
396. Damle, R. N. et al. Telomere length and telomerase activity delineate distinctive replicative features of the B-CLL subgroups defined by immunoglobulin V gene mutations. *Blood* **103**, 375–382 (2004).
397. Sellmann, L. et al. Telomeres and prognosis in patients with chronic lymphocytic leukaemia. *Int. J. Hematol.* **93**, 74–82 (2011).
398. Grabowski, P. et al. Telomere length as a prognostic parameter in chronic lymphocytic leukemia with special reference to VH gene mutation status. *Blood* **105**, 4807–4812 (2005).
399. Ramsay, A. J. et al. POT1 mutations cause telomere dysfunction in chronic lymphocytic leukemia. *Nat. Genet.* **45**, 526–530 (2013).
400. Speedy, H. E. et al. Germ line mutations in shelterin complex genes are associated with familial chronic lymphocytic leukemia. *Blood* **128**, 2319–2326 (2016).
401. Huang, S. J. et al. Clonal evolution as detected by interphase fluorescence in situ hybridization is associated with worse overall survival in a population-based analysis of patients with chronic lymphocytic leukemia in British Columbia, Canada. *Cancer Genet* **210**, 1–8 (2017).
402. Landau, D. A. et al. Mutations driving CLL and their evolution in progression and relapse. *Nature* **526**, 525–530 (2015).
403. Matsuda, F. et al. Structure and physical map of 64 variable segments in the 3'0.8-megabase region of the human immunoglobulin heavy-chain locus. *Nat. Genet.* **3**, 88–94 (1993).
404. Jain, P. et al. The absolute percent deviation of IGHV mutation rather than a 98% cut-off predicts survival of chronic lymphocytic leukaemia patients treated with fludarabine, cyclophosphamide and rituximab. *Br. J. Haematol.* **180**, 33–40 (2018).
405. Schroeder, H. W. & Dighiero, G. The pathogenesis of chronic lymphocytic leukemia: analysis of the antibody repertoire. *Immunol. Today* **15**, 288–294 (1994).



406. Hamblin, T. J., Davis, Z., Gardiner, A., Oscier, D. G. & Stevenson, F. K. Unmutated Ig V(H) genes are associated with a more aggressive form of chronic lymphocytic leukemia. *Blood* **94**, 1848–1854 (1999).
407. Damle, R. N. et al. Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. *Blood* **94**, 1840–1847 (1999).
408. Tobin, G. et al. Somatic mutated Ig V(H)3-21 genes characterize a new subset of chronic lymphocytic leukemia. *Blood* **99**, 2262–2264 (2002).
409. Oscier, D. G., Thompsett, A., Zhu, D. & Stevenson, F. K. Differential rates of somatic hypermutation in V(H) genes among subsets of chronic lymphocytic leukemia defined by chromosomal abnormalities. *Blood* **89**, 4153–4160 (1997).
410. Tobin, G. et al. Chronic lymphocytic leukemias utilizing the VH3-21 gene display highly restricted Vlambda2-14 gene use and homologous CDR3s: implicating recognition of a common antigen epitope. *Blood* **101**, 4952–4957 (2003).
411. Baliakas, P. et al. Not all IGHV3-21 chronic lymphocytic leukemias are equal: prognostic considerations. *Blood* **125**, 856–859 (2015).
412. Sutton, L.-A. et al. Extensive intraclonal diversification in a subgroup of chronic lymphocytic leukemia patients with stereotyped IGHV4-34 receptors: implications for ongoing interactions with antigen. *Blood* **114**, 4460–4468 (2009).
413. Widhopf, G. F. et al. Chronic lymphocytic leukemia B cells of more than 1% of patients express virtually identical immunoglobulins. *Blood* **104**, 2499–2504 (2004).
414. Ghiotto, F. et al. Remarkably similar antigen receptors among a subset of patients with chronic lymphocytic leukemia. *J. Clin. Invest.* **113**, 1008–1016 (2004).
415. Stamatopoulos, K. et al. Over 20% of patients with chronic lymphocytic leukemia carry stereotyped receptors: Pathogenetic implications and clinical correlations. *Blood* **109**, 259–270 (2007).
416. Kirkham, P. M., Mortari, F., Newton, J. A. & Schroeder, H. W. Immunoglobulin VH clan and family identity predicts variable domain structure and may influence antigen binding. *EMBO J.* **11**, 603–609 (1992).
417. Agathangelidis, A. et al. Stereotyped B-cell receptors in one-third of chronic lymphocytic leukemia: a molecular classification with implications for targeted therapies. *Blood* **119**, 4467–4475 (2012).



418. Darzentas, N. et al. A different ontogenesis for chronic lymphocytic leukemia cases carrying stereotyped antigen receptors: molecular and computational evidence. *Leukemia* **24**, 125–132 (2010).
419. Stamatopoulos, K., Agathangelidis, A., Rosenquist, R. & Ghia, P. Antigen receptor stereotypy in chronic lymphocytic leukemia. *Leukemia* **31**, 282–291 (2017).
420. Bomben, R. et al. Molecular and clinical features of chronic lymphocytic leukaemia with stereotyped B cell receptors: results from an Italian multicentre study. *Br. J. Haematol.* **144**, 492–506 (2009).
421. Rossi, D. et al. Stereotyped B-cell receptor is an independent risk factor of chronic lymphocytic leukemia transformation to Richter syndrome. *Clin. Cancer Res.* **15**, 4415–4422 (2009).
422. Rossi, D. et al. The prognostic value of TP53 mutations in chronic lymphocytic leukemia is independent of Del17p13: implications for overall survival and chemorefractoriness. *Clin. Cancer Res.* **15**, 995–1004 (2009).
423. Malcikova, J. et al. Detailed analysis of therapy-driven clonal evolution of TP53 mutations in chronic lymphocytic leukemia. *Leukemia* **29**, 877–885 (2015).
424. Landau, D. A. et al. Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. *Cell* **152**, 714–726 (2013).
425. Nadeu, F. et al. Clinical impact of clonal and subclonal TP53, SF3B1, BIRC3, NOTCH1, and ATM mutations in chronic lymphocytic leukemia. *Blood* **127**, 2122–2130 (2016).
426. Welcome - IARC TP53 Database. Available at: <http://p53.iarc.fr/>. (Accessed: 1st October 2018)
427. The TP53 Website - TP53 Database. Available at: <https://p53.fr/tp53-database>. (Accessed: 1st October 2018)
428. Varley, J. M. et al. Characterization of germline TP53 splicing mutations and their genetic and functional analysis. *Oncogene* **20**, 2647–2654 (2001).
429. Stankovic, T. et al. Microarray analysis reveals that TP53- and ATM-mutant B-CLLs share a defect in activating proapoptotic responses after DNA damage but are distinguished by major differences in activating prosurvival responses. *Blood* **103**, 291–300 (2004).
430. Navrkalova, V. et al. ATM mutations uniformly lead to ATM dysfunction in chronic lymphocytic leukemia: application of functional test using doxorubicin. *Haematologica* **98**, 1124–1131 (2013).



431. Mansouri, L. et al. Feasibility of targeted next-generation sequencing of the TP53 and ATM genes in chronic lymphocytic leukemia. *Leukemia* **28**, 694–696 (2014).
432. Austen, B. et al. Mutations in the ATM gene lead to impaired overall and treatment-free survival that is independent of IGVH mutation status in patients with B-CLL. *Blood* **106**, 3175–3182 (2005).
433. Skowronska, A. et al. ATM germline heterozygosity does not play a role in chronic lymphocytic leukemia initiation but influences rapid disease progression through loss of the remaining ATM allele. *Haematologica* **97**, 142–146 (2012).
434. Lozano-Santos, C. et al. Biallelic ATM alterations detected at diagnosis identify a subset of treatment-naïve chronic lymphocytic leukemia patients with reduced overall survival similar to patients with p53 deletion. *Leuk. Lymphoma* **58**, 859–865 (2017).
435. Stankovic, T. et al. ATM mutations and phenotypes in ataxia-telangiectasia families in the British Isles: expression of mutant ATM and the risk of leukemia, lymphoma, and breast cancer. *Am. J. Hum. Genet.* **62**, 334–345 (1998).
436. Taylor, A. M. R. & Byrd, P. J. Molecular pathology of ataxia telangiectasia. *J. Clin. Pathol.* **58**, 1009–1015 (2005).
437. Stankovic, T. et al. Ataxia telangiectasia mutated-deficient B-cell chronic lymphocytic leukemia occurs in pregerminal center cells and results in defective damage response and unrepaired chromosome damage. *Blood* **99**, 300–309 (2002).
438. Carter, A., Lin, K., Sherrington, P. D. & Pettitt, A. R. Detection of p53 dysfunction by flow cytometry in chronic lymphocytic leukaemia. *Br. J. Haematol.* **127**, 425–428 (2004).
439. Best, O. G. et al. A novel functional assay using etoposide plus nutlin-3a detects and distinguishes between ATM and TP53 mutations in CLL. *Leukemia* **22**, 1456–1459 (2008).
440. te Raa, G. D. et al. Assessment of p53 and ATM functionality in chronic lymphocytic leukemia by multiplex ligation-dependent probe amplification. *Cell Death Dis* **6**, e1852 (2015).
441. Rothenberg, E. V. T cell lineage commitment: identity and renunciation. *J. Immunol.* **186**, 6649–6655 (2011).
442. Santos, M. A. et al. Notch1 engagement by Delta-like-1 promotes differentiation of B lymphocytes to antibody-secreting cells. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 15454–15459 (2007).
443. Fabbri, G. et al. Analysis of the chronic lymphocytic leukemia coding genome: role of NOTCH1 mutational activation. *J. Exp. Med.* **208**, 1389–1401 (2011).



444. Puente, X. S. et al. Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. *Nature* **475**, 101–105 (2011).
445. Rossi, D. et al. Mutations of NOTCH1 are an independent predictor of survival in chronic lymphocytic leukemia. *Blood* **119**, 521–529 (2012).
446. Chiaretti, S. et al. NOTCH1, SF3B1, BIRC3 and TP53 mutations in patients with chronic lymphocytic leukemia undergoing first-line treatment: correlation with biological parameters and response to treatment. *Leuk. Lymphoma* **55**, 2785–2792 (2014).
447. Rosati, E. et al. Constitutively activated Notch signaling is involved in survival and apoptosis resistance of B-CLL cells. *Blood* **113**, 856–865 (2009).
448. Rossi, D. et al. Mutations of the SF3B1 splicing factor in chronic lymphocytic leukemia: association with progression and fludarabine-refractoriness. *Blood* **118**, 6904–6908 (2011).
449. Wang, L. et al. SF3B1 and other novel cancer genes in chronic lymphocytic leukemia. *New England Journal of Medicine* **365**, 2497–2506 (2011).
450. Schwaederlé, M. et al. Subclonal evolution involving SF3B1 mutations in chronic lymphocytic leukemia. *Leukemia* **27**, 1214–1217 (2013).
451. Raponi, S. et al. Biallelic BIRC3 inactivation in chronic lymphocytic leukaemia patients with 11q deletion identifies a subgroup with very aggressive disease. *Br. J. Haematol.* (2018). doi:10.1111/bjh.15405
452. Ngo, V. N. et al. Oncogenically active MYD88 mutations in human lymphoma. *Nature* **470**, 115–119 (2011).
453. Treon, S. P. et al. MYD88 L265P somatic mutation in Waldenström's macroglobulinemia. *N. Engl. J. Med.* **367**, 826–833 (2012).
454. Gachard, N. et al. IGHV gene features and MYD88 L265P mutation separate the three marginal zone lymphoma entities and Waldenström macroglobulinemia/lymphoplasmacytic lymphomas. *Leukemia* **27**, 183–189 (2013).
455. O'Neill, L. A. J. & Bowie, A. G. The family of five: TIR-domain-containing adaptors in Toll-like receptor signalling. *Nat. Rev. Immunol.* **7**, 353–364 (2007).
456. Qin, S.-C. et al. MYD88 mutations predict unfavorable prognosis in Chronic Lymphocytic Leukemia patients with mutated IGHV gene. *Blood Cancer J* **7**, (2017).
457. Rizzo, D. et al. IgM peak independently predicts treatment-free survival in chronic lymphocytic leukemia and correlates with accumulation of adverse oncogenetic events. *Leukemia* **29**, 337–345 (2015).



458. Martínez-Trillos, A. et al. Mutations in TLR/MYD88 pathway identify a subset of young chronic lymphocytic leukemia patients with favorable outcome. *Blood* **123**, 3790–3796 (2014).
459. Rigolin, G. M. et al. Extensive next-generation sequencing analysis in chronic lymphocytic leukemia at diagnosis: clinical and biological correlations. *J Hematol Oncol* **9**, 88 (2016).
460. Baliakas, P. et al. Prognostic relevance of MYD88 mutations in CLL: the jury is still out. *Blood* **126**, 1043–1044 (2015).
461. Ojha, J. et al. Identification of recurrent truncated DDX3X mutations in chronic lymphocytic leukaemia. *Br. J. Haematol.* **169**, 445–448 (2015).
462. Takahashi, K. et al. Clinical implications of cancer gene mutations in patients with chronic lymphocytic leukemia treated with lenalidomide. *Blood* **131**, 1820–1832 (2018).
463. Mansouri, L. et al. Functional loss of I κ B ϵ leads to NF- κ B deregulation in aggressive chronic lymphocytic leukemia. *J. Exp. Med.* **212**, 833–843 (2015).
464. Doménech, E. et al. New mutations in chronic lymphocytic leukemia identified by target enrichment and deep sequencing. *PLoS ONE* **7**, e38158 (2012).
465. Damm, F. et al. Acquired initiating mutations in early hematopoietic cells of CLL patients. *Cancer Discov* **4**, 1088–1101 (2014).
466. Ljungström, V. et al. Whole-exome sequencing in relapsing chronic lymphocytic leukemia: clinical impact of recurrent RPS15 mutations. *Blood* **127**, 1007–1016 (2016).
467. Wu, B. et al. MED12 mutations and NOTCH signalling in chronic lymphocytic leukaemia. *Br. J. Haematol.* **179**, 421–429 (2017).
468. Young, E. et al. EGR2 mutations define a new clinically aggressive subgroup of chronic lymphocytic leukemia. *Leukemia* **31**, 1547–1554 (2017).
469. Jeromin, S. et al. SF3B1 mutations correlated to cytogenetics and mutations in NOTCH1, FBXW7, MYD88, XPO1 and TP53 in 1160 untreated CLL patients. *Leukemia* **28**, 108–117 (2014).
470. Jain, P. et al. Clinical and molecular characteristics of XPO1 mutations in patients with chronic lymphocytic leukemia. *Am. J. Hematol.* **91**, E478–E479 (2016).
471. Tiacci, E. et al. BRAF mutations in hairy-cell leukemia. *N. Engl. J. Med.* **364**, 2305–2315 (2011).



472. Wan, P. T. C. et al. Mechanism of activation of the RAF-ERK signaling pathway by oncogenic mutations of B-RAF. *Cell* **116**, 855–867 (2004).
473. Jebaraj, B. M. C. et al. BRAF mutations in chronic lymphocytic leukemia. *Leuk. Lymphoma* **54**, 1177–1182 (2013).
474. Sellar, R. S. et al. BRAF(V600E) mutations are found in Richter syndrome and may allow targeted therapy in a subset of patients. *Br. J. Haematol.* **170**, 282–285 (2015).
475. O’Neil, J. et al. FBW7 mutations in leukemic cells mediate NOTCH pathway activation and resistance to gamma-secretase inhibitors. *J. Exp. Med.* **204**, 1813–1824 (2007).
476. Falisi, E. et al. B-cell receptor configuration and mutational analysis of patients with chronic lymphocytic leukaemia and trisomy 12 reveal recurrent molecular abnormalities. *Hematol Oncol* **32**, 22–30 (2014).
477. Amin, N. A. et al. A Quantitative Analysis of Subclonal and Clonal Gene Mutations before and after Therapy in Chronic Lymphocytic Leukemia. *Clin. Cancer Res.* **22**, 4525–4535 (2016).
478. Quijada-Álamo, M. et al. Next-generation sequencing and FISH studies reveal the appearance of gene mutations and chromosomal abnormalities in hematopoietic progenitors in chronic lymphocytic leukemia. *J Hematol Oncol* **10**, 83 (2017).
479. Leeksma, A. C. et al. Clonal diversity predicts adverse outcome in chronic lymphocytic leukemia. *Leukemia* **1** (2018). doi:10.1038/s41375-018-0215-9
480. Sutton, L.-A. et al. Targeted next-generation sequencing in chronic lymphocytic leukemia: a high-throughput yet tailored approach will facilitate implementation in a clinical setting. *Haematologica* **100**, 370–376 (2015).
481. Herling, C. D. et al. Complex karyotypes and KRAS and POT1 mutations impact outcome in CLL after chlorambucil-based chemotherapy or chemoimmunotherapy. *Blood* **128**, 395–404 (2016).
482. Clifford, R. et al. SAMHD1 is mutated recurrently in chronic lymphocytic leukemia and is involved in response to DNA damage. *Blood* **123**, 1021–1031 (2014).
483. Rossi, D. SAMHD1: a new gene for CLL. *Blood* **123**, 951–952 (2014).
484. Schuh, A. et al. Monitoring chronic lymphocytic leukemia progression by whole genome sequencing reveals heterogeneous clonal evolution patterns. *Blood* **120**, 4191–4196 (2012).



485. Kroll, J. et al. The BTB-kelch protein KLHL6 is involved in B-lymphocyte antigen receptor signaling and germinal center formation. *Mol. Cell. Biol.* **25**, 8531–8540 (2005).
486. Delgado, J., Villamor, N., López-Guillermo, A. & Campo, E. Genetic evolution in chronic lymphocytic leukaemia. *Best Pract Res Clin Haematol* **29**, 67–78 (2016).
487. Wu, C. J. CLL clonal heterogeneity: an ecology of competing subpopulations. *Blood* **120**, 4117–4118 (2012).
488. Ojha, J. et al. Deep sequencing identifies genetic heterogeneity and recurrent convergent evolution in chronic lymphocytic leukemia. *Blood* **125**, 492–498 (2015).
489. Rossi, D. et al. Clinical impact of small TP53 mutated subclones in chronic lymphocytic leukemia. *Blood* **123**, 2139–2147 (2014).
490. Sagatys, E. M. & Zhang, L. Clinical and Laboratory Prognostic Indicators in Chronic Lymphocytic Leukemia. *Cancer Control* **19**, 18–25 (2012).
491. Simonsson, B., Wibell, L. & Nilsson, K. Beta 2-microglobulin in chronic lymphocytic leukaemia. *Scand J Haematol* **24**, 174–180 (1980).
492. Di Giovanni, S., Valentini, G., Carducci, P. & Giallonardo, P. Beta-2-microglobulin is a reliable tumor marker in chronic lymphocytic leukemia. *Acta Haematol.* **81**, 181–185 (1989).
493. Gentile, M. et al. Predictive value of β 2-microglobulin (β 2-m) levels in chronic lymphocytic leukemia since Binet A stages. *Haematologica* **94**, 887–888 (2009).
494. Hallek, M. et al. Serum beta(2)-microglobulin and serum thymidine kinase are independent predictors of progression-free survival in chronic lymphocytic leukemia and immunocytoma. *Leuk. Lymphoma* **22**, 439–447 (1996).
495. Tsimberidou, A. M. et al. Beta-2 microglobulin (B2M) is an independent prognostic factor for clinical outcomes in patients with CLL treated with frontline fludarabine, cyclophosphamide, and rituximab (FCR) regardless of age, creatinine clearance (CrCl). *JCO* **25**, 7034–7034 (2007).
496. Delgado, J. et al. Beta2-microglobulin is a better predictor of treatment-free survival in patients with chronic lymphocytic leukaemia if adjusted according to glomerular filtration rate. *Br. J. Haematol.* **145**, 801–805 (2009).
497. International CLL-IPI working group. An international prognostic index for patients with chronic lymphocytic leukaemia (CLL-IPI): a meta-analysis of individual patient data. *Lancet Oncol.* **17**, 779–790 (2016).



498. Pan, L., Beverley, P. C. & Isaacson, P. G. Lactate dehydrogenase (LDH) isoenzymes and proliferative activity of lymphoid cells--an immunocytochemical study. *Clin. Exp. Immunol.* **86**, 240–245 (1991).
499. Weinberg, J. B. et al. Clinical and molecular predictors of disease severity and survival in chronic lymphocytic leukemia. *Am. J. Hematol.* **82**, 1063–1070 (2007).
500. Martin, W. et al. Serum-free light chain-a new biomarker for patients with B-cell non-Hodgkin lymphoma and chronic lymphocytic leukemia. *Transl Res* **149**, 231–235 (2007).
501. Sarris, K. et al. Prognostic significance of serum free light chains in chronic lymphocytic leukemia. *Adv Hematol* **2013**, 359071 (2013).
502. Pratt, G. et al. Abnormal serum free light chain ratios are associated with poor survival and may reflect biological subgroups in patients with chronic lymphocytic leukaemia. *Br. J. Haematol.* **144**, 217–222 (2009).
503. Maurer, M. J. et al. Monoclonal and polyclonal serum free light chains and clinical outcome in chronic lymphocytic leukemia. *Blood* **118**, 2821–2826 (2011).
504. Xu, W. et al. Prognostic significance of serum immunoglobulin paraprotein in patients with chronic lymphocytic leukemia. *Leuk. Res.* **35**, 1060–1065 (2011).
505. Matutes, E. & Polliack, A. A matter of life or early death in CLL: do serum immunoglobulin levels really count? *Leukemia & Lymphoma* **57**, 1501–1502 (2016).
506. Freeman, J. A. et al. Immunoglobulin G subclass deficiency and infection risk in 150 patients with chronic lymphocytic leukemia. *Leuk. Lymphoma* **54**, 99–104 (2013).
507. Svensson, T., Höglund, M. & Cherif, H. Clinical significance of serum immunoglobulin G subclass deficiency in patients with chronic lymphocytic leukemia. *Scand. J. Infect. Dis.* **45**, 537–542 (2013).
508. Shvidel, L. et al. Serum immunoglobulin levels at diagnosis have no prognostic significance in stage A chronic lymphocytic leukemia: a study of 1113 cases from the Israeli CLL Study Group. *Eur. J. Haematol.* **93**, 29–33 (2014).
509. Andersen, M. A. et al. Hypogammaglobulinemia in newly diagnosed chronic lymphocytic leukemia is a predictor of early death. *Leuk. Lymphoma* **57**, 1592–1599 (2016).
510. Sarfati, M. et al. Prognostic importance of serum soluble CD23 level in chronic lymphocytic leukemia. *Blood* **88**, 4259–4264 (1996).



511. Lowe, J., Brown, B., Hardie, D., Richardson, P. & Ling, N. Soluble forms of CD21 and CD23 antigens in the serum in B cell chronic lymphocytic leukaemia. *Immunol. Lett.* **20**, 103–109 (1989).
512. Beguin, Y. et al. Soluble CD23 and other receptors (CD4, CD8, CD25, CD71) in serum of patients with chronic lymphocytic leukemia. *Leukemia* **7**, 2019–2025 (1993).
513. Reinisch, W. et al. Soluble CD23 reliably reflects disease activity in B-cell chronic lymphocytic leukemia. *J. Clin. Oncol.* **12**, 2146–2152 (1994).
514. Knauf, W. U. et al. Serum levels of soluble CD23, but not soluble CD25, predict disease progression in early stage B-cell chronic lymphocytic leukemia. *Leuk. Lymphoma* **27**, 523–532 (1997).
515. Molica, S. et al. Cellular expression and serum circulating levels of CD23 in B-cell chronic lymphocytic leukemia. Implications for prognosis. *Haematologica* **81**, 428–433 (1996).
516. Schwarzmeier, J. D. et al. The role of soluble CD23 in distinguishing stable and progressive forms of B-chronic lymphocytic leukemia. *Leuk. Lymphoma* **43**, 549–554 (2002).
517. Meuleman, N. et al. Doubling time of soluble CD23: a powerful prognostic factor for newly diagnosed and untreated stage A chronic lymphocytic leukemia patients. *Leukemia* **22**, 1882–1890 (2008).
518. Kaaks, R. et al. Lag times between lymphoproliferative disorder and clinical diagnosis of chronic lymphocytic leukemia: a prospective analysis using plasma soluble CD23. *Cancer Epidemiol. Biomarkers Prev.* **24**, 538–545 (2015).
519. Grelier, A., Le Garff-Tavernier, M., Nauwelaers, F., Sarfati, M. & Merle-Beral, H. Soluble CD23 measurement by CBA: a convenient and reliable quantification method in chronic lymphocytic leukemia. *Cytometry B Clin Cytom* **86**, 91–97 (2014).
520. Ellims, P. H., Gan, T. E. & Van der Weyden, M. B. Thymidine kinase isoenzymes in chronic lymphocytic leukaemia. *Br. J. Haematol.* **49**, 479–481 (1981).
521. Di Raimondo, F. et al. Retrospective study of the prognostic role of serum thymidine kinase level in CLL patients with active disease treated with fludarabine. *Ann. Oncol.* **12**, 621–625 (2001).
522. Hallek, M. et al. Elevated Serum Thymidine Kinase Levels Identify a Subgroup at High Risk of Disease Progression in Early, Nonsmoldering Chronic Lymphocytic Leukemia. *Blood* **93**, 1732–1737 (1999).
523. Magnac, C. et al. Predictive value of serum thymidine kinase level for Ig-V mutational status in B-CLL. *Leukemia* **17**, 133–137 (2003).



524. Xu, W. et al. Serum thymidine kinase 1 concentration in Chinese patients with chronic lymphocytic leukemia and its correlation with other prognostic factors. *Int. J. Hematol.* **90**, 205–211 (2009).
525. Matthews, C. et al. Serum TK levels in CLL identify Binet stage A patients within biologically defined prognostic subgroups most likely to undergo disease progression. *Eur. J. Haematol.* **77**, 309–317 (2006).
526. Szánthó, E. et al. Serum thymidine kinase activity: analytical performance, age-related reference ranges and validation in chronic lymphocytic leukemia. *PLoS ONE* **9**, e91647 (2014).
527. Ibrahim, S. et al. CD38 expression as an important prognostic factor in B-cell chronic lymphocytic leukemia. *Blood* **98**, 181–186 (2001).
528. Damle, R. N. et al. CD38 expression labels an activated subset within chronic lymphocytic leukemia clones enriched in proliferating B cells. *Blood* **110**, 3352–3359 (2007).
529. Hamblin, T. J. et al. CD38 expression and immunoglobulin variable region mutations are independent prognostic variables in chronic lymphocytic leukemia, but CD38 expression may vary during the course of the disease. *Blood* **99**, 1023–1029 (2002).
530. Chan, A. C., Iwashima, M., Turck, C. W. & Weiss, A. ZAP-70: a 70 kd protein-tyrosine kinase that associates with the TCR zeta chain. *Cell* **71**, 649–662 (1992).
531. Scielzo, C. et al. ZAP-70 is expressed by normal and malignant human B-cell subsets of different maturational stage. *Leukemia* **20**, 689–695 (2006).
532. Chen, L. et al. Expression of ZAP-70 is associated with increased B-cell receptor signaling in chronic lymphocytic leukemia. *Blood* **100**, 4609–4614 (2002).
533. Efremov, D. G., Gobessi, S. & Longo, P. G. Signaling pathways activated by antigen-receptor engagement in chronic lymphocytic leukemia B-cells. *Autoimmun Rev* **7**, 102–108 (2007).
534. Crespo, M. et al. ZAP-70 expression as a surrogate for immunoglobulin-variable-region mutations in chronic lymphocytic leukemia. *N. Engl. J. Med.* **348**, 1764–1775 (2003).
535. Orchard, J. A. et al. ZAP-70 expression and prognosis in chronic lymphocytic leukaemia. *Lancet* **363**, 105–111 (2004).
536. Le Garff-Tavernier, M. et al. National standardization of ZAP-70 determination by flow cytometry: the French experience. *Cytometry B Clin Cytom* **72**, 103–108 (2007).



537. Gachard, N. et al. Multicenter study of ZAP-70 expression in patients with B-cell chronic lymphocytic leukemia using an optimized flow cytometry method. *Haematologica* **93**, 215–223 (2008).
538. Wiestner, A. et al. ZAP-70 expression identifies a chronic lymphocytic leukemia subtype with unmutated immunoglobulin genes, inferior clinical outcome, and distinct gene expression profile. *Blood* **101**, 4944–4951 (2003).
539. Bulian, P. et al. CD49d is the strongest flow cytometry-based predictor of overall survival in chronic lymphocytic leukemia. *J. Clin. Oncol.* **32**, 897–904 (2014).
540. Gattei, V. et al. Relevance of CD49d protein expression as overall survival and progressive disease prognosticator in chronic lymphocytic leukemia. *Blood* **111**, 865–873 (2008).
541. Rossi, D. et al. CD49d expression is an independent risk factor of progressive disease in early stage chronic lymphocytic leukemia. *Haematologica* **93**, 1575–1579 (2008).
542. Shanafelt, T. D. et al. CD49d expression is an independent predictor of overall survival in patients with chronic lymphocytic leukaemia: a prognostic parameter with therapeutic potential. *Br. J. Haematol.* **140**, 537–546 (2008).
543. Montserrat, E., Sanchez-Bisono, J., Viñolas, N. & Rozman, C. Lymphocyte doubling time in chronic lymphocytic leukaemia: analysis of its prognostic significance. *Br. J. Haematol.* **62**, 567–575 (1986).
544. Molica, S. & Alberti, A. Prognostic value of the lymphocyte doubling time in chronic lymphocytic leukemia. *Cancer* **60**, 2712–2716 (1987).
545. Viñolas, N., Reverter, J. C., Urbano-Ispizua, A., Montserrat, E. & Rozman, C. Lymphocyte doubling time in chronic lymphocytic leukemia: an update of its prognostic significance. *Blood Cells* **12**, 457–470 (1987).
546. Rai, K. R. et al. Clinical staging of chronic lymphocytic leukemia. *Blood* **46**, 219–234 (1975).
547. Hallek, M. et al. iwCLL guidelines for diagnosis, indications for treatment, response assessment, and supportive management of CLL. *Blood* **131**, 2745–2760 (2018).
548. Binet, J. L. et al. A new prognostic classification of chronic lymphocytic leukemia derived from a multivariate survival analysis. *Cancer* **48**, 198–206 (1981).
549. Gentile, M. et al. Validation of the CLL-IPI and comparison with the MDACC prognostic index in newly diagnosed patients. *Blood* **128**, 2093–2095 (2016).



550. da Cunha-Bang, C., Christiansen, I. & Niemann, C. U. The CLL-IPI applied in a population-based cohort. *Blood* **128**, 2181–2183 (2016).
551. Rawstron, A. C. et al. A complementary role of multiparameter flow cytometry and high-throughput sequencing for minimal residual disease detection in chronic lymphocytic leukemia: an European Research Initiative on CLL study. *Leukemia* **30**, 929–936 (2016).
552. Böttcher, S. et al. Minimal residual disease quantification is an independent predictor of progression-free and overall survival in chronic lymphocytic leukemia: a multivariate analysis from the randomized GCLLSG CLL8 trial. *J. Clin. Oncol.* **30**, 980–988 (2012).
553. Kovacs, G. et al. Minimal Residual Disease Assessment Improves Prediction of Outcome in Patients With Chronic Lymphocytic Leukemia (CLL) Who Achieve Partial Response: Comprehensive Analysis of Two Phase III Studies of the German CLL Study Group. *J. Clin. Oncol.* **34**, 3758–3765 (2016).
554. Rai, K. R. et al. Fludarabine compared with chlorambucil as primary therapy for chronic lymphocytic leukemia. *N. Engl. J. Med.* **343**, 1750–1757 (2000).
555. Johnson, S. et al. Multicentre prospective randomised trial of fludarabine versus cyclophosphamide, doxorubicin, and prednisone (CAP) for treatment of advanced-stage chronic lymphocytic leukaemia. The French Cooperative Group on CLL. *Lancet* **347**, 1432–1438 (1996).
556. Lepage, M. et al. Randomized comparison of fludarabine, CAP, and ChOP in 938 previously untreated stage B and C chronic lymphocytic leukemia patients. *Blood* **98**, 2319–2325 (2001).
557. Knauf, W. U. et al. Phase III randomized study of bendamustine compared with chlorambucil in previously untreated patients with chronic lymphocytic leukemia. *J. Clin. Oncol.* **27**, 4378–4384 (2009).
558. Knauf, W. U. et al. Bendamustine compared with chlorambucil in previously untreated patients with chronic lymphocytic leukaemia: updated results of a randomized phase III trial. *Br. J. Haematol.* **159**, 67–77 (2012).
559. Niederle, N. et al. Bendamustine compared to fludarabine as second-line treatment in chronic lymphocytic leukemia. *Ann. Hematol.* **92**, 653–660 (2013).
560. Rai, K. R. et al. Alemtuzumab in previously treated chronic lymphocytic leukemia patients who also had received fludarabine. *J. Clin. Oncol.* **20**, 3891–3897 (2002).
561. Keating, M. J. et al. Therapeutic role of alemtuzumab (Campath-1H) in patients who have failed fludarabine: results of a large international study. *Blood* **99**, 3554–3561 (2002).



562. Stilgenbauer, S. & Döhner, H. Campath-1H-induced complete remission of chronic lymphocytic leukemia despite p53 gene mutation and resistance to chemotherapy. *N. Engl. J. Med.* **347**, 452–453 (2002).
563. Lozanski, G. et al. Alemtuzumab is an effective therapy for chronic lymphocytic leukemia with p53 mutations and deletions. *Blood* **103**, 3278–3281 (2004).
564. Brown, J. R. et al. Idelalisib, an inhibitor of phosphatidylinositol 3-kinase p110 δ , for relapsed/refractory chronic lymphocytic leukemia. *Blood* **123**, 3390–3397 (2014).
565. Advani, R. H. et al. Bruton tyrosine kinase inhibitor ibrutinib (PCI-32765) has significant activity in patients with relapsed/refractory B-cell malignancies. *J. Clin. Oncol.* **31**, 88–94 (2013).
566. Byrd, J. C. et al. Targeting BTK with ibrutinib in relapsed chronic lymphocytic leukemia. *N. Engl. J. Med.* **369**, 32–42 (2013).
567. Woyach, J. A. et al. Resistance mechanisms for the Bruton's tyrosine kinase inhibitor ibrutinib. *N. Engl. J. Med.* **370**, 2286–2294 (2014).
568. Maddocks, K. J. et al. Etiology of Ibrutinib Therapy Discontinuation and Outcomes in Patients With Chronic Lymphocytic Leukemia. *JAMA Oncol* **1**, 80–87 (2015).
569. Woyach, J. A. et al. BTKC481S-Mediated Resistance to Ibrutinib in Chronic Lymphocytic Leukemia. *J. Clin. Oncol.* **35**, 1437–1443 (2017).
570. Souers, A. J. et al. ABT-199, a potent and selective BCL-2 inhibitor, achieves antitumor activity while sparing platelets. *Nat. Med.* **19**, 202–208 (2013).
571. Roberts, A. W. et al. Targeting BCL2 with Venetoclax in Relapsed Chronic Lymphocytic Leukemia. *N. Engl. J. Med.* **374**, 311–322 (2016).
572. Stilgenbauer, S. et al. Venetoclax in relapsed or refractory chronic lymphocytic leukaemia with 17p deletion: a multicentre, open-label, phase 2 study. *Lancet Oncol.* **17**, 768–778 (2016).
573. Strati, P. et al. Lenalidomide induces long-lasting responses in elderly patients with chronic lymphocytic leukemia. *Blood* **122**, 734–737 (2013).
574. Maddocks, K. et al. Reduced occurrence of tumor flare with flavopiridol followed by combined flavopiridol and lenalidomide in patients with relapsed chronic lymphocytic leukemia (CLL). *Am. J. Hematol.* **90**, 327–333 (2015).



575. Fink, A. M. et al. Lenalidomide maintenance after first-line therapy for high-risk chronic lymphocytic leukaemia (CLLM1): final results from a randomised, double-blind, phase 3 study. *Lancet Haematol* **4**, e475–e486 (2017).
576. Chanan-Khan, A. A. et al. Lenalidomide maintenance therapy in previously treated chronic lymphocytic leukaemia (CONTINUUM): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet Haematol* **4**, e534–e543 (2017).
577. Ding, W. et al. Pembrolizumab in patients with CLL and Richter transformation or with relapsed CLL. *Blood* **129**, 3419–3427 (2017).
578. Porter, D. L., Levine, B. L., Kalos, M., Bagg, A. & June, C. H. Chimeric antigen receptor-modified T cells in chronic lymphoid leukemia. *N. Engl. J. Med.* **365**, 725–733 (2011).
579. Bellosillo, B. et al. In vitro evaluation of fludarabine in combination with cyclophosphamide and/or mitoxantrone in B-cell chronic lymphocytic leukemia. *Blood* **94**, 2836–2843 (1999).
580. Byrd, J. C. et al. Addition of rituximab to fludarabine may prolong progression-free survival and overall survival in patients with previously untreated chronic lymphocytic leukemia: an updated retrospective comparative analysis of CALGB 9712 and CALGB 9011. *Blood* **105**, 49–53 (2005).
581. Hallek, M. et al. Addition of rituximab to fludarabine and cyclophosphamide in patients with chronic lymphocytic leukaemia: a randomised, open-label, phase 3 trial. *Lancet* **376**, 1164–1174 (2010).
582. Robak, T. et al. Rituximab plus fludarabine and cyclophosphamide prolongs progression-free survival compared with fludarabine and cyclophosphamide alone in previously treated chronic lymphocytic leukemia. *J. Clin. Oncol.* **28**, 1756–1765 (2010).
583. Fischer, K. et al. Bendamustine combined with rituximab in patients with relapsed and/or refractory chronic lymphocytic leukemia: a multicenter phase II trial of the German Chronic Lymphocytic Leukemia Study Group. *J. Clin. Oncol.* **29**, 3559–3566 (2011).
584. Eichhorst, B. et al. First-line chemoimmunotherapy with bendamustine and rituximab versus fludarabine, cyclophosphamide, and rituximab in patients with advanced chronic lymphocytic leukaemia (CLL10): an international, open-label, randomised, phase 3, non-inferiority trial. *Lancet Oncol.* **17**, 928–942 (2016).
585. Goede, V. et al. Chemoimmunotherapy with GA101 plus chlorambucil in patients with chronic lymphocytic leukemia and comorbidity: results of the CLL11 (BO21004) safety run-in. *Leukemia* **27**, 1172–1174 (2013).



586. Hillmen, P. et al. Chlorambucil plus ofatumumab versus chlorambucil alone in previously untreated patients with chronic lymphocytic leukaemia (COMPLEMENT 1): a randomised, multicentre, open-label phase 3 trial. *Lancet* **385**, 1873–1883 (2015).
587. Lepage, S. et al. Excess mortality after treatment with fludarabine and cyclophosphamide in combination with alemtuzumab in previously untreated patients with chronic lymphocytic leukemia in a randomized phase 3 trial. *Blood* **119**, 5104–5110 (2012).
588. Faderl, S. et al. Experience with alemtuzumab plus rituximab in patients with relapsed and refractory lymphoid malignancies. *Blood* **101**, 3413–3415 (2003).
589. Furman, R. R. et al. Idelalisib and rituximab in relapsed chronic lymphocytic leukemia. *N. Engl. J. Med.* **370**, 997–1007 (2014).
590. Jones, J. A. et al. Efficacy and safety of idelalisib in combination with ofatumumab for previously treated chronic lymphocytic leukaemia: an open-label, randomised phase 3 trial. *Lancet Haematol* **4**, e114–e126 (2017).
591. Burger, J. A. et al. Safety and activity of ibrutinib plus rituximab for patients with high-risk chronic lymphocytic leukaemia: a single-arm, phase 2 study. *Lancet Oncol.* **15**, 1090–1099 (2014).
592. Jain, P. et al. Long-term Follow-up of Treatment with Ibrutinib and Rituximab in Patients with High-Risk Chronic Lymphocytic Leukemia. *Clin. Cancer Res.* **23**, 2154–2158 (2017).
593. Jaglowski, S. M. et al. Safety and activity of BTK inhibitor ibrutinib combined with ofatumumab in chronic lymphocytic leukemia: a phase 1b/2 study. *Blood* **126**, 842–850 (2015).
594. Chanan-Khan, A. et al. Ibrutinib combined with bendamustine and rituximab compared with placebo, bendamustine, and rituximab for previously treated chronic lymphocytic leukaemia or small lymphocytic lymphoma (HELIOS): a randomised, double-blind, phase 3 study. *Lancet Oncol.* **17**, 200–211 (2016).
595. Seymour, J. F. et al. Venetoclax plus rituximab in relapsed or refractory chronic lymphocytic leukaemia: a phase 1b study. *Lancet Oncol.* **18**, 230–240 (2017).
596. Fischer, K. et al. Venetoclax and obinutuzumab in chronic lymphocytic leukemia. *Blood* **129**, 2702–2705 (2017).
597. Böttcher, S., Ritgen, M. & Dreger, P. Allogeneic stem cell transplantation for chronic lymphocytic leukemia: lessons to be learned from minimal residual disease studies. *Blood Rev.* **25**, 91–96 (2011).



598. Tsigrelis, C. & Ljungman, P. Vaccinations in patients with hematological malignancies. *Blood Rev.* **30**, 139–147 (2016).
599. Sánchez-Ramón, S., Dhalla, F. & Chapel, H. Challenges in the Role of Gammaglobulin Replacement Therapy and Vaccination Strategies for Hematological Malignancy. *Front Immunol* **7**, 317 (2016).
600. Hallek, M., Shanafelt, T. D. & Eichhorst, B. Chronic lymphocytic leukaemia. *Lancet* **391**, 1524–1537 (2018).
601. Dreger, P. et al. Managing high-risk CLL during transition to a new treatment era: stem cell transplantation or novel agents? *Blood* **124**, 3841–3849 (2014).
602. Bertrand, P. et al. Mapping of MYC breakpoints in 8q24 rearrangements involving non-immunoglobulin partners in B-cell lymphomas. *Leukemia* **21**, 515–523 (2007).
603. BLAST: Basic Local Alignment Search Tool. Available at: <https://blast.ncbi.nlm.nih.gov/Blast.cgi>. (Accessed: 2nd October 2018)
604. Le Noir, S. et al. Extensive molecular mapping of TCR α / δ - and TCR β -involved chromosomal translocations reveals distinct mechanisms of oncogene activation in T-ALL. *Blood* **120**, 3298–3309 (2012).
605. Chen, X. et al. Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. *Bioinformatics* **32**, 1220–1222 (2016).
606. Derouault, P. et al. ‘COV’COP’ allows to detect CNVs responsible for inherited diseases among amplicons sequencing data. *Bioinformatics* **33**, 1586–1588 (2017).
607. Fujiwara, T. et al. Primary effusion lymphoma of the pericardial cavity carrying t(1;22)(q21;q11) and t(14;17)(q32;q23). *Cancer Genet. Cytogenet.* **156**, 49–53 (2005).
608. IGH (Immunoglobulin Heavy). Available at: <http://atlasgeneticsoncology.org/Genes/IgHID40.html>. (Accessed: 29th August 2018)
609. Gu, G., Sederberg, M. C., Drachenberg, M. R. & South, S. T. IGF2BP1: a novel IGH translocation partner in B acute lymphoblastic leukemia. *Cancer Genet* **207**, 332–334 (2014).
610. Burmeister, T. et al. Erroneous class switching and false VDJ recombination: molecular dissection of t(8;14)/MYC-IGH translocations in Burkitt-type lymphoblastic leukemia/B-cell lymphoma. *Mol Oncol* **7**, 850–858 (2013).
611. Chaganti, S. R. et al. Deregulation of BCL6 in non-Hodgkin lymphoma by insertion of IGH sequences in complex translocations involving band 3q27. *Genes Chromosomes Cancer* **23**, 328–336 (1998).



612. Peterson, J. F. et al. Integration of microarray analysis into the clinical diagnosis of hematological malignancies: How much can we improve cytogenetic testing? *Oncotarget* **6**, 18845–18862 (2015).
613. RSS Database. Available at: <https://www.itb.cnr.it/rss/index.html>. (Accessed: 19th September 2018)
614. Merelli, I. et al. RSSsite: a reference database and prediction tool for the identification of cryptic Recombination Signal Sequences in human and murine genomes. *Nucleic Acids Res.* **38**, W262–267 (2010).
615. Klijn, C. et al. A comprehensive transcriptional portrait of human cancer cell lines. *Nature Biotechnology* **33**, 306–312 (2015).
616. Yoshihara, K. et al. The landscape and therapeutic relevance of cancer-associated transcript fusions. *Oncogene* **34**, 4845–4854 (2015).
617. Miki, T. et al. Molecular cloning of the breakpoint for 3q27 translocation in B-cell lymphomas and leukemias. *Blood* **83**, 217–222 (1994).
618. Shimanuki, M. et al. Molecular cloning of IG λ rearrangements using long-distance inverse PCR (LDI-PCR). *Eur. J. Haematol.* **90**, 59–67 (2013).
619. Champagne, D. P. & Shockett, P. E. Illegitimate V(D)J recombination-mediated deletions in Notch1 and Bcl11b are not sufficient for extensive clonal expansion and show minimal age or sex bias in frequency or junctional processing. *Mutat Res* **761**, 34–48 (2014).
620. Papaemmanuil, E. et al. RAG-mediated recombination is the predominant driver of oncogenic rearrangement in ETV6-RUNX1 acute lymphoblastic leukemia. *Nat. Genet.* **46**, 116–125 (2014).
621. Mijušković, M. et al. Off target V(D)J recombination drives lymphomagenesis and is escalated by loss of the Rag2 C-terminus. *Cell Rep* **12**, 1842–1852 (2015).
622. Löffler, D. et al. METRNL decreases during adipogenesis and inhibits adipocyte differentiation leading to adipocyte hypertrophy in humans. *Int J Obes (Lond)* **41**, 112–119 (2017).
623. Graham, M. & Adams, J. M. Chromosome 8 breakpoint far 3' of the c-myc oncogene in a Burkitt's lymphoma 2;8 variant translocation is equivalent to the murine pvt-1 locus. *EMBO J* **5**, 2845–2851 (1986).
624. Wang, J. J. et al. Expression and Function of Long Non-coding RNA CASC19 in Colorectal Cancer. *Zhongguo Yi Xue Ke Xue Yuan Xue Bao* **39**, 756–761 (2017).



625. Ozawa, T. et al. CCAT1 and CCAT2 long noncoding RNAs, located within the 8q.24.21 'gene desert', serve as important prognostic biomarkers in colorectal cancer. *Ann. Oncol.* **28**, 1882–1888 (2017).
626. Bastard, C. et al. Comparison of a quantitative PCR method with FISH for the assessment of the four aneuploidies commonly evaluated in CLL patients. *Leukemia* **21**, 1460–1463 (2007).
627. Lehmann, S. et al. Molecular allelokaryotyping of early-stage, untreated chronic lymphocytic leukemia. *Cancer* **112**, 1296–1305 (2008).
628. Zenz, T. et al. Detailed analysis of p53 pathway defects in fludarabine-refractory chronic lymphocytic leukemia (CLL): dissecting the contribution of 17p deletion, TP53 mutation, p53-p21 dysfunction, and miR34a in a prospective clinical trial. *Blood* **114**, 2589–2597 (2009).
629. Hagenkord, J. M. et al. Array-based karyotyping for prognostic assessment in chronic lymphocytic leukemia: performance comparison of Affymetrix 10K2.0, 250K Nsp, and SNP6.0 arrays. *J Mol Diagn* **12**, 184–196 (2010).
630. Grygalewicz, B. et al. Monoallelic and biallelic deletions of 13q14 in a group of CLL/SLL patients investigated by CGH Haematological Cancer and SNP array (8x60K). *Mol Cytogenet* **9**, 1 (2016).
631. Hartmann, L. et al. Detection of clonal evolution in hematopoietic malignancies by combining comparative genomic hybridization and single nucleotide polymorphism arrays. *Clin. Chem.* **60**, 1558–1568 (2014).
632. Sellmann, L. et al. Shorter telomeres correlate with an increase in the number of uniparental disomies in patients with chronic lymphocytic leukemia. *Leuk. Lymphoma* **57**, 590–595 (2016).
633. Zinovieva, O. L. et al. Expression of long non-coding RNA LINC00973 is consistently increased upon treatment of colon cancer cells with different chemotherapeutic drugs. *Biochimie* **151**, 67–72 (2018).
634. Huppi, K., Pitt, J. J., Wahlberg, B. M. & Caplen, N. J. The 8q24 gene desert: an oasis of non-coding transcriptional activity. *Front Genet* **3**, 69 (2012).
635. Colombo, T., Farina, L., Macino, G. & Paci, P. PVT1: A Rising Star among Oncogenic Long Noncoding RNAs. *Biomed Res Int* **2015**, (2015).
636. Tseng, Y.-Y. et al. PVT1 dependence in cancer with MYC copy-number increase. *Nature* **512**, 82–86 (2014).



637. Houshmand, M. et al. Long non-coding RNA PVT1 as a novel candidate for targeted therapy in hematologic malignancies. *Int. J. Biochem. Cell Biol.* **98**, 54–64 (2018).
638. Morrison, V. A. Infectious complications of chronic lymphocytic leukaemia: pathogenesis, spectrum of infection, preventive approaches. *Best Practice & Research Clinical Haematology* **23**, 145–153 (2010).
639. Rizzo, D. Étude des paramètres biologiques influençant l’histoire naturelle de la leucémie lymphoïde chronique : expression du CD45, sécrétion d’une IgM monoclonale et dérégulation de ZAP70. (Limoges, 2014).



Sommaire

Sommaire	3
Introduction.....	24
Chapitre 1 : Oncogénèse des lymphomes B.....	26
A- Physiologie de la maturation B normale	26
1- Introduction	26
2- Structure des immunoglobulines.....	27
3- Structure des gènes des immunoglobulines.....	29
4- Le récepteur de la cellule B ou BCR	42
5- Lymphopoïèse B normale	45
B- Origine cellulaire des lymphomes B.....	55
C- Voies d'oncogénèse dans les lymphomes B.....	56
1- Dérégulation oncogénique via translocation	57
2- Mutations géniques	69
3- Autres mécanismes d'activation de proto-oncogènes	70
4- Inactivation de gènes suppresseurs de tumeur	70
5- Transformation par les virus.....	71
6- Activation du BCR	72
7- Inactivation du BCR.....	72
8- Activation par l'antigène.....	73
9- Rôle du microenvironnement.....	74
10- Rôle des ARN longs non codants	75



11- Modèles de lymphomagénèse.....	77
Chapitre 2 : la leucémie lymphoïde chronique	79
A- Définition.....	79
B- Epidémiologie.....	79
C- Etiologies	80
1- Facteurs de risque génétiques.....	80
2- Facteurs de risque environnementaux.....	81
3- Facteurs liés au mode de vie	82
4- Antécédents médicaux	82
D- Physiopathologie.....	82
1- Origine cellulaire.....	82
2- Sélection par l'antigène	83
3- Altérations génétiques.....	83
4- Microenvironnement.....	89
E- Clinique.....	94
1- Présentation clinique classique	94
2- Les variants	95
3- Les complications.....	96
F- Diagnostic	99
1- Numération-Formule sanguine.....	99
2- Cytométrie en flux	100
3- Autres.....	100
G- Facteurs pronostiques	101
	290



1- Cytogénétique.....	101
2- Génétique de la partie variable des chaînes lourdes des immunoglobulines (IGHV)	108
3- Mutations récurrentes hors <i>IGHV</i>	111
4- Marqueurs sériques.....	119
5- Marqueurs immunophénotypiques.....	122
6- Temps de doublement de la lymphocytose	124
7- Classifications pronostiques	124
H- Traitement	127
1- Critères de mise en place thérapeutique	127
2- Evaluation de la réponse au traitement	129
3- Mesure de la maladie résiduelle.....	130
4- Les différentes options thérapeutiques disponibles	130
5- Choix du traitement en fonction des patients.....	137
Chapitre 3 : Problématiques et objectifs	139
Chapitre 4 : Matériels et méthodes	140
A- Patients et échantillons.....	140
1- Etude de translocations impliquant le gène des immunoglobulines	140
2- LLC et électrophorèse des protéines sériques.....	142
3- Evaluation du logiciel Cov'Cop dans la LLC	142
B- Techniques conventionnelles.....	144
C- Techniques spécialisées	146
1- FISH à façon.....	146
2- LM-PCR et séquençage	148
	291



3- CGH-array.....	150
4- Séquençage haut débit (NGS).....	151
Chapitre 5 : Résultats	158
A- Etude de translocations atypiques dans les lymphomes B.....	158
1- Patient 1.....	158
2- Patient 2.....	177
3- Patient 3.....	183
B- Article 1 : Normal serum protein electrophoresis and mutated IGHV genes detect very slowly evolving chronic lymphocytic leukemia patients, accepté le 29 Mars 2018 dans Cancer medicine	193
1- Introduction	193
2- Article.....	194
3- Discussion	202
C- Article 2 : « Updates to « COV’COP » to easily detect CNVs in inherited and cancer diseases », manuscrit soumis dans Bioinformatics le 11/06/2018 et refusé le 18/06/2018, en attente de resoumission dans un autre journal	203
1- Introduction	203
2- Article.....	204
3- Discussion	208
D- Evaluation du logiciel CovCop dans la LLC	219
1- Introduction	219
2- Résultats.....	219
3- Courte discussion.....	227
Chapitre 7 : Discussion générale et perspectives.....	229



A- Remaniements atypiques impliquant les gènes des immunoglobulines dans la LLC/lymphome lymphocytaire.....	229
1- Mécanisme de survenue des remaniements impliquant les gènes des immunoglobulines	229
2- Mécanisme de dérégulation oncogénique	232
3- Modèles de lymphomagénèse.....	233
4- Perspectives	234
B- Impact pronostique de l'électrophorèse des protéines sériques dans la LLC	235
C- Utilisation du séquençage haut débit pour la détection d'anomalies de nombre	236
Conclusion	238
Références bibliographiques.....	239
Sommaire	289
Résumé.....	238



Résumé

La leucémie lymphoïde chronique (LLC) est le lymphome avec phase circulante le plus fréquent chez l'adulte dans les pays occidentaux. Elle est caractérisée par une grande hétérogénéité dans son évolution naturelle avec des formes indolentes ne nécessitant jamais de traitement spécifique et des formes agressives requérant une chimiothérapie rapidement après le diagnostic. Dans ce travail de thèse, nous avons posé la question du rôle des anomalies moléculaires et en particulier des gènes des immunoglobulines dans l'histoire naturelle de la maladie, tant au plan mécanistique que pour le pronostic. Nous avons étudié trois remaniements atypiques impliquant un gène des immunoglobulines et un partenaire inconnu dans la LLC/lymphome lymphocytaire. Les points de cassure ont pu être identifiés et nous ont permis de mettre en évidence l'implication dans 2 cas d'ARN longs non codants en 17q25 et 8q24. De plus, deux cas ont des points de cassure dans une région chromosomique restreinte (espacés de 200 kb en 17q25). Il pourrait s'agir d'un locus important dans la lymphomagenèse, de même que pour la région 8q24 contenant *MYC* et de nombreux gènes non codants jusqu'ici peu explorés. Par ailleurs, dans l'ère du séquençage haut débit, de nombreux marqueurs pronostiques moléculaires sont décrits dans la LLC. Nous avons démontré que l'électrophorèse des protéines sériques normale (immunoglobulines polyclonales sans hypogammaglobulinémie) au diagnostic, un marqueur simple et peu coûteux, reste dans l'ère du NGS, un marqueur indépendant de bon pronostic dans la LLC. Sa combinaison avec un statut IGHV muté identifie un groupe de patients qui n'auront probablement jamais besoin de traitement spécifique. Ce travail nous a conduit à mettre au point un outil performant de détection des anomalies de nombre de copies par séquençage haut débit. Celui-ci permet la mise en évidence de disomies uniparentales dont la signification pronostique n'est actuellement pas établie dans la LLC. Ces anomalies pourraient être le reflet d'une instabilité chromosomique et il serait intéressant d'étudier leur impact pronostique dans la LLC.

Mots clés : leucémie lymphoïde chronique, translocation, immunoglobuline, lncRNA, séquençage haut débit, électrophorèse des protéines sériques

Abstract

Chronic lymphocytic leukemia (CLL) is the most frequent lymphoma with leukemic phase in the elderly in Western countries. It is characterized by a great heterogeneity in its natural history with indolent forms never requiring any specific treatment and aggressive forms needing chemotherapy rapidly after diagnosis. In this work of thesis, we asked the question of the role of molecular abnormalities, and in particular of the immunoglobulin genes in the natural history of the disease, at mechanistic level and for prognosis. We studied three atypical rearrangements implicating an immunoglobulin gene and an unknown partner in CLL/lymphocytic lymphoma. The breakpoints have been identified and the implication of long non coding RNA was highlighted in two cases in 17q25 and 8q24. Moreover, two cases harboured breakpoints in a restricted chromosomal region (200 kb spaced in 17q25). It could be an important locus in lymphomagenesis, as is 8q24 region containing *MYC* and numerous other non coding genes poorly characterized by now. Furthermore, in high throughput sequencing (HTS) era, many molecular prognosis markers have been described in CLL. We demonstrated that normal serum protein electrophoresis (polyclonal immunoglobulin without hypogammaglobulinemia) at diagnosis, a simple and unexpensive marker, stays in HTS era an independent good prognosis marker in CLL. Its combination with unmutated IGHV genes status identifies a group of patients who will probably never require any specific treatment. This work led us to develop an efficient tool to detect copy number variations by HTS. This tool allows to highlight uniparental disomy whose prognosis signification is not established in CLL. These abnormalities could reflect chromosomal instability and it could be interesting to study their prognosis impact in CLL.

Keywords : Chronic lymphocytic leukemia, translocation, immunoglobulin, lncRNA, high throughput sequencing, serum protein electrophoresis