



**HAL**  
open science

# Développement de nouvelles approches pour l'analyses des interactions cellulaires dans les maladies immunitaires

Thomas Laurent

► **To cite this version:**

Thomas Laurent. Développement de nouvelles approches pour l'analyses des interactions cellulaires dans les maladies immunitaires. Médecine humaine et pathologie. Nantes Université, 2023. Français. NNT : 2023NANU1017 . tel-04208535

**HAL Id: tel-04208535**

**<https://theses.hal.science/tel-04208535>**

Submitted on 15 Sep 2023

**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.

# THESE DE DOCTORAT DE

NANTES UNIVERSITE

ECOLE DOCTORALE N° 605

*Biologie-Santé*

Spécialité : Immunologie

Par

**Thomas LAURENT**

## **Développement de nouvelles approches pour l'analyse des interactions cellulaires dans les maladies immunitaires**

**Thèse présentée et soutenue à Nantes, le 06 avril 2023**

**Unité de recherche :** Centre de Recherche en Transplantation et Immunologie Translationnelle (CR2TI)

INSERM UMR1064 - CHU Nantes - Hôtel Dieu, 30 Bd Jean Monnet, 44093 Nantes Cedex 01

### **Rapporteurs avant soutenance :**

Florence APPARAILLY  
Barbara SEITZ-POLSKI

Directrice de recherche, INSERM U1183 - IRMB, Montpellier  
Professeur d'université – Praticien Hospitalier, CRMN SNI, Nice

### **Composition du Jury :**

Président : Barbara SEITZ-POLSKI  
Examineurs : Florence APPARAILLY  
Barbara SEITZ-POLSKI  
Julie HELFT

Directrice de recherche, INSERM U1183 - IRMB, Montpellier  
Professeur d'université – Praticien Hospitalier, CRMN SNI, Nice  
Chargée de recherche, Institut Cochin, Paris

Dir. de thèse : Régis JOSIEN  
Co-Encadrant : Jérémie POSCHMANN  
Co-Encadrant : Jérôme MARTIN

Professeur d'université – Praticien Hospitalier, UMR 1064 – CR2TI, Nantes  
Chargé de recherche, UMR 1064 – CR2TI, Nantes  
Maître de Conférences des Universités – Praticien Hospitalier,  
UMR 1064 – CR2TI, Nantes

Je tiens tous d'abord à remercier sincèrement le Dr. Florence APPARAILLY, le Dr. Barbara SEITZ-POLSKI et le Dr. Julie HELFT d'avoir accepté de faire partie de mon jury de thèse ainsi que pour leur temps accordé à lire et évaluer mon travail.

Je remercie aussi les membres de mon comité de suivi de thèses, le Dr. Vassili SOUMELIS, le Dr. Ephraïm KENIGSBURG et le Dr. Mathieu GIRAUD pour avoir suivi mes travaux sur les trois années et demies et pour leurs divers conseils d'un point de vue extérieur.

Je remercie ensuite très chaleureusement le Dr. Jérémie POSCHMANN pour m'avoir ouvert les portes du monde de la recherche scientifique en m'acceptant en stage de Master et de m'avoir fait confiance et soutenu depuis, jusqu'à cette fin de thèse. Ce fut un plaisir de découvrir de nouvelles technologies et méthodes qui me donnaient l'impression grisante d'être à la pointe de l'innovation.

Je remercie tout aussi chaleureusement le Dr. Jérôme MARTIN pour m'avoir proposé de rejoindre son équipe et d'avoir toujours su me pousser à m'améliorer, et pour m'avoir aiguillé tout du long pour fournir un travail scientifique rigoureux et pour m'avoir fait aimer le fait de creuser toujours plus loin dans la biologie du vivant.

Merci à tous les deux pour ces deux projets de recherche très stimulants qui me font aimer la recherche et me rendent fier de la petite pierre que j'apporte à l'édifice de la recherche scientifique. Merci pour les nombreuses connaissances que j'ai acquises grâce à vous ainsi que pour tous les conseils et remarques qui ont permis de construire mon expertise et ma réflexion scientifique.

Merci de m'avoir accompagné tout au long de ma thèse, pour votre confiance et pour toujours valoriser mon travail, même en période de doute, je suis vraiment reconnaissant d'avoir eu la chance de travailler avec vous et d'avoir fait partie de deux équipes bienveillantes et je suis très fier de tout ce que j'ai pu accomplir à vos côtés durant ces trois ans et demi.

Je remercie également chaleureusement le Dr Régis JOSIEN, à la fois en tant que directeur de thèse et pour son suivi de mes travaux ainsi que ces conseils pour améliorer mes travaux et affiner mon raisonnement scientifique.

Un grand merci aussi à tous les membres de mes deux équipes qui ont contribué à la réussite de cette thèse, notamment Cynthia sur qui j'ai toujours pu compter lorsque je quittais mon ordinateur pour aller mener les quelques manip essentielles à mon projet, ainsi que Debajyoti qui m'a aidé à mieux comprendre de nombreux concepts informatique, mathématiques et tout ce qui pouvait me manquer. Merci aussi à Alice, Pierre, Lucas, Martin, Victor, Aurélie, Camille, Gaëlle, Aurélie, Simon, Monika et tous ceux qui ont participé de près ou de loin aux deux projets. Merci aussi à tous les membres de l'équipe 1 et 6 pour vos retours et vos challenges sur mes projets, merci entre autres à Laurence pour son aide et ses précieux conseils.

Et plus largement, merci à tous les inconnus des divers forums qui m'ont permis de construire mes connaissances en analyse transcriptomique.

Un énorme merci à Florian, Céline, Sita et Marie qui ont rendu cette thèse tellement plus facile à supporter lors des moments les plus intenses. Merci pour votre très précieuse aide, les réflexions scientifiques à bâtons rompus mais aussi pour tous les bons moments passés au labo comme à l'extérieur, je suis très heureux de vous avoir eu à mes côtés !

Merci à Xavier et Camille, mes deux « aqua », pour leurs oreilles attentives et les discussions interminables à parler de tout et de rien, être toujours là même à distance lors des victoires et des épreuves

Un grand merci aussi à tous mes autres amis qui ont donné plus de saveur à ces trois années et demies et pour m'avoir apporté ce que le travail n'apporte pas, merci Jeanne, Jérémie, Aziliz, Camille, Lucien, Marion, Carlos, Joyce, et tous ceux que je ne pourrais pas citer.

Et enfin merci à ma famille qui a toujours été là pour moi, un havre de confort et d'encouragement sans failles qui m'a grandement aidé à avancer sereinement en vous sachant toujours là pour moi. Merci Pierre pour tout ce qu'on partage, pour m'avoir supporté même dans les longues soirées de travail. Merci papa et maman pour votre soutien, vos conseils, votre suivi, et avec Marie et Louise pour les moments partagés qui me permettaient de souffler un peu. Merci aussi à tous les autres membres de la famille qui ont toujours montré de l'intérêt pour mes recherches.



# Sommaire

<b>Table des abréviations</b>	6
<b>Introduction : Application du Single-Cell RNA sequencing à la caractérisation des mécanismes immunopathologiques impliqués dans la transplantation et les maladies inflammatoires chroniques</b>	8
<b>Le séquençage d'ARN en cellules uniques</b>	8
De l'implantation de la génomique au séquençage de l'ARN	8
Séquençage d'ARNm	10
La révolution du scRNA-seq au service de la transcriptomique	11
Le séquençage en cellule unique	11
La microfluidique au service de la transcriptomique	13
Enjeux techniques et analytiques du scRNA-seq	16
Encapsulation et multiplets	16
Vue d'ensemble de l'analyse des données de scRNA-seq	18
La problématique du "Batch Effect"	20
Le choix d'une méthode de correction du batch effect	21
La méthode Metacells	23
Objectifs du scRNA-seq	25
<b>Chapitre 2 : Le rejet chronique et la transplantation rénale</b>	26
Rein et néphropathie	26
Le recours à la greffe de rein	27
Compatibilité HLA	28
Rejets aigus et rejets chroniques	29
TCMR	30
ABMR	30
Suivi et immunodépression chez les patients transplantés	32
Les traitements immunosuppresseurs dans la transplantation rénale	33
Les corticostéroïdes	33
Les inhibiteurs de la calcineurine (signal 1)	33
Les bloqueurs du signal de costimulation (signal 2)	33
Les antiprolifératifs inhibiteurs de mTOR (signal 3)	34
Les antiprolifératifs inhibiteurs de métabolites (signal 3)	34
<b>Chapitre 3 : Les Maladies Inflammatoires Chroniques de l'Intestin</b>	35
Les fonctions de l'intestin	35
L'intestin, un rôle dans la digestion et dans la protection	35
La muqueuse	35
La sous-muqueuse	37
La musculature	37
La membrane séreuse	37

Le microbiote intestinal et ses interactions	37
Organisation du système immunitaire intestinal	39
L'immunité innée	39
L'immunité adaptative	42
Barrière épithéliale et signature IL-17/IL-22	43
Maladie de Crohn et Rectocolite hémorragique : causes et traitements actuels	44
Les maladies inflammatoires chroniques de l'intestin	44
Diagnostic clinique et suivi des patients	44
Complications dans les MICI	45
L'origine imprécise des MICI et leurs prédispositions génétiques	46
L'impact des facteurs environnementaux et la dysbiose intestinale	48
Stratégies de traitements dans les MICI	48
Inhibiteurs du TNF $\alpha$	49
Anticorps anti- $\alpha$ 4 $\beta$ 7	50
Anticorps anti-interleukines IL-12/IL-23p40	50
Autres stratégies	51
Stratégie générale de prise en charge des patients	53
Résistance aux traitements	54
Utilisation du single cell dans la caractérisation de la pathologie	56
Le scRNA-seq dans la rectocolite hémorragique	56
La signature GIMATS, une nouvelle piste pour comprendre la résistance à l'anti-TNF $\alpha$ dans la maladie de Crohn	57
Metacells	58
<b>Objectifs de recherche</b>	59
<b>Résultats</b>	59
1) Le rejet de greffe de rein	59
2) La résistance aux traitements dans les MICI	89
3) Projets annexes	105
<b>Discussion</b>	106
Partie Transplantation	106
Partie MICI	109
<b>Bibliographie</b>	113
<b>Annexes</b>	136
Annexe 1 : Liste des 150 modules identifiés dans le modèle supergut	136
Annexe 2 - Metadata de la cohorte SuperGut	138
Annexe 3 - Distribution des compartiments de cellules myéloïdes et stromales dans la cohorte SuperGut	139
Annexe 4 - Characterization of Rat ILCs Reveals ILC2 as the Dominant Intestinal Subset	139

# Table des abréviations

ABMR : Rejet médié par anticorps (Antibody-Mediated Rejection)  
ADCC : Cytotoxicité à médiation Cellulaire Dépendante des Anticorps  
ADN : Acide DésoxyriboNucléique  
ADNc : ADN complémentaire  
AGCC : Acides Gras à Chaînes Courtes  
AMP : Peptides anti-microbiens (Anti Microbial Peptides)  
APC : Cellules présentatrices d'antigènes (Antigen Presenting Cells)  
ARN : Acides RiboNucléique  
ARNInc : ARN long non-codant  
ARNm : ARN messenger  
ARNnc : ARN non-codant  
ATAC-seq : Séquençage de la chromatine accessible par transposase (Assay for Transposase-Accessible Chromatin)  
CCA : Analyse Canonique des Corrélations (Canonical correlation Analysis)  
cDC : Cellules dendritique conventionnelle  
CMH : Complexe Majeur d'Histocompatibilité  
CNI : Inhibiteurs à la CalciNeurine  
DC : Cellule dendritique  
DEG : Gènes différenciellement exprimés (Differentially Expressed Genes)  
DSA : Anticorps spécifiques du donneur (Donor Specific Antigen)  
FACS : Tri cellulaire activé par fluorescence (Fluorescence Activated Cell Sorting)  
FT : Facteur de transcription  
GALT : Tissu lymphoïde associé au tube digestif (Gut-Associated Lymphoid Tissue)  
GEM : Billes de Gel en Émulsion (Gel Emulsions-in-Beads)  
GSEA : Gene Set Enrichment Analysis  
GWAS : Genome Wide Association Study  
HLA : Antigènes des Leucocytes Humains  
IEC : Cellules épithéliales intestinales (Intestinal Epithelial Cells)  
IFN : Interféron  
IFTA : Fibrose Interstitielle et Atrophie Tubulaire  
Ig : Immunoglobuline  
IL : Interleukine  
ILC : Cellule lymphoïde Innée (Innate Lymphoid Cell)  
IRC : Insuffisance Rénale Chronique  
LB : Lymphocyte B  
LT : Lymphocyte T  
MAIT : Cellule T invariante associée aux muqueuses (Mucosal Associated Invariant T)  
MC : Maladie de Crohn  
MEC : Matrice Extra-Cellulaire  
MMF : Mycophenolate Mofetil  
MMP : Matrix MetalloProteinases  
MNN : Mutual Nearest Neighbourhood  
MNP : Phagocytes Mononucléés (Mononuclear phagocytes)  
NF-κB : Nuclear Factor-Kappa B  
NGS : Next-generation sequencing  
NK : Cellules tueuses naturelles (Natural Killer)  
OLS : Organes lymphoïdes secondaires

PBMC : Cellules mononuclées du sang périphérique (Peripheral Blood Mononuclear Cells)  
PCA : Analyse en composantes principales (Principal Component Analysis)  
PCR : Réaction de polymérisation en chaîne  
PRR : Récepteurs de reconnaissance de motifs moléculaires (Pattern Recognition Receptor)  
pDC : Cellule dendritique plasmacytoïde  
RCH : RectoColite Hémorragique  
RNA-seq : Séquençage d'ARN  
ROS : Espèces réactives de l'oxygène (Reactive oxygen species)  
RT : Transcriptase inverse (Reverse Transcriptase)  
Sc : Single-cell  
SNP : Polymorphismes mononucléotidique (Single-Nucleotide Polymorphism)  
TCMR : Rejet Médié par les Cellules T (T Cell Mediated Rejection)  
TCR : Récepteur des Cellules T  
Teff : Lymphocyte T effecteur  
TNF : Facteur de nécrose tumorale  
Treg : Lymphocyte T régulateur  
TRM : Cellule T mémoire résidente  
TSO : Template Switching Oligo  
TSS : Site de départ de la transcription (Transcription Start Site)  
UMAP : Uniform Manifold Approximation and Projection  
UMI : Identifiants moléculaires uniques  
WGCNA : Analyse de réseau de co-expression génique pondérée (Weighted Correlation Network Analysis)

# Introduction : Application du Single-Cell RNA sequencing à la caractérisation des mécanismes immunopathologiques impliqués dans la transplantation et les maladies inflammatoires chroniques

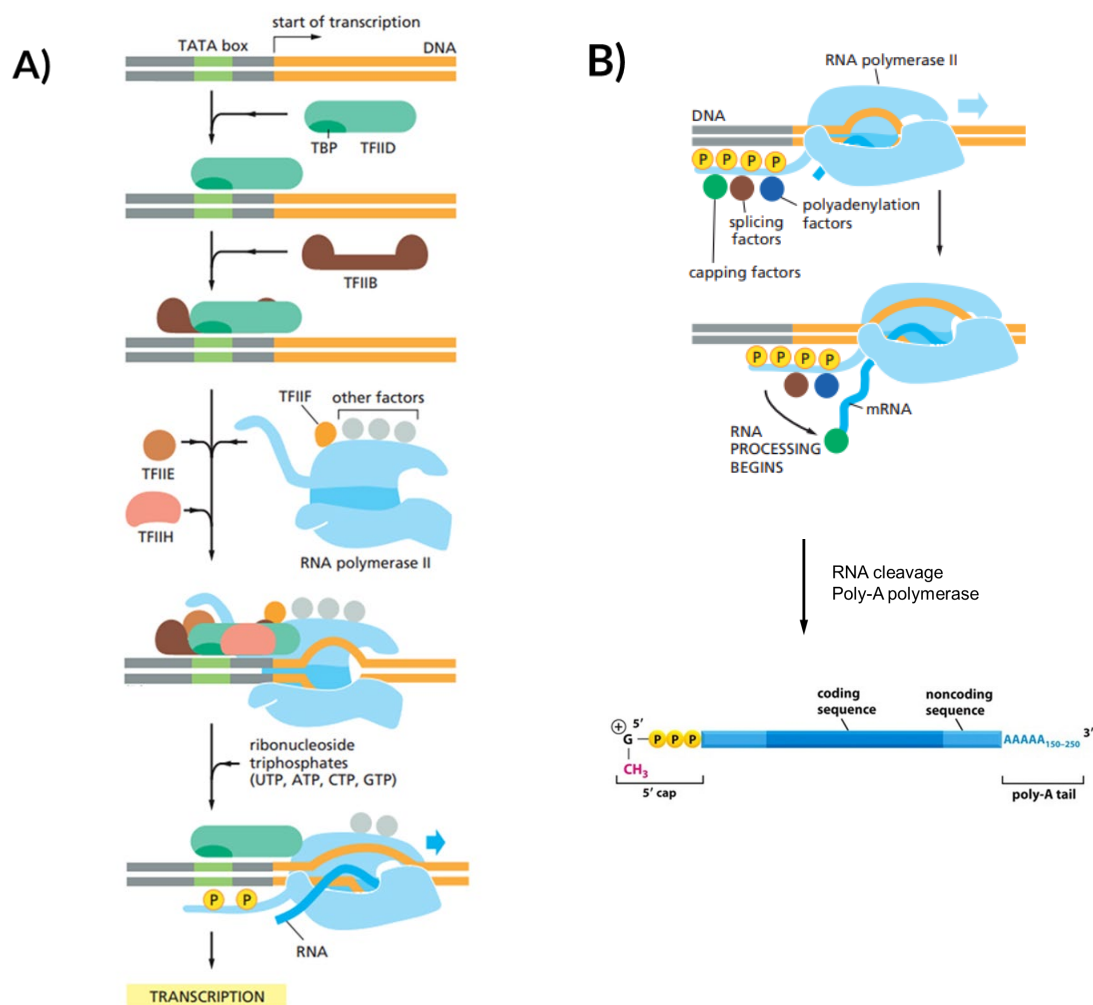
## Le séquençage d'ARN en cellules uniques

### De l'implantation de la génomique au séquençage de l'ARN

La compréhension du corps humain n'a cessé de s'affiner à travers les siècles, passant de l'étude uniquement phénotypique, à une compréhension plus précise de la biologie cellulaire puis moléculaire prenant place dans les organes. Plus récemment, la compréhension du rôle de l'ADN dans la régulation de l'organisme a permis le développement d'une discipline se concentrant sur la caractérisation du transcriptome et du protéome afin de déterminer les fonctions intrinsèques des cellules : la génomique fonctionnelle. L'ADN représente le champ des possibles, l'ensemble des éléments que la cellule peut créer. Cependant, l'ADN n'est pas directement traduit en protéine, il subit de nombreuses régulations lors de la transcription, de la traduction et même en post-transcription : son étude seule représenterait une vision limitée des processus biologiques en cours dans la cellule. On estime à 80% la part du génome qui peut être transcrite (Li et Liu 2019), mais à un temps donné, seule une faible partie de l'ADN est accessible aux facteurs de transcriptions et aux molécules de régulation de la transcription à cause des marques épigénétiques qui contrôlent la compaction de l'ADN. Une plus faible partie encore sera exprimée en protéines, puisque d'après le consortium ENCODE (Encyclopedia of DNA Elements), seuls 2% du génome correspondent aux exons de gènes codants (Dunham et al. 2012), le reste étant majoritairement transcrit en ARN non-codants qui vont réguler la transcription (Hon et al. 2017) mais dont les rôles divers sont encore loin d'être entièrement compris. Malgré la relative fragilité des ARN messagers (ARNm), l'étude de ces transcrits codants permet d'avoir une idée plus précise de l'activité cellulaire en regard du séquençage de l'ADN (Z. Wang, Gerstein, et Snyder 2009; Casamassimi et al. 2017; Lowe et al. 2017).

Pour production d'une séquence d'ARNm durant la transcription, qui se déroule en 3 étapes : l'initiation, l'élongation et la terminaison. La structure de tous les gènes est sensiblement la même chez tous les eucaryotes, et c'est dans la région du promoteur que la transcription va démarrer. Environ 25 nucléotides en amont du site de départ de la transcription (TSS) se situe la TATA box, appelée ainsi en référence à sa séquence nucléotidique enrichie en adénosines (A) et thymidines (T). Elle est reconnue par les facteurs généraux de transcriptions dont font partie TFIIA, TFIIB, TFIID avec sa sous-unité TBP (TATA binding protein), TFIIIE, TFIIF, et TFIIF, qui vont

venir se fixer sur l'ADN (Hampsey 1998), puis recruter et stabiliser l'ARN polymérase II afin de former le complexe de pré-initiation de la transcription (**Figure 1A**). Ce complexe va permettre la séparation des brins d'ADN et positionne l'ADN dans le site actif de transcription de l'ARN polymérase II. En plus de la matrice d'ADN et de l'ARN polymérase, des sels de magnésiums et les nucléosides triphosphates sont nécessaires pour débiter la synthèse de l'ARN. L'élongation du brin d'ARNm s'effectue toujours dans le sens 5'→3', 5' faisant référence au carbone 5' du désoxyribose qui se retrouve libre à l'extrémité. La transcription s'arrête lorsque le complexe de transcription reconnaît une séquence de terminaison de la transcription.



**Figure 1. Initiation de la transcription chez les eucaryotes.**

A) La première enzyme du complexe de transcription qui se lie au promoteur proximal est TFIID, en reconnaissant la TATA box. Elle induit un changement de conformation de l'ADN pour pouvoir recruter les sous-unités TFIIA et TFIIB qui stabilisent les interactions avec l'ADN. TFIIB recrute ensuite TFIIF et l'ADN Polymérase II puis TFIIE et TFIIH.

B) Terminaison de la transcription : fixation d'une coiffe à l'extrémité 5'. Clivage à la reconnaissance du site de polyadénylation et synthèse d'une queue poly-A à l'extrémité 3' du brin d'ARN. D'après (Alberts et al. 2015)

Un processus de maturation est ensuite mis en jeu afin de protéger le brin nouvellement transcrit (**Figure 1B**). Cette étape de maturation est spécifique aux cellules eucaryotes. En effet, chez les organismes eucaryotes, la transcription et la traduction ne sont pas simultanées puisque la transcription se déroule dans le noyau et la traduction dans le cytoplasme. La protection de l'ARNm passe dans un premier temps par la fixation d'une coiffe de méthyl-7-guanylate à l'extrémité 5', qui permet de protéger l'extrémité des nucléases et phosphatases, mais aussi plus tard pour initier la traduction. À l'autre extrémité, en 3', on retrouve une séquence de signal de polyadénylation AAUAAA qui va être reconnue et induire le clivage du brin par une endonucléase avant qu'une poly-A polymérase vienne incorporer jusqu'à 250 nucléotides A en fin de chaîne. Cette queue poly A permet la protection de l'extrémité 3' (un ARNm nu serait vite dégradé par les RNAses extracellulaires) et permet aussi la migration de l'ARNm vers le cytoplasme. Avant de sortir du noyau, l'ARNm subit une autre étape, celle de l'épissage, qui va exciser les introns de la séquence pour aboutir à l'ARN mature final.

Les gènes ainsi exprimés sont ensuite exposés à de nombreux facteurs qui peuvent réguler leur traduction en protéine. Capturer les ARN messager permet donc d'avoir une idée relativement précise des mécanismes en cours dans la cellule et des protéines en production. C'est l'étude quantitative de l'ensemble des transcrits produits dans une cellule que l'on appelle transcriptomique.

## Séquençage d'ARNm

Depuis la résolution de la première séquence ARN d'un gène complet séquencé dans le bactériophage MS2 des bactéries du genre *Escherichia* (Jou et al. 1972), l'essor de la transcriptomique a beaucoup dû à l'amélioration des méthodes de séquençage et de capture de l'ARN.

Avant le développement de l'analyse en cellules uniques, c'est la technique de puces à ADN (microarrays), développée au début des années 90 (Fodor et al. 1991; Schena et al. 1995), qui était la plus répandue pour permettre l'étude du transcriptome cellulaire. Son principe repose sur l'hybridation des ARN messagers des cellules déposées sur la puce, avec des sondes d'ADN simple brin connues qui sont fixées à une plaque. Grâce à un marquage par des fluorochromes ou des radioéléments, on peut mesurer l'intensité de fixation pour chaque sonde et en déduire le niveau d'expression de l'ARN dans la solution de cellules. Dans cette technique le séquençage n'est pas nécessaire, ce qui permet de réduire les coûts, mais on fait face à plusieurs limites : des gènes aux séquences similaires peuvent être confondus en se fixant aux mêmes sondes, la séquence exacte des ARNm capturés n'est pas connue et enfin seules les séquences correspondant à celles présentes sur la puce peuvent être détectées (Tang et al. 2010). À ce manque de précision et de spécificité s'ajoute le fait que l'information mesurée reflète une moyenne de l'ensemble des cellules présentes dans l'échantillon analysé, ce qui a conduit à la définition du terme de "bulk" RNAseq pour désigner ce type d'analyse.

## La révolution du scRNA-seq au service de la transcriptomique

Le besoin d'étudier les cellules séparément s'appuie sur le constat des différentes limites liées à l'analyse en "bulk". En premier lieu ce type d'analyse s'appuie sur une observation générale des ARNm produits au sein d'un échantillon, sans tenir compte de la composition de celui-ci. Si une analyse d'expression différentielle met en avant l'enrichissement d'un gène entre deux échantillons, on ne pourra pas l'attribuer avec certitude uniquement à une variation dans la quantité d'ARN transcrite puisqu'une augmentation du nombre de cellules exprimant ce gène pourrait aussi en être à l'origine. Cette confusion peut aussi s'appliquer à un sous-type cellulaire particulier, des cellules originaires du même tissu et caractérisées comme appartenant au même type cellulaire ont dévoilé des dynamiques et des profils d'expression distincts qui ajoutent un degré de complexité dans l'interprétation des résultats (Rosenfeld et al. 2005). Jusqu'alors les analyses se basaient sur une représentation moyennée de cet ensemble de cellules hétérogènes. Or, pour comprendre le fonctionnement d'un système biologique dans toute sa complexité, il est nécessaire d'analyser chaque acteur, leur évolution et leurs interactions de façon individuelle. Et c'est pour répondre à ce besoin que le séquençage d'ARN en cellule unique a pu émerger. Son principe repose sur l'isolement de cellules individuelles, que l'on lyse dans le but de capturer puis d'amplifier les ARNm présents dans leur cytoplasme à la suite d'une transcription inverse (RT). La librairie de fragments d'ADN complémentaires (ADNc) générés est ensuite amplifiée puis séquencée. Il faut aussi prendre en compte que dans une cellule, on dénombre de 50000 à 300000 séquences d'ARNm (Marinov et al. 2014), ce qui ne représente que quelques picogrammes et impose donc des méthodes efficaces pour amplifier cette faible quantité de matériel, tel que l'amplification exponentielle en chaîne par polymérase (PCR).

### Le séquençage en cellule unique

Le premier séquençage en cellule unique a été effectué sur un blastomère d'un embryon de souris au stade 4 cellules (Tang et al. 2009). Leur étude du transcriptome complet par l'amplification des ARNm issus d'une seule cellule (aussi appelé WTA pour "whole transcriptome amplification") a pavé la voie à un portage des techniques existantes à l'échelle d'une cellule unique. Ce n'est pas un hasard si l'expérience pionnière de ce domaine est appliquée sur un blastomère puisque d'une part le contexte facilite l'isolement de cellules, et d'autre part c'est dans l'étude du développement embryonnaire que l'analyse unicellulaire est la plus logique, la compréhension du développement de l'individu reposant en effet sur les phénomènes prenant part dans un nombre très limité de cellules pluripotentes.

Les années ayant suivi ces expériences initiales ont vu fleurir un nombre grandissant de méthodes (Zhu et al. 2017; Svensson, Vento-Tormo, et Teichmann 2018; Chen, Ye, et Guo 2019) pour permettre d'améliorer la capture des ARNm et augmenter le nombre de cellules pouvant être analysées simultanément (**Figure 2**). On trouve tout d'abord des techniques reposant sur le tri unicellulaire en plaque, pour lesquelles la

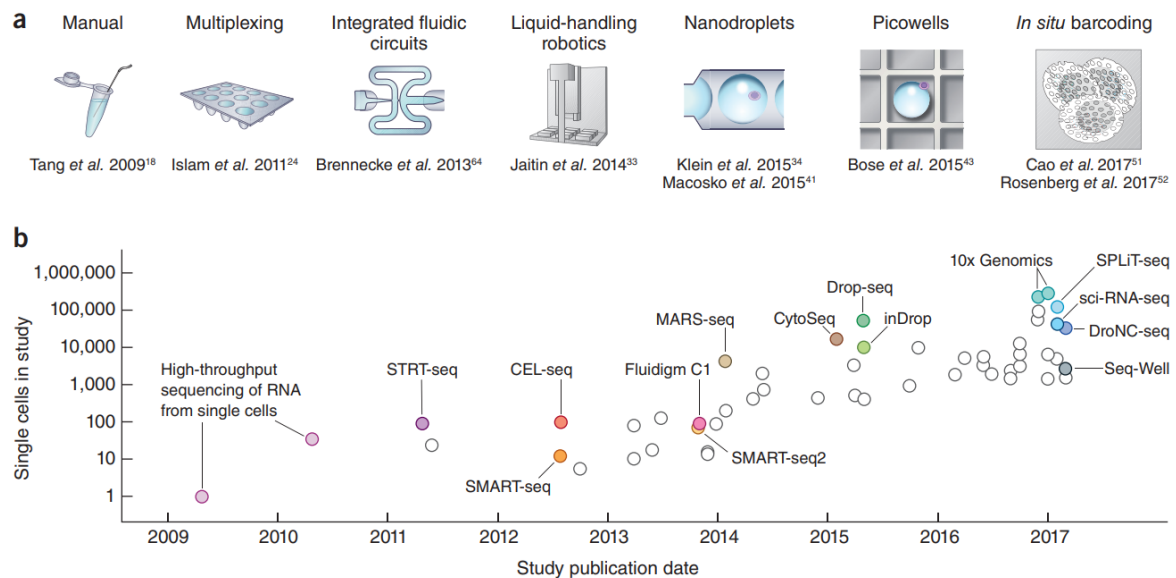


réaction d'amplification d'ADNc a lieu dans chaque puits. Parmi les techniques pionnières que l'on peut citer pour les avancées qu'elles ont apportées, on peut d'abord citer le STRT-Seq (Islam et al. 2011), avec l'introduction de codes-barres ("*barcodes*") uniques à chaque cellule isolée. Dans cette méthode, un oligonucléotide contenant une séquence poly-T (oligo-dT) se fixe sur la queue poly-A à l'extrémité 3' des ARNm et permet la transcription inverse de l'ARN en ADNc dans la direction 3'→5'. Arrivé en fin de séquence, la transcriptase inverse va ajouter 3 à 6 cytidines supplémentaires en continuité du brin d'ADNc nouvellement synthétisé. C'est sur cette séquence additionnelle que vient se fixer un oligonucléotide contenant une séquence de fixation aux nucléotides excédents (template switching oligo, TSO), le code-barre unique au puit et une séquence d'amorce universelle pour l'amplification par PCR. Les barre-codes permettent de multiplexer les différentes cellules avant la préparation de la librairie et donc d'amplifier les fragments de plusieurs cellules ensemble, ce qui réduit les coûts et les produits nécessaires.

À peine trois ans plus tard, c'est l'ajout d'un second niveau de séquences uniques, appelées Identifiant Moléculaire Unique (UMI), qui a permis d'améliorer l'estimation d'expression des cellules analysées en parallèle (Islam et al. 2014). Ces UMI ajoutés aux ADNc sont des séquences aléatoires, théoriquement uniques à chaque ADNc synthétisé lors de la rétrotranscription. Lors de l'amplification, pendant la préparation de la librairie, chaque brin amplifié porte l'information du transcrit d'origine en plus du code-barre associé à la cellule d'origine. À l'étape d'analyse du séquençage, on peut alors identifier les biais d'amplification et retrouver le nombre de transcrits capturés par gènes, ce qui réduit le bruit de fond et fait gagner en sensibilité, c'est-à-dire la capacité à détecter un gène exprimé (qui dépend du niveau d'expression du gène et de l'efficacité de capture).

Deux autres méthodes, le CEL-seq (Hashimshony et al. 2012) puis MARS-seq (Jaitin et al. 2014), appliquées respectivement à l'analyse de 7 cellules de blastomères, et à celle de 1500 cellules réparties sur une plaque 384 puits (**Figure 2**) se distinguent par l'utilisation d'une autre méthode d'amplification de l'ADNc avec la transcription in vitro (ITV). Contrairement à la PCR, l'ITV utilise l'ARN polymérase T7 et permet une amplification linéaire se voulant plus sensible et reproductible. En parallèle, le Smart-seq2 (Picelli et al. 2013) fut développé pour permettre le séquençage d'ARNm en *full-length*, c'est-à-dire que la séquence entière est amplifiée et séquencée. Cette technique utilise aussi un TSO pour fixer les cytidines qui terminent le brin d'ADNc synthétisé mais contrairement au STRT-seq qui ne permet de séquencer qu'une centaine de paires de bases depuis l'extrémité 5', on ajoute ici une étape de tagmentation. Une transposase Tn5 est utilisée pour tagmenter, c'est-à-dire fragmenter l'ADN et insérer des amorces à chaque extrémité libre par ligation. Ainsi, chaque fragment de gène peut être amplifié pour garantir une couverture complète des gènes exprimés lors du séquençage. Bien que n'utilisant pas d'UMI et étant difficilement adaptable à un grand nombre de cellules, le Smart-seq et le séquençage en *full-length* représentent un intérêt pour améliorer l'alignement des séquences, et pour l'étude des isoformes et variants alléliques. L'objectif suivant a ensuite été de

trouver les moyens d'étendre la technique à un plus grand nombre de cellules en s'affranchissant du tri, et c'est la microfluidique qui s'est imposée comme solution la plus adaptée.

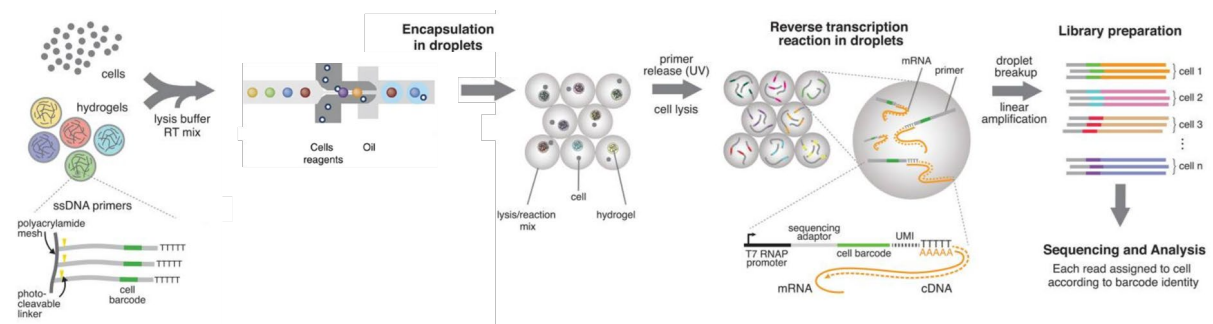


**Figure 2 : Évolution des méthodes de scRNA-seq.** (A) Expériences clés qui ont défini l'amélioration des capacités de séquençage en cellules uniques. (B) Représentation du nombre de cellules capturées par chaque méthode selon l'année de publication (Svensson, Vento-Tormo, et Teichmann 2018)

## La microfluidique au service de la transcriptomique

Les systèmes microfluidiques permettent d'isoler les cellules soit en puits soit en microgouttelettes. Le système Fluidigm C1 (Fluidigm, South San Francisco, CA, USA) est un des premiers systèmes de fluidique développé (**Figure 2**), et propose un isolement des cellules sur une plaque de 96 puits. Cette technique permet de vérifier par microscopie la présence des cellules dans chacun des puits, et souffre du manque de flexibilité pour être appliquée à des centaines voire des milliers de cellules. Pour remédier à cette limite, de nouvelles techniques de micro-fluidique à haut débit ont été développées afin de permettre un isolement automatisé des cellules. Elles se basent sur deux méthodes de scRNA-seq en gouttelettes : le DropSeq (Macosko et al. 2015), et inDrop (Klein et al. 2015) qui utilise des micro-billes d'hydrogel pour introduire les oligonucléotides, et qui a été adaptée par 10X Genomics (Zheng et al. 2017) pour une utilisation commerciale plus répandue. C'est la méthode de 10X Genomics que j'ai utilisée durant ma thèse pour sa mise en pratique aisée et sa polyvalence, assurant une meilleure capture de l'information transcriptomique en regard des technologies concurrentes (Mereu et al. 2020). Dans ce système de microfluidique, un premier canal distribue les microbilles de gel où sont fixés les code-barres dans des microgouttelettes contenant les réactifs de lyse et de RT (**Figure 3**). Il est estimé que de cette manière on obtient moins de 16% des gouttelettes sans bille de gel et parmi celles qui en ont, moins de 5% se retrouvent avec plus d'une bille de gel. Les cellules sont ensuite distribuées via un second canal en densité limitante, suivant une loi de

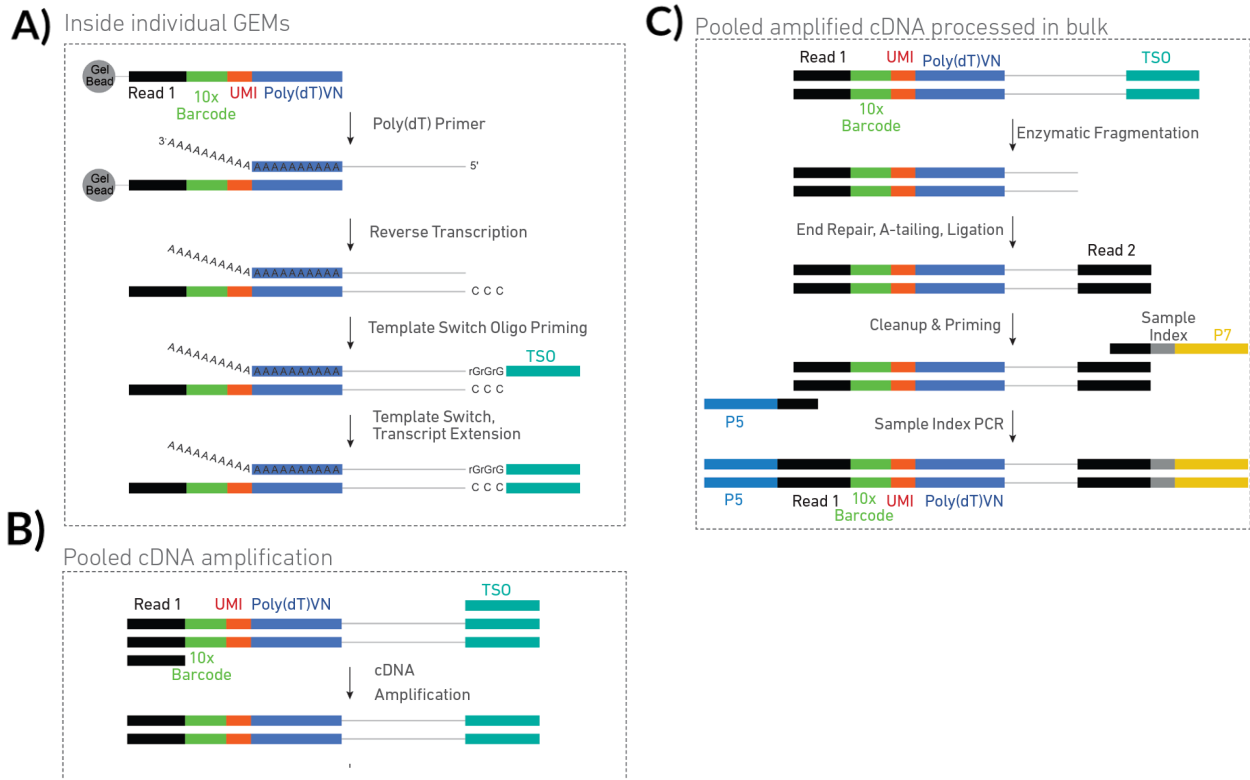
Poisson pour éviter de surcharger les microgouttelettes. Pour les cellules, leur dilution limitante induit que près de 95% des gouttelettes se retrouvent sans cellules, mais il garantit un faible taux de multipléts c'est-à-dire une gouttelette où le contenu de deux cellules différentes ou plus a été capturé.



**Figure 3 : Encapsulation des cellules uniques et capture des transcrits.**

Les cellules, les billes de gel annotées et les réactifs sont encapsulés par microfluidiques en microgouttelettes. Les cellules sont lysées dans l'émulsion et les réactifs initient la transcription inverse des ARNm puis leur amplification. Les ADNc barcodés sont ensuite amplifiés en bibliothèques pour le séquençage. Adapté de (Klein et al. 2015)

La réaction ayant lieu au sein de la gouttelette se déroule en plusieurs étapes. Premièrement, une fois encapsulées, les cellules sont lysées par les détergents contenus dans le tampon et répandent leur contenu cytoplasmique. En parallèle, les microbilles de gel sont dissoutes et libèrent les oligonucléotides contenant une séquence poly-T (oligo-dT) qui recouvraient sa surface. L'utilisation d'oligo-dT permet de capturer spécifiquement les ARN coiffés d'une queue poly-A, constitués majoritairement d'ARNm mais aussi en partie d'ARN longs non-codants (ARNlnc) (Y. Zhang, Yang, et Chen 2014). En contrepartie, d'autres types d'ARN ne codant pas pour des protéines (ARNnc) mais présentant un intérêt biologique ne sont pas analysés. Ces ARNnc contiennent entre autres les ARN ribosomiques (ARNr) constituant les sous-unités du ribosome, les ARN de transfert (ARNt) qui apportent les acides aminés au ribosome pour la synthèse des protéines, ou encore les microRNA qui sont des régulateurs post-transcriptionnels pouvant s'hybrider ou cliver un ARNm afin d'inhiber la traduction de la protéine. Les ARNnc représenteraient plus de 90% des ARN capturés (Slack et Chinnaiyan 2019) et leur capture réduirait la qualité des analyses transcriptomiques. Dans cette méthode de scRNA-seq, tout comme dans la méthode de Islam et al. (2014), les oligonucléotides utilisés en amorce contiennent un code-barre caractéristique de la bille, ainsi qu'un deuxième code-barre unique à chaque oligonucléotide, la séquence UMI (**Figure 4**). Cependant contrairement à la précédente technique où les deux code-barres sont situés dans le TSO à l'extrémité 5', on fait ici un séquençage en 3'. Les code-barres sont donc situés dans la séquence nucléotidique portée par la bille, avec la queue poly-dT.



**Figure 4 : Détail de la construction des séquences lors de la capture et l'amplification des transcrits au sein d'une microgouttelette**

A) Génération des Capture des ARNm et retrotranscription à l'intérieur des gouttelettes. TSO = template switching oligo. B) Amplification des ADNc suivant l'éclatement des gouttelettes. C) Construction de la librairie d'expression en 3'. Les amorces P5 et P7 servent à l'amplification PCR pour le séquençage Illumina. D'après le Guide Utilisateur Chromium Next GEM Single Cell 3' Reagent Kits v3.1 (CG000204 RevD)

Le premier code-barre permettra d'identifier tous les ARNm capturés dans la même gouttelette lors de l'étape d'analyse, tandis que les séquences UMI vont permettre de retrouver le nombre de transcrits capturés avant l'amplification des fragments. L'équipe à l'origine de cette méthode (Zheng et al. 2017) a démontré l'efficacité de la séparation des cellules dans ces billes de gel en émulsion (GEM) en regroupant dans une même expérience 1200 cellules humaines (293T) et 1200 cellules de souris (3T3). Après séquençage des librairies en alignant les brins aux génomes des deux espèces, moins de 2% des cellules capturées peuvent être considérées comme multiplètes. Dans ce cas, seuls les multiplètes inter-espèces sont retrouvés et pas ceux provenant d'un même individu, mais ils permettent déjà de confirmer l'efficacité de séparation par émulsion.

Lors des étapes suivantes, la microbille est dissoute et une première amplification est réalisée grâce à l'enzyme de transcriptase inverse (RT) pour synthétiser le brin complémentaire à la séquence capturée. L'émulsion est ensuite cassée et une seconde amplification s'effectue ensuite au moyen de la technique d'amplification en chaîne par polymérase (PCR), qui permet d'amplifier l'ADN complémentaire ainsi

généralisé en une multitude de copies afin d'avoir une quantité de matériel suffisant à la construction d'une librairie.

Au final, le taux de capture des cellules est estimé à environ 65 % du total de cellules chargées au départ et la quantité de transcrits d'ARNm capturés par cellule avoisine les 30% dans la version la plus récente (chimie v3)

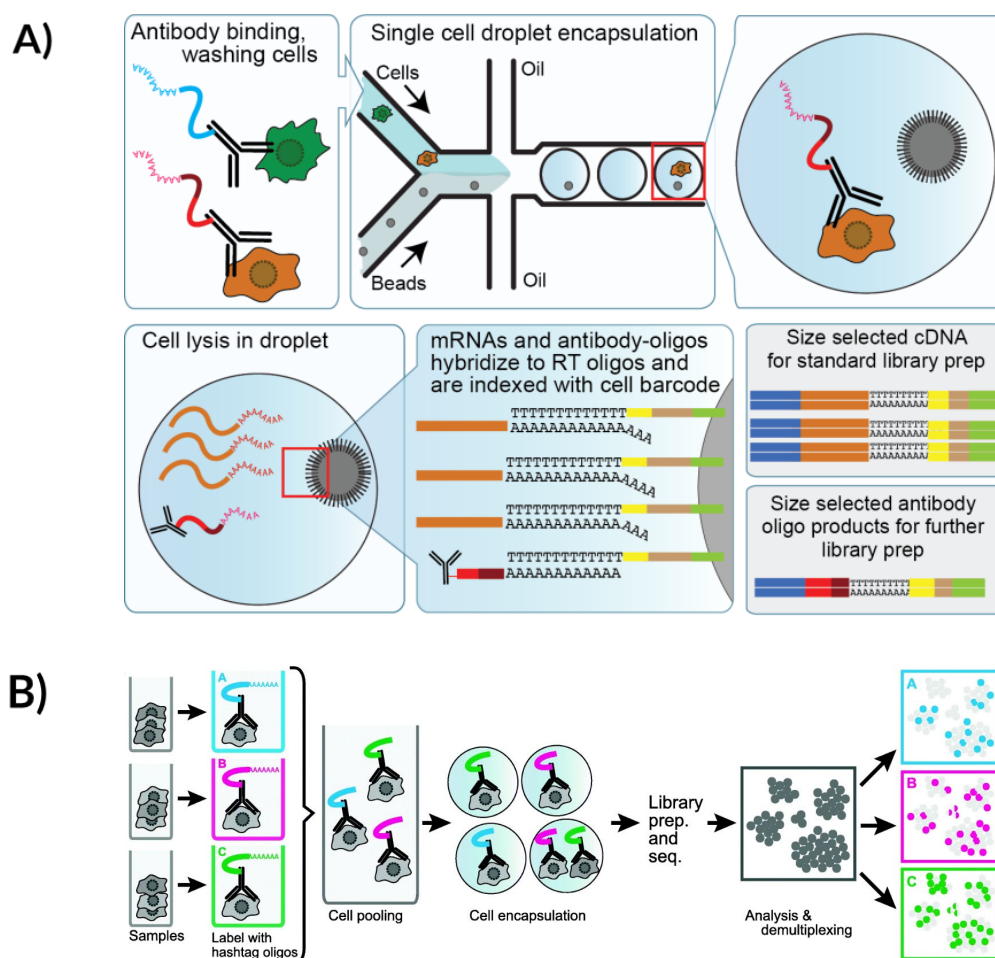
## Enjeux techniques et analytiques du scRNA-seq

### Encapsulation et multipléts

La question de l'encapsulation est un des points fondamentaux pour garantir l'unicité de la cellule capturée. La capture de multiples cellules au sein d'une même gouttelette provient majoritairement d'un risque statistique qui varie en fonction du nombre de cellules chargées lors de l'émulsion (Bloom 2018), mais peut aussi être lié à une mauvaise dissociation des tissus qui induit des agrégats cellulaires ou encore plus rarement à des interactions physiques entre cellules. Bien qu'il soit possible d'identifier des doublets en cas de présence simultanée de gènes marqueurs de deux types cellulaires distincts, il ne sera pas possible de déterminer la présence de doublets dans le cas de deux cellules proches en termes de fonction. En effet, la séquence qui permet la capture des ARNs est équipée d'un code-barre unique à l'unité de capture de la cellule, mais si deux cellules sont encapsulées dans la même bulle, chaque fragment capturé des deux cellules sera annoté du même identifiant spécifique à l'unité de capture.

Afin d'y remédier, il existe plusieurs solutions. La plus simple consiste simplement à charger un nombre plus faible de cellules pour garantir un nombre plus faible de doublets en suivant la loi de Poisson. Une autre possibilité est d'utiliser des outils informatiques pour démultiplexer les échantillons lors de l'analyse, par exemple en se basant sur les polymorphismes nucléotidiques (SNP) retrouvés dans les fragments séquencés, pour caractériser chaque cellule tel que le fait le programme Demuxlet (Kang et al. 2018). Enfin, une autre solution consiste à indexer les cellules afin de pouvoir les identifier séparément malgré leur grand nombre. Cette indexation peut se faire sans avoir besoin d'isoler chaque cellule, par l'utilisation d'amorces étiquetées dans des cellules ou noyau de cellules fixées (J. Cao et al. 2017). La fixation des cellules permet d'utiliser la cellule elle-même comme lieu de réaction. Les cellules sont réparties sur une plaque à puits, où chacun des puits contient une amorce connue qui lui est propre. Ces amorces sont des oligonucléotides qui fixeront les ARNm des cellules du puits grâce à leur queue poly-T, et dont la séquence de "code-barre" est connue pour chaque puits et servira à identifier l'origine de la cellule lors de l'analyse. Dans la méthode de Cao et al, les cellules sont ensuite mélangées et redistribuées dans les puits pour un second marquage en suivant la même méthode, et où le 2nd code-barre fait office d'UMI pour identifier le nombre de séquences capturées avant amplification. En n'utilisant que des plaques de 96 ou 384 puits et en ne chargeant que 10 à 100 cellules dans chaque puits grâce à un tri en cytométrie FACS

(Fluorescence Activated Cell Sorting), on obtient une dilution limitante qui permet de garantir une combinaison des deux indexes unique. Cette méthode peut être adaptée pour augmenter le nombre d'étapes de marquage comme avec la méthode de SPLIT-seq (A. B. Rosenberg et al. 2018) qui passe par 4 étapes de "barcodage" pour permettre de multiplexer un plus grand nombre de cellules. Une autre méthode de multiplexage plus adaptée à la microfluidique et dont ont bénéficié mes propres expériences de single cell durant ma thèse, consiste à utiliser des anticorps dits de "hachage" (Stoeckius et al. 2018) (**Figure 5**). Ces anticorps ciblent des protéines ubiquitaires (CD298 et la  $\beta$ -2-microglobuline) retrouvées sur une grande majorité des cellules étudiées en single-cell. Marquer les cellules en surface ne nécessite donc pas de perméabiliser les cellules au préalable. Les anticorps sont hybridés avec une séquence oligonucléotidique comprenant un code barre partagé et une séquence poly-A.



**Figure 5 : Méthode de microfluidique couplée à des anticorps porteurs de séquences code-barre d'identification.** (A) L'utilisation d'anticorps code-barre permet de garder l'identité de la cellule une fois celle-ci encapsulée par micro-fluidique. Les ARNm libérés par la lyse de la cellule et les anticorps code-barre qu'elle portait vont se fixer aux amorces portées par les billes de gel, avant d'être amplifiée par PCR. (B) Multiplexage d'échantillons permis par l'utilisation d'un anticorps barre-code spécifique à chaque échantillon. À l'étape d'analyse des résultats de séquençage, la séquence des code-barres capturée en parallèle des transcrits permet la réattribution des cellules à leur échantillon d'origine ainsi que l'identification de doublets. (Stoeckius et al. 2018)

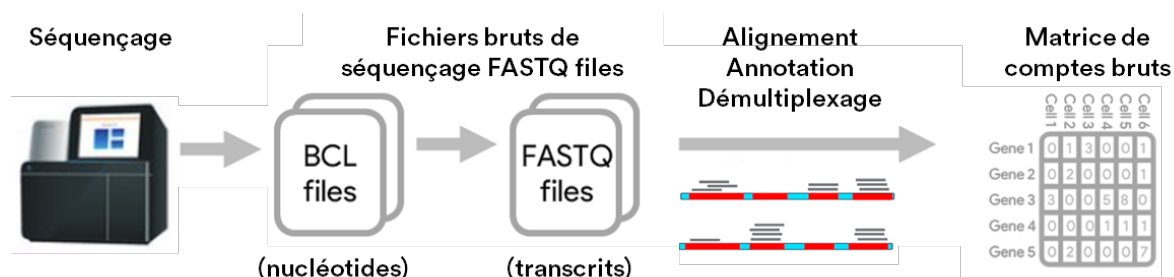
La séquence poly-A permet la capture de ce code-barre en parallèle des ARNm de la cellule marquée par anticorps à la suite de l'encapsulation. On peut dès lors identifier différentes sources de cellules en marquant chaque échantillon d'un anticorps différent, et les rassembler avant l'encapsulation pour permettre un multiplexage de plusieurs expériences en une. Dans ce cas, si deux cellules de source différentes sont encapsulées dans la même bulle, le code-barre de hachage permettra d'identifier le doublet pour son exclusion de l'analyse. Néanmoins, l'encapsulation de deux cellules provenant du même échantillon ne pourra pas être détectée, et il faudra effectuer une étape de contrôle qualité pour identifier les "cellules" présentant un nombre anormalement élevé de transcrits, ou en vérifiant que le profil cellulaire ne mélange pas des marqueurs de deux lignées indépendantes, auquel cas, il faudra exclure cette « cellule » de l'analyse. La méthode du multiplexage permet ainsi un certain degré de « surcharge » des cellules, puisqu'on optimise la reconnaissance des doublets, mais aussi de combiner les cellules de plusieurs expériences différentes en une seule.

### Vue d'ensemble de l'analyse des données de scRNA-seq

Malgré les différences dans les techniques utilisées pour préparer et capturer les cellules d'un échantillon, le résultat de séquençage sera sensiblement le même : on obtient des fichiers FASTQ qui contiennent les séquences nucléotidiques capturées et amplifiées auxquelles un score de confiance est associé selon la qualité du séquençage (**Figure 6**). La seule différence peut provenir de la nature de l'ARN capturé, si on étudie des ARN complets ou des ARN dont seule une extrémité de quelques dizaines de bases a été capturée, en 3' ou en 5'. Bien qu'un séquençage de transcrite entier inclurait l'information sur les différents isoformes présents, les SNPs ou les mutations des gènes, on se limite à séquencer une dizaine de paires de bases par transcrite car séquencer les transcrits complets réduirait la quantité de transcrits que l'on peut séquencer et ne permettrait pas d'analyser des cellules à haut débit. Dans la méthode la plus courante actuelle proposée par 10X Genomics (Zheng et al. 2017), le traitement des fichiers FASTQ passe dans un premier temps par l'extraction des différentes informations : la séquence en ARN elle-même, son score de qualité, les séquences de code-barre associées, et la séquence UMI du transcrite. Pour maximiser le nombre de séquences gardées, les codes-barres sont corrigés dans les cas où leur séquence est proche d'une séquence code-barre connue de 1 nucléotide de différence, qui peut être dû à un manque de qualité ou une erreur de transcription. Les séquences d'ARN sont ensuite alignées sur un génome de référence via le logiciel STAR (Dobin et al. 2013) pour attribuer chaque séquence à un gène connu, la longueur de la séquence capturée permet en général de s'assurer qu'une seule position peut correspondre à la séquence exacte. Le nombre moyen de séquences alignées sur une position du génome permet d'estimer la profondeur de séquençage, et pour s'assurer une bonne qualité d'information et une bonne puissance statistique pour interpréter les valeurs d'expression, il est nécessaire de séquencer une quantité suffisante de transcrits par cellules. Il est estimé qu'utiliser entre 10000 à 50000 transcrits par cellules est suffisant pour identifier le type cellulaire (Haque et al. 2017). On prend ensuite en compte les séquences UMI qui vont permettre de différencier les séquences attribuées à la même position qui proviennent soit de transcrits différents



soit de l'amplification des transcrits pendant la création de la librairie. On détermine à partir de cette analyse une matrice de comptes bruts où pour chaque gène de chaque cellule, on associe un nombre de séquences observées. Dans le cas où des anticorps code-barre ont été capturés en parallèle, une matrice des comptes par cellules pour ces séquences d'oligonucléotides sera aussi générée.



**Figure 6. Sommaire des transformations de données brutes post-séquençage**

À partir de cette matrice d'expression, on effectue un premier contrôle qualité pour restreindre l'analyse aux cellules contenant assez d'information. On exclura alors les cellules qui ont un nombre d'UMI trop faible (j'utilise un seuil de 800 UMI déterminé par expérience), pour lesquelles il serait trop compliqué d'attribuer un profil précis. À l'inverse, on peut aussi exclure les cellules dont le nombre de transcrits est anormalement haut, ce qui est certainement permis par la capture de deux cellules ensemble. Un autre paramètre lié à la qualité des cellules est le taux de transcrits liés aux ARN mitochondriaux qui sont retrouvés. Une cellule en cours d'apoptose exprime un nombre beaucoup plus important de gènes mitochondriaux et ces transcrits sont exportés vers le cytoplasme. On exclut alors les cellules qui ont un ratio de gènes mitochondriaux trop élevé afin de ne garder que les cellules de qualité suffisante à l'analyse. L'une des lacunes du scRNA-seq réside dans la capture parcellaire des transcrits d'une cellule. Pour pouvoir comparer les différentes cellules, il est alors nécessaire d'appliquer une méthode de mise à niveau entre les cellules. Cela peut passer par un sous-échantillonnage aléatoire des transcrits pour obtenir une profondeur commune de capture par cellule. Cela peut aussi se faire avec une normalisation des valeurs d'expression des gènes en fonction de la profondeur de séquençage (Luecken et Theis 2019). Il existe différentes méthodes de normalisations avec en premier lieu les normalisations de transcrits par millions, où l'on divise l'expression d'un gène d'une cellule par le total des transcrits de la cellule, que l'on multiplie par 1 million pour avoir une valeur compréhensible. Cependant cette méthode est sensible aux gènes fortement exprimés qui vont éclipser les autres gènes de la cellule, on préférera alors une méthode de normalisation en échelle logarithmique : on divise aussi l'expression d'un gène par le total des transcrits de la cellule, multiplié par un facteur choisi (a), mais on passe en échelle logarithmique (b) ensuite (en ajoutant 1 pour éviter les erreurs liées aux gènes non capturés).

$$a) \quad x_i = \frac{\text{Transcrits du gène } X \text{ dans la cellule}_i}{\text{Transcrits totaux dans la cellule}_i} \times 10^5$$

$$b) \quad f(x_i) = \log(x_i + 1)$$



À partir de ces valeurs on peut passer à l'étape d'analyse secondaire qui englobe la visualisation des cellules, l'identification des types cellulaires, puis toutes les analyses descriptives des mécanismes moléculaires comme la recherche de gènes différentiels ou de modules.

### La problématique du "Batch Effect"

Lors de l'intégration de multiples échantillons, que ces échantillons proviennent d'expériences séparées ou qu'ils soient multiplexés lors d'une unique expérience grâce aux anticorps code-barres, on rencontrera dans tous les cas des différences plus ou moins prononcées durant l'analyse, qui peuvent provenir de deux origines : d'une variabilité technique que l'on nomme "*Batch effect*" et de la variabilité biologique intrinsèque aux échantillons qui doit être prise en compte (âge, sexe, condition).

La variabilité technique comprend toutes les différences expérimentales induites de la préparation des cellules au séquençage. Certains facteurs techniques peuvent être contrôlés pour limiter les biais entre expériences, comme le fait d'étudier les mêmes cellules, suivre un protocole unique ou utiliser les mêmes outils d'analyse. Mais on y trouve aussi des variables qui sont des facteurs de biais sur lesquels l'utilisateur n'a pas de contrôle : l'efficacité d'encapsulation des cellules en micro-fluidique, l'efficacité de la lyse cellulaire et des enzymes utilisées, l'amplification inégale des ADNc pendant la PCR, l'échantillonnage aléatoire des brins qui sont séquencés ou encore des facteurs inconnus qu'il faudra identifier si un biais est observé. Un moyen de mesurer cette variabilité a d'abord été développé pour le séquençage en bulk puis adapté au scRNA-seq en plaque, via l'utilisation d'ARN synthétiques (Baker et al. 2005; External RNA Controls Consortium 2005; Lun et al. 2017). Ces ARN synthétiques aussi appelés "*spike-in*" sont des transcrits dont la séquence et la quantité est connue et qui est ajoutée à chaque puits de manière homogène afin d'établir un contrôle externe de l'efficacité de capture des transcrit pour chaque cellule unique qui sera dans le même puits. Cette expérience a permis de quantifier la variabilité de capture des ARN et de normaliser les résultats avec, cependant ce contrôle ne peut pas être utilisé en microfluidique à cause de la faible distribution de cellules. Il y a près de 90% des gouttelettes sans cellules dans lesquelles on amplifierait uniquement ces ARN synthétiques, ce qui noierait l'information transcriptomique réelle des cellules. C'est pourquoi on utilise des UMI en micro-fluidique, ces séquences uniques aléatoires permettent d'estimer le nombre absolu de transcrits capturés pour chaque cellule et de réduire grandement le biais d'amplification induits lors de la préparation de la librairie. La variabilité biologique qui n'est pas au centre de l'étude peut aussi compliquer les comparaisons, puisque chaque individu a une hétérogénéité cellulaire qui lui est propre, et qui fluctue au cours du temps (Kanodia et al. 2019; Lakshmikanth et al. 2020). À cela viennent s'ajouter des variables comme le phénotype, le sexe, l'âge, l'ethnie ou le statut physiologique du patient qui vont aussi influencer le profil transcriptomique des cellules (Márquez et al. 2020; Z. Huang et al. 2021). Il est donc préférable d'identifier ces sources de variations, et les garder en co-variables puisqu'elles peuvent se révéler importantes pour expliquer la variabilité biologique d'intérêt, que ça soit un impact sur les mécanismes d'action d'un traitement, ou sur la pathogenèse comme dans les maladies inflammatoires chroniques de l'intestin où la

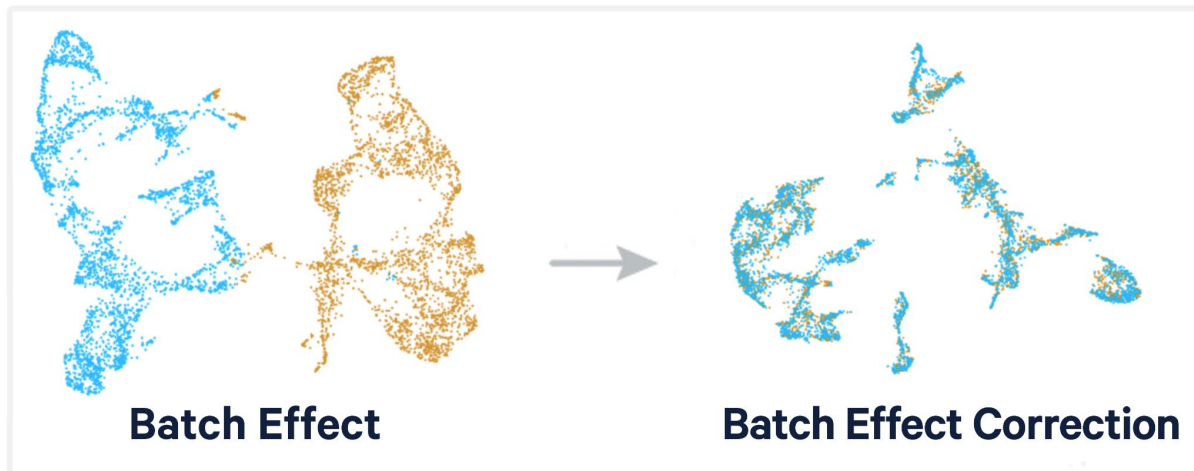
maladie de Crohn est plus fréquente chez les femmes tandis que les hommes ont de plus forts risques d'être atteints par une rectocolite hémorragique (Goodman, Erkkila, et Pizarro 2020).

## Le choix d'une méthode de correction du batch effect

La correction de la variabilité indésirable est un enjeu dont l'application ne fait pas consensus. Son objectif est de permettre une comparaison de différents échantillons ou de différentes expériences en identifiant et réduisant les différences dues à la variance technique pour tenter de ne garder que les variations biologiques. Ainsi, il est possible d'entreprendre l'analyse jointe de différents jeux de données à des fins de clustering et de visualisation.

En scRNA-seq, pour chaque cellule, on a des dizaines de milliers de gènes exprimés qui comptent chacun pour une dimension. Pour avoir une vue compréhensible des données, et essayer d'estimer la présence d'un batch effect, on est alors obligés de synthétiser les observations tout en essayant de garder le maximum d'information. C'est ce qu'on appelle la "réduction de dimension", qui peut se visualiser via des projections en deux dimensions formant des nuages de points où chaque point résume le transcriptome d'une cellule. Pour rentrer plus dans le détail, la méthode la plus simple de réduction est l'analyse en composante principale, ou PCA (Hotelling 1933; Jolliffe et Cadima 2016), qui va transformer de façon linéaire des gènes potentiellement corrélés pour former des variables artificielles non corrélées appelées composantes principales (PC). Pour maximiser la variance expliquée La variance la plus importante existant dans les données va être résumée dans la première coordonnée, la PC1, puis la seconde variance la plus forte sera expliquée par la PC2, et ainsi de suite. On obtient alors un espace avec un nombre fortement diminué de dimensions, tout en préservant le maximum d'information statistique sur la variabilité des données. Toutefois, une limite de la PCA est que son modèle linéaire ne permet pas de préserver une information locale, c'est-à-dire de garder proche les points au profil complet très similaires. C'est dans ce but qu'a été développée la méthode non-linéaire de visualisation appelée "t-SNE", pour t-distributed stochastic neighborhood embedding (van der Maaten et Hinton 2008; Maaten 2014). Cet algorithme cherche à simplifier les données en les ajustant à une distribution t de Student, tout en gardant la structure locale des données. C'est-à-dire que les cellules qui ont des profils d'expression mutuellement proches, que l'on appelle plus proches voisins, vont aussi se retrouver proches sur le plan (2D ou 3D) simplifié de la t-SNE. Plus récemment encore, une nouvelle méthode non-linéaire, la technique UMAP pour uniform manifold approximation and projection (McInnes, Healy, et Melville 2018), s'est imposée. Elle permet à la fois de garder la structure locale des données, mais aussi la structure globale tout en réduisant les temps de calculs en comparaison à la technique de t-SNE. C'est-à-dire que dans la représentation des données en UMAP, deux points ne seront pas éloignés seulement parce qu'ils ne se ressemblent pas comme en t-SNE, mais leur distance dépendra de la valeur de leur différence. Ces représentations 2D permettent d'estimer rapidement les différences liées au batch effect et leur importance car elles vont impacter la distribution de chaque cellule sur le plan. En

effet, si les différences entre échantillons sont assez importantes, une représentation en dimensions réduites (PCA, t-SNE ou UMAP) montrera des groupes de cellules séparés par échantillon au lieu de seulement les distribuer par leur phénotype. Corriger la variabilité entre échantillons va donc permettre de s'assurer une distribution comparable (**Figure 7**) ou au contraire d'identifier des populations de cellules qui se détachent de par leur profil transcriptomique unique.



**Figure 7 : La correction du batch effect permet d'homogénéiser la distribution des cellules en dimensions réduites.**

Dans le déroulé de l'analyse secondaire proposée avec le package Seurat, une méthode d'intégration à des fins de visualisation et d'annotations est proposée (Stuart et al. 2019). Pour alléger les calculs, on réduit la matrice de comptes à une liste de 2000 gènes les plus variables qui sont suffisants pour résumer la majorité de la variabilité des données tout en ignorant les gènes peu exprimés qui causent du bruit de fond aléatoire. La sélection de ces gènes aléatoire se base sur les valeurs d'expression moyenne et la variance de chaque gène pour ne garder que les gènes dont le coefficient de variation est le plus élevé. La correction de la matrice de comptes est ensuite effectuée par une analyse canonique des corrélations (CCA) (Butler et al. 2018) combinée à une normalisation des valeurs de CCA. La CCA cherche à réduire les dimensions pour comparer les corrélations entre groupes de variables quantitatives et trouver les groupes qui seront le plus corrélés à l'effet batch. On recherche ensuite des paires de cellules avec un profil commun, une de chaque jeu de données, que l'on appelle voisins mutuels (ou MNN pour *mutual nearest neighbours*) (Haghverdi et al. 2018), qui vont servir "d'ancres" pour calculer un score d'effet batch par MNN. Les MNN et leurs scores sont ensuite utilisés pour transformer la matrice d'expression brute de ces cellules pour que les cellules qui étaient proches dans l'espace réduit deviennent aussi proches quand on regarde l'ensemble de l'expression de leurs gènes. De cette manière, on s'attend à faciliter l'étape de regroupement des cellules pour identifier les populations cellulaires.

Il est toutefois souvent déconseillé d'utiliser les données transformées des méthodes d'intégration pour des analyses plus poussées comme une recherche de gènes différentiellement exprimés (gènes DE) si on ne connaît pas très bien le

fonctionnement de la correction. En effet, la correction risque d'induire un nouveau biais chez les gènes qui n'étaient pas affectés par l'effet batch précédemment. On peut aussi se heurter à des problèmes de correction abusive qui gommerait des variations d'origines biologiques en plus des variations techniques. Certaines méthodes sont aussi limitées par la puissance de calcul nécessaire qui augmente le nombre de cellules à corriger.

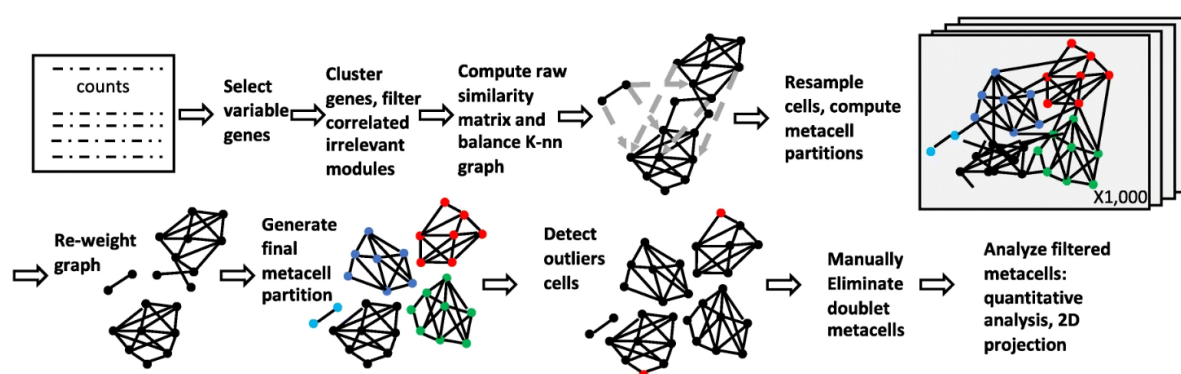
Des projets comme le Human cell Atlas (Regev et al. 2017) ou le Human Tumor Atlas Network (Rozenblatt-Rosen et al. 2020) disposent de jeux de données qui s'élèvent désormais à plusieurs millions de cellules, et même à l'échelle d'une équipe, le séquençage de plusieurs patients peut facilement amener à une centaine de milliers de cellules à analyser. Or à cette échelle, certaines méthodes d'analyse ou d'intégration commencent à poser des limitations techniques de puissance de calcul ou de quantité de mémoire nécessaire. Des méthodes plus optimisées ont été proposées, comme la RPCA (*Reciprocal principal component analysis*), qui comme la méthode CCA va calculer des "ancres" pour corriger les données, mais cette fois ci en utilisant la PCA comme méthode de réduction de dimension, en projetant par paires un échantillon sur la PCA d'un autre échantillon. De plus, les données utilisées pour la PCA ne sont pas mises à l'échelle (c'est-à-dire ramener la moyenne d'expression à 0 et l'écart-type à 1), ce qui permet d'éviter de corriger à l'excès des différences biologiques. Cette méthode permet ainsi de réduire les temps de calcul et de mieux préserver les populations ou états d'activations de cellules qui ne sont pas partagés entre tous les échantillons.

## La méthode Metacells

Un autre moyen d'optimiser l'intégration d'un grand nombre de cellules est de les regrouper par profils similaires. C'est le cas de la méthode des Metacells (Baran et al. 2019; Ben-Kiki et al. 2022), dont l'objectif est de créer des "méta-cellules" (metacells) à partir d'un groupement de cellules au profil transcriptomique proche dont l'expression serait fusionnée. En scRNA-seq, on ne capture qu'une fraction des transcrits exprimés par cellules, environ 30% pour les méthodes les plus récentes. On a donc dans chaque cellule une information incomplète de son profil avec une certaine quantité de gènes ne sont pas capturés ou que de façon parcellaire. Ce phénomène de "données éparses" induit une variance d'expression liée à l'échantillonnage et forme un bruit de fond qui gêne l'analyse de la variance biologique. Plutôt que de créer un modèle statistique pour tenter de réduire cet effet ou de faire de l'imputation de données (estimer les transcrits manquants pour les rajouter), la méthode de metacells garde l'information disponible, et part de l'hypothèse que les cellules ayant un profil transcriptomique très proches pourraient représenter chacune un sous-échantillonnage d'une même et unique cellule, qui est représentée par la metacell. Pour approximer l'expression de cette metacell, l'expression d'un gène de la metacell est le résultat d'une moyenne géométrique régularisée de l'expression de ce gène dans les cellules incluses. La moyenne géométrique permet de réduire le poids des valeurs extrêmes dans la moyenne et la régularisation empêche un ajustement excessif des données. Ainsi, cette méthode permet de gagner en robustesse en

réduisant l'effet d'échantillonnage sous-optimal des données et permet d'identifier des états transcriptomiques rares avec une meilleure sensibilité.

Pour la création des metacells (**Figure 8**), l'algorithme utilise les valeurs d'expression des gènes variables pour regrouper les gènes et calculer une matrice de similarité entre cellules. Il calcule ensuite un graphe des plus proches voisins (ou graphe Knn pour *K-nearest neighbour*), où chaque cellule représente un nœud, et la connectivité entre nœuds est estimée à partir de la matrice de similarité. À partir de ce graphe, l'algorithme va chercher à effectuer une séparation optimale pour grouper les cellules très proches transcriptomiquement en des metacells. Le graphe Knn est ensuite sous-échantillonné de multiple fois pour construire des sous graphes. Les cellules qui se retrouvent le plus connectées dans les sous-graphes vont se retrouver dans les même metacells. Pour pouvoir intégrer de grands jeux de données dans l'algorithme de Metacell sans devoir comparer toutes les cellules entre elles, une étape permet de diviser ce problème complexe en plusieurs problèmes plus simples : les cellules sont divisées en groupements aléatoires de 10000 cellules environ, sur lesquels sont calculés une première version de metacells. Les cellules aux valeurs aberrantes sont gardées de côté pour ensuite relancer le calcul de metacells sur les cellules aberrantes récupérées des différents groupements. Une fois toutes les metacells calculées, l'algorithme considère ces metacells comme des observations uniques pour relancer un second niveau de calcul et grouper les metacells par profil. Une fois ces "meta-groupes" identifiés, on revient au niveau des cellules, où cette fois ci on a toujours des groupements de cellules de taille adaptées aux analyses, mais désormais avec des profils proches et pas aléatoires, et donc des cellules avec une plus grande connectivité et des distributions d'expression plus homogènes ce qui permet une plus grande sensibilité pour regrouper les cellules. Un dernier calcul de metacells est donc appliqué dans chaque méta-groupe pour obtenir les metacells finales



**Figure 8. Schéma d'explication du principe de Metacell.**

Dans un premier temps, les gènes variables sont sélectionnés pour calculer une matrice de similarité entre cellules selon leur profil transcriptomique. De cette matrice est créé un graphe dont les connections dépendent de la similarité d'expression. Le sous échantillonnage aléatoire de ce graphe permet de recalculer des sous graphes qui identifient les cellules les plus connectées. Ce poids permet la ségrégation des cellules en metacells qui sont ensuite "nettoyées" des valeurs extrêmes. D'après (Baran et al. 2019)

Cette méthode est donc plus spécifique pour détecter des profils transcriptomiques rares, et plus sensible puisqu'elle maintient une cohésion locale même dans de grands jeux de données. Les outils d'intégrations basés sur la correction d'expression permettent souvent de ne comparer que ce qui est commun et peuvent corriger à l'excès des indices biologiques pertinents. Une méthode qui offre plus de souplesse comme Metacells permet alors de s'en affranchir et de regarder des programmes moléculaires qui se distribuent différemment dans des cellules qui se ressemblent.

## Objectifs du scRNA-seq

Le séquençage d'ARN en cellule uniques permet donc d'analyser l'expression des gènes de manière quantifiable et spécifique au sein d'une population de cellules. L'utilisation de cette technique ouvre de nouvelles fonctionnalités pour analyser le vivant et notamment caractériser les processus immunopathologiques. En premier lieu, le séquençage en cellules unique permet de combler une des lacunes de l'analyse en bulk puisqu'il est désormais possible de caractériser le transcriptome de populations cellulaires rares (Grün et al. 2015) et même d'en identifier de nouvelles à partir de leur profil transcriptomique comme l'identification de nouveaux sous-types cellulaires dans les îlots pancréatiques dans les premières expériences notables de scRNA-seq (Segerstolpe et al. 2016) ou encore de caractériser dans le sang une nouvelle population immunitaire de cellules dendritiques AXL+ SIGLEC6+ (Villani et al. 2017). Cette étude de la diversité cellulaire s'étend aussi à l'analyse de pathologies, comme dans le cancer où l'approche en scRNA-seq permet d'établir un historique des clones tumoraux (Kuipers, Jahn, et Beerenwinkel 2017). En plus de caractériser la diversité cellulaire, il est possible d'observer des différences d'expression et d'en découler les fonctions assurées par les cellules. Dans des maladies auto-immunes comme la polyarthrite rhumatoïde (PR), le scRNA-seq a permis de définir des sous-types de fibroblastes spécifiquement enrichis dans les patients atteints de PR et qui participent au recrutement de cellules immunitaires (Mizoguchi et al. 2018). En comparant des phénotypes sains et pathologiques, il est alors possible de mieux comprendre l'étiologie et les fonctions impactées par une maladie, et ce, de manière propre au patient (Y. Cao et al. 2020; Shalek et Benson 2017). Les maladies auto-immunes sont complexes et souvent, plusieurs mécanismes pathologiques différents peuvent aboutir à des manifestations histologiques ou cliniques semblables, ce qui complique la découverte et l'application de traitements. Le profilage transcriptomique des individus pourrait ainsi permettre d'identifier des voies pathogéniques en s'affranchissant du diagnostic clinique et histologique afin de proposer aux patients des thérapies plus personnalisées en lien direct avec les mécanismes moléculaires en jeu.

# Chapitre 2 : Le rejet chronique et la transplantation rénale

Le scRNA-seq offre un nouvel outil exploratoire pour caractériser plus en détails les mécanismes intra-cellulaires à l'origine de pathologies et l'objet de ma thèse a principalement porté sur l'application de cette méthode émergente à des sujets médicaux concrets. Le premier projet auquel j'ai pris part m'a permis de générer et d'analyser des données de scRNA-seq basées sur une cohorte de patients transplantés du rein qui regroupe des rejets chroniques et des patients stables.

## Rein et néphropathie

Les reins sont une paire d'organes vitaux situés de part et d'autre de la colonne vertébrale dans la partie postérieure de l'abdomen assurant des rôles multiples. Ils permettent en premier lieu la filtration du sang par excrétion et réabsorption tubulaire afin d'éliminer les divers déchets transportés par le sang, comme l'urée ou la créatinine. Ils sont constitués d'un million d'unités appelées néphrons, dans lesquelles on retrouve le glomérule, qui reçoit et filtre le plasma par un système de capillaires, et le tubule, qui permet la réabsorption et sécrétion pour produire l'urine finale. La réabsorption permet d'assurer l'homéostasie en régulant la quantité d'eau et de sels minéraux de l'organisme, et tout surplus sera éliminé dans les urines. Et par sa fonction excrétrice, les reins sécrètent les vitamines et hormones qu'ils produisent, comme la rénine pour contrôler la pression artérielle, ou l'érythropoïétine, pour stimuler la production d'érythrocytes au niveau de la moelle osseuse.

Une atteinte au bon fonctionnement de cet organe est donc dangereuse, et peut entraîner une destruction irréversible du rein. L'insuffisance rénale chronique (IRC) est une diminution progressive et irréversible de la fonction du rein et représente un enjeu important puisqu'elle touche plus de 3 millions de français d'après la Haute Autorité de Santé. L'allongement de la durée de vie, l'augmentation de l'incidence du diabète et de l'hypertension artérielle favorisent l'apparition de cette maladie et participent à expliquer l'augmentation du nombre de cas années après années. Parmi ces patients, l'agence de la biomédecine dénombrait en 2020 un nombre de 91 875 patients traités pour insuffisance rénale chronique terminale, dont 55% étaient traités par dialyse, et les 45% restants par greffe en France.

En clinique, pour suivre la dégradation du rein, on mesure la fonction rénale par le dosage des éléments filtrés par le rein, notamment le dosage de la créatinine et par la protéinurie. La créatinine est un déchet métabolique issu de la créatine stockée dans les muscles lors de son utilisation pour la production d'énergie, et son élimination est effectuée par les reins. On mesure dans le sang le taux de créatinine pour déterminer le débit de filtration glomérulaire (DFG), et l'augmentation de son taux permet de révéler une insuffisance rénale. La protéinurie est quant à elle dosée dans les urines. En effet, la barrière de filtration glomérulaire n'est pas censée pouvoir laisser passer les protéines comme l'albumine, et une forte protéinurie révèle une

potentielle une dégradation de la barrière de filtration glomérulaire. La mesure de ces niveaux permet donc d'établir la gravité de l'IRC jusqu'à l'insuffisance rénale terminale. Dans ce cas, une suppléance du rein est nécessaire pour compenser la perte de fonction rénale. Le premier moyen de suppléance est l'hémodialyse, où le sang du patient va être filtré par une machine. Ce traitement est contraignant pour le patient puisqu'il doit faire en moyenne trois séances par semaine, le plus souvent à l'hôpital. Une meilleure solution consiste à effectuer une transplantation de rein, qui bien que nécessitant un suivi au début et un traitement immunosuppresseur à vie, va permettre une amélioration de la qualité de vie du patient et une plus grande chance de survie (Wolfe et al. 1999; Heldal et al. 2010)

## Le recours à la greffe de rein

En 2021, l'agence de la biomédecine a recensé 3215 greffes rénales dont 502 à partir de donneurs vivants, ce qui fait de la transplantation rénale la greffe d'organes la plus commune. Le nombre de transplantation de rein est en constante augmentation à part ces 3 dernières années à cause des déprogrammations d'opérations dues à l'épidémie de COVID-19. Cependant, la quantité de candidats à la greffe de rein augmente aussi, en 2017 on dénombrait 18 793 personnes en liste d'attente et chaque année le nombre de nouveaux inscrits excède le nombre de transplantations effectuées, ce qui incite non seulement à améliorer le recensement des donneurs, la compatibilité de greffes, mais aussi de réduire l'incidence des rejets de greffes, surtout dans les rejets chroniques.

En effet, bien que la survie à court terme ait été grandement améliorée pour atteindre une survie de près de 90% des greffons après 1 an, la survie des greffons sur le long terme montre une amélioration assez lente et plutôt récente (Poggio et al. 2021), grâce à une meilleure prise en charge des comorbidités et une meilleure gestion des greffons provenant de patients décédés. Actuellement la majorité des greffons (85%) proviennent de patients décédés en état de mort encéphalique, mais il a été montré qu'un greffon provenant d'un donneur vivant augmentait la survie du greffon (Nemati et al. 2014). En France on estime à 14 ans la demie vie d'un greffon issu d'un patient décédé et de 20 ans quand il est issu d'un donneur vivant. Il reste donc encore des améliorations à apporter à la gestion des greffons de patients décédés, mais il serait aussi intéressant de développer le recrutement de donneurs vivants pour garantir les conditions optimales à la survie du greffon. On pourra aussi noter que pour faire face à la pénurie de greffons, l'amélioration de la survie des greffons déjà transplantés n'est pas la seule solution, de nouvelles approches que l'on pourrait croire tirées de la science-fiction commencent à émerger. La xénotransplantation par exemple, est un moyen prometteur pour garantir un accès facilité à un plus grand nombre de greffons. Des premiers tests de xéno greffe de reins de porcs sur deux humains en état de mort cérébrale ont montré une acceptation de la greffe sans rejet hyper aigu, grâce à un alpha-1,3-galactosyltransferase, qui est le xénoantigène porcine majeur reconnu par les lymphocytes T (Montgomery et al. 2022).



Et enfin une autre piste est la génération en laboratoire d'un organoïde, une structure 3D mimant un organe ou un tissu, généré à partir de cellules du patient. En premier lieu un organoïde pourrait servir pour faire de la médecine personnalisée, puisqu'il permettrait de tester des traitements en parallèle pour les adapter par patient. Dans un deuxième temps, l'amélioration de la technique laisse entrevoir l'utilisation des cellules cultivées in-vitro pour repeupler les tissus lésés ou permettre une réparation sûre d'une partie de l'organe sans risque de rejet. Toutefois, ces nouvelles stratégies sont encore loin de rivaliser avec les conditions de greffe actuelles, et il reste plus stratégique de se concentrer sur la compréhension des différentes causes de rejet et une amélioration du suivi de greffe afin de garantir une survie sur le long terme des greffons.

## Compatibilité HLA

L'acceptation du greffon par le receveur, qu'il provienne d'un donneur décédé ou vivant, repose majoritairement sur une bonne compatibilité du groupe sanguin ainsi que du complexe majeur d'histocompatibilité (CMH) afin d'éviter la réponse allo-immune contre les CMH du donneur (Opelz et Döhler 2007).

Les glycoprotéines du CMH, aussi appelées HLA chez l'homme (pour "human leukocyte antigen"), sont codées par un complexe de gènes polymorphiques du chromosome 6 qui se répartissent en plus de 25000 allèles (Robinson et al. 2020) et sont réparties en deux classes. Les CMH de classe I ont un rôle de présentation des antigènes endogènes, que ça soit des antigènes du soi ou des antigènes du non-soi, en cas d'infection virale par exemple. Le CMH de classe I est présent sur toutes les cellules nucléées et est reconnu par le récepteur des cellules T (TCR) des lymphocytes T CD8+. Il existe 3 loci classiques pour les HLA de classe I : HLA-A, -B et -C. Le CMH de classe II est lui restreint aux cellules présentatrices d'antigènes (APC). Ils permettent de présenter des antigènes exogènes, internalisés par phagocytose ou pinocytose, qui sont reconnus par le TCR des lymphocytes T CD4+. Il existe 3 loci classiques pour les HLA de classe II : HLA-DP, -DQ et -DR, séparés en une chaîne alpha et une chaîne bêta pour former un hétérodimère.

Dans le cas d'une greffe de rein, on réalise un typage en amont des HLA-A, -B, -DRB1 et -DQB1 pour chercher une compatibilité du donneur avec le receveur. Plus les épitopes des HLA sont partagés entre le donneur et le receveur, moins il y aura de risque d'émergence d'anticorps dirigés contre les HLA du donneur : les DSA (donor-specific antibodies).

L'Agence de la Biomédecine appelle "hyperimmunisés" les patients dont le taux de greffons incompatibles dépasse 85% pendant la phase d'attente. En 2020 ils représentaient 11% des greffés rénaux et 24% des patients en liste d'attente. Il est compliqué de garantir un accès à la transplantation rénale pour les patients hyperimmunisés ou porteurs d'allèles HLA rares mais un système de points a été mis en place pour prendre en compte cet élément en plus de l'ancienneté d'inscription sur liste d'attente, l'âge du patient et son taux de compatibilité au greffon prélevé pour garantir une plus grande équité.

## Rejets aigus et rejets chroniques

Dans la transplantation d'organes, on distingue trois types de rejets selon la temporalité durant laquelle ils ont lieu.

Le rejet le plus précoce est le rejet hyper aigu, qui survient dès les premières minutes suivant la greffe et provoque un infarctus du greffon. Ce rejet est lié à la présence d'anticorps anti-HLA (ou DSA) du donneurs pré-existants (Kissmeyer-Nielsen et al. 1966). En cas d'incompatibilité, les DSA vont se fixer à l'endothélium vasculaire du greffon et induire des lésions par activation de la voie du complément. Ces anticorps vont aussi pousser les cellules endothéliales à acquérir un phénotype pro-coagulant, qui va provoquer des thrombi microvasculaires par agrégation plaquettaire et donc une perte du greffon par nécrose des tissus. Il peut être prévenu par un test de compatibilité lymphocytaire appelé "cross-match" pour vérifier que les lymphocytes du donneur ne soient pas attaqués par le sérum du receveur.

Le second est le rejet aigu, il intervient souvent dans les premières semaines suivant la greffe, une fois que les cellules immunitaires ont repeuplé les tissus du greffon. Ce rejet peut être évité grâce à une adaptation précoce du traitement en réponse aux premiers signes de rejet. La classification de Banff a été établie par un consortium de cliniciens et pathologistes en 1991 afin d'établir une base commune au diagnostic des rejets de greffe de rein et d'autres lésions observées et est maintenue à jour tous les deux ans pour prendre en compte l'évolution des connaissances et améliorer les critères de diagnostics. On distingue deux types majeurs de rejets dans les rejets aigus : les rejets médiés par cellules T (TCMR), et les rejets médiés par anticorps (ABMR), tous deux inclus dans la classification la plus récente de Banff, qui découpe le diagnostic en 6 catégories (**Table 1**) (Roufosse et al. 2018; Haas et al. 2018; Alexandre Loupy, Mengel, et Haas 2022).

<b>Banff Diagnostic Categories</b>	
<b>Catégorie 1</b>	Greffon normal, changements non spécifiques
<b>Catégorie 2</b>	ABMR, qui selon les lésions et la temporalité peut être aiguë ou chronique active
<b>Catégorie 3</b>	Suspicieux, ou Borderline, dans le cas ambigu où des lésions ressemblant à un TCMR aigu sont observées, mais à un degré trop faible
<b>Catégorie 4</b>	TCMR, qui selon les lésions et la temporalité peut être aiguë ou chronique active
<b>Catégorie 5</b>	Fibrose interstitielle et atrophie tubulaire (IFTA)
<b>Catégorie 6</b>	Changements autres non liés à un rejet aigu ou chronique (néphropathie lié au virus BK, Néphrite interstitielle causée par les traitements, lésion tubulaire aiguë, etc)

**Table 1** : Catégories des biopsies de greffon de rein selon la classification de Banff de 2017.

Pour établir le diagnostic à partir de biopsie, il est donc nécessaire d'avoir une vision assez large des structures tissulaires pour une observation fiable. On parle de biopsie "adéquate" quand plus de 10 glomérules et 2 artères de moyen calibre ou plus sont visible, de biopsie "limite" si 7 à 9 glomérules et 1 artère de moyen calibre est visible, et en dessous, la biopsie est jugée "insuffisante" à l'établissement d'un diagnostic.

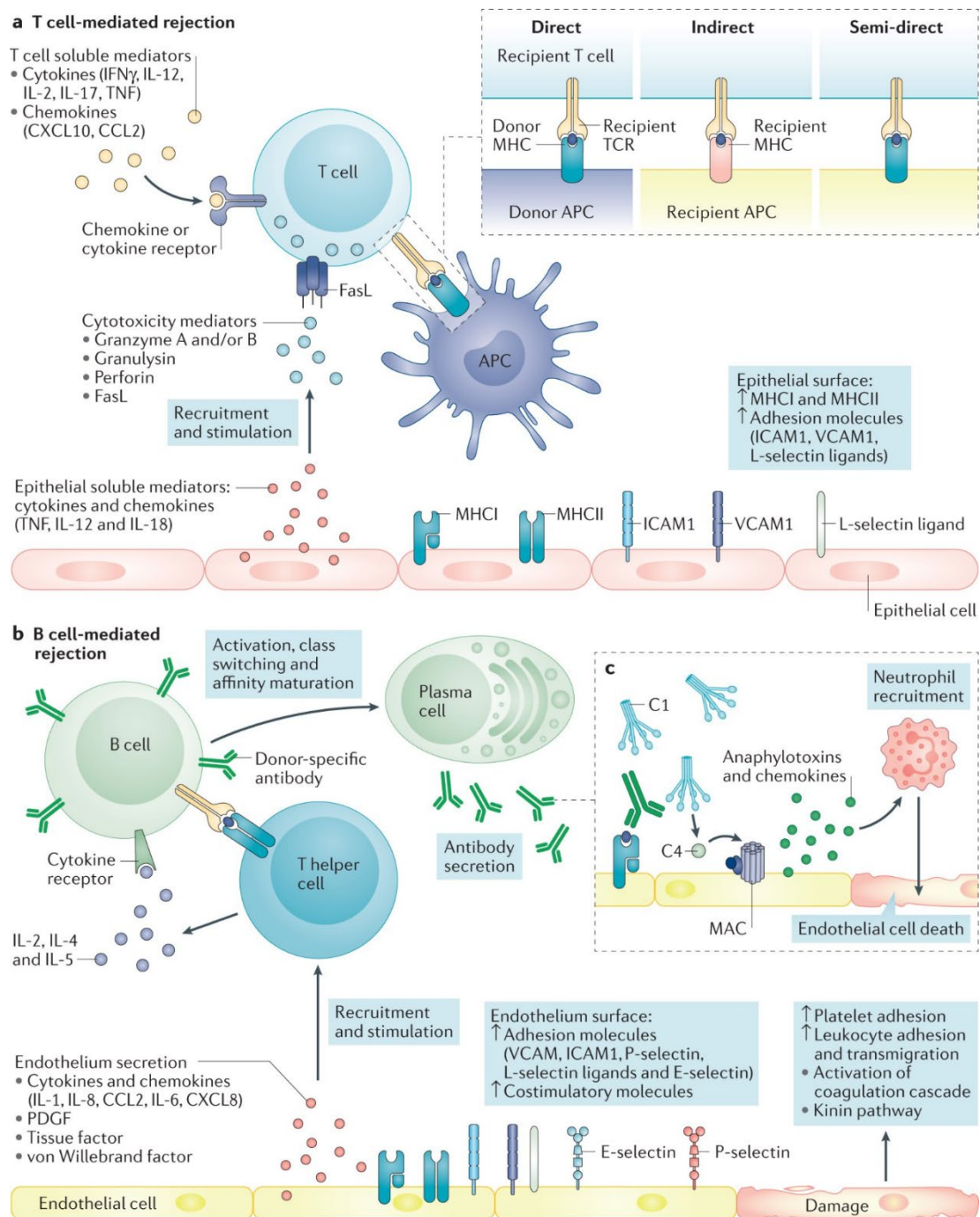
## TCMR

Les rejets cellulaires représentent la majorité des rejets aigus liés à une réponse immunitaire. Ils sont décrits en histologie par la présence d'une inflammation tubulo-interstitielle (tubulite) liée à l'infiltration interstitielle et des tubules rénaux par des lymphocytes T et monocytes, mais aussi par la présence de lésions vasculaires dans les artères (artérite). Dans les TCMR aiguës, les lymphocytes T CD4 et CD8 vont être activés par la reconnaissance des antigènes HLA du greffon, que ça soit de manière directe quand ils sont présentés par les APC du donneur ou indirecte quand ces antigènes sont récupérés et présentés par des APC du receveur (**Figure 9A**). Les APC activées migrent ensuite vers les organes lymphoïdes secondaires (OLS) où elles vont présenter l'antigène HLA reconnu aux lymphocytes T naïfs et mémoire pour les activer à leur tour. En parallèle, les cellules endothéliales activées par les cytokines pro-inflammatoires vont exprimer des molécules d'adhésion afin de faciliter l'infiltration des cellules immunitaires dans le tissu. Les cellules T CD8<sup>+</sup> effectrices sont majoritaires dans l'infiltrat du greffon et vont provoquer l'apoptose des cellules du greffon en les reconnaissant via leur CMH de classe I et en sécrétant des granules cytotoxiques contenant du granzyme B et des perforines pour les attaquer. Les TCMR sont régulièrement accompagnés d'une ABMR.

## ABMR

Le rejet humoral, ou ABMR, est la cause majeure des rejets chroniques après 1 an, mais peut aussi être identifié lors d'un rejet aigu. Il est caractérisé par une inflammation importante des capillaires péri-tubulaires et de la microcirculation, (A. Loupy et al. 2017). Lors du rejet médié par anticorps, les lymphocytes B jouent un rôle majeur puisqu'elles sont à l'origine de la production des anticorps spécifiques au donneur (DSA). Ces allo-antigènes sécrétés dans la circulation vont venir se fixer aux HLA portés par les cellules endothéliales du greffon, reconnues comme un corps étranger, et cette fixation va entraîner l'activation de la voie classique du complément (**Figure 9B**). La voie du complément va entraîner le clivage de la protéine C4 dont le fragment C4b va se fixer à la surface de la cellule endothéliale. Une protéase clive ensuite le C4b pour stopper son activité et ne laisse que le fragment C4d en surface. Le système du complément forme un système d'attaque membranaire qui va endommager les cellules endothéliales et recruter des cellules immunitaires. Les cellules endommagées vont sécréter le facteur de von Willebrand (FvW) qui attire et permet l'adhésion de plaquettes au lieu d'inflammation, entraînant une thrombose. C'est pourquoi dans de nombreux cas d'ABMR, on mesure l'accumulation diffuse de C4d dans les capillaires péri-tubulaires pour établir le diagnostic. Une partie des ABMR étant aussi C4d négative, il a été découvert que les DSA pouvaient provoquer un rejet

indépendant du complément par fixation directe sur les cellules reconnues appelé cytotoxicité à médiation cellulaire dépendante des anticorps (ADCC). Le fragment Fc des allo-anticorps va recruter les cellules ayant un récepteur pour le Fc des immunoglobulines comme les macrophages et les cellules NK. Les macrophages vont tuer la cellule par phagocytose tandis que les cellules NK vont sécréter des granzymes et des perforines pour induire l'apoptose de la cellule.



**Figure 9 : Mécanismes du rejet de greffe de rein.**

A) Mécanisme de TCMR : Les lymphocytes T sont activés par la reconnaissance du HLA du donneur et sécrètent des molécules cytotoxiques en réponse. B) Mécanismes d'ABMR : Les lymphocytes B sont activés par la reconnaissance du HLA du donneur porté par un T effecteur et maturent en plasmocytes producteur d'IgA en réponse. D'après (Oellerich et al. 2021)

L'espacement du suivi qui suit la première année post-transplantation est souvent un facteur aggravant du rejet chronique car il diminue la probabilité d'une détection et d'une prise en charge précoce du rejet.

Dans le rejet chronique, on retrouve la même dichotomie qu'en rejet aigu avec le TCMR, l'ABMR ou un rejet mixte lorsque les deux sont reconnus histologiquement. Dans le TCMR, les APC du donneur ont laissé place aux APC du receveur au cours du temps. On observe aussi un certain degré d'atteinte du cortex dans le parenchyme rénal, avec la présence de fibrose interstitielle et d'atrophie tubulaire (IFTA). On trouve aussi des cas de IFTA qui ne sont pas d'origine immunologique comme dans les cas vus dans la catégorie 6 (**Table 1**)

## Suivi et immunodépression chez les patients transplantés

Ces dernières années, les progrès en termes de gestion des traitements immunosuppresseurs promulgués ont permis de réduire les risques de rejets aigus et amélioré le pronostic sur le long terme des patients transplantés (M. A. Lim, Kohli, et Bloom 2017).

L'enjeu et la difficulté de l'immunothérapie est de remplir son objectif principal qui consiste à contrôler le système immunitaire pour que l'organisme accepte plus facilement le greffon, sans pour autant rendre le patient trop sensible aux pathogènes. En effet, le greffon étant vu comme un corps étranger, limiter l'inflammation et l'activation des cellules immunitaires qui en découlent est crucial à sa survie. L'usage de thérapies immunosuppressives suivant une greffe de rein est donc un équilibre fragile puisqu'une trop faible immunosuppression induirait un rejet par une réponse auto-immune, avec l'activation des lymphocytes T et B qui attaqueraient l'organe et une trop forte immunosuppression serait contre-productive en plus d'augmenter les risques secondaires. Plus précisément, l'immunosuppression promulguée n'est pas spécifique, son effet est généralisé à l'organisme entier, celle-ci peut donc entraîner une immunodéficience qui peut favoriser l'apparition de cellules cancéreuses, notamment avec une augmentation des risques de mélanomes de la peau (Adami et al. 2003; Garrett et al. 2017). Ces cancers peuvent émerger à cause d'une surveillance immunitaire affaiblie contre les cellules néoplasique à laquelle s'ajoute une vulnérabilité accrue aux infections virales et aux réactivations virales de virus dormants comme le virus d'Epstein-Barr (EBV) (Rama et Grinyó 2010) ou le polyomavirus BK (Hirsch et Snyderman 2005; Dall et Hariharan 2008). L'EBV est en effet connu pour être inactif dans la population générale mais en cas d'immunodéficience prolongée, il favorise une lymphoprolifération des cellules B non-contrôlée pouvant mener à des lymphomes non-Hodgkiniens (Liebowitz 1998). L'immunodéficience entraîne aussi la prévalence des infections nosocomiales du site chirurgical et des voies urinaires (Dantas et al. 2006; Dorschner, McElroy, et Ison 2014), et une augmentation des risques de diabète sucrés post-transplantatoires

(Shivaswamy, Boerner, et Larsen 2016) qui viennent augmenter la morbidité et la mortalité chez les patients transplantés (W. H. Lim et al. 2017).

## Les traitements immunosuppresseurs dans la transplantation rénale

De par la nature des rejets, les traitements existants ciblent majoritairement la réponse allo-immune médiée par les lymphocytes T. On peut regrouper les traitements immunosuppresseurs en 4 branches (Sitruk et al. 2018):

### Les corticostéroïdes

Les glucocorticoïdes (GCs) ont été dans les premiers traitements à être utilisés lors de la greffe de rein. Leur rôle est d'inhiber l'inflammation en passant notamment par un effet génomique. Les GCs entrent dans la cellule puis dans le noyau via les récepteurs aux glucocorticoïdes cytosoliques. D'ici, ils vont induire des gènes anti-inflammatoires mais aussi réprimer l'action de facteurs pro-inflammatoires comme le NF- $\kappa$ B. Les GCs vont impacter la plupart des cellules de l'immunité innée (Dashti-Khavidaki, Saidi, et Lu 2021): dans les lymphocytes T, ils répriment l'expression de cytokines pro-inflammatoires (IL-1, IL-2, IL-3, IL-6, TNF $\alpha$ , IFN $\gamma$ ), tandis que dans les monocytes et cellules dendritiques (DC) ce sont l'IL-1 $\beta$ , le TNF $\alpha$  et l'IL-12 qui sont réprimées. Ils réduisent aussi l'activité cytotoxique des cellules NK, bien que cela augmente leur production de cytokines pro-inflammatoires (Eddy et al. 2014). Parmi les GCs utilisés en traitement, on peut citer la Prédnisone et la Prédnisolone, des dérivés de synthèse du cortisol.

### Les inhibiteurs de la calcineurine (signal 1)

Les inhibiteurs de la calcineurine (CNI) comme la Tacrolimus ou la Ciclosporine-A, vont comme leur nom l'indique inhiber l'action de la calcineurine. La calcineurine entre en jeu lors du premier signal dans les lymphocytes T, quand le TCR reconnaît un antigène porté par une APC, la cascade d'activation augmente le taux de calcium intracellulaire, ce qui permet l'activation de la calcineurine. L'action de phosphatase de la calcineurine activée permet de déphosphoryler le facteur nucléaire des cellules T activées (NF-AT), qui se transloque ensuite dans le noyau pour induire l'expression d'IL-2 (qui promeut la différenciation des lymphocytes T). Le tacrolimus va se complexer avec la protéine FKBP-12 et inhiber la calcineurine, ce qui inhibe l'activation et la prolifération des lymphocytes T.

### Les bloqueurs du signal de costimulation (signal 2)

Les bloqueurs du signal de costimulation sont des médicaments plus récents, et leur rôle est de cibler les molécules de CD80 et CD86 des APCs afin d'empêcher les lymphocytes T de recevoir le signal de costimulation. Dans ce cas l'engagement du TCR sans signal de costimulation peut mener à une forme de tolérance, par l'anergie de la cellule T. Le bloqueur de ce signal le plus utilisé actuellement est le Belatacept, une protéine de fusion qui combine le domaine Fc d'une IgG1 avec la partie extracellulaire du CTLA-4 des lymphocytes T, qui semble garantir une réduction des DSA et une meilleure survie et fonction du greffon sur le long terme qu'avec un

traitement par CNI qui peut s'avérer néphrotoxique (Vincenti et al. 2016; Durrbach et al. 2010).

#### Les antiprolifératifs inhibiteurs de mTOR (signal 3)

La protéine mTOR (mammalian Target of rapamycin) est une protéine cytosolique impliquée dans la cascade du signal de croissance et de prolifération des cellules T provoqué par les cytokines sécrétées par les DCs ou l'IL-2. L'utilisation d'inhibiteurs de mTOR comme le Sirolimus ou l'Évérolimus va bloquer l'effet stimulateur de mTOR sur le cycle cellulaire, et empêcher la prolifération des lymphocytes T en bloquant le passage de la phase G1 à S.

#### Les antiprolifératifs inhibiteurs de métabolites (signal 3)

Dans cette dernière classe d'antiprolifératifs, on retrouve le mycophénolate mofétil (MMF), qui est une prodrogue de l'acide mycophénolique (MPA). Le MPA va dépléter la cellule en guanosine en bloquant l'enzyme responsable de sa voie de synthèse de novo, et il inhibe en conséquence à la fois le cycle cellulaire et la synthèse d'ADN.

Dans les cas de TCMR pur, c'est à dire sans ABMR sous-jacent, les traitements aux stéroïdes ou aux anticorps déplétant les lymphocytes T ont montré de très bons résultats pour un retour de la fonction rénale sans perte du greffon (Famulski et al. 2010)

La greffe de rein permet une amélioration de survie et de la qualité de vie des patients souffrant d'IRC et le développement des traitements immunosuppresseurs ainsi que l'amélioration des méthodes de sélection de greffons ont permis de réduire l'impact des rejets aigus au cours des années. Cependant les rejets chroniques ne restent encore que partiellement compris et difficiles à identifier en amont, et c'est sur ces prémices que mon travail de thèse sur les rejets de greffes de rein repose.

# Chapitre 3 : Les Maladies Inflammatoires Chroniques de l'Intestin

Le second projet de ma thèse a porté sur l'analyse d'une cohorte de patients atteints de MICI à partir de données de scRNA-seq pour mieux caractériser les mécanismes immunitaires impliqués et les causes de résistances aux traitements existants. J'aborderais donc dans cette partie l'état actuel des connaissances au sujet de la physiopathologie des MICI, l'organisation du système immunitaire et ses changements dans les MICI en lien avec la dysbiose intestinale, et enfin un résumé des stratégies de traitements et leurs limites.

## Les fonctions de l'intestin

### L'intestin, un rôle dans la digestion et dans la protection

L'intestin est un organe aux fonctions multiples et parfois paradoxales, puisqu'il est la porte d'entrée des nutriments et éléments indispensables au fonctionnement du corps humain, tout en devant mettre en place des fonctions de barrière aux microorganismes. L'intestin offre un lieu propice au développement de la flore commensale dans sa lumière mais reste cependant capable d'éliminer les agents pathogènes.

L'intestin grêle est subdivisé en 3 portions successives que sont le duodénum, le jéjunum et l'iléon, qui atteignent une longueur moyenne cumulée de 6 mètres, ce qui en fait l'organe le plus étendu du tractus gastro-intestinal. Le duodénum est le seul segment fixe de l'intestin et fait directement suite à l'estomac. Le duodénum permet de neutraliser l'acidité de la nourriture digérée par l'estomac via les glandes de Brunner, et il est aussi le lieu où sont sécrétés la bile et les enzymes digestives du pancréas qui participent à digestion. Ces enzymes permettent la dégradation chimique des protéines en acides aminés, de l'amidon en glucose et des lipides en acides gras, les rendant absorbables par la muqueuse intestinale. A la suite du duodénum, on retrouve le jéjunum qui forme, avec l'iléon, la portion flottante de l'intestin, recourbé en "anses" intestinales de par ses nombreux replis. Cette portion occupe une majorité de la cavité abdominale et sa longueur assure une optimisation du temps de digestion et d'absorption des éléments nutritifs. Concernant son organisation tissulaire, l'intestin grêle est constitué de 4 couches majeures : la muqueuse, la sous-muqueuse, la musculuse et la membrane séreuse, aussi appelée péritoine

### La muqueuse

La muqueuse assure un rôle de sécrétion, d'absorption et de protection. Afin d'optimiser sa fonction d'absorption, la muqueuse utilise une structure en replis pour augmenter considérablement sa surface en contact avec la lumière intestinale. La muqueuse se replie ensuite en des villosités intestinales en forme de doigts d'une dizaine de millimètres dans lesquels passent des capillaires sanguins et lymphatiques. Le dernier niveau de complexité structurelle de la muqueuse est situé à la face apicale



des entérocytes, dont la membrane cytoplasmique forme des microvillosités. Au total chez un adulte, la surface dépliée de la muqueuse représente près de 200m<sup>2</sup> au contact de la lumière intestinale.

Cette muqueuse est couverte de mucus, qui empêche les microorganismes d'être en contact direct avec les cellules épithéliales, et qui permet de ne laisser diffuser que les petites particules à absorber.

Le mucus couvre une monocouche de cellules épithéliales qui assurent un rôle de rempart entre l'environnement extérieur et l'organisme, et dont l'étanchéité est assurée par les jonctions serrées qui lient ces cellules. Les cellules épithéliales se renouvellent régulièrement grâce aux cryptes intestinales ou cryptes de Lieberkühn qui sont des invaginations dans lesquelles sont situées des cellules souches progénitrices, regroupées en zones germinatives. Les cellules souches intestinales se divisent et produisent des cellules à amplification transitoires (van der Flier et Clevers 2009; Clevers 2013). Ces cellules transitoires vont ensuite se différencier pour générer les différents types cellulaires de la barrière intestinale. Ils produiront entre autres les cellules sécrétrices, comme les cellules en gobelet ou caliciformes qui sécrètent le mucus, ou les cellules entéro-endocrines qui produisent différentes hormones. Il est à noter que les cellules caliciformes pourraient aussi jouer un rôle dans la réponse immunitaire et la tolérance orale en présentant aux cellules dendritiques de la lamina propria des antigènes provenant des aliments ou de la flore commensale (McDole et al. 2012). Néanmoins, la majorité de la couche épithéliale est constituée des entérocytes, des cellules sécrétrices d'enzymes digestives, et dont le rôle est d'absorber et faire transiter les nutriments de la lumière à la circulation générale. Les nouvelles générations successives de cellules vont pousser les cellules plus vieilles à migrer vers le haut avant de se différencier. Les cellules au sommet des villosités vont être éliminées dans la lumière par desquamation après une semaine environ. Au fond des cryptes intestinales, on retrouve aussi les cellules de Paneth, qui ont été décrites comme sécrétrices de défensines, des peptides antimicrobiens, et de lysozyme pour assurer une fonction bactéricide (Peeters et Vantrappen 1975; Selsted et al. 1992). Les défensines alpha et beta produites par les cellules de Paneth ne ciblent pas les bactéries commensales, uniquement certaines souches pathogènes. Additionnellement aux quatre types cellulaires principaux décrits précédemment, on retrouve aussi dans cette monocouche deux autres types cellulaires spécialisés : les cellules de Tuft, et les cellules à plis microscopiques ou cellules microfold. Les cellules Tuft sont très éparses dans l'épithélium et de description récente, pour leur rôle de sentinelle, notamment dans la réponse immunitaire contre les parasites helminthes (Gerbe et al. 2016; Hendel et al. 2022). Les cellules microfold sont quant à elles retrouvées quasi uniquement dans l'iléon et le côlon puisqu'elles colonisent la surface des plaques de Peyer. Ces plaques sont des organes lymphoïdes secondaires ancrés dans la muqueuse qui permettent l'induction de la réponse immunitaire dans l'intestin. Au-dessus de ces zones, la couche de mucus est plus légère et permet aux cellules microfold de transporter des antigènes et des bactéries aux DCs présentes dans ces plaques.

Sous la couche de cellules épithéliales, on retrouve du tissu conjonctif appelé lamina propria, dans laquelle les nombreux capillaires sanguins et lymphatiques vont

permettre l'absorption des nutriments digérés. La *lamina propria* est principalement constituée de fibroblastes et de nombreuses cellules immunitaires.

Enfin on retrouve à l'interface avec la sous-muqueuse, une multicouche de muscle lisse appelée *muscularis mucosae*, qui garde la muqueuse en mouvement et aide les cryptes à relarguer leur contenu

#### La sous-muqueuse

La sous-muqueuse est la couche du tissu dans laquelle circulent les vaisseaux sanguins et lymphatiques. On y trouve aussi des glandes qui produisent le mucus en supplément des cellules caliciformes, et dans le duodénum, la sous-muqueuse renferme aussi les glandes de Brünner dont la production de fluide alcalin contenant de la mucine permet de neutraliser l'acidité du chyme sortant de l'estomac (Sedano et al. 2015).

#### La musculature

La musculature est découpée en deux couches distinctes dans l'intestin. Les contractions phasiques coordonnées de ces deux couches sont contrôlées par le réseau nerveux du plexus myentérique, et vont assurer l'avancement du contenu de la lumière intestinale, c'est le péristaltisme.

#### La membrane séreuse

La membrane est composée de deux feuillets (un pariétal et un viscéral) et du mésentère qui relie les différentes anses intestinales. Elle permet aux différentes portions de l'intestin de glisser entre elles lors des mouvements péristaltiques, tandis que le mésentère renferme un réseau de veines et artères qui participent à l'irrigation des tissus intestinaux.

L'intestin est donc un organe dynamique optimisé pour la digestion de par ses capacités sécrétoires, ses parties mobiles et sa grande surface d'échange avec la lumière intestinale. Cette interface est aussi un lieu clef dans la défense immunitaire envers les entités qui composent le microbiote intestinal.

## Le microbiote intestinal et ses interactions

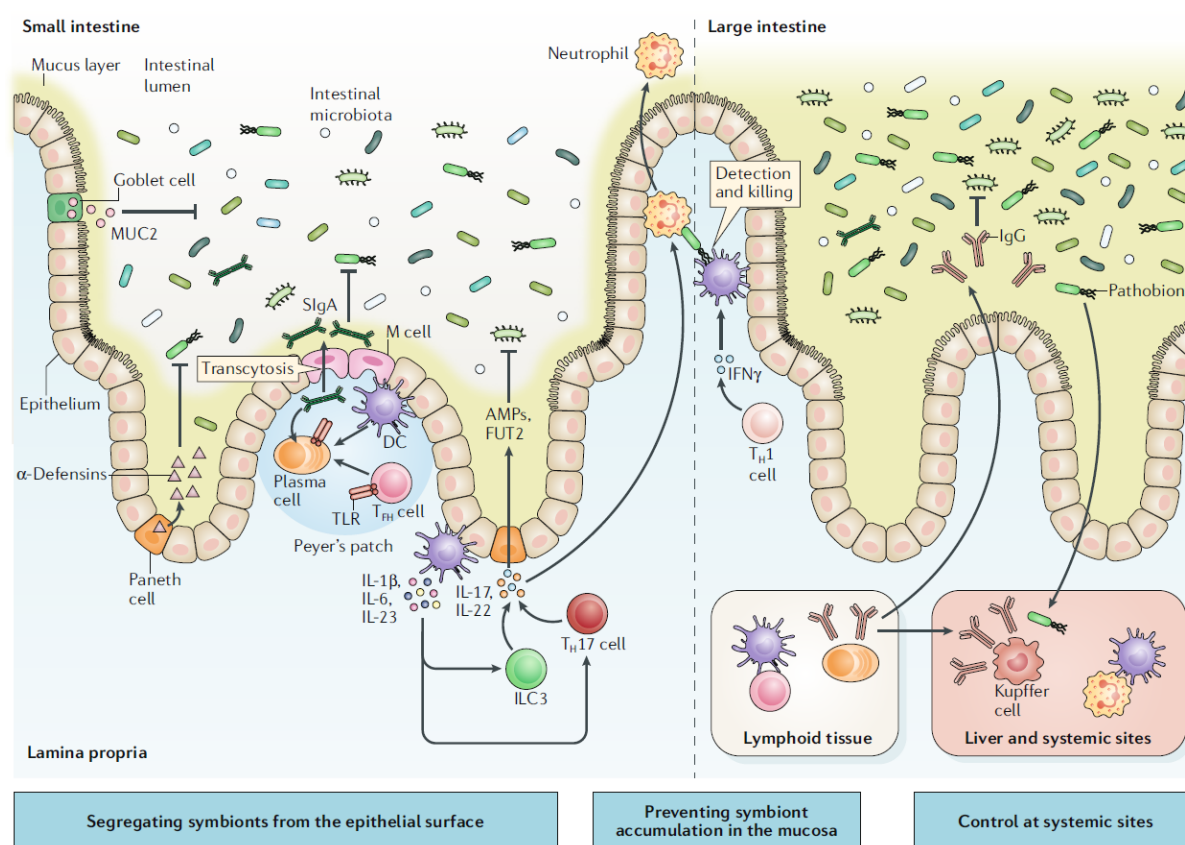
Le microbiote intestinal englobe l'ensemble des bactéries, virus, levures et autres micro-organismes se situant dans la lumière du tube digestif et dans la couche de mucus qui tapisse l'intestin grêle et le côlon. On dénombre plus de 100 billions ( $10^{13}$ ) de ces micro-organismes, dont la composition et le métabolisme influencent l'homéostasie du système digestif et par ricochet, l'organisme entier (Valdes et al. 2018). C'est durant l'enfance que se déroule la majeure partie de la colonisation des muqueuses intestinales par le microbiote, et sa composition est grandement influencée par les facteurs environnementaux comme le type de lait (maternel ou industriel), l'hygiène, l'utilisation d'antibiotiques (van den Elsen et al. 2019;

Dominguez-Bello et al. 2011). Au cours de la vie adulte, le microbiote reste plus stable, mais connaît quand même des évolutions en réponse aux facteurs environnementaux. Le microbiote intestinal établit une sorte de symbiose avec l'organisme de par ses différents rôles bénéfiques pour le maintien de l'homéostasie. Les microbiotes du cæcum et du côlon ont par exemple un rôle métabolique puisque certaines fibres alimentaires, des protéines et glucides complexes ne peuvent pas être digérés et assimilés avec les enzymes du tractus gastro-intestinal supérieur. Certaines bactéries prennent le relai pour les digérer et leur activité de fermentation va notamment produire des acides gras à chaînes courtes (AGCC), des molécules composées d'une chaîne de 6 atomes de carbone maximum (Koh et al. 2016). Ces AGCC vont agir en tant que source d'énergie des colonocytes avec le butyrate, et permettre la croissance d'autres bactéries avec l'acétate.

La construction du microbiote intestinal participe aussi à façonner l'immunité protectrice de la barrière intestinale par le développement et la maturation du système immunitaire. Dans des modèles de souris élevées sans germes, il a été montré qu'en l'absence d'une colonisation microbienne de la paroi intestinale, les follicules lymphoïdes isolés ne peuvent pas se développer (C. E. West, Jenmalm, et Prescott 2015), montrant ainsi l'importance du microbiote dans la promotion de la barrière muqueuse. Le rôle de protection propre au microbiote passe aussi par une compétition physique, le peuplement des muqueuses par des bactéries commensales protège la paroi intestinale en empêchant la colonisation de pathogènes via la compétition pour les nutriments disponibles (Kamada et al. 2013). Le microbiote agit aussi en lien avec les cellules immunitaires elles-même, puisque le système immunitaire muqueux intestinal s'est développé pour faire face aux complications possibles. En effet, les réponses immunitaires adaptatives ont lieu en permanence en réponse aux micro-organismes commensaux avec le recrutement et la différenciation de plasmocytes sécrétant des IgA, des lymphocytes helper Th17 ou des lymphocytes T régulateurs (Treg). Et certaines bactéries commensales peuvent influencer sur ces générations comme des *E. Coli* invasives et adhérentes (AIEC) ou des *Bifidobacteria adolescentis* peuvent induire la différenciation et l'accumulation de lymphocytes en T helper 17 (Th17) pro-inflammatoires dans les tissus (Tan et al. 2016; Viladomiu et al. 2017), tandis que d'autres comme celles de la classe des *Clostridia* vont accroître le recrutement des Treg dans les tissus (Nagano, Itoh, et Honda 2012) qui contrôlent ces réponses pour éviter que leur potentiel pro-inflammatoire ne s'exacerbe et devienne pathologique. Les bactéries pathogènes du microbiote sont aussi appelées pathobiontes et leur présence s'accompagne des motifs moléculaires associés aux pathogènes (PAMPs, pour *Pathogen-associated molecular patterns*) qui sont reconnus par le système immunitaire inné via des récepteurs de reconnaissance de motifs moléculaires (PRR, pour *Pattern recognition receptor*). En retour, le système immunitaire va contrôler le microbiote pour maintenir l'homéostasie en sécrétant plus de mucus ou des molécules pour sa protection comme les défensines ou les immunoglobulines A (IgA) (de Souza et Fiocchi 2016). Les IgA sont sécrétées dans la lumière et s'accumulent dans le mucus où elles vont neutraliser les micro-organismes pour les empêcher de passer la barrière tout en leur laissant

## Organisation du système immunitaire intestinal

Les muqueuses intestinales sont le lieu d'une régulation continue de l'immunité à l'homéostasie. Cette régulation se révèle complexe due à la proximité entre le système immunitaire et le microbiote intestinal. Les réponses immunitaires doivent être mesurées pour éviter d'impacter les nombreux micro-organismes commensaux et bénéfiques à la physiologie intestinale. Le système immunitaire intestinal regroupe des sites inducteurs de réponses immunitaires adaptatives : ganglions lymphatiques mésentériques et GALT (plaques de Peyer, les follicules lymphoïdes isolés, appendice...); et des sites effecteurs de ces réponses : lamina propria et épithélium (Figure 10).



**Figure 10 : Organisation du système immunitaire.**

Les défenses se partagent en trois strates, avec dans la lumière intestinale en surface de la muqueuse la couche de mucus contenant les peptides anti-microbiens (AMPs) sécrétés par la seconde strate, la barrière épithéliale. Cette monocouche de cellules comprend des cellules spécialisées comme les cellules caliciformes (goblet cells) productrices de mucus, les cellules microfold (M cells) qui facilitent la récupération d'antigènes et les cellules de Paneth au fond des cryptes qui produisent une diversité d'AMPs. Figure d'après (Caruso, Lo, et Núñez 2020)

### L'immunité innée

La première ligne de défense est constituée par les cellules de l'immunité innée. Les neutrophiles sont les premiers leucocytes sanguins à infiltrer le tissu et migrer vers la zone lésionnelle. Ils jouent un rôle d'élimination des pathogènes extracellulaires par

phagocytose ou par dégranulation à la suite de la reconnaissance de micro-organismes. Suite à leur activation, les neutrophiles peuvent aussi produire des espèces réactives de l'oxygène (ROS) qui aident à éliminer les pathogènes présents (Zhou et Liu 2017). Une fois leur rôle effectué au site de l'inflammation, les neutrophiles meurent par apoptose et sont éliminés par les macrophages résidents selon le processus d'efférocytose, ce qui induit leur production de signaux anti-inflammatoires avec la sécrétion d'IL-10, et de résolvines (Greenlee-Wacker 2016). Cependant, dans les MICI et de manière plus prononcée dans la RCH, l'activation continue des neutrophiles et leur recrutement excessif les rends délétères. Leur production de ROS est augmentée et participe à la dégradation de la barrière intestinale, leur activité apoptotique diminue et la clairance des microorganismes n'est plus complète (Wéra, Lancellotti, et Oury 2016).

Avec les cellules dendritiques, les macrophages correspondent aux phagocytes mononucléés (MNP pour Mononuclear phagocytes) et assurent un rôle de présentation d'antigènes dans l'intestin afin de maintenir la tolérance envers le microbiote ou permettre l'élimination de pathobiontes. En conditions inflammatoires, l'élimination de pathogènes passe aussi par l'infiltration de monocytes circulants dans les tissus qui se différencient en macrophages pro-inflammatoires sécrétant de TNF $\alpha$  et d'IL-1 $\beta$  (Coillard et Segura 2019). Pour remplir le rôle de présentation de l'antigène, les MNP détectent les motifs moléculaires associés aux microbes (MAMPs) grâce à leur pathogen-recognizing receptors (PRR). Les PRR impliqués sont notamment les récepteurs Toll-like (TLR) qui se divisent en TLR membranaires de surface et TLR endosomaux, ainsi que les récepteurs NOD-like (NLR) cytoplasmiques (Kawai et Akira 2009). Le récepteur NOD2 est un NLR localisé dans le cytoplasme qui reconnaît le muramyl dipeptide (MDP), le constituant principal des peptidoglycanes dont sont constituées les parois de bactéries gram positives. L'activation de NOD2 dans les MNP induit le recrutement de la protéine ATG16L1 au site d'entrée de la bactérie qui induit l'autophagie de la bactérie (Travassos et al. 2010). La mutation de *ATG16L1* qui est un des variants associés à la MC est non seulement responsable d'un défaut d'élimination des composants bactériens causant l'inflammation, mais aussi de l'augmentation de la production de cytokines pro-inflammatoires TNF $\alpha$  via la voie IL-22 dans les cellules endothéliales qui participent aussi à la phagocytose des pathogènes par autophagie à l'homéostasie (Okai et al. 2022). L'autophagie est un des outils des MNP dans la défense contre les bactéries pathogènes et ce mécanisme induit aussi des changements au niveau des cytokines puisqu'il inhibe la sécrétion d'IL-1 $\beta$  et d'IL-18, deux cytokines liées à l'inflammasome (Kabat, Pott, et Maloy 2016). Dans le cas d'une déficience d'ATG16L1 due à un variant, il a été observé une suractivation de l'inflammasome, de la caspase-1 qui y est associé, et par conséquent d'une sécrétion augmentée d'IL-1 $\beta$  et d'IL-18 par les MNP (Fritz et al. 2011), ce qui semble confirmer une régulation de l'inflammasome par les mécanismes d'autophagie. De plus, l'augmentation d'IL-1 $\beta$  dans les MICI induit l'augmentation de la sécrétion d'IL-23, une autre cytokine pro-inflammatoire importante (Peral de Castro et al. 2012). Les cellules dendritiques (DC) exercent en outre une fonction présentatrice des antigènes en condition homéostatique dans les plaques de Peyer

ou en migrant dans les ganglions lymphatiques mésentériques. Elles interagissent avec les lymphocytes T et B pour induire leur différenciation en Th17, Treg et plasmocytes producteurs d'IgA et promouvoir la tolérance et le maintien de l'homéostasie, ce qui en fait des acteurs à la frontière entre l'immunité innée et adaptative. En cas de reconnaissance de pathobiontes par leur PRR, les DC vont sécréter d'autres cytokines pour induire la différenciation des lymphocytes T CD4+ en cellules T effectrices aussi appelées T helpers (Th) comme les Th1 ou dans le cas d'une reconnaissance par les PRR d'AIEC, induire la différenciation en Th17 (Viladomiu et al. 2017).

Les lymphocytes de l'immunité innée (ILC pour *Innate Lymphoid Cell*) sont des cellules ayant une plasticité fonctionnelle similaire aux cellules T effectrices (de Souza et Fiocchi 2016) qui ont divisées en 3 catégories (**Figure 11**). Ces cellules participent à l'homéostasie en régulant les défenses contre les microbes et la réparation des tissus (Tait Wojno et Artis 2012).

- Les ILC1 sont équivalentes aux Th1, sécrètent de l'IFN $\gamma$  et participent aux défenses antimicrobiennes
- Les ILC2 sont équivalentes aux Th2, sécrètent de l'IL-4, IL-5, IL-9 et IL-13 pour participer à la défense antiparasitaire et les allergies.
- Les ILC3 sont les ILC les plus nombreuses dans l'intestin. Elles sont équivalentes aux Th17, sécrètent de l'IL-22, IL-17A et l'IL17F, des cytokines impliquées dans la réparation des tissus et dans l'immunité de la barrière, mais aussi des cytokines inflammatoires comme le Facteur de stimulation des colonies de granulocytes et de macrophages (GM-CSF) qui permet de recruter des monocytes vers le site inflammatoire (Friedrich, Pohin, et Powrie 2019).

Dans l'intestin inflammatoire de patients de MICI, on observe par exemple une activation des ILC3 qui semble conduire à une transdifférenciation en ILC1 productrices d'IFN $\gamma$  en présence des cytokines pro-inflammatoires IL-1 $\beta$ , IL-15 et IL-12 (Vivier et al. 2018). Un retour possible en ILC3 en présence d'IL-23, d'IL-2 et d'acide rétinoïque (Bernink et al. 2015).

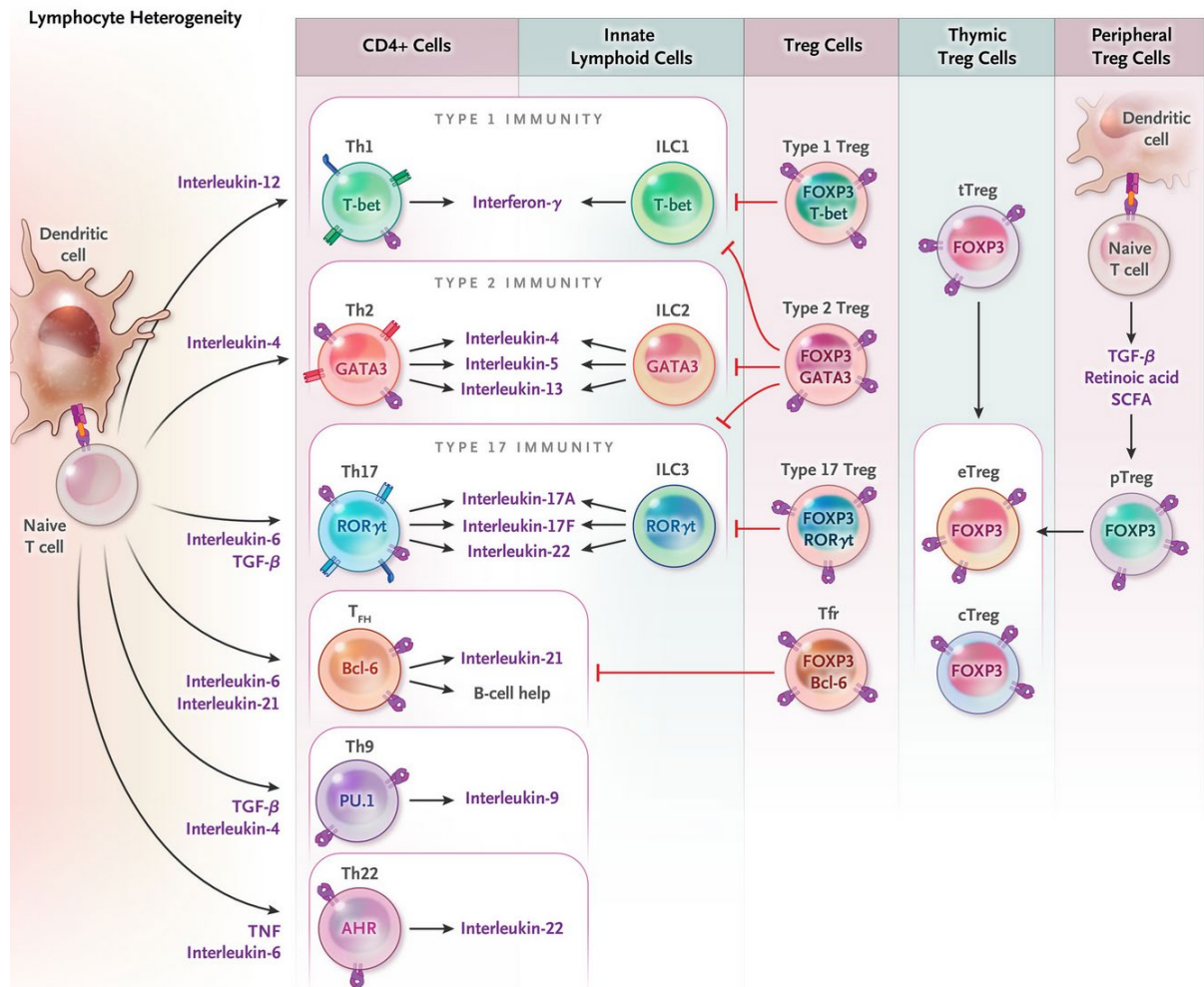
Parmi les cellules qui participent à l'immunité innée, on distingue aussi les cellules non-immunitaires. Les cellules épithéliales intestinales (IEC) forment la barrière physique de cellules étroitement jointes entre la lumière intestinale et la lamina propria, et cette monocouche est constituée de cellules spécialisées. Les cellules caliciformes vont notamment produire la mucine 2 (MUC2) qui est une protéine qui polymérise avec d'autres protéines liées au mucus pour former le gel de mucus qui recouvre la barrière intestinale. Les IEC sécrètent divers peptides anti-microbiens (AMPs) dans le mucus pour empêcher les divers micro-organismes de venir au contact de la membrane. Au fond des cryptes intestinales, ce sont les cellules de Paneth qui sont les cellules sécrétrices majeures, la stimulation de leurs TLR induit par exemple la production de RegIII $\gamma$ , une AMP ciblant les bactéries à Gram positif (Vaishnava et al. 2011), et par la stimulation de leur récepteurs NOD2, elles sécrètent des  $\alpha$ -défensine.

## L'immunité adaptative

Une particularité de l'endothélium intestinal est que les cellules endothéliales qui le composent expriment spécifiquement une molécule pour recruter les lymphocytes dans les muqueuses : la molécule d'adhérence cellulaire d'adressine de muqueuses (MAdCAM-1) (Berlin et al. 1993; Briskin et al. 1997; Gorfu, Rivera-Nieves, et Ley 2009). Les lymphocytes T et B activés par les DC dans les GALT (Plaques de Peyer, ganglions mésentériques, ...) vont exprimer l'intégrine  $\alpha 4\beta 7$  en surface, qui permet aux cellules de se fixer à MAdCAM-1 dans les capillaires intestinaux pour ralentir leur circulation et leur permettre de migrer vers les tissus. Les lymphocytes activés vont aussi exprimer le récepteur de chimiokine CCR9 qui contrôle leur migration et permet un homing spécialisé vers l'intestin. Dans le cas des lymphocytes B activés, CCR10 est exprimé dans le cas d'un homing vers le colon (Habtezion et al. 2016).

Les lymphocytes B activés dans les GALT vont se différencier en plasmocytes et revenir dans le sang par homing vers les tissus à l'homéostasie. Ils sont responsables de la sécrétion d'IgA spécifiques de l'antigène. Ces IgA se dimérisent et passent ensuite dans la lumière intestinale en surface de la muqueuse où elles reconnaissent les antigènes pour empêcher le passage dans les tissus, et piègent les micro-organismes dans le mucus.

Les lymphocytes T naïfs activés par les DC après la présentation des antigènes sont majoritairement des CD4<sup>+</sup> et peuvent se différencier en cellule auxiliaire (Th), régulatrice (Treg) ou effecteur mémoire (Tmem). La différenciation est dépendante des cytokines présentes et de la condition (**Figure 11**). Les taux élevés d'IL-12 et d'IFN $\gamma$  induisent le facteur de transcription T-bet, qui conduit à une différenciation en Th1. Les Th1 participent à la réponse aux infections bactériennes intracellulaires, recrutent les LT CD8<sup>+</sup>, les NK et les macrophages au site inflammatoire et produisent de l'IFN $\gamma$ . Les taux élevés d'IL-6 et TGF $\beta$  induisent quant à eux le facteur de transcription ROR $\gamma$ t qui permet la différenciation en Th17. Les Th17 permettent le recrutement de neutrophiles et sécrètent diverses interleukines : IL-17A, IL-17F, IL-21 et IL-22 pour combattre les infections bactériennes extracellulaires (Chang 2020). Les Th2 sont quant à elles spécifiques aux réponses contre l'infection par des helminthes. Les Tregs participent à la tolérance et maintiennent l'homéostasie en inhibant l'immunopathologie induite par les Th1 et Th17. Dans les MICI, les conditions inflammatoires chroniques et la dégradation de la barrière intestinale font pencher la balance des T effecteurs vers une différenciation en Th1 et Th17 et une diminution de Treg.



**Figure 11. Hétérogénéité des Lymphocytes T dans les tissus intestinaux.** Résumé des cytokines nécessaire à la différenciation des LT naïfs, des facteurs de transcriptions associés à chaque sous-type et les cytokines produites par ces cellules. D'après (Chang 2020)

## Barrière épithéliale et signature IL-17/IL-22

L'IL-22 est une cytokine de la famille des IL-10 (IL-10, IL-19, IL-20, IL-24, and IL-26) qui a été découverte récemment en 2000 (Dumoutier, Louahed, et Renaud 2000). Cette protéine a premièrement été décrite comme étant produite par les lymphocytes T et cellules NK activées (Wolk et al. 2002), puis elle a été observée dans les lymphocytes Th17, où elle est coexprimée avec IL-17A et IL-17F (Liang et al. 2006). Elle a aussi été décrite comme étant produite par les cellules lymphoïdes innées de groupe 3 (ILC3) (Geremia et al. 2011), ce qui lui permet d'activer l'immunité innée dans la barrière intestinale, via les defensin 2 notamment, sans impliquer le système immunitaire adaptatif. Sa production montre un rôle de protection des muqueuses, en induisant l'expression de peptides anti-microbiens et en stimulant la sécrétion de l'IL-10 par les cellules épithéliales via la voie STAT3 (Cella et al. 2009; Nagalakshmi et al. 2004). La production de l'IL-22 est donc majoritairement assurée par les cellules immunitaires, tandis que l'expression de son récepteur, l'IL-22R est restreinte à des cellules non-immunitaires, comme les kératinocytes, les hépatocytes, ou dans notre cas les cellules épithéliales intestinales (Wolk et al. 2004). Cela en fait un messager



intéressant dans la régulation de l'homéostasie de la barrière épithéliale. L'IL-22 semble jouer un double rôle, d'un côté elle est anti-inflammatoire par son rôle anti-microbien et de réparation dans les tissus de la barrière, mais elle jouerait aussi un rôle pro-inflammatoire en cas de surproduction durant les poussées inflammatoires des MICI (Mizoguchi et al. 2018; Yan et al. 2021). L'IL-22 a aussi un antagoniste soluble, l'IL-22BP qui est constitué des domaines extracellulaires du récepteur à l'IL-22 (Dumoutier et al. 2001) et dont la sécrétion accrue en cas de MICI bloque les actions protectives de l'IL-22 (J. C. Martin et al. 2016).

Dernièrement, une nouvelle théorie a été apportée concernant le rôle de l'IL-22 dans la résistance à l'ustekinumab (un bloqueur de l'IL-12/IL-23p40) : l'augmentation de l'IL-22 dans la RCH semble augmenter la production de chimiokines attirant les neutrophiles dans la muqueuse (CXCL1,2,3,5,6 et 8) et cette infiltration de neutrophiles semble corrélérer avec la résistance à l'ustekinumab (Pavlidis et al. 2022)

## Maladie de Crohn et Rectocolite hémorragique : causes et traitements actuels

### Les maladies inflammatoires chroniques de l'intestin

Les maladies inflammatoires chroniques de l'intestin (MICI) comprennent deux pathologies proches mais à l'évolution et à la localisation différentes. La première est la maladie de Crohn (MC), qui peut être localisée tout au long du tractus intestinal mais se déclenche majoritairement au niveau de l'iléon terminal. La seconde est la rectocolite hémorragique (RCH), qui ne touche que le rectum et le côlon. La RCH apparaît principalement chez les patients entre 30 et 40 ans, et sa prévalence en Europe est estimée autour de 24,3 pour 100000 personnes (Ungaro et al. 2017) Ces deux pathologies se caractérisent par des inflammations transitoires du tube digestif, des poussées inflammatoires entrecoupées de phases de rémission (Torres et al. 2017). Les rechutes accumulées vont causer des dégâts irréversibles qui peuvent conduire à une hospitalisation pour près de 40% des patients voire même une chirurgie pour près de 20% des patients (Burisch et al. 2019). Dans la RCH, les atteintes sont localisées dans le rectum et se propagent seulement dans le côlon à proximité, avec une inflammation très visible et marquée des tissus et des lésions souvent superficielles. Les lésions de la MC sont plus discontinues, avec de nombreux ulcères le long de l'intestin.

### Diagnostic clinique et suivi des patients

L'hyperactivation du système immunitaire qui mène à cette inflammation va entraîner différentes complications et ulcérations de la paroi intestinale qui permettent de classer cliniquement les patients en fonction de la gravité des atteintes comme avec la classification de Montréal (Silverberg et al. 2005; Satsangi et al. 2006), et d'adapter le traitement en conséquence. Dans la maladie de Crohn, la classification comprend trois paramètres : l'âge (A), la localisation (L) et le comportement (B pour "Behaviour"). La localisation se classe en quatre catégories : L1 pour les atteintes iléales, L2 les

atteintes coliques, L3 les atteintes iléo-coliques et L4 les atteintes digestives supérieures isolées. Les complications dans les MICI peuvent mener à la formation de fistules, qui sont des connexions avec un autre organe que le tube digestif, ou à une sténose, un rétrécissement du tube digestif causé par la présence de tissus cicatriciel sur les parois de l'intestin. La classification du comportement prend donc en compte ces complications : B1 pour les atteintes non sténosantes ni pénétrantes, B2 les atteintes sténosantes, B3 les atteintes pénétrantes. Dans la RCH, les symptômes décrits couvrent aussi des diarrhées régulières et la présence de sang dans les selles. Dans la classification de Montréal, la RCH est classifiée selon son étendue : E1 ou rectite si elle est limitée au rectum (chez 30 à 60 % des patients), E2 ou rectocolite gauche si elle touche toute la partie gauche du côlon (chez 16 à 45 % des patients) et E3 ou pancolite si elle est étendue à tout le côlon (chez 15 à 35% des patients) (Ungaro et al. 2017).

Pour établir le premier diagnostic chez les patients, on utilise une combinaison des observations cliniques, endoscopiques, histologiques, biologiques et radiologiques. Les lésions sont caractérisées par endoscopie et permettent de différencier correctement la MC de la RCH dans plus de 90% des cas, depuis des décennies (Pera et al. 1987). Au niveau histologique, la RCH montre des malformations des cryptes, une infiltration accrue de cellules immunitaires et une diminution de mucine. Au niveau biologique, le diagnostic d'une maladie inflammatoire de l'intestin repose aussi sur la mesure non invasive de la calprotectine fécale. La calprotectine est un hétérodimère formé des protéines S100A8 et S100A9 que l'on retrouve dans le cytoplasme des macrophages et des polynucléaires neutrophiles (Fukunaga et al. 2018). En cas de lésions de la muqueuse intestinale, l'infiltration leucocytaire va faire augmenter la quantité de calprotectine dans les fèces, et ce de manière proportionnelle à la gravité des lésions, ce qui en fait un marqueur utile et sensible au diagnostic des MICI (Torres et al. 2017). La concentration habituelle de la calprotectine est d'environ 20-30 µg/g, et une valeur au-delà de 50 µg/g indique une inflammation de la muqueuse.

## Complications dans les MICI

Les MICI sont caractérisées par des poussées inflammatoires par intermittence, entrecoupées de périodes de rémission plus ou moins longues. On définit la rémission à différents niveaux avec les mêmes outils que pour le diagnostic : au niveau clinique le patient n'a presque plus de symptômes, au niveau biologique on observe un retour à la normale des niveaux de calprotectine fécale, et au niveau endoscopique on n'observe plus d'inflammation des parois. Cependant, malgré une rémission apparente en clinique, il peut toujours subsister une inflammation infraclinique de la muqueuse, comme chez les patients atteints de RCH où dans une cohorte de patients en rémission clinique, on a pu observer une inflammation par endoscopie chez 45% d'entre eux (L. Rosenberg et al. 2013). Cette inflammation chronique va causer des complications localisées qui endommagent les parois de manière irréversible et progressive. Les conséquences cliniques sont d'abord digestives, avec des diarrhées, des nausées accompagnées de douleurs abdominales et parfois des rectorragies, et des symptômes plus généraux comme une fatigue chronique ou de la fièvre. Au niveau histologique, la complication la plus fréquente est la survenue de fibrose. La

fibrose est une conséquence de l'inflammation continue qui empêche une réparation efficace des lésions de la muqueuse. La complication la plus courante est la sténose intestinale, une diminution du diamètre du tube intestinal liée à la fibrose des tissus qui impacte le bon passage des aliments et touche plus de la moitié des patients 5 ans après le diagnostic de la MC (Mager et al. 2021). Cette fibrose est liée à un déséquilibre dans la production et la dégradation de matrice extracellulaire (MEC), notamment par la sur-expression de matrix métalloprotéinases (MMP), qui dégradent la MEC et ainsi participent à la régulation des cytokines (O'Sullivan, Gilmer, et Medina 2015). Cette sténose est la cause principale des recours à la chirurgie dans la maladie de Crohn. En effet, les progrès récents dans la prise en charge thérapeutique des patients n'empêchent cependant pas d'avoir recours à des chirurgies localisées dans les cas les plus sévères (Hwang et Varma 2008). Les résections de la seule partie endommagée sont préférées afin de garder le maximum de surface intestinale après opération pour tenter de préserver au maximum les fonctions d'absorption des aliments du tube digestif.

Au-delà de l'inflammation localisée, des atteintes extra-intestinales peuvent aussi toucher les patients avec des inflammations du système musculo-squelettique, de la peau, du foie ou des yeux, chez 25 à 40% des patients. On retrouve parmi les manifestations les plus fréquentes l'arthrite (inflammation des articulations) qui touche un tiers des patients MC et 20% des patients RCH, des ulcérations buccales chez 10% des patients ou encore des uvéites (inflammation des yeux) chez 5% d'entre eux (Vavricka et al. 2011; Levine et Burakoff 2011). Il existe aussi des patients asymptomatiques chez qui une MICI peut être identifiée, mais dont l'évolution de la maladie est plus dure à suivre puisque le patient n'est pas inquiété.

## L'origine imprécise des MICI et leurs prédispositions génétiques

Les origines de l'apparition de MICI ne restent à l'heure actuelle encore que partiellement identifiées, mais le consensus semble se porter sur un rôle des facteurs environnementaux dans le déclenchement de réponses inflammatoires dérégulées contre le microbiote intestinal chez des individus prédisposés génétiquement.

L'impact de l'environnement sur l'apparition de ces pathologies se traduit par un taux d'incidence des maladies intestinales fortement augmenté dans les pays les plus développés durant les dernières décennies (Molodecky et al. 2012; Alatab et al. 2020). En Europe de l'ouest, la prévalence moyenne des MICI est estimée à 137 pour 100000 personnes contre environ 85 pour 100000 personnes dans le monde (M. Zhao et al. 2021). Cet accroissement peut être dû à une diversité de facteurs environnementaux parmi lesquels les études épidémiologiques ont identifié la pollution de l'air, la cigarette, les antibiotiques ou le régime alimentaire (Frolkis et al. 2013; de Souza et Fiocchi 2016; Piovani et al. 2019). Les facteurs environnementaux semblent être à de nombreux égards déclencheurs, ou au minimum aggravants de la pathologie mais on retrouve aussi la participation de la composante génétique dans l'émergence de ces pathologies. Dans la maladie de Crohn, plusieurs études ont permis de montrer un lien entre l'apparition de la maladie et le chromosome 16 (Cavanaugh et IBD International Genetics Consortium 2001)

Des prédispositions aux MICI ont aussi été découvertes à l'aide d'études d'association pangénomique (GWAS), une technique d'analyse des variations génétiques retrouvées dans une population large d'individus, afin d'étudier leurs corrélations avec une pathologie ou des traits phénotypique. Une GWAS sert à chercher les mutations d'un nucléotide donné, un SNP, comparé au reste d'une population et estimer si certaines de ces variations sont plus fréquentes dans une population donnée. Ces SNPs peuvent être localisés dans un même locus, qu'il soit situé dans la séquence d'un gène unique et connu ou dans une région non annotée. Il n'est donc pas toujours aisé de déterminer l'impact de la mutation sur la fonction cellulaire, et dans le cas d'une GWAS appliquée à une pathologie, de trouver la raison de l'association entre le SNP et la pathologie. Une GWAS appliquée à plus de 75000 cas de MICI et de contrôles a permis d'identifier 163 loci, dont 110 sont communs à la MC et la RCH, indiquant que les deux pathologies partagent de nombreux mécanismes biologiques (Jostins et al. 2012). Dans ces variants identifiés, l'analyse fonctionnelle a révélé un enrichissement de la voie de signalisation JAK-STAT ainsi qu'une régulation accrue de la production de cytokines, notamment de l'IFN $\gamma$ , IL-12, TNF $\alpha$ , et l'IL-10, qui participent à l'activation des lymphocytes T et B, et des cellules NK. Depuis, les études d'association génétiques n'ont cessé de s'accumuler (Liu et al. 2015; de Lange et al. 2017; H. Huang et al. 2017) en incluant plus de patients afin de définir des variants associés à la maladie qui seraient présents à plus faible fréquence. On dénombre désormais plus de 250 loci associés aux MICI. Une de ces études a aussi permis de lier ces variants à une distribution en 3 catégories plutôt qu'une dichotomie entre MC et RCH, en définissant une atteinte de la MC iléale, une atteinte de la MC colique et la RCH qui auraient des origines génétiques différentes de par leur localisation (Cleyne et al. 2016).

Les SNPs les plus importants ont été retrouvés dans la région du gène *NOD2*, un récepteur intracellulaire qui induit la réponse inflammatoire lors de l'entrée de bactéries dans les tissus via la reconnaissance de leurs peptidoglycanes. Il participe ainsi à la régulation de la flore intestinale via la réponse antimicrobienne et son expression dépend de la flore commensale des parois de l'intestin, il a un rôle dans le maintien de l'homéostasie de la barrière intestinale. Des mutations de *NOD2* peuvent induire une suractivation de la réponse inflammatoire et une diminution de la fonction protectrice de la barrière intestinale (Hugot et al. 2001; Cho et Abraham 2007, 2) comme par exemple en altérant l'autophagie par les MNPs, ou en impactant la production d' $\alpha$ -défensines par les cellules de Paneth (Wehkamp et al. 2004). Une partie importante des gènes de susceptibilité sont impliqués dans la régulation des réponses hôte-microbiote, ce qui confirme le lien très étroit entre la dysbiose et le fonctionnement du système immunitaire intestinal et a permis de mieux aiguiller les voies fonctionnelles à cibler en thérapeutique ou de mieux comprendre le fonctionnement des traitements existants (Graham et Xavier 2020)

Les prédispositions génétiques peuvent aussi être populationnelles puisqu'il a été montré dans la maladie de Crohn que la population de juifs Ashkénaze comparée à d'autres individus européens non juifs (Kenny et al. 2012) faisait ressortir de nouveaux variants génétiques significativement associés à l'apparition de la maladie de Crohn

qui n'avaient pas été retrouvées précédemment dans les populations européennes non-juives, en plus de valider les variants déjà décrits précédemment tel que *NOD2*. Dans les MICI, il semble aussi y avoir une influence familiale dans la prévalence de la maladie. Dans les cas de RCH, c'est près de 12% des patients qui en étaient atteints qui avaient des antécédents familiaux de MICI, notamment de RCH à hauteur de 9% (Childers et al. 2014). Dans la maladie de Crohn, l'observation est la même, sur 553 patients, 17 % d'entre eux avaient des antécédents familiaux de MICI (Polito et al. 1996). Une étude française a montré que les MICI apparaissent à un âge plus précoce notamment quand plusieurs membres de la famille étaient affectés (Colombel et al. 1996)

## L'impact des facteurs environnementaux et la dysbiose intestinale

La dysbiose joue un rôle majeur dans la virulence des MICI (Tremaroli et Bäckhed 2012), et cette observation a été démontrée dans des modèles murins. L'IL-10 est un régulateur important contre les réponses pro-inflammatoires, et la délétion du gène de l'IL-10 permet de développer une entérocite chez des souris classiques, et semble lié à un défaut de signalisation dans les macrophages (Ip et al. 2017). Cependant, la même délétion du gène de l'IL-10 dans des souris sans germes et donc sans microbiote fonctionnel n'entraînait plus d'inflammation contrairement aux souris ayant des bactéries entériques (Sellon et al. 1998). Depuis, cette observation est étayée par de nombreuses études sur le sujet et la comparaison du microbiote intestinal provenant d'individus sains et de patients Crohn a montré une baisse significative de la diversité bactérienne chez ces derniers (Dicksved et al. 2008).

Il est cependant compliqué pour le moment de statuer si le rôle central de la dysbiose intestinale est une cause ou une conséquence des MICI puisque les dérégulations du microbiote et du système immunitaire sont étroitement liées.

## Stratégies de traitements dans les MICI

Il n'existe pas de traitement curatif pour la maladie de Crohn ou la rectocolite hémorragique à l'heure actuelle. Les traitements proposés ont pour objectifs principaux de soulager des symptômes et réduire les lésions puis dans un second temps de maintenir la rémission et limiter les risques de rechute. Les premiers traitements proposés dans les cas de maladies inflammatoires chroniques ont longtemps été restreints à des traitements non-spécifiques comme les corticostéroïdes pour leurs propriétés anti-inflammatoires fortes et sont encore utilisés à l'heure actuelle lors des poussées inflammatoires. L'arrivée de traitements plus ciblés (**Figure 13**) a toutefois permis d'améliorer la prise en charge des patients sur le long terme et d'améliorer leur qualité de vie (Sherman et al. 2014). Dans la RCH, le traitement de première ligne se fait majoritairement avec des dérivés aminosalicylés comme le 5-ASA en tant qu'anti-inflammatoire local pour les poussées faibles et modérées, et les corticostéroïdes ou la chirurgie sont utilisées dans les cas les plus graves. En cas de non-réponse, il est envisagé de passer à des thérapies plus ciblées, comprenant des traitements par anticorps anti-TNF, des anti-intégrines, des anti-IL-

12/IL-23p40 ou des inhibiteurs de la voie JAK (Kobayashi et al. 2020). Dans la MC, les corticostéroïdes sont utilisés en première ligne, et en cas de résistance, c'est majoritairement le traitement aux anti-TNF qui est envisagé (Torres et al. 2017). Les traitements anti-adhésion ou anti IL-12/IL-23p40 peuvent aussi être proposés, mais pas le 5-ASA et les inhibiteurs de la voie JAK qui sont moins efficaces dans la MC que le RCH (W. Lim et al. 2016; Rogler 2020). Le rationnel de l'utilisation des thérapies plus avancées est de tenter de cibler plus finement les malfunctions du système immunitaire qui participent à la physiopathologie des MICI.

### Inhibiteurs du TNF $\alpha$

Le TNF $\alpha$  est une cytokine pro-inflammatoire produite majoritairement par les lymphocytes T, cellules NK activées et les macrophages mais aussi à plus petite échelle par d'autres cellules immunitaires ou non comme les lymphocytes B, les fibroblastes et mastocytes. À l'homéostasie, le TNF $\alpha$  assure l'intégrité de la barrière intestinale en médiant l'équilibre entre la survie et la mort programmée des cellules. La survie passe par son action sur l'activation de la voie NF $\kappa$ B et l'expression de gènes anti-apoptotiques qui en découle. À l'opposé, dans certaines conditions le TNF $\alpha$  va induire l'apoptose des cellules épithéliales intestinales ou leur nécroptose via la caspase 8 (Günther et al. 2011; 2013). Bien qu'il semble jouer un rôle protecteur en lien avec l'immunité innée dans la phase précoce de l'inflammation avec l'immunité innée, mais devient un facteur aggravant dans les phases tardives, avec un rôle pro-inflammatoire très important. Il joue un rôle pour recruter et activer des lymphocytes et des granulocytes. Il augmente aussi l'adhésion cellulaire en induisant une surexpression de MAdCAM-1, ce qui induit l'infiltration excessive des lymphocytes (Sasaki et al. 2003). Le TNF $\alpha$  participe aussi à la dérégulation la barrière en impactant l'efficacité des jonctions serrées (Al-Sadi et al. 2016)

Étant donné l'importance du TNF $\alpha$  dans la phase chronique des MICI, il est logique que son blocage ait été envisagé depuis longtemps. Le premier traitement approuvé dans le traitement des MICI est l'Infliximab, un anticorps monoclonal chimérique anti-TNF $\alpha$  (Targan et al. 1997) en réponse à la découverte que le TNF $\alpha$  était augmenté dans les muqueuses de patients Crohn (MacDonald et al. 1990; Murch et al. 1993). L'infliximab peut se fixer au TNF $\alpha$  dans sa forme soluble, ce qui l'empêche sa fixation aux récepteurs TNFR1 et TNFR2. Mais il peut aussi cibler la forme transmembranaire qui est exprimée à la surface des lymphocytes T et monocytes activés, et ainsi induire leur apoptose (Atreya et al. 2011). Le traitement anti-TNF $\alpha$  a fait ses preuves non seulement pour réduire les symptômes mais aussi la reformation de la muqueuse et une réduction des besoins de chirurgie (Van Assche, Vermeire, et Rutgeerts 2010). L'Infliximab a aussi montré des preuves d'efficacité pour les patients atteints de RCH dans la réduction des atteintes pour les patients traités (Rutgeerts et al. 2005). Cependant, près d'un tiers des patients ne répondent pas au traitement initial par anti-TNF $\alpha$ , et près d'un tiers des répondeurs vont devenir insensibles au traitement au cours du temps (Allez et al. 2010; Eder et al. 2015). De nouvelles stratégies thérapeutiques sont peu à peu arrivées pour tenter de combler cette efficacité partielle.

## Anticorps anti- $\alpha 4\beta 7$

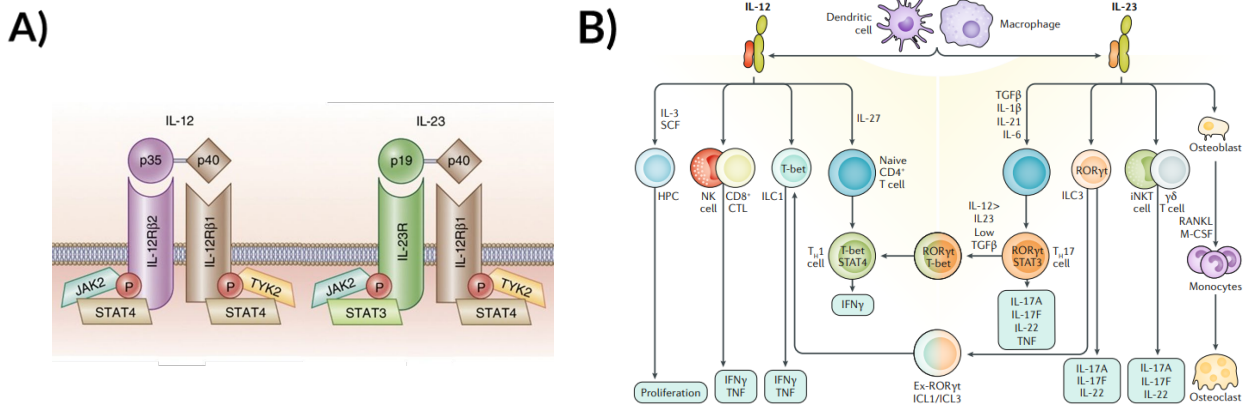
Une autre hypothèse repose sur la présence accrue des lymphocytes dans les tissus de l'intestin et les tissus lymphoïdes associés à l'intestin qui aggravent l'inflammation dans les MICI. Bloquer les molécules qui sont responsables de l'adhérence à l'endothélium et leur entrée dans les tissus permettrait donc de réduire cette inflammation. Une particularité de l'endothélium intestinal est que les cellules endothéliales qui le composent expriment spécifiquement une molécule pour recruter les lymphocytes dans les muqueuses : le MAdCAM-1 (Berlin et al. 1993; Briskin et al. 1997; Gorfu, Rivera-Nieves, et Ley 2009). Et à la surface des lymphocytes, c'est le complexe de l'intégrine  $\alpha 4\beta 7$  qui est le ligand associé à MAdCAM-1. Un anticorps monoclonal humanisé IgG1 a été développé pour se fixer à l'intégrine  $\alpha 4\beta 7$  afin d'empêcher l'infiltration des lymphocytes spécifiquement dans l'intestin (Picarella et al. 1997). Malgré des résultats encourageants en essais cliniques (Feagan et al. 2016; Sands, Peyrin-Biroulet, et al. 2019) le blocage d' $\alpha 4\beta 7$  s'est avéré avoir un impact aussi pour bloquer l'infiltration d'autres leucocytes comme les monocytes non classiques vers les muqueuses intestinales. Or le déficit d'infiltration de ces monocytes non classiques semble réduire la fonction de réparation tissulaire et d'homéostasie, soulignant par la même occasion l'hypothèse que ces monocytes non classiques se différencieraient préférentiellement en macrophage ayant un profil de résolution de l'inflammation (Schleier et al. 2020)

## Anticorps anti-interleukines IL-12/IL-23p40

Les deux interleukines IL-12 et IL-23 sont deux cytokines appartenant à la famille des IL-12, et à la superfamille des IL-6. Elles sont toutes deux des hétérodimères et partagent une sous unité en commun : la p40. L'IL-12 est un hétérodimère p40-IL12p35 tandis que l'IL-23 est formée par le dimère p40-IL23p19 (**Figure 12**). Cette similarité de structure leur confère des rôles proches tout en ayant chacune leur spécificité (Oppmann et al. 2000). L'IL-12, aidé de l'IL-27, induit la différenciation des lymphocytes naïfs CD4+ en T helper type 1 (Th1) (Trinchieri 1995). Ces Th1 sont des cellules sécrétrices d'INF $\gamma$ , d'IL-2 et de TNF $\alpha$  qui servent d'une part à stimuler la production d'IgG opsonisantes pour recruter le complément et effectuer une lyse des pathogènes, et d'autre part à stimuler les cellules cytotoxiques (LT, NK) et activer les cellules phagocytaires comme les macrophages. Les Th1 sont donc des acteurs de la réponse contre les pathogènes intracellulaires à l'homéostasie. Cependant dans les MICI, ces Th1 ont une auto-expansion dérégulée et s'accumulent dans les tissus de l'intestin et participent à l'inflammation par la sécrétion exacerbée de leurs cytokines pro-inflammatoires. De son côté, l'IL-23 n'induit pas de différenciations par manque de récepteur à l'IL-23 (IL-23R) sur les LT CD4+ naïfs, mais joue un rôle dans le maintien de la population de T helper 17 (Th17). Les Th17 vont exprimer l'IL-17A et l'IL-23R suite à leur différenciation, les rendant sensibles à l'IL-23. La fixation de l'interleukine à son récepteur active la voie STAT3 qui déclenche une boucle de rétrocontrôle positive grâce à la transcription de ROR $\gamma$ t et d'IL-23R, permettant de maintenir la pathogénicité des Th17 et leur expansion (Bunte et Beikler 2019). L'IL-17A qui est connue pour son rôle pro-inflammatoire est donc très liée à l'expression

du récepteur IL-23R et à la présence d'IL-23 au sein des tissus intestinaux, ce qui fait de cette voie une cible thérapeutique intéressante de la réponse immunitaire adaptative, surtout que des SNPs touchant la région du gène IL23R font partie des loci associés aux MICI (Duerr et al. 2006).

Des antagonistes d'IL-12 et 23 ont donc vu le jour, dont l'Ustekinumab, un anticorps monoclonal IgG1k bloquant la sous unité p40 des interleukines IL-12 et 23, et qui empêche leurs fixations aux complexes IL-12Rβ1-IL12Rβ2 et IL-23R-IL12Rβ2 respectivement (**Figure 12**).



**Figure 12. Structure et biologie de l'IL-12 et l'IL-23.**

A) L'IL-12 est constituée de la sous unité p40 et p35 et l'IL-23 de p40 et p19.

B) L'IL-12 et IL-23 sont produites par les DC et les macrophages. L'IL-12 induit la différenciation des T CD4+ Naïfs en Th1 et induit la production d'IFNγ tandis que l'IL-23 stabilise les Th17 et induit la production d'IL-17A, IL-17F et IL-22. D'après (Moschen, Tilg, et Raine 2019)

L'Ustekinumab était précédemment utilisé en clinique pour traiter le psoriasis, une maladie inflammatoire de la peau liée à la voie Th17 et la surproduction d'IL-23. Cette relation sur l'immunogénicité de la voie Th17 explique pourquoi il est désormais étudié dans des essais cliniques appliqués aux MICI (Feagan et al. 2016), surtout que la présence d'une des deux pathologies est un facteur de risque significatif à l'apparition de l'autre (Vlachos et al. 2016)

### Autres stratégies

La dysbiose étant un élément central dans les MICI, des stratégies d'adaptation du régime alimentaire ont été proposées pour remodeler le microbiote intestinal. Une autre méthode plus complexe est d'effectuer un transfert fécal pour modifier directement la flore bactérienne. Mais cette transplantation n'est proposée que dans le cas d'une infection par *Clostridium difficile*, une souche bactérienne très agressive (Rohlke et Stollman 2012; van Nood et al. 2013). On peut aussi introduire plus spécifiquement certaines souches bactériennes connues pour leur rôle protecteur pour modifier la composition du microbiote intestinal, c'est ce qu'on appelle un probiotique. On peut citer par exemple le rôle de la souche de *Roseburia intestinalis* qui semble participer à l'inhibition de l'inflammasome, à la différenciation en Treg et au maintien des jonction serrées (Nie et al. 2021; Shen et al. 2022), ou

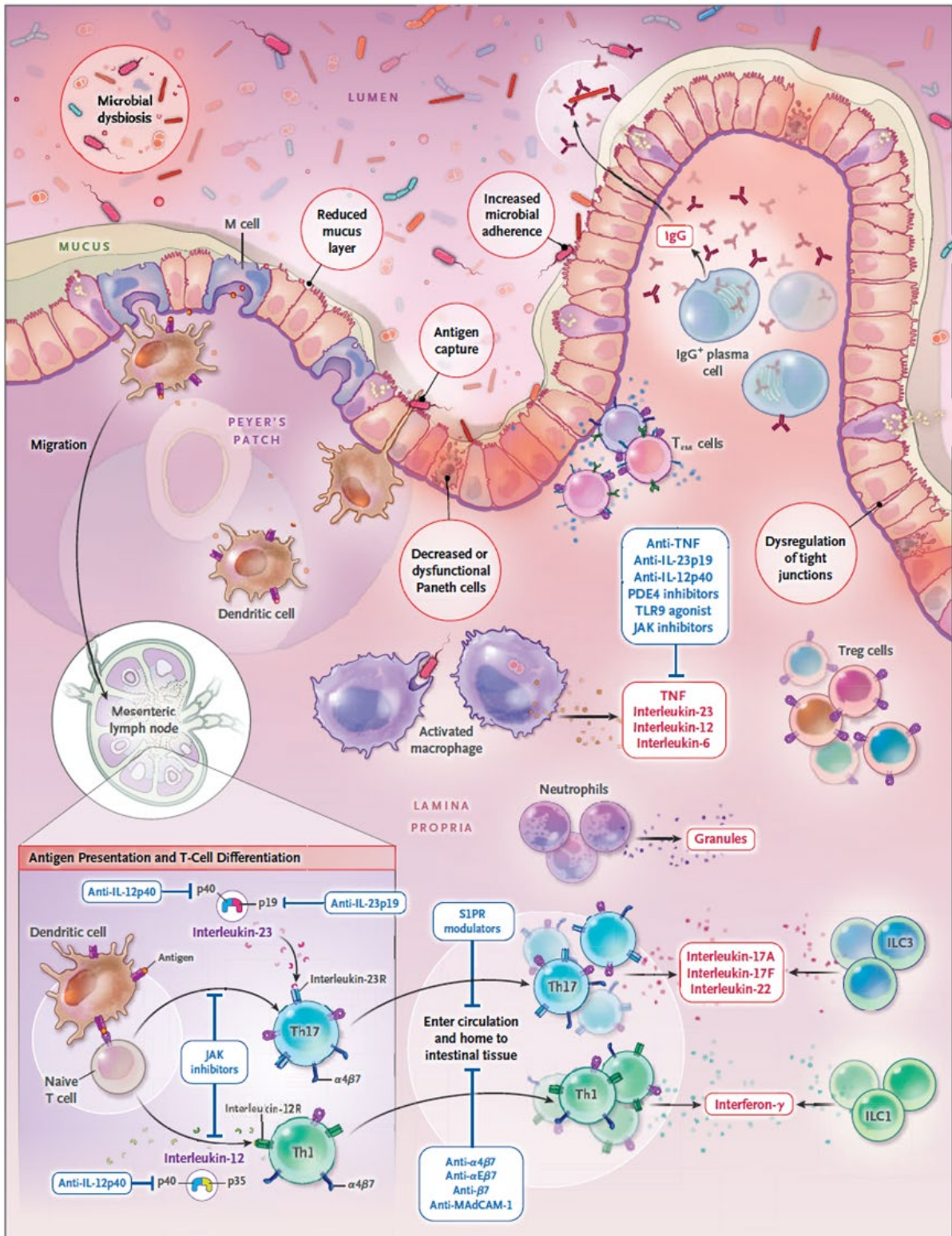


*Faecalibacterium prausnitzii*, une bactérie qui produit des peptides anti-inflammatoires et induit la production de l'IL-10 anti-inflammatoire et dont le nombre diminue fortement en cas de MICI (Sokol et al. 2009; Quévrain et al. 2016).

## Stratégie générale de prise en charge des patients

Comme décrit précédemment, les objectifs des traitements des MICI reposent sur l'amélioration de la condition du patient, et le maintien de la rémission. Cependant, les essais cliniques se concentrant sur la prévention de la progression se font attendre et il manquait un consensus sur la définition de la progression de la maladie dans les MICI. À cela s'ajoute le manque de critères pour caractériser la réponse induite par les traitements dans les études cliniques car on a difficilement les preuves de leur fonctionnement direct, mais aussi des questions comme le choix de prise en charge dans les cas où aucun traitement ne semble fonctionner. Ces problématiques ont poussé des experts du domaine à proposer une stratégie de prise en charge commune des patients MICI reposant sur deux consensus : le consensus STRIDE-II pour *Selecting Therapeutic Targets in Inflammatory Bowel Disease*, qui établit des objectifs de traitement à court terme (Peyrin-Biroulet et al. 2015; D. Turner et al. 2021), et le consensus SPIRIT pour *Selecting End Points for Disease-Modification Trials* qui établit des objectifs à long terme sur le maintien de la rémission et les changements de thérapies (Le Berre, Peyrin-Biroulet, et SPIRIT-IOIBD study group 2021)

Les objectifs thérapeutiques fixés par le consensus (**Figure 14**) comprennent entre autres l'amélioration de la qualité de vie du patient, avec une réponse clinique rapide et une rémission clinique définie par une baisse des douleurs abdominale et un retour à la normale de la fréquence des selles. A moyen terme, l'objectif porte sur la prévention des complications comme les dommages tissulaires à l'intestin pour la MC, l'extension de la pathologie dans la RCH, le recours à de la chirurgie ou à des hospitalisations.

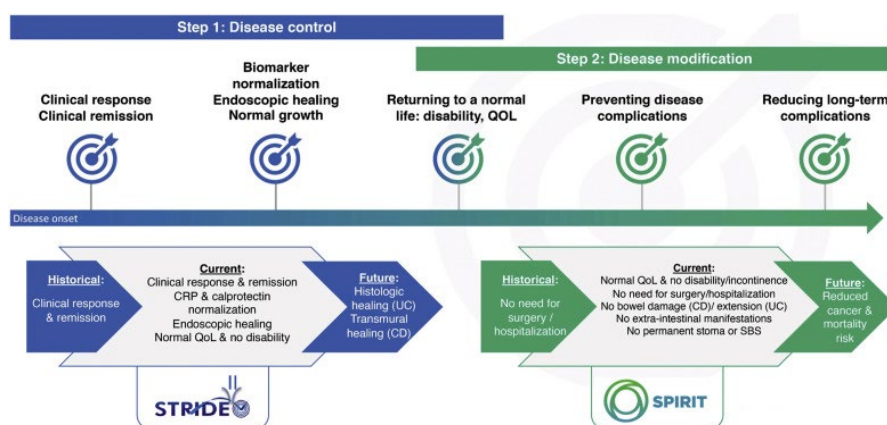


**Figure 13. Système immunitaire des muqueuses intestinales durant la dysbiose.**  
 En rouge : les cytokines et molécules sécrétées. En bleu : les traitements anti-inflammatoires.

Un autre objectif à moyen terme est de retrouver des taux normaux de protéine C-réactive, une protéine sécrétée par le foie dans le sang en cas d'inflammation dans l'organisme, et un taux normal de calprotectine, dont les taux élevés sont liés à une dégradation des parois intestinales. Sur le long terme, on cherchera à obtenir une rémission endoscopique, attestée par colonoscopie ou endoscopie. La rémission

observée en histologie ou la cicatrisation transmurale vue par IRM ne font pas partie des objectifs cruciaux à atteindre dû à la difficulté de définir une rémission complète de la muqueuse, mais permettent d'attester un niveau de rémission plus profond. Cette rémission endoscopique ou histologique est aussi associée à des intervalles sans poussées plus longs et une réduction des complications et besoins d'hospitalisation (Cazzato et al. 2021). Enfin, le dernier objectif est l'absence de handicap ou de complications sur le long terme comme l'émergence de cancers ou le décès du patient.

L'utilisation de thérapies avancées s'avère efficace surtout en tant que première ligne de traitement, et l'adaptation à un nouveau traitement en cas de résistance s'accompagne souvent d'une baisse d'efficacité comparé à quand utilisé en premier (Baumgart et Le Berre 2021). Cette baisse d'efficacité pourrait être directement liée à l'impact du premier traitement, mais aussi à une évolution de la maladie dont l'augmentation de la gravité réduit l'efficacité quand une nouvelle thérapie est appliquée. Et il est probable aussi que les traitements actuels qui ciblent surtout la voie lymphocytaire bénéficient le plus aux mêmes patients, avec un certain chevauchement des voies ciblées : les anti-TNF qui participent à la déplétion des lymphocytes T, les anti- $\alpha 4\beta 7$  réduisent l'infiltration de ces LT et les anti-IL12p40 inhibent l'activation des LT.



**Figure 14. Vue d'ensemble de la stratégie clinique de sélection de thérapie en MICI.** Objectifs à court terme (STRIDE-II) récapitulant les étapes de décision nécessaires au choix de traitement à appliquer pour optimiser une rémission et objectifs à long terme (SPIRIT) pour optimiser le maintien de cette rémission. D'après (Baumgart et Le Berre 2021)

## Résistance aux traitements

Malgré la diversité de traitements existant, il n'y a pas moyen de prévoir à l'avance avec certitude lequel sera efficace, le seul moyen étant souvent de les tester. La résistance à un traitement va être définie par l'échec d'obtenir une rémission durable suite au traitement, mais les raisons des résistances primaires et progressives à ce traitement qui touchent une part importante des patients restent mal définies (Atreya, Neurath, et Siegmund 2020). On privilégiera alors les traitements qui fonctionnent pour le plus grand nombre en premier et selon la progression, le dosage ou le

traitement sera changé. Les recherches en génétiques n'ont pas permis d'identifier d'association entre les variants connus et la résistance aux traitements existants, et des variants identifiés comme *NOD2* ne sont informatifs que sur la localisation de la maladie. Ils ne peuvent donc pas être utilisés comme marqueurs pour faire un choix de traitement (Cleyne et al. 2016). Il existe cependant des profils transcriptomiques identifiés par séquençage d'ARN qui permettent d'associer l'expression de certains gènes à une résistance primaire ou secondaire au traitement. Dans la MC, c'est par exemple la sur-représentation de *TNFAIP6*, *S100A8*, *IL11*, *GOS2* et *S100A9* dans une cohorte de patients MC avec une atteinte colique qui permet d'identifier les patients non-répondeurs primaires à l'Infliximab à partir de biopsies prises avant traitement (Arijs et al. 2010). Cette signature ne permettait pas de prédire la réponse dans les MC à atteinte iléale, mais il a plus tard été montré que dans les deux cas, l'enrichissement en ARNm d'*IL13RA2* dans les muqueuses est corrélé à la résistance primaire à l'Infliximab, probablement dû à une plus forte présence de TNF dans ces muqueuses (Verstockt et al. 2019). Dans la RCH, certaines signatures de gènes ont été associées à la résistance au traitement par anti-TNF $\alpha$ , comme la signature définie par la même équipe que précédemment, avec les gènes *IL13RA2*, *TNFRSF11B*, *STC1*, *IL-6* et *IL-11* enrichis chez les patients RCH non-répondeurs à l'Infliximab. Pour rentrer plus en détail, *S100A8* et *S100A9* qui se dimérisent pour former la calprotectine, sont tous les deux induits par des cytokines pro-inflammatoires comme TNF $\alpha$  et l'IL-6. L'IL-6 est un des régulateurs importants dans l'inflammation chronique puisqu'il régule la différenciation et l'activation des LT : il participe avec TGF $\beta$  à la différenciation des LT CD4<sup>+</sup> naïfs en Th17 et en contrepartie inhibe la différenciation en Treg (Kimura et Kishimoto 2010; Waldner et Neurath 2014). L'IL-11 est une interleukine majoritairement exprimée par les myofibroblastes activés, qui semble participer aux complications fibrotiques dans les MICI (Schafer et al. 2017; W.-W. Lim et al. 2020). L'Ostéoprotégérine (OPG), codée par le gène *TNFRSF11B*, est une protéine enrichie dans les MICI et induite par le TNF qui semble jouer un double rôle, pro-inflammatoire en induisant la production de cytokines et protecteur en inhibant l'apoptose de certaines cellules épithéliales (De Voogd et al. 2016).

Néanmoins, ces gènes restent des gènes marqueurs et ne permettent pas de connaître en détail les raisons causant l'émergence de l'inflammation ou les raisons pour lesquelles ils sont enrichis dans les MICI. Il est donc important d'améliorer la compréhension des MICI afin d'apporter le traitement adapté au plus tôt et d'identifier les causes de résistances aux traitements actuels pour réduire les recours aux chirurgies et leurs effets secondaires.

Une proposition récente pour adapter la stratégie de traitement au patient est de développer des organoïdes intestinaux, c'est à dire une version miniature et simplifiée de l'intestin que l'on développe à partir des cellules prélevées directement sur le patient, parfois dans la zone inflammée afin de tenter d'émuler un organoïde inflammé aussi. Cette technique permettrait donc d'obtenir un modèle pour tester l'efficacité in vitro de différents traitements en parallèles en étant plus proches de la condition tissulaire du patient, bien que limités par l'absence du microbiote intestinal (Angus et al. 2020). Mais à l'heure actuelle, il reste plus accessible de se concentrer sur l'étude

des molécules, des cellules et de leurs interactions dans les muqueuses intestinales pour dresser un portrait plus précis des causes de résistances.

Une difficulté supplémentaire pour déterminer l'efficacité du traitement est due à la nature de la pathologie : il est difficile de savoir si la rémission d'un patient précis suite à un traitement est due à l'effet direct de ce traitement ou uniquement dû à l'évolution naturelle de la pathologie qui est caractérisée par des phases de poussées suivies de phases de rémission. La génétique n'a pas permis non plus d'identifier clairement des marqueurs génétiques liés à la résistance aux traitements : aucun variant identifié en GWAS n'a pu montrer d'association avec l'efficacité thérapeutique, la seule association étant celle du gène *NOD2* avec une localisation iléale de la maladie. Cela souligne le manque d'outils pour aiguiller le choix des traitements et stratifier les patients en amont. C'est pour cette raison que l'utilisation de la transcriptomique est une perspective encourageante pour identifier de nouveaux processus moléculaires impliqués dans la résistance. La différence de réponse thérapeutique à des symptômes similaires doit provenir d'une origine cellulaire et moléculaire précise, et dans ce contexte, le scRNA-seq peut se révéler très utile pour analyser la composition cellulaire des tissus et l'ensemble des gènes exprimés par chaque cellule.

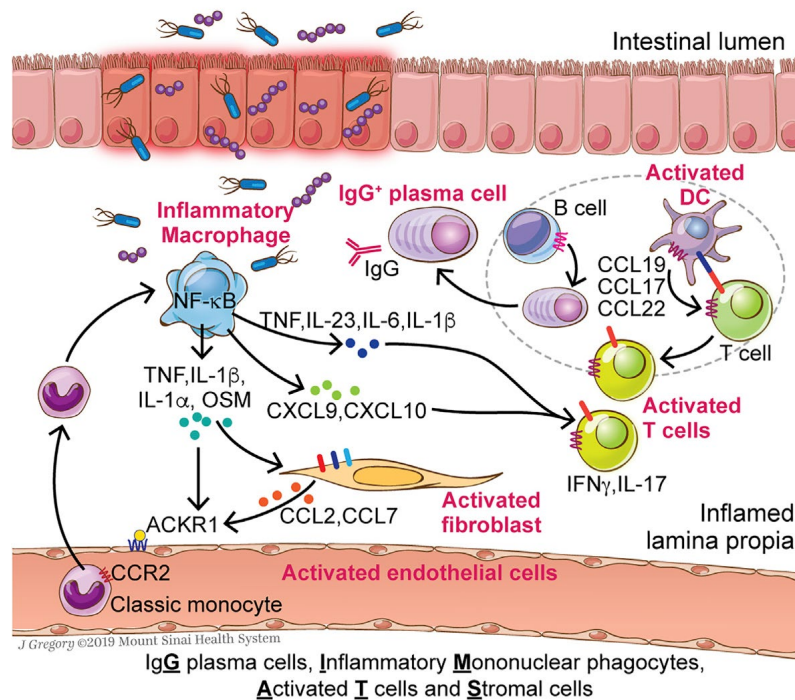
## Utilisation du single cell dans la caractérisation de la pathologie

### Le scRNA-seq dans la rectocolite hémorragique

Dans la RCH, le scRNA-seq a été utilisé dans un premier temps pour étayer les résultats de GWAS et pour décrire les populations différentes trouvées spécifiquement dans la RCH (Smillie et al. 2019). Cela a permis d'observer en cellule unique l'expression des gènes liés à une augmentation de risques par de précédentes études GWAS. Sur les 57 gènes de risque, 36 d'entre eux se retrouvent différenciellement exprimés lors de la maladie, et l'étude en cellule unique a permis de définir pour certains dans quel type cellulaire précis leur expression était modifiée. C'est le cas de *CCL20*, *NR5A2* et *JAK2* qui sont surexprimés dans les cellules Microfold-like du colon des patients atteints de RCH, et la dérégulation de plusieurs gènes de risques dans le même type cellulaire semble indiquer un rôle important de ces cellules dans la maladie. De plus cette étude a montré l'enrichissement d'une population de fibroblastes inflammatoires spécifique aux tissus inflammés de RCH. Une seconde étude centrée sur les lymphocytes B à laquelle mon équipe a participé (Uzzan et al. 2022) a aussi permis de mettre en lumière une dérégulation forte avec une augmentation des LB naïfs imprimés par l'IFN et des plasmocytes producteurs d'IgG. Cette étude a notamment démontré la présence de clones B autoréactifs ciblant l'intégrine  $\alpha 5\beta 6$  nécessaires à l'activation de la cytokine immunorégulatrice TGF $\beta$  par les cellules épithéliales intestinales. Les résultats de l'étude suggéraient également une activation des macrophages inflammatoires favorisée par l'engagement de leurs récepteurs pour le fragment Fc des IgG. La présence de cette réponse B dérégulée dans l'intestin inflammatoire des patients était associée à des formes plus compliquées de la maladie.



## La signature GIMATS, une nouvelle piste pour comprendre la résistance à l'anti-TNF $\alpha$ dans la maladie de Crohn



**Figure 15. Composition de la réponse GIMATS dans les tissus inflammatoires de Crohn.** Les Macrophages inflammatoires récemment infiltrés jouent un rôle central dans la structuration de la réponse GIMATS par l'activation des lymphocytes T et des fibroblastes et participent à une réponse inflammatoire excessive. D'après (Jerome C. Martin et al. 2019)

Le scRNA-seq peut aussi être permettre de définir les cadres cellulaires et moléculaires pouvant stratifier les patients en sous-groupes s'intéret en regard de paramètres cliniques relevant comme la réponse aux traitements ciblés. C'est le cas des précédents travaux de notre équipe qui ont montré une hétérogénéité cellulaire entre les iléons inflammatoires de patients Crohn (Jerome C. Martin et al. 2019). L'analyse de la distribution cellulaire permise par la précision du scRNA-seq a mené à définir un ensemble de types cellulaires enrichis dans un sous-groupe de patient peu à même de répondre au traitement par anticorps anti-TNF $\alpha$ . Ce module a été nommé GIMATS d'après les types cellulaires enrichis le constituant : des plasmocytes IgG, des phagocytes Mononucléés Inflammatoires, des lymphocytes T activés et des cellules Stromales Activées (**Figure 15**). Leurs rôles dans le processus inflammatoire sont les suivants : les anticorps IgA et IgM ont un rôle protecteur de la muqueuse vis-à-vis des microorganismes du tractus intestinal, mais les IgG qui sont sécrétées par les plasmablastes IgA<sup>-</sup> CXCR4<sup>+</sup> ont précédemment été décrites dans les MICI comme engageant les récepteurs Fc $\gamma$ R des macrophages, ce qui participe ensuite à la génération de Th17 pathologiques (Uo et al. 2013; Castro-Dopico et Clatworthy, s. d.). Les macrophages inflammatoires sont définis par un profil qui se distingue des macrophages résidents, avec notamment une sur-expression de cytokines inflammatoires, et une expression accrue du facteur de transcription NF- $\kappa$ B. Ces macrophages expriment de nombreuses cytokines et chimiokines inflammatoires

comme CCL3 et CCL4, des ligands de CCR5 qui permettent de recruter des monocytes et des lymphocytes T effecteurs (D. Wang, DuBois, et Richmond 2009) ainsi que les chimiokines CXCL2, CXCL3 et CXCL8 pour recruter des neutrophiles dans le site inflammatoire. Ces macrophages inflammatoires expriment aussi des cytokines qui peuvent activer les LT et les cellules stromales dont TNF, IL-6, IL-1 $\beta$ , IL-23 et OSM (Neurath 2014). L'expression excessive d'IL-23 est proposée pour expliquer l'échappement à l'apoptose des LT dans les patients non répondeurs à l'anti-TNF $\alpha$  (Schmitt et al. 2019) et l'Oncostatine M (OSM) est aussi associée à une résistance quand elle est retrouvée élevée dans les échantillons avant traitement (N. R. West et al. 2017). Des DC activées ont aussi été observées dans des agrégats de lymphocytes T activés. Les DC activées expriment de manière élevée les chimiokines CCL17, CCL19, CCL22, connues pour être des ligands des récepteurs de chimiokines CCR4 et CCR7, pouvant être à la membrane des lymphocytes T, ce qui semble indiquer leur rôle direct dans la formation des agrégats. Dans le compartiment des fibroblastes, les fibroblastes activés sont définis par leur expression accrue de *THY1*, *PDPN* et *CHI3L1*, des gènes marqueurs d'activation. Ces fibroblastes produisent les chimiokines CCL2 et CCL7, qui sont captées par les cellules endothéliales exprimant ACKR1 (Pruenster et al. 2009). Ces cellules endothéliales facilitent ensuite l'infiltration des leucocytes dans les tissus, incluant notamment les monocytes exprimant le récepteur CCR2.

Dans ce module GIMATS dont les multiples acteurs participent à la fonction pro-inflammatoire, ce sont les macrophages inflammatoires qui semblent jouer un rôle central dans la pathogénicité des MICI, de par l'activation des LT et des fibroblastes, et par la boucle indirecte pour le recrutement de monocytes additionnels qui vont ensuite se différencier en macrophages inflammatoires pathogéniques.

Cette hétérogénéité de profils cellulaires permet une stratification des patients Crohn entre patients enrichis dans le module : GIMATS<sup>high</sup> et ceux appelés GIMATS<sup>low</sup> qui ont majoritairement des LT mémoires, des macrophages résidents et des plasmocytes IgA. L'association du module à la réponse aux anti-TNF $\alpha$  est intéressante sur le plan thérapeutique puisqu'elle montre une sensibilité différente au traitement qui serait liée à des réponses cellulaires et moléculaires distincts dans les tissus intestinaux.

## Metacells

C'est dans l'optique d'utiliser des outils plus adaptés à cette résolution que j'ai utilisé un algorithme de partition des données de scRNA-seq : Metacells (Baran et al. 2019; Ben-Kiki et al. 2022, 2), présenté plus tôt dans la première partie. Cet algorithme codé en langage Python permet d'identifier les cellules, provenant de différents échantillons ou non, qui ont un profil d'expression similaire pour les regrouper en des unités synthétisant l'expression de ces cellules, appelées metacells.

# Objectifs de recherche

Ces dernières années, l'émergence de toutes les technologies ayant trait au séquençage de nouvelle génération (NGS) a modifié la façon de faire de la science, avec une importance grandissante de l'analyse des grandes quantités de données générées. Durant ma formation j'ai eu l'occasion de couvrir les bases de l'immunologie en ayant en plus une introduction aux méthodes d'analyse informatiques des données scientifiques. Et c'est ce cursus qui mélange biologie et informatique que je souhaite mettre à profit de l'analyse des processus immunologiques au sein de mes projets, unis par la technique et les méthodes de recherche.

Mon premier projet s'inscrit dans une démarche exploratoire, pour faire un profilage des cellules immunitaires au cours du temps chez les patients transplantés pour suivre les altérations visibles dans le sang en réponse à l'émergence d'une réponse inflammatoire localisée afin d'identifier des marqueurs précoces de rejet dans le sang et mieux comprendre les mécanismes de rejet chronique.

Mon second projet s'inscrit en continuité des recherches de mon équipe (Jerome C. Martin et al. 2019) qui ont montré l'intérêt d'appliquer l'analyse transcriptomique là où les études portant sur l'efficacité des traitements manquaient de profondeur. Les macrophages inflammatoires apparaissaient comme des acteurs centraux de l'organisation de la réponse GIMATS dans les tissus inflammatoires de patients ne répondant pas au traitement par anticorps anti-TNF. Mes objectifs dans ce projet sont

1. D'approfondir la caractérisation moléculaire des populations de macrophages rencontrées dans l'intestin inflammatoire des patients avec maladie de Crohn iléale ;
2. D'analyser la régulation de des programmes moléculaires d'intérêt identifiés dans ces populations ;
3. d'explorer les relations entre ces programmes moléculaires et ceux exprimés par les populations de fibroblastes inflammatoires.

## Résultats

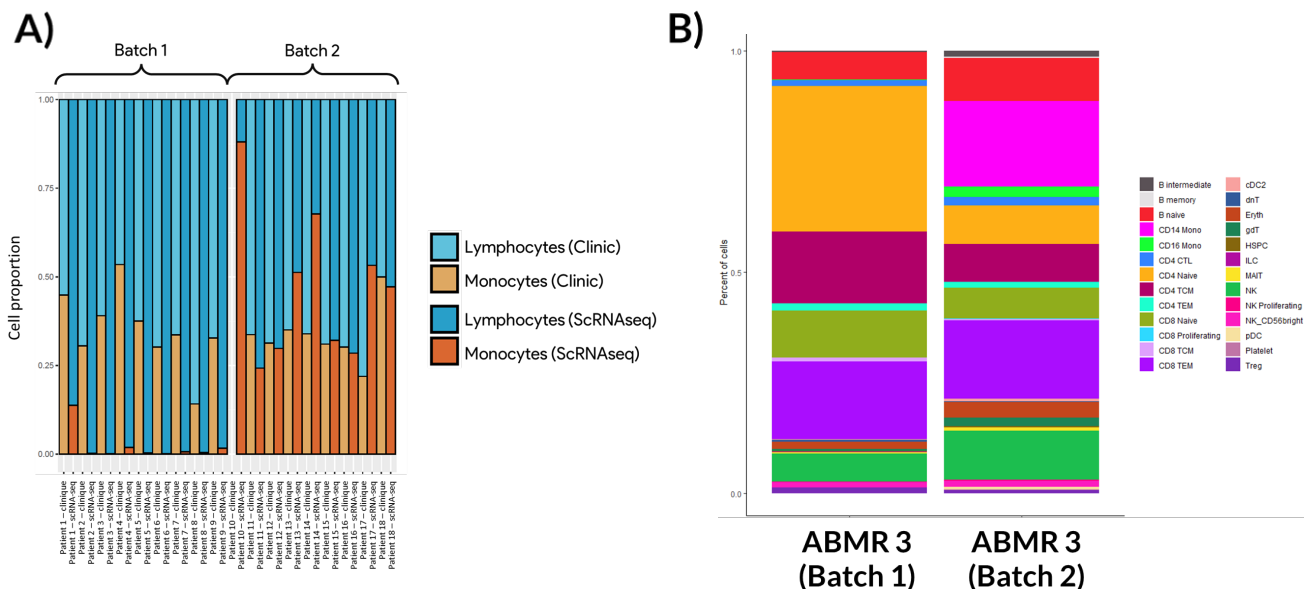
### 1) Le rejet de greffe de rein

Sur ce projet, j'ai contribué à la partie expérimentale dans le choix de la cohorte et la préparation des échantillons, et à la partie portant sur l'analyse bio-informatique du transcriptome des patients séquençés en scRNA-seq pour intégrer les données, identifier les populations cellulaires et caractériser les différences longitudinales et liées au phénotype.

Ce projet a commencé en amont de ma thèse, durant mon stage de 2e année de Master, pendant lequel je me suis familiarisé aux analyses de données issues de scRNA-seq, et où j'ai pris part à la mise en place et à l'optimisation de l'étape de préparation des échantillons pour les passer en scRNA-seq. Les échantillons de la cohorte de patients transplantés que nous avons utilisés sont des échantillons de PBMC congelés dans l'azote depuis 1 à 10 ans. Pour garantir une bonne qualité d'encapsulation en microfluidique, il est nécessaire d'avoir un nombre de cellules suffisant au chargement mais surtout une bonne viabilité (>70% sur les



recommandations de 10X Genomics). Un nombre trop élevé de cellules mortes et de débris résiduels peut altérer la formation de microgouttelettes mais aussi produire un bruit de fond qui gênera la bonne analyse des transcrits cellulaires. Nous avons donc optimisé notre méthode pour atteindre des taux de viabilité entre 75 et 90%, et environ 12000 cellules vivantes chargées par expérience (voir les Méthodes du papier). Cependant, l'analyse des résultats de cette première cohorte a révélé un problème dans notre méthode de préparation, la proportion de monocytes obtenues en scRNA-seq est bien plus faible qu'attendue, et largement sous-estimée en comparaison des résultats de numération cellulaire effectués en clinique sur ces mêmes patients. Nous avons identifié l'origine du problème comme venant des colonnes de filtration de cellules mortes (Dead Cell Removal Kit - Miltenyi Biotec) qui semblent retenir par erreur les monocytes avec la fraction de cellules mortes. Nous sommes donc passé à une exclusion des cellules mortes par cytométrie en flux avec un marquage au DAPI, et pour notre nouvelle cohorte les résultats de proportions Lymphocytes/Monocytes dans les PBMC correspondent à la numération cellulaire (**Figure 16 A**). Une preuve supplémentaire est l'utilisation d'échantillons en duplicats d'un patient commun entre les deux expériences, où pour des échantillons prélevés au même moment, les résultats de proportions cellulaires en scRNA-seq indiquent une forte diminution de monocytes classiques et non classiques capturés, au profit d'un plus fort taux de lymphocytes T (**Figure 17**). L'analyse de cette cohorte est présentée dans le papier qui suit.



**Figure 16. Disparités de composition cellulaire selon la méthode d'élimination des cellules mortes en scRNA-seq.**

A) Proportion de Monocytes et de Lymphocytes dans les échantillons de PBMC, les couleurs claires correspondent aux données de numération cellulaire en clinique, et en foncé les valeurs issues du scRNA-seq du patient apparié. L'élimination de cellules mortes est effectuée par colonne d'éluion dans le batch 1 et par tri en FACS dans le batch 2.

B) Comparaison des proportions cellulaires pour dans les échantillons d'un patient commun aux deux batches. Les annotations cellulaires sont effectuées par annotation automatique d'Azimuth.

## **Title:** A gene-expression module in circulating immune cells regulates cellular activation and migration during immuno-pathologies

Thomas Laurent\*, Debajyoti Sinha\*, Cynthia Fourgeux, Simon Ville, Thibault Letellier, + Laurence Bouchet-Delbos, Clarisse Kerleau, Magalie Giral, Regis Josien, Sophie Brouard & Jeremie Poschmann

### **Abstract**

Here we report the longitudinal single-cell transcriptome profiling of peripheral blood mononuclear cells in patients undergoing kidney transplantation rejection. We identify a gene expression module which is associated to transcriptional regulation, homing and early activation in multiple cell types. The cells expressing this module are reduced in patients undergoing rejection when compared to stable kidney transplantation patients. We confirm this phenomenon of module reduction in a pig model of acute kidney transplantation rejection. Furthermore, we find that the module gene expression is increased in kidney biopsies during rejection as compared to stable biopsies, indicating that the cells expressing this module are recruited into the graft. This is supported by the membership of the chemokine receptor CXCR4 gene in this module, providing a molecular communication mechanism between circulating cells and a CXCL12 gradient originating from injured graft. We then explore publicly available scRNA-seq data and show that this module is also associated to various other immune states such as Covid-19 disease and systemic inflammatory response syndrome as mimicked by LPS injection. In summary, we show that not all circulating cells are equal, but that cells expressing this module are preferentially recruited to the tissue to mediate their effector activities.

### **Introduction**

Circulating immune cells are critical to replenish dying cells and to be recruited to the site of inflammation, infection, and cancer. This compartment which includes the peripheral blood mononuclear cells (PBMCs) thus offers an attractive resource for precision medicine as in a single experiment, diverse cell types, including the CD4<sup>+</sup> T cells, CD8<sup>+</sup> T cells, B cells, NK cells and monocytes are profiled. Especially since the introduction of scRNA-seq, profiling PBMCs has been highly successful in identifying gene expression signatures and cell-types associated with immune-related diseases<sup>1-4</sup>. For example, a monocyte signature associated with sepsis was discovered in circulating cells<sup>1</sup>. Recently, a study on systemic lupus erythematosus revealed gene expression changes with disease state and genetic variation<sup>2</sup>. In addition distinct Covid-19 studies revealed signatures associated to infection and disease severity in blood<sup>5,6</sup>. Transcriptomics profiling enables the characterization of genes groups (i.e., modules) which perform critical cellular functions such as maintaining a cell identity, homeostasis & metabolism and respond to external signals. Notably, in circulating immune cells we have previously shown that monocytes express a gene module associated with Herpes Simplex virus reactivation after traumatic brain injury<sup>7</sup>. While gene expression programs of circulating immune cells are likely to be distinct from the same cells which migrated into the tissue,

identifying modules in circulating cells may reveal early immune activation programs or modules associated with homing and migration. For example, in a previous study we identified large gene regulatory and gene expression alterations in circulating monocytes during active Tuberculosis which improved pathogen clearance for these cells <sup>8</sup>.

Leveraging on scRNA-seq and the availability of recent module identification approaches tailored for single cell transcriptomics <sup>9</sup> we aimed to identify gene expression programs associated with kidney transplantation rejection. Currently, rejection status is monitored via clinical practice via blood and urine metabolites like creatinine to assess the renal function and then validated via pathology analysis of kidney biopsies. But metabolite monitoring is nonspecific to rejection and can be approximative. Pathology analysis of kidney biopsies thus remains the main diagnostic tool with a more reliable result, but even if it is of limited risk for the patients, it is an invasive procedure which can't be performed on a regular basis. Thus identifying gene expression signatures in circulating cells may improve precision medicine diagnostics of kidney rejection <sup>10</sup>. In addition, the rejection is characterized by the infiltration of immune cells into the graft, which then mediate the inflammation response in the tissue, causing the rejection of the graft. Blood thus constitutes an easily accessible compartment to identify gene expression modules associated with homing and early activation. Two archetypes of rejection are prominent according to the Banff classification<sup>11</sup>, either antibody mediated rejection (ABMR) or T-cell mediated rejection (TCMR) and which can also arise in a mixed form. The infiltration of immune cells into the graft is typically mediated via the blood stream either through donor specific antibodies (DSA) which bind to the endothelium of the graft in the case of ABMR or via cytokine and homing signals in the case of TCMR <sup>12</sup>

In this study, our aim was to identify putative modules in circulating cells which may be associated to kidney transplantation rejection. We profiled a longitudinal patient cohort (3 stable patients (no rejection), 3 TCMR and 3 ABMR patients across three timepoints) for which PBMCs were collected at the same time of graft biopsy, allowing us to characterize relationship of graft rejection status with gene expression modules of peripheral immune cell-types. We identified a module associated with transcriptional regulation and early activation and used an acute pig-transplantation model to validate its association with rejection. Further characterization of this module was carried out in transcriptomics data from over 1500 kidney biopsies revealing a cytokine-receptor interaction between the graft and circulating cells, respectively. Finally, we demonstrate that this module is not specific to graft rejection but is implicated and predictive of a variety of immuno-pathologies.

## Results

### Single cell transcriptome analysis of circulating immune cells in a longitudinal kidney transplantation cohort

We performed scRNA-seq on PBMC isolated from 3 stable (STA) patients, 3 antibody mediated rejection (ABMR) patients and 3 T-cell mediated rejection (TCMR) patients (Figure 1A). The patients were selected based on their treatment, age, sex, collection time etc. (see Table 1). For each patient, three time points were profiled; T0 which happened within the 1st days after the transplantation, T1 at 90-150 days after T0 and T2 which was sampled at time of rejection for ABMR and TCMR or for STA patients at 1 year after transplantation. Rejection status was defined by clinical pathology assessment of kidney biopsies performed at T1 and

T2 for all patients. For ABMR and TCMR patients, T2 was associated with a kidney rejection state, whereas for STA patients T1 and T2 were defined as stable kidney states. In addition, PBMC samples from 3 kidney transplantation patients which maintained allograft function in the absence of immunosuppression (i.e., operational tolerant (TOL)) were included in the cohort (Table 1). To minimize scRNA-seq related experimental variation, we performed CITE-seq (Cell Hashing)<sup>13</sup> using hashtag oligo-conjugated antibodies (HTO) to label each patient and time point and then pooled 10 samples (1 ABMR, 1 TCMR and 1 STA patient across T0, T1 & T2 and 1 TOL patient) into a single experiment. We thus generated the complete transcriptomic data in 3 balanced batches (Figure 1B). After removing doublets, data cleaning, normalization, and batch correction we obtained 50,507 cells across the three batches (see Methods). Cells were automatically annotated using Azimuth<sup>14,15</sup> and manually validated for cell-type specific markers (Figure 1C and Supplemental Figure S1A). Cell-type proportions varied minimally between conditions (Supplemental Figure S1B). We inspected the distribution of cells across patients and by time points (Supplemental Figure S1C) and across disease states (Figure 1D). Neither of these variables formed unique clusters suggesting that the clusters were driven by cell-type specific expression rather than by condition or cellular states as was also observed in other PBMC studies in patient cohorts<sup>1,2,5</sup>. We then explored whether the clusters were affected by merging the three experimental batches (Figure 1E). To precisely quantify such possible batch effects not visible by eye, the KBET metric was used to evaluate the batch correction (Supplemental Figure S1D)<sup>16</sup>. The acceptance rate for the complete data set was 0.969, indicating that batch integration was successful. In summary, the pooling strategy and subsequent bioinformatics analysis resulted in a robust dataset of 50,507 cells to be analyzed for time and disease state specific gene expression.

### **Gene co-expression analysis identifies a module associated to rejection state**

We used a gene co-expression analysis to identify modules (i.e., co-expressed groups of genes) across

all three batches independently (Figure 2A). This approach was chosen to avoid potential signal alterations induced during the batch correction step. We applied consensus non-negative matrix factorization (cNMF)<sup>9</sup> to identify gene expression programs which may either be associated to cell-type specific gene expression programs or to cellular activity (see methods). For each batch 11 to 14 modules were identified, and Jacquard score was used to select modules with overlapping genes across the three batches. This resulted in nine overlapping modules which were evenly distributed across the three batches (Figure 2B, Supplemental Figure 2A). These modules mostly revealed cell-type specific expressions, notably three of these modules were associated to monocytes (Mod\_1-Mod\_3), and two modules were mostly expressed in a specific cell type such as B cells and pDC cells (Mod\_4 and Mod\_5 respectively). Three modules were enriched for the CD4 & CD8 T lymphocytes (Mod\_6) and/or NK cells (Mod\_7, Mod\_8, Figure 2C). Interestingly, the Mod\_9 was present in multiple cell-types (B, CD4, CD8, DC and NK cells) as revealed by its co-expression score (module score) in various cell types (Figure 2D). To test whether this cell-activity module or any of the cell-specific modules were associated to disease state, i.e. (rejection or stable) and if it would vary throughout time, we estimated the module score for each cell type and calculated the trend of the module score longitudinally in each cell-type (Figure 2E). A positive or negative slope thus indicates whether

a module changes over time across the stable, humoral & cellular rejection. While neither of the cell-type specific module revealed a significant trend between time points, or any differences between conditions, the activity- module revealed a negative trend in the rejection conditions but not in the stable condition. Further inspection of this module trend revealed that there was indeed a reduction in ABMR patients, while in TCMR patients the module score reduced sharply during T1 and restored to prior levels (T0) during T2. Interestingly, the 3 stable patient's module scores remained consistent across time in most cell-types NK, CD4 and CD8 cells (Figure 2F).

### **Characterization of the module**

To explore the function of Mod\_9 we first investigated whether it was enriched for ribosomal, proliferation or cell cycle genes which are characteristic gene sets altered across cell types in various scRNA-seq dataset. There was not a single gene overlap of any of such genes which were previously described for such functions and are provided by SEURAT<sup>17</sup> implying that this module is involved in other cellular functions (data not shown). We thus explored the genes within this module by performing gene ontology analysis. Enrichment of the module genes was quantified for molecular function (MF) and biological processes (BP) compared to the combined set of 4000 most variable genes from the three batches (Figure 2G). MF could be associated to 24 genes and was significantly enriched (FDR <0.05) for helicase activity, ubiquitin-like protein binding, RNA catalytic activity, transcription activator activity and ribonucleoprotein binding. The BP (20 genes) was associated to NF- $\kappa$ B signaling, response to peptide and oxidative stress, regulation of RNA metabolic processes and viral gene expression. Of the 56 annotated module genes, 30 genes were linked to a gene ontology. This indicates that this module is likely to be involved in multiple molecular functions implicated in transcription, mRNA process and ubiquitination. The BP suggest an involvement in the response to stress (i.e., viral gene expression, NF- $\kappa$ B signaling, oxidative stress). This was supported by the 5-fold enrichment of Transcription factor genes in this module (OR: 4.9; Fisher-Test P-value  $2.5e^{-5}$ ), such as the AP-1 complex (JUN, JUND, FOS), REL (NF-KB subunit), MAFF and NR4A2 (Figure 2G). Further manual examination exposed the early activation marker CD69, a cell surface type II lectin. This receptor was described to be rapidly expressed at the membrane in T cells upon TCR activation<sup>18</sup>. In addition, CD69 has been described as a marker of tissue retention of activated T cells of chronic inflammatory states<sup>19-21</sup>. Interestingly, CD69 gene promoter is controlled by AP-1 TF complex and NF- $\kappa$ B, both of which are also members of this module<sup>22,23</sup>. In addition, we revealed the presence of the cell surface marker CXCR4 in this module, a chemokine receptor known to play a role for recruiting leukocytes in the kidney after an ischemic injury via the chemokine CXCL12<sup>24</sup>. The role of the CXCR4/CXCL12 axis in kidney rejection is still unclear<sup>25</sup> but an elevated expression of CXCL12 has been described in chronic kidney rejection<sup>26</sup> compared to control allograft kidney tissues, suggesting may act as a chemotactic signal to recruit immune cells in the inflamed tissues. Next, we investigated the cohesion of Mod\_9 by quantifying its gene co-expression using pair-wise Pearson correlation coefficient R. As the module was identified using cNMF, Pearson correlation coefficient R is thus an independent evaluation of gene co-expression. We note that pairwise gene correlation analysis is used for weighted gene co-expression analysis (WGCNA), a prominent method to identify modules in bulk and single cell transcriptomics<sup>27</sup>. We compared the avg. Pearson correlation per cell for Mod\_9 genes to a set of 63 genes randomly picked in

the same cells. While ALARM genes displayed a robust and significant correlation in all cell types (for rejection and stable), random genes picked 1000 times revealed a Pearson correlation of 0 (Supplemental figure S2B). That the Mod\_9 genes are co-expressed by chance is thus highly unlikely.

In summary, the Mod\_9 module is composed of genes implicated in gene expression, response to stress, mRNA processing and early activation and homing. We thus chose the acronym ALARM for eArLy Activation trAnscription factoR Module.

### **Circulating ALARM expressing cells are recruited to the graft during acute graft rejection in a pig kidney transplantation model**

We found that circulating cells expressing ALARM are reduced in humoral & cellular rejection in a timely dependent manner (Figure 2). To test this phenomenon, whether ALARM expressing cells are reduced in circulating cells during kidney transplantation rejection, we set up a controlled experiment of acute rejection in a pig kidney transplantation model. Two HLA-mismatched pigs were subjected to kidney transplantation, keeping one of their own kidneys (see methods). This model typically results in an acute rejection within a few days after transplantation as no immunosuppressive treatment is given (Figure 3A). Kidney biopsies and PBMCs were collected daily before and after the transplantation. Microscopical analysis of the biopsies at D2, D4 and D6 indicated a time-dependent infiltration of immune cells which culminated at day 7 when the animal was sacrificed (Figure 3B). We note that the second pig did not display any signs of rejection, possibly due to under perfusion of the transplanted kidney and was thus discarded from the subsequent analysis. On the PBMCs collected at D0, D2, D4 and D6, cell hashing was performed to pool them in a single scRNA-seq experiment, resulting in a total of 4411 annotated cells across cell-types and time-points (Figure 3C and Supplemental figure 3). We note that cellular proportions within the PBMC compartment drastically changed from D0 to D2 after the transplantation characterized by a drastic increase in monocytes concomitant with a reduction of B, CD4+ and CD8+ T cells (Figure 3C). This suggests that the lymphocytes are rapidly recruited to the kidney graft and accumulate there, as demonstrated by the cellular invasion observed in the biopsies at the same time (Figure 3B). To test whether the decline of blood lymphocytes is accompanied by a reduction of ALARM expressing cells, we quantified ALARM expression across the time-points in CD4+, CD8+, B cells and NK cells (Figure 3D). Interestingly, as soon as D2 the levels of ALARM cells drastically decreased when compared to before transplantation (D0) and remained low in the blood until sacrifice of the animal (D6). Taken together, this controlled time-course experiment reveals a drastic immune cell infiltration in the graft associated with the depletion of ALARM expressing T and B lymphocytes and NK cells in the circulation. This experiment mirrors the reduction of circulating ALARM cells observed during the kidney transplantation rejection (Figure 2). It is thus likely that ALARM expressing cells are preferentially recruited to the tissue during rejection.

### **ALARM gene expression increases in kidney grafts undergoing rejection**

To confirm the hypothesis that ALARM expressing cells are recruited to the graft during rejection from the blood stream, we evaluated the expression of ALARM genes in graft biopsies from kidney transplantation patients. For this, we used a previously published

transcriptomics analysis of 1,208 biopsies collected from 13 clinical sites and with a patient classification into STA, TCMR, ABMR and mixed rejection (TCMR and ABMR)<sup>28</sup>. After precleaning and QC controls of the available microarray data (Methods), we quantified the ALARM gene expression for each sample (Figure 4A). The ALARM score was consistently and significantly increased in the three rejections as compared to the stable biopsies. Heatmap analysis of all the ALARM genes further revealed that this score is driven by the upregulation of a large fraction of the ALARM genes, most notably, by a cluster of genes containing CD69, CXCR4 and IRF1 (Figure 4B). Differential gene expression analysis confirmed an upregulation of most ALARM genes suggesting a concerted upregulation of this module. To corroborate these results, we exploited a second similar transcriptomics study performed on kidney biopsies in 224 patients which were either stable (168 patients) or undergoing ABMR<sup>29</sup>. Consistent with the above result, there was a significant upregulation of ALARM genes in the graft during rejection as compared to the stable patients (Figure 4C and 4D). Collectively, these results indicate that ALARM genes are systematically increased during rejection in the grafts. Given that graft rejection is defined by immune cell infiltration and that circulating ALARM cells are depleted at the same time, it is possible that ALARM cells are preferentially recruited to the graft to mediate the rejection.

To investigate how ALARM cells could be preferentially recruited to the graft, we explored the cytokine expression in the graft and receptor expression in the circulating cells. We first identified all possible ligand- receptor pairs and then tested whether these pairs were differentially expressed. We found 10 differentially expressed cytokines in the graft pairing with 7 receptors upregulated in circulating immune cells (Figure 4E). The most prominent receptor was CXCR4, expressed in CD4, NK, CD8 and B cells. As mentioned above, CXCR4 is a member of ALARM genes, indicating a likely mechanism of signaling from the graft via CXCL12 and recruitment of ALARM cells expressing CXCR4. This cytokine receptor pair has been previously described as a homing mechanism in various distinct tissues, such as bone marrow<sup>30,31</sup> and in cancer<sup>32</sup>. This receptor-ligand interaction was independently validated in the 2<sup>nd</sup> cohort of kidney rejection biopsies (Figure 4F). In summary, ALARM cells are recruited to the kidney graft from circulation likely via the homing signaling pathway of CXCL12 and CXCR4.

### **ALARM cells are implicated in distinct disease states**

So far, the role of ALARM cells has been explored with regards to kidney transplantation and thus we investigated whether this module may also be involved in other disease conditions which implicate circulating immune cells. To test this hypothesis, we exploited a publicly available scRNA-seq data on PBMCs in which healthy individuals were intravenously injected with the endotoxin lipopolysaccharide (LPS), a component of the cell wall of Gram-negative bacteria<sup>5</sup>. LPS in the bloodstream causes an immediate systemic release of a variety of inflammatory mediators, a fever and a rapid but transient leukopenia<sup>33</sup>. This experiment is thought to mimic an acute systemic inflammatory response (SIRS)<sup>5</sup>, and thus provides an ideal proxy of how ALARM expressing cells are responding to LPS-induced SIRS and leukopenia. We used the preprocessed available scRNA-seq data which contained the 0-time point (before LPS injection), 90 minutes and 10h after the LPS injection and first evaluated how the cellular proportions changed over time. We then measured the ALARM score in the three conditions across the CD4, CD8, NK and B cells (Figure 5A). There was rapid and significant

downregulation of ALARM expressing cells as soon as 90 minutes and which further decreased until 10h after the LPS injection. We evaluated whether this downregulation was apparent in the same individuals. For this, we had data from three individuals for which both time points (90min and 10h) were available for all cell-types (Figure 5B). This drastic change of ALARM expression in such a short time frame suggests that cells which highly express this module rapidly egress from circulation. Furthermore, this response to LPS which is thought to be initiated via TLR4 receptor signaling expressed on circulating monocytes<sup>34,35</sup> signifies that ALARM is not solely implicated in transplantation rejection or kidney immune cell invasion, but appears to also be involved in the inflammatory response to endotoxin. To extend on this notion, we analyzed an additional scRNA-seq datasets generated in PBMCs from Covid-19 patients<sup>5</sup>. Interestingly, while moderate and severe Covid-19 patients had an ALARM decrease concomitant with disease severity, non-symptomatic Covid-19 patients displayed a drastic increase of ALARM expression as compared to healthy volunteers (Figure 5C). Of note, there were cell-type specific differences in ALARM expression, notably NK cells were reduced for ALARM in non-symptomatic Covid-19 as opposed to the other lymphocytes. This suggests that ALARM is a general module but used differently in cell-types with distinct functions. We observed in various inflammatory states (LPS-induced and Covid-19) that ALARM cells are drastically changing, and this is dependent on the condition (symptomatic vs non-symptomatic), location of inflammation (systemic vs local) and disease severity (moderate vs severe).

## Methods

The scRNA-seq was performed in 3 different experiments following the same protocol. Each experiment included longitudinal samples from three patients (one stable patient, one humoral rejection and one cellular rejection) as well as one late sample of a tolerant patient.

### 1.1 Cell preparation

Frozen vials of peripheral blood were obtained from multiple transplanted patients followed through time, dating back more than 9 years for the oldest samples. The PBMC samples were frozen with DMSO 10%, but due to its toxicity for cells, the thawing must be performed quickly to wash the DMSO away. The frozen PBMC samples were successively immersed in a pre-warmed 37°C water bath for a few seconds until there was only a little block of ice left. Each sample is then transferred in a 50mL conical tube and diluted with 24mL of complete Roswell Park Memorial Institute (RPMI) 1640 media (Invitrogen, Carlsbad, CA) with 5% FBS, pre-heated at 37°C.

### 1.2 Cell purification



Following washing steps in PBS+0.04%BSA, cell pellets were resuspended in 200 $\mu$ L FACS buffer (1X PBS supplemented with 2mM EDTA, 2% FBS) in which dead cells were labelled by adding 0.1  $\mu$ g/mL DAPI (Invitrogen, Carlsbad, CA). Cells were filtered on 70  $\mu$ m cell strainer and living cells were then sorted using a Fluorescence-activated cell sorting (FACS) Aria II cell sorter (BD Biosciences, Mountain View, CA)

### 1.3 Single Cell preparation

We used the same method as previously described<sup>37</sup> for the single cell preparation and analysis. One million cells were kept by sample and resuspended in 100 $\mu$ L of Staining Buffer (PBS,2%BSA,0.01% Tween) according to the cell hashing protocol<sup>13</sup> recommendations. Cells were incubated for 10 min with 10 $\mu$ L of human Fc blocking reagent. Each sample was then mixed with 1 $\mu$ L of a specific TotalSeq-A hashtag antibody (BioLegend, San Diego, CA) and incubated on ice for 30 min. Following 3 washing steps with the Staining Buffer, cells were counted and their viability measured using an ADAM-MC automatic cell counter (NanoEntek, Seoul, South Korea) to ensure a viability above the recommended 70%. All the samples were pooled at an equal cell concentration in a single vial, centrifugated and resuspended in PBS to obtain a concentration of 700 cells/ $\mu$ L, to match the targeted cell recovery of 32,000 cells. Encapsulation of single cells was performed on a 10XChromium (10X Genomics, San Francisco, CA) with the Chromium Single Cell 3' Library and Next GEM reagent kit v3. The libraries were sequenced twice for each of the three experiments on the NovaSeq 6000 (Illumina, San Diego, CA) with S1 flow cells. The sequenced libraries were aligned to the GRCh38-2020-A reference genome with CellRanger v5.0.0 (10X Genomics, San Francisco, CA)

### 1.4 Method demultiplexing and Seurat analysis

The count matrices were analyzed in R 4.0.3 using the Seurat R package (v4.0.2, Satija Lab<sup>14</sup>). Each experiment was first processed separately, with the same workflow. First, following the standard workflow recommendations, cells with less than 200 unique feature counts were removed (potential empty droplets). Cells with a percentage of mitochondrial genes greater than 15% were excluded as it results from mitochondria degradation from dead or dying cells. The hashing antibody sequences were then collected to demultiplex and assign each cell to its sample using the MULTIseqDemux function. Cells with too little labels information were called “Negative” while cells with a high count of two or more different oligo-conjugated antibody sequences were called “Doublets”. Only cells with a unique HTO were kept for downstream analysis. Singlet cells were annotated automatically with the Azimuth workflow within Seurat, by mapping the query cells on an annotated reference of 162,000 PBMC measured with 228 antibodies<sup>15</sup>.

All the runs are then merged in a single Seurat object. Using the scWGCNA<sup>38</sup> R package, gene modules were computed on each major cell type to more easily identify contaminant cells and doublets to exclude. After normalization of the global object, the 2,000 most variable genes in the data were selected to compute the correction using the reciprocal principal component analysis (RPCA). RPCA is suited to find anchors between multiple datasets from a similar platform. The final annotated and corrected object gathering the 12 samples is composed of 50,507 cells

## 2.1 Animals

The study protocol was approved by the French Ministry of Higher Education, Research and Innovation (APAFiS #30136). The experiment was performed on 60 to 80kg male pigs (*Sus scrofa*). Test card with pre-applied antibodies from Serafol (Berlin, Germany) were used to identify the pigs' ABO blood groups. The alloreactivity was performed by mixed lymphocyte reaction assay between donor and recipient. The donor pig was selected from a different breed as inbred pigs might escape rejection. Donor and recipients blood groups were ensured to be compatible to avoid hyper-acute rejection, and mixed lymphocyte reaction assays were positive thus proving their alloreactivity.

## 2.2 Allogeneic transplantation

Unilateral nephrectomies were performed on two recipient pigs under general anesthesia with a premedication by Zolazepam/Tiletamine (Zoletil® Virbac, Carros, France) 15 mg/kg IM, before intubation and a maintained ventilation with a mixture of 49% oxygen, 49% nitrous oxide and 2% isoflurane. The two kidneys from a third donor pig were harvested in the same operating time. The two recipient pigs received one collected kidney each for an orthotopic transplantation. During surgery, a central venous catheter (CVC) was peripherally inserted into the internal jugular vein for hydration and medication. Post-operative analgesia was performed every day with intravenous injections of Nalbuphine (Nubain®, Mylan, Canonsburg, Pennsylvania) and Paracetamol at a dose of 25 mg/kg. Prophylactic antimicrobial therapy was conducted with Cefazolin 1 g (Cefovet®, Dopharma, Ancenis, France).

## 2.3 Sample collection

Blood samples were collected daily through the CVC and frozen in a CoolCell® container (Corning®, Corning, NY, USA) at -80°C following the PBMC isolation. Kidney biopsies were collected daily from both sides using automated biopsy needles of 16 gauges under ultrasound guidance while pigs were sedated by Zolazepam/Tiletamine and locally anesthetized with Lidocaine. Kidney samples were then placed in cryovials with 1mL fetal bovine serum (FBS) and 10% dimethylsulfoxide (DMSO) for gradual cooling in a CoolCell chamber.

## 2.4 Tissue processing

Cryovials were thawed in a pre-warmed water bath at 37°C for one minute before being poured in a well of a 12-well plate. The kidney biopsies were rinsed twice with Roswell Park Memorial Institute medium with 10% fetal bovine serum (RPMI/10%FBS) and incubated in it for 5 minutes. Tissue dissociation was performed in a 1.5 mL microcentrifuge tube by adding 445 µL RPMI, 50 µL Liberase TL (250 µg/mL) and 5 µL DNase I (100 µg/mL). Samples were put in an orbital shaker at 400 rpm and 37°C for 12 minutes, with a manual resuspension after 6 minutes. Enzyme digestion was stopped by adding 500 µL of RPMI/10%FBS. Cells are then filtered with a 70 µm cell strain and washed with additional RPMI/10%FBS. Blood sample thawing also performed in a warm water bath followed by a dilution in RPMI/10%FBS to neutralize DMSO cytotoxic effect.

## 2.5 Single Cell preparation

After filtering, cells were centrifugated at 300g for 10 minutes at 4°C and resuspended in staining buffer for the HTO antibody labeling (see Methods 1.3). For kidney biopsies, HTO1

= day 0, HTO2 = day 2, HTO3 = day 4, and for blood samples an additional HTO9 was used to label day 6. The single cell suspension was prepared according to the Chromium Single Cell 3' Reagent Kits v3 user manual. 10000 cells were loaded for each experiment on the 10XChromium (10X Genomics, San Francisco, CA) with the Chromium Single Cell 3' Library and Next GEM reagent kit v3 according to the user manual. The libraries were sequenced twice for each of the three experiments on the NovaSeq 6000 (Illumina, San Diego, CA). The sequenced libraries were aligned to the Sscrofa 11.1 (February 2017 release) reference genome with CellRanger v5.0.0.

### 2.6 Gene Module Identification

Consensus Non-Negative Matrix Factorization (cNMF) was used to decompose the cell vs gene expression matrix into cell vs module and usage vs gene low-rank matrices. Different values of K ranging from 7 to 14) were explored to arrive at the optimal number of modules. For each K, the *stability* and *error* metrics were examined. The best K was determined such that the error was minimum and the stability was maximum.

In order to mitigate batch effects from the three different runs, each batch was independently processed. All modules from the three batches were pooled and clustered to group identical modules together. Jaccard similarity score was used to define the distance between two modules.

Finally, only those clusters representing modules from all three batches were retained. A unique gene set was then defined for each of these consistent clusters. The common genes from the top 200 genes from all three modules within the cluster were intersected.

### 3. Pig Biopsy Analysis

Kidney biopsies fixed Carnoy's solution for 30 minutes followed by a fixation in formaldehyde for 24h for optical microscopy purpose. A second batch of kidney biopsies was prepared for immunostaining purpose: biopsies were placed in cryomold, covered with optimal cutting temperature (OCT) compound and immersed in cold isopentane. Following their solidification, cryomolds were stored in azote at -80°C. The slides of kidney biopsies were stained with periodic acid-Schiff (PAS) and Masson's trichrome stains (TM) and were evaluated by a nephropathologist.

## **Discussion**

In this study, we gathered a cohort of matching ABMR, TCMR and stable patients and generated a comprehensive scRNA-seq atlas of circulating immune cells across time and conditions. We then identified multiple gene co-expression modules across batches. Five out of nine modules were nearly exclusively expressed in a single cell-type while three were expressed in closely related cells (CD4+, CD8+ and NK cells) and only the ALARM module was prominent in multiple cell-types. The observation that mostly cell-type specific modules were discovered was also described by Kotliar et al.,<sup>9</sup> in which they distinguish between identity (i.e. cell-type specific) gene expression programs (GEP) and activity GEPs. It is possible that cell-type specific co-expressed genes are better detected as they show a greater coherence within a well-defined group of cells. This is also notable in the presented data as

the module scores show less variance than the ALARM module. Nevertheless, cell-type specific gene expression may not necessarily imply that it remains constant across conditions. For example, we noted that some cell type specific modules were associated to disease state (figure 2 regression), notably the NK cell and monocyte specific modules were increased during rejection. It is also possible that because of the relatively low number of genes per cell detected when compared to bulk transcriptomics, cell identity programs are preferentially detected, and more subtle condition specific modules are not robustly detected. Indeed, while cNMF identified revealed additional modules in the separate batches, only the ALARM module was consistently identified across the three batches.

Its genes were highly enriched for transcription factors and gene ontology pathways associated with the gene expression machinery including transcription, mRNA processing and ubiquitination. Prominent transcription factors included the AP-1 complex and the NFkB subunit REL which are both associated with stress responses and inflammation. The membership of CD69 in the ALARM module also suggests a role of stress response. CD69 is a classical early activation marker of lymphocytes, as demonstrated by its rapid display on the surface of T cells after TCR stimulation<sup>39</sup>. In addition, CD69 is also known to be a tissue retention marker as it is expressed on resident memory T cells in distinct tissues. In blood, this gene has been associated with chronic inflammation in various diseases including rheumatoid arthritis<sup>40</sup> and systemic lupus erythematosus<sup>41</sup>. More specifically in the case of an allograft rejection, the increase of CD69 surface antigen has been observed at the protein level on circulating lymphocytes during the acute rejection of pediatric heart transplant<sup>42</sup>. In kidney graft rejection, Posselt et al. showed a similar increase on the CD4+ and CD8+ T cells during acute rejection<sup>43</sup>. The CD69 membership in this module thus indicates that this module is important to prepare circulating cells for TCR stimulation and for tissue retention once moved into a tissue, i.e., to become T resident effector cells.

Intriguingly, within this module we also see the cell surface receptor CXCR4, expressed mostly by lymphocytes as well as monocytes and through which the CXCL12 ligand promotes chemotaxis to tissues via a concentration gradient<sup>44</sup>. CXCL12 is expressed in multiple tissues including the kidney and is altered during pathophysiological responses including immune diseases. Indeed, an alteration of CXCL12 expression was observed in the kidney transplantation biopsies. It is possible that the CXCR4-CXCL12 axis in this module contributes to the recruitment of ALARM expressing cells in the case of kidney transplantation rejection and other immune diseases. This is also consistent with the observation that cells expressing the ALARM module decrease in the circulation during kidney graft rejection. This observation was confirmed by scRNA-seq and histological study of pig kidney biopsies, thus, via CXCR4-CXCL12 leading ALARM cells to infiltrate the tissues during rejection.

This CXCR4-CXCL12 axis also highlights the notion that the ALARM module is not necessarily specific to transplantation rejection or the kidney. Indeed, reanalysis of circulating immune cells from publicly available scRNA-seq data showed dynamic ALARM expression in two distinct conditions (LPS and Covid-19). Intriguingly, ALARM expression was associated with disease severity in covid-19. In both cases as severity increased the ALARM expressing cells were further depleted in the blood stream this may be explained by an increased recruitment of circulating immune cells to the inflamed tissue possibly via a CXCL12 gradient which increases with severity.

There are multiple limitations of this study, first it is solely based on gene transcription and thus remains to be explored for protein expression. While the module expression was shown to low in the blood and high in the kidney tissue, the direct evidence of the recruitment of these cells to the tissue remains to be confirmed. For this, notably trans well assays may be suitable for demonstrating the preferential recruitment of ALARM cells via the CXCL12-CXCR4 axis. A further limitation is that as of now we do not have a protein surface marker panel which would be associated to ALARM expression. Such markers would enable the purification of ALARM cells enabling further molecular and cellular characterization. Modulating the recruitment of ALARM cells to the graft or during Covid19 may thus impact disease severity. Nevertheless, our study remains important in terms of precision medicine, highlighting the notion that not all circulating cells are equal and that a pool of ALARM expressing cells is preferentially recruited to the inflamed tissue. This notion is likely to open up novel strategies of disease monitoring and disease intervention.

## References

1. Reyes, M. *et al.* An immune cell signature of bacterial sepsis. *Nat. Med.* **26**, 333–340 (2020).
2. Perez, R. K. *et al.* Single-cell RNA-seq reveals cell type-specific molecular and genetic associations to lupus. *Science* **376**, eabf1970 (2022).
3. Kazer, S. W. *et al.* Integrated single-cell analysis of multicellular immune dynamics during hyperacute HIV-1 infection. *Nat. Med.* **26**, 511–518 (2020).
4. Zhang, F. *et al.* Defining inflammatory cell states in rheumatoid arthritis joint synovial tissues by integrating single-cell transcriptomics and mass cytometry. *Nat. Immunol.* **20**, 928–942 (2019).
5. Stephenson, E. *et al.* Single-cell multi-omics analysis of the immune response in COVID-19. *Nat. Med.* **27**, 904–916 (2021).
6. Bernardes, J. P. *et al.* Longitudinal Multi-omics Analyses Identify Responses of Megakaryocytes, Erythroid Cells, and Plasmablasts as Hallmarks of Severe COVID-19. *Immunity* **53**, 1296-1314.e9 (2020).
7. Chaumette, T. *et al.* Monocyte Signature Associated with Herpes Simplex Virus Reactivation and Neurological Recovery After Brain Injury. *Am. J. Respir. Crit. Care Med.* (2022) doi:10.1164/rccm.202110-2324OC.
8. del Rosario, R. C. H. *et al.* Histone acetylome-wide associations in immune cells from individuals with active Mycobacterium tuberculosis infection. *Nat. Microbiol.* **7**, 312–326 (2022).
9. Kotliar, D. *et al.* Identifying gene expression programs of cell-type identity and cellular activity with single-cell RNA-Seq. *eLife* **8**, e43803 (2019).
10. Sirota, M. & Sarwal, M. M. Transplantomics: Toward Precision Medicine in Transplantation Research. *Transplantation* **101**, 1777–1782 (2017).
11. Loupy, A., Mengel, M. & Haas, M. Thirty years of the International Banff Classification for Allograft Pathology: the past, present, and future of kidney transplant diagnostics. *Kidney Int.* **101**, 678–691 (2022).
12. Moreau, A., Varey, E., Anegon, I. & Cuturi, M.-C. Effector Mechanisms of Rejection. *Cold Spring Harb. Perspect. Med.* **3**, a015461 (2013).

13. Stoeckius, M. *et al.* Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. *Genome Biol.* **19**, 224 (2018).
14. Stuart, T. *et al.* Comprehensive integration of single-cell data. *Cell* **177**, 1888-1902.e21 (2019).
15. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587.e29 (2021).
16. Büttner, M., Miao, Z., Wolf, F. A., Teichmann, S. A. & Theis, F. J. A test metric for assessing single-cell RNA-seq batch correction. *Nat. Methods* **16**, 43–49 (2019).
17. Nestorowa, S. *et al.* A single-cell resolution map of mouse hematopoietic stem and progenitor cell differentiation. *Blood* **128**, e20–e31 (2016).
18. Yamashita, I., Nagata, T., Tada, T. & Nakayama, T. CD69 cell surface expression identifies developing thymocytes which audition for T cell antigen receptor-mediated positive selection. *Int. Immunol.* **5**, 1139–1150 (1993).
19. Ziegler, S. F., Ramsdell, F. & Alderson, M. R. The activation antigen CD69. *Stem Cells* **12**, 456–465 (1994).
20. Werfel, T., Boeker, M. & Kapp, A. Rapid expression of the CD69 antigen on T cells and natural killer cells upon antigenic stimulation of peripheral blood mononuclear cell suspensions. *Allergy* **52**, 465–469 (1997).
21. Turner, J.-E., Becker, M., Mittrücker, H.-W. & Panzer, U. Tissue-Resident Lymphocytes in the Kidney. *J. Am. Soc. Nephrol. JASN* **29**, 389–399 (2018).
22. López-Cabrera, M. *et al.* Transcriptional Regulation of the Gene Encoding the Human C-type Lectin Leukocyte Receptor AIM/CD69 and Functional Characterization of Its Tumor Necrosis Factor- $\alpha$ -responsive Elements (\*). *J. Biol. Chem.* **270**, 21545–21551 (1995).
23. Castellanos, M. C. *et al.* Expression of the leukocyte early activation antigen CD69 is regulated by the transcription factor AP-1. *J. Immunol. Baltim. Md 1950* **159**, 5463–5473 (1997).
24. Tögel, F., Isaac, J., Hu, Z., Weiss, K. & Westenfelder, C. Renal SDF-1 signals mobilization and homing of CXCR4-positive cells to the kidney after ischemic injury. *Kidney Int.* **67**, 1772–1784 (2005).
25. Song, A., Jiang, A., Xiong, W. & Zhang, C. The Role of CXCL12 in Kidney Diseases: A Friend or Foe? *Kidney Dis.* **7**, 167–176 (2021).
26. Hoffmann, U. *et al.* SDF-1 expression is elevated in chronic human renal allograft rejection. *Clin. Transplant.* **20**, 712–718 (2006).
27. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559 (2008).
28. Reeve, J. *et al.* Assessing rejection-related disease in kidney transplant biopsies based on archetypal analysis of molecular phenotypes. *JCI Insight* **2**, 94197 (2017).
29. Callemeyn, J. *et al.* Transcriptional Changes in Kidney Allografts with Histology of Antibody-Mediated Rejection without Anti-HLA Donor-Specific Antibodies. *J. Am. Soc. Nephrol.* **31**, 2168–2183 (2020).
30. Moll, N. M. & Ransohoff, R. M. CXCL12 and CXCR4 in bone marrow physiology. *Expert Rev. Hematol.* **3**, 315–322 (2010).
31. Yellowley, C. CXCL12/CXCR4 signaling and other recruitment and homing pathways in fracture repair. *BoneKEy Rep.* **2**, 300 (2013).
32. Burger, J. A. & Kipps, T. J. CXCR4: a key receptor in the crosstalk between tumor cells and their microenvironment. *Blood* **107**, 1761–1767 (2006).
33. Vos, A. F. de *et al.* In Vivo Lipopolysaccharide Exposure of Human Blood Leukocytes Induces Cross-Tolerance to Multiple TLR Ligands. *J. Immunol.* **183**, 533–542 (2009).
34. Beutler, B. TLR4 as the mammalian endotoxin sensor. *Curr. Top. Microbiol. Immunol.* **270**, 109–120 (2002).

35. Beutler, B., Du, X. & Poltorak, A. Identification of Toll-like receptor 4 (Tlr4) as the sole conduit for LPS signal transduction: genetic and evolutionary studies. *J. Endotoxin Res.* **7**, 277–280 (2001).
36. Melms, J. C. *et al.* A molecular single-cell lung atlas of lethal COVID-19. *Nature* **595**, 114–119 (2021).
37. Abidi, A. *et al.* Characterization of Rat ILCs Reveals ILC2 as the Dominant Intestinal Subset. *Front. Immunol.* **11**, (2020).
38. Feregrino, C. & Tschopp, P. Assessing evolutionary and developmental transcriptome dynamics in homologous cell types. *Dev. Dyn.* **251**, 1472–1489 (2022).
39. Cibrián, D. & Sánchez-Madrid, F. CD69: from activation marker to metabolic gatekeeper. *Eur. J. Immunol.* **47**, 946–953 (2017).
40. Afeltra, A. *et al.* Expression of CD69 antigen on synovial fluid T cells in patients with rheumatoid arthritis and other chronic synovitis. *Ann. Rheum. Dis.* **52**, 457–460 (1993).
41. Su, C. C., Shau, W. Y., Wang, C. R., Chuang, C. Y. & Chen, C. Y. CD69 to CD3 ratio of peripheral blood mononuclear cells as a marker to monitor systemic lupus erythematosus disease activity. *Lupus* **6**, 449–454 (1997).
42. Schowengerdt, K. O., Fricker, F. J., Bahjat, K. S. & Kuntz, S. T. Increased expression of the lymphocyte early activation marker CD69 in peripheral blood correlates with histologic evidence of cardiac allograft rejection. *Transplantation* **69**, 2102–2107 (2000).
43. Posselt, A. M. *et al.* CD69 expression on peripheral CD8 T cells correlates with acute rejection in renal transplant recipients. *Transplantation* **76**, 190–195 (2003).
44. Britton, C., Poznansky, M. C. & Reeves, P. Polyfunctionality of the CXCR4/CXCL12 axis in health and disease: Implications for therapeutic interventions in cancer and immune-mediated diseases. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* **35**, e21260 (2021).

Table1. Clinical informations of the cohort study

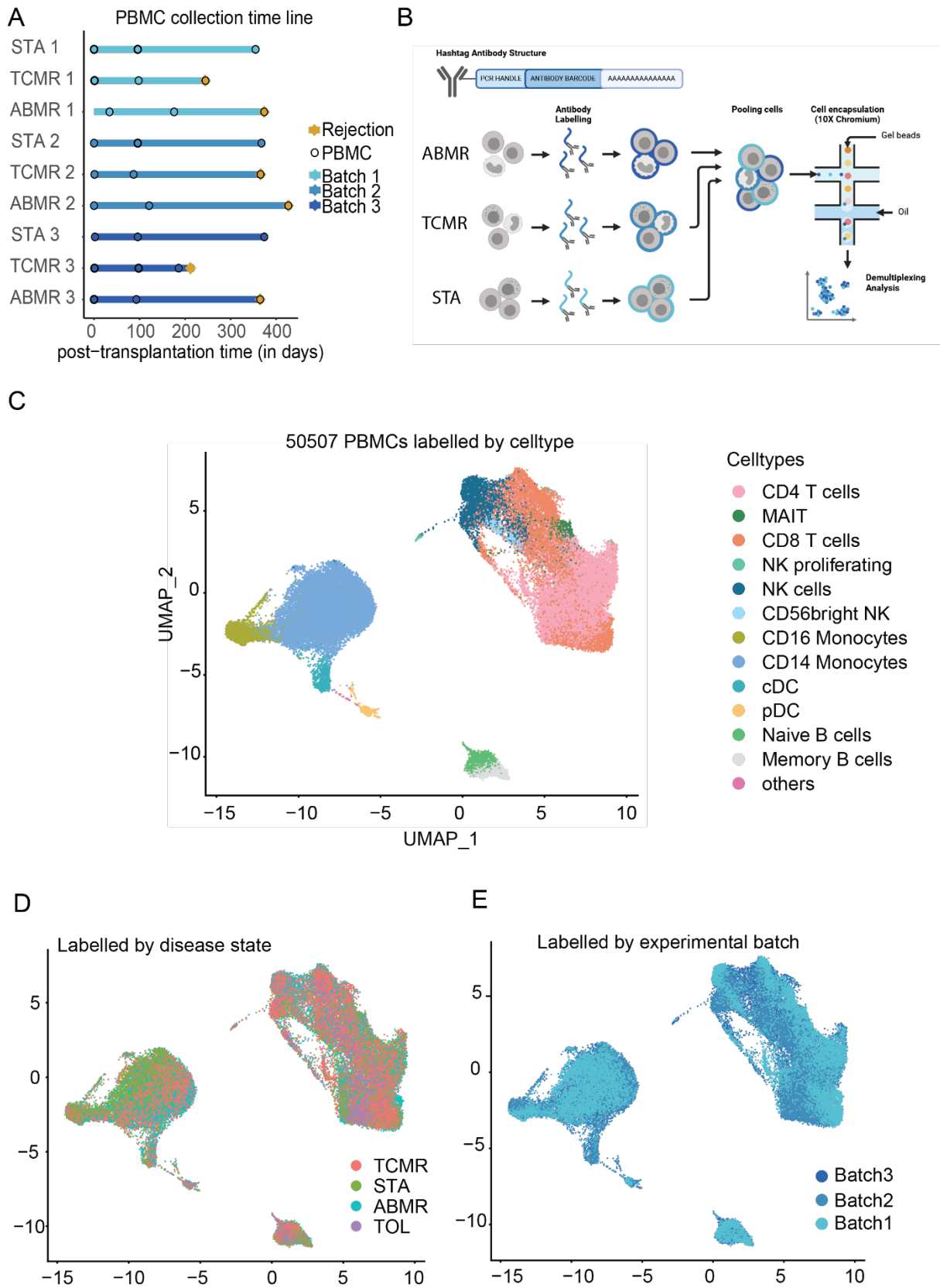
Table 1: Clinical summary of the cohort composition.

Patients characteristics						
	Name	Age	Sex	Treatment	Rejection time (in days)	
Run 1	CT1	35	Male	CNI, Tacrolimus, MMF, Corticoids	–	
	CR1	70	Male	CNI, Tacrolimus, MMF, Corticoids	244	
	HR1	51	Female	CNI, Tacrolimus, MMF, Corticoids	374	
	TT1	75	Male	–	–	
Run 2	CT2	69	Male	CNI, Tacrolimus, MMF	–	
	CR2	35	Male	CNI, Tacrolimus, MMF, Corticoids	369	
	HR2	56	Male	CNI, Tacrolimus, MMF, Corticoids	427	
	TT2	45	Male	–	–	
Run 3	CT3	25	Female	CNI, Tacrolimus, MMF, Corticoids	–	
	CR3	24	Male	CNI, Tacrolimus, MMF, Corticoids	213	
	HR3	61	Female	CNI, Tacrolimus, MMF, Corticoids	365	
	TT3	39	Female	–	–	
Mean Age Sex Ratio						
		48.75	0.66			

\* CNI=Calcineurin inhibitors, MMF=Mycophenolate mofetil



Figure 1



**Figure 1. Comprehensive longitudinal single-cell RNA-sequencing of circulating immune cells in a cohort of kidney allograft recipient.**

A) Timeline of the blood sampling points post-transplantation for the patients followed longitudinally. STA=Stable patient (n=3), TCMR=T cell mediated rejection (n=3), ABMR=Antibody-mediated rejection (n=3). Tolerant patients are not shown.

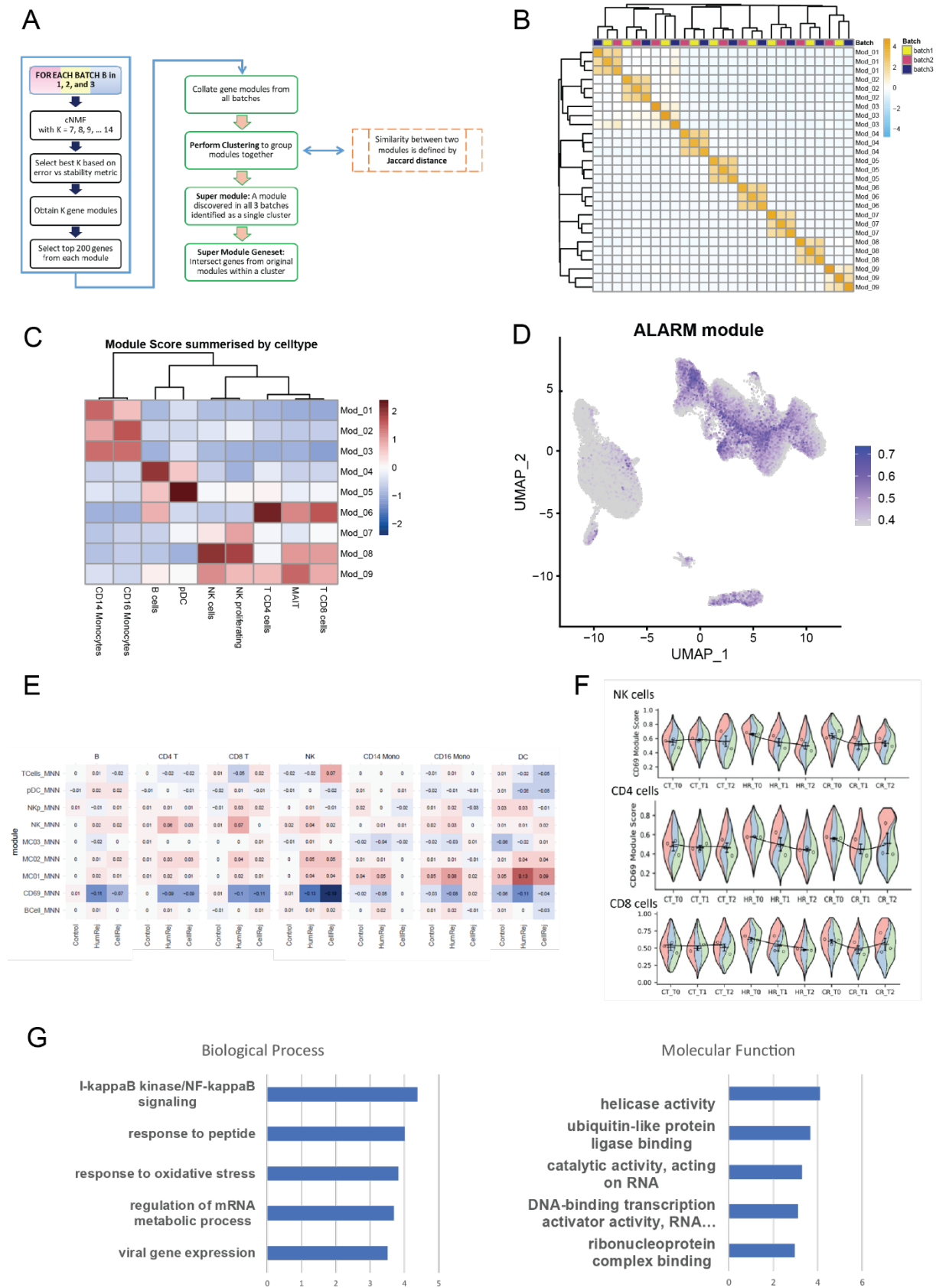
B) Schematic diagram of the scRNA-seq preparation workflow using the 10X Genomics platform. Peripheral blood mononuclear cells (PBMCs) were collected from stable, ABMR and TCMR patients and then stained with one different oligo-conjugated antibody before being pooled and processed using the 10X Genomics platform.

C) UMAP dimensional-reduction embedding of the integrated samples (n=30). Each colour represents a different cell subpopulation, adapted and manually curated from the automatic Azimuth annotation.

D) UMAP projection showing the disease state distribution, TOL=Tolerant patients.

E) UMAP projection coloured according to experimental batch of origin

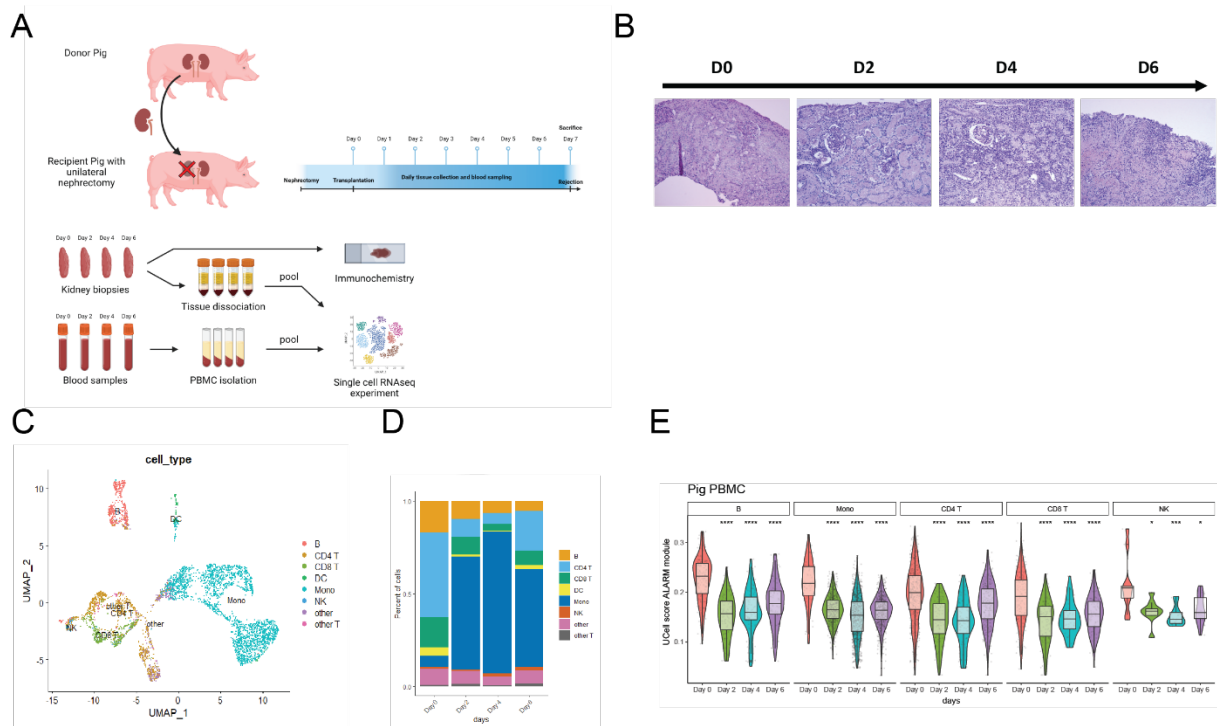
Figure 2



## **Figure 2. ALARM Module identification**

- A) Schematic workflow of the gene co-expression analysis with cNMF module detection. The module selection is based on the overlapping genes between the three batches.
- B) Overlapping scores between the 9 identified modules across the 3 experiment batches, using the Jaccard distance.
- C) Enrichment score of the  $r$  modules computed for each distinct cell-type clusters.
- D) UMAP projection of the ALARM (Module 9) score distribution across cell-types, minimum threshold for ALARM score = 0.4.
- E) Regression analysis of of the 9 modules by cell types. Heatmap shows the beta values (trend) for each cell-type and disease state. Negative values correspond to a decrease of the module score across the three time points, positive values to an increase.
- F) Gene Ontology analysis of biological process (BP) and molecular function (MF) identified by comparing the ALARM gene enrichment using the 4000 most variable genes as background.
- G) Super violin plots showing the longitudinal trend of ALARM module in NK (top), CD4 Tcells (middle) and CD8 Tcells (bottom), stratified by individuals (Red= Batch1, Green= Batch2, Blue= Batch3).

Figure 3



**Figure 3. Pig Model scRNA-seq analysis**

A) Schematic diagram of the acute kidney allograft rejection in pig model. Recipient pig with a unilateral nephrectomy received a kidney graft from a second healthy pig. Biopsies were then collected daily and observed in immunohistochemistry. Immune cells from the tissues were extracted for a scRNA-seq analysis and in parallel, circulating immune cells were prepared for a second scRNA-seq analysis.

B) Immunohistochemistry of pig biopsies stained with Periodic Acid Schiff (PAS), to stain polysaccharides, glycoproteins and glycolipids at day 0, day 2, day 4 and day 6

C) UMAP dimensional-reduction projection of the circulating immune cell types identified in the scRNA-seq experiment after PBMC isolation.

D) Circulating immune cells proportions throughout the different time points (D0-D6) in the recipient transplanted pig.

E) Violin plot of the ALARM module score in B cells, monocytes, CD4 T cells, CD8 T cells and NK cells throughout acute rejection in the kidney tissues.  $*P < 0.05$ ,  $**P < 0.01$ ,  $***P < 0.001$ ,  $****P < 0.0001$ .



**Figure 4. ALARM gene expression increases in kidney tissues during rejection**

A) Violin plot of the ALARM module score across conditions (Data from Reeve et al., 2017) NR=Stable patient, ABMR=Antibody-mediated rejection, TCMR=T cell mediated rejection, Mixed=Graft undergoing ABMR with T cell infiltrates and T-cell mediated lesions.

B) Heatmap showing relative expression of ALARM genes across graft biopsies classified by condition. (Data from Reeve et al., 2017)

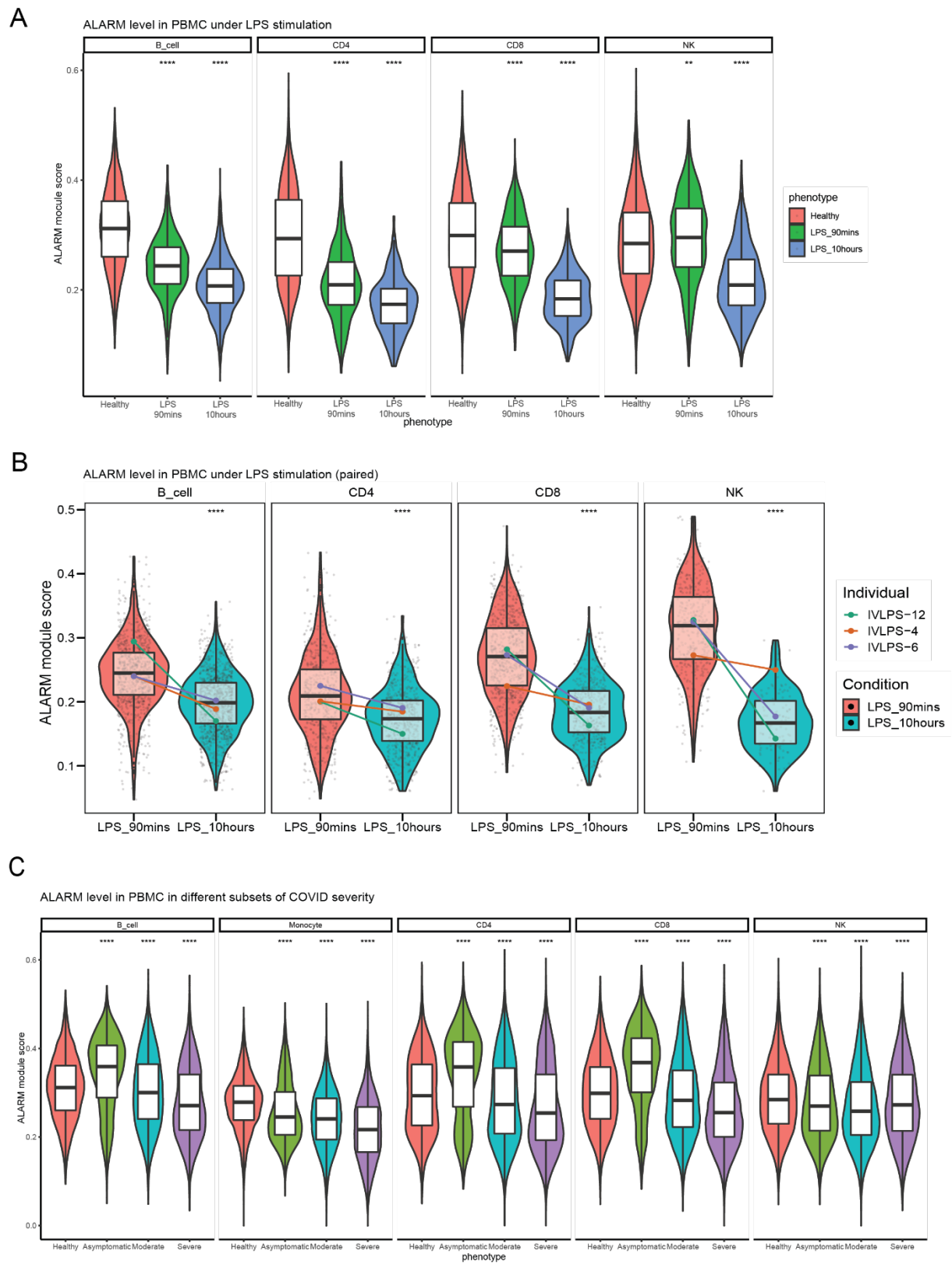
C) Violin plot of the ALARM module score across conditions (Data from Callemeyn et al., 2020) No\_ABMR=Stable patient, DSAnegABMR=Donor specific antibody negative ABMR, DSaposABMR= Donor specific antibody positive ABMR.

D) Heatmap showing relative expression of ALARM genes across graft biopsies classified by condition. (Data from Callemeyn et al., 2020)

E) Ligand-Receptor analysis between receptors genes identified in circulating immune cells and cytokines genes expressed in the allograft kidney tissue. Width of the arrow is proportional to the linkage enrichment score. Only top 20 differentially expressed receptors (PBMC scRNA-seq) and cytokines (Biopsy microarray) are shown.

F) Ligand-Receptor analysis between receptors genes identified in circulating immune cells and cytokines genes expressed in the allograft kidney tissue from Callemeyn et al 2020. Width of the arrow is proportional to the linkage enrichment score. Only top 20 differentially expressed receptors (PBMC scRNA-seq) and cytokines (Biopsy microarray) are shown.

Figure 5





**Figure 5. ALARM gene expression is altered in distinct immune-conditions**

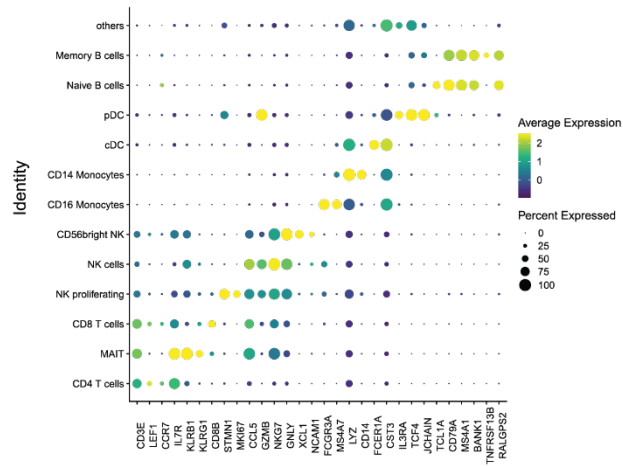
A) Violin plot of the ALARM module score across time points of after LPS injection in healthy volunteers. Data was obtained from Stephenson et al., 2021.

B) Violin plot at 90 min and 10h of ALARM module score in three individuals in different cell types.

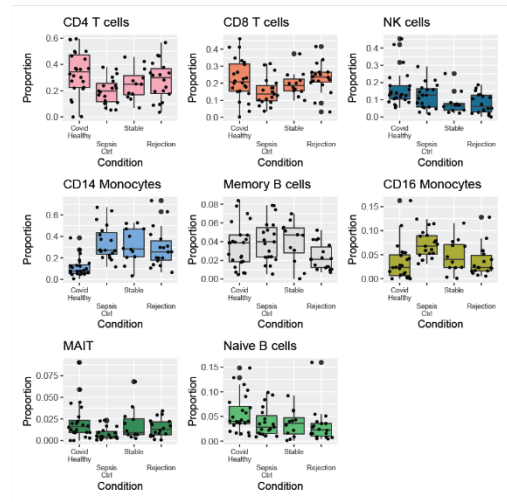
C) Violin plots of the ALARM module score in different cell-types, across distinct Covid-19 disease conditions. Either, healthy, asymptomatic Covid-19, moderate or severe disease state.

# Supplementary figure 1

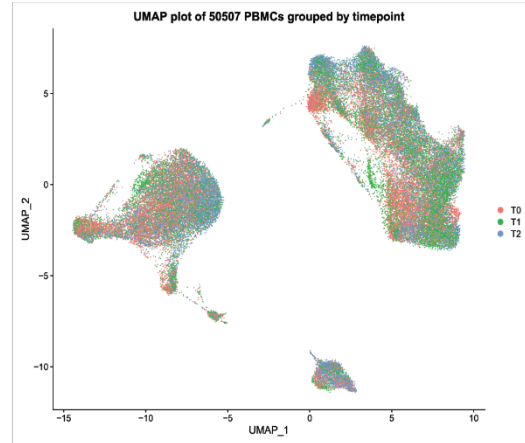
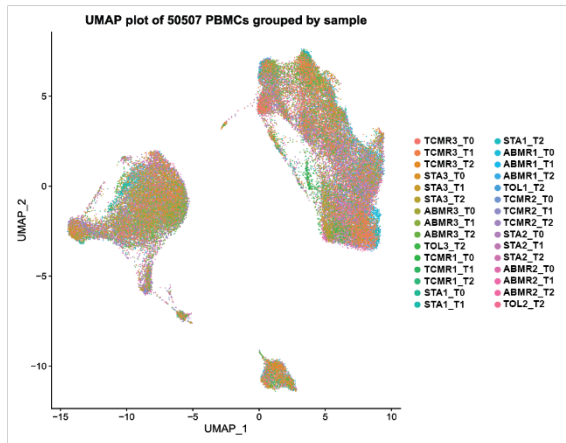
A



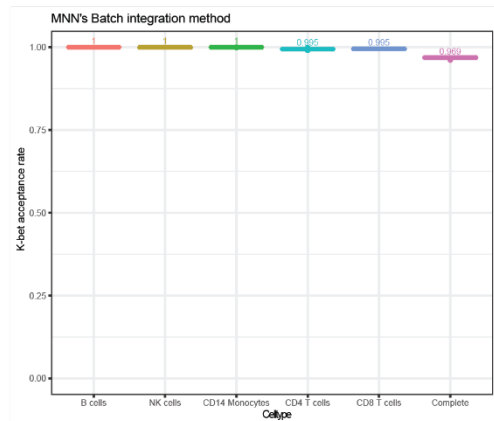
B



C



D



## **Supplementary figure 1. ScRNA-seq integration of a cohort of kidney allograft recipients**

- A) Expression profiles of cell-specific markers distinguish the PBMC populations. Average expression is the log-normalized expression average of the cells by cell type, size of the dots are associated to the fraction of cells of the cluster in which the gene is detected.
- B) Cell type proportion by cluster in the overall PBMC population. Comparison include Stable patients, Rejection (ABMR+TCMR) and two public datasets: healthy volunteers from Stephenson et al, 2021, and healthy volunteers from Reyes et al, 2020.
- C) UMAP projection showing the sample distribution, ABMR=Humoral rejection, TCMR=Cellular rejection, STA=Stable, TOL=Tolerant. Second UMAP shows the timepoint distribution, T0=Time point 0 (Graft), T1=Time point 1, T2=Time point 2.
- D) K-bet acceptance rate by cell types following the MNN batch integration. Complete k-bet acceptance rate was computed on the overall PBMC population.



## Supplementary figure 3

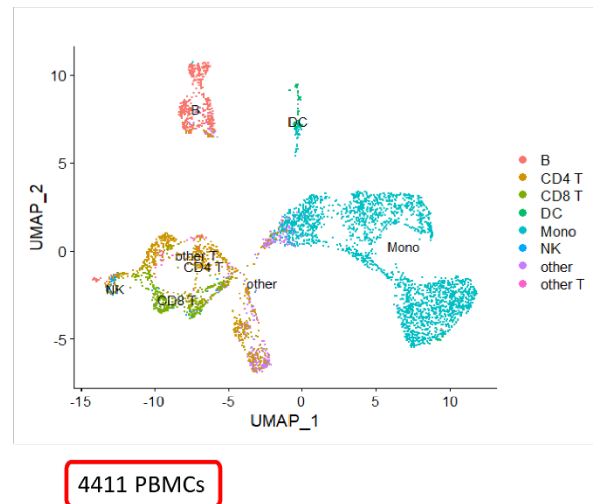
### A Immune cells annotation process in pig model



### B Differentially expressed cell type markers



### C Immune cells from pig PBMC samples



## Supplementary figure 3. Cellular characterization of pig PBMC by scRNA-seq analysis

A) Pig data were annotated using the Sscrofa reference genome and associated to their human homolog.

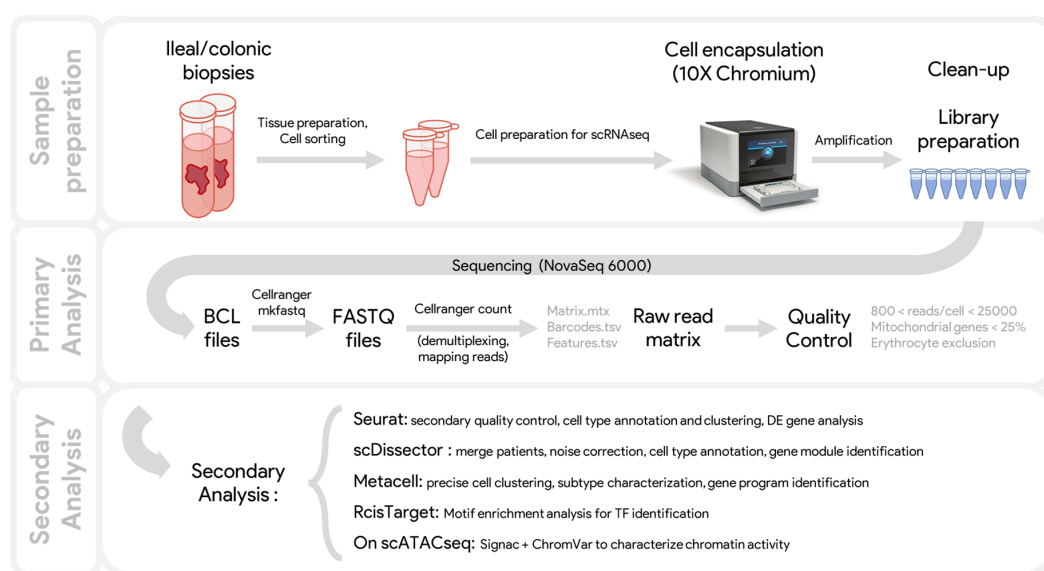
B) The expression profiles of cell type-specific markers is shown in the dotplot of relative expression by cell type. Size of the dots are associated to the fraction of cells of the cluster in which the gene is detected.

C) The UMAP embeds the 4 samples (D0, D2, D4, D6) with their corresponding cell type annotation.

## 2) La résistance aux traitements dans les MICI

Sur ce projet, j'ai contribué à l'analyse transcriptomique en scRNA-seq des échantillons et à l'intégration des patients inclus à Nantes et à New York, puis à l'intégration de jeux de données publics additionnels, pour identifier l'hétérogénéité cellulaire présente et les différentes voies fonctionnelles impliquées.

Sur un autre pan, j'ai aussi participé à l'analyse de données de microarray en bulk RNAseq de jeux de données publics sur des cohortes de test clinique de traitements contre les MICI pour compléter des résultats expérimentaux sur la barrière intestinale. Ces données vont être intégrées dans une étude collaborative en cours de finalisation portée par le Dr Saurabh Mehandru (ISMMS), qui vise à explorer les altérations moléculaires de la fraction épithéliale de la muqueuse inflammatoire de patients atteints de rectocolite hémorragique.



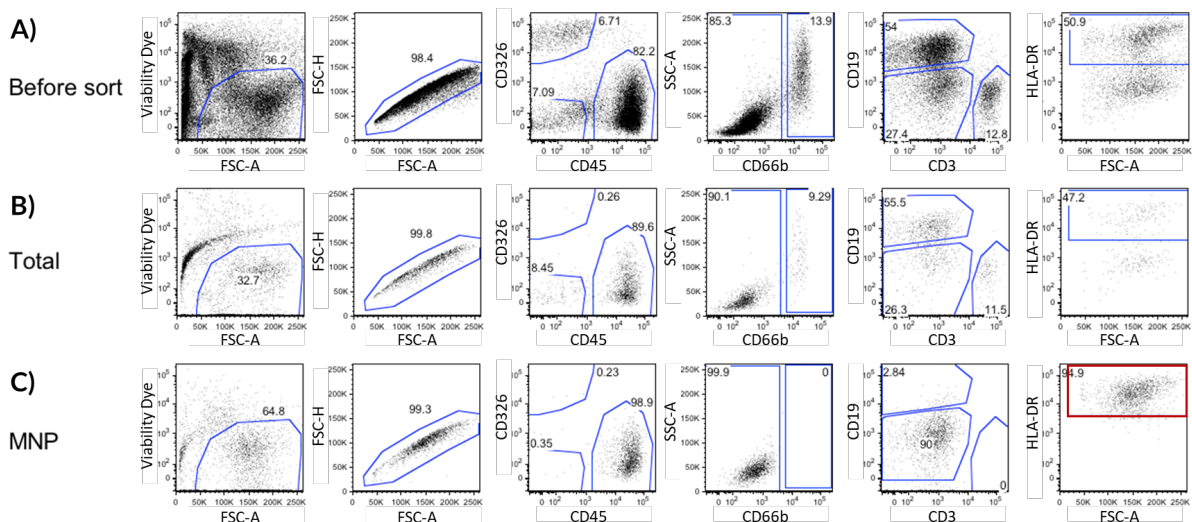
**Figure 17. Vue d'ensemble de la préparation d'échantillons de tissus intestinaux provenant de patients MICI de leur analyse en scRNA-seq.**

Les résections chirurgicales ou biopsies sont digérées, et les cellules de la lamina propria isolées par cytométrie en flux avant d'être préparées pour l'encapsulation et la création de librairie d'ADNc. Les fichiers de séquençages sont ensuite transformés en matrice d'expression contenant les UMI associés par gènes et par cellules.

L'identification d'une stratification de patients atteints de la maladie de Crohn selon l'hétérogénéité cellulaire (Jerome C. Martin et al. 2019) a pavé la voie des recherches de ma thèse. La corrélation de la réponse GIMATS avec une résistance au traitement par anticorps anti-TNF $\alpha$  était une première piste pour comprendre comment les différents profils de patients pouvaient expliquer cette résistance mais manquait de résolution pour permettre une exploration approfondie des programmes moléculaires des MNPs. De nouveaux patients ont donc été inclus à la cohorte en appliquant une méthode commune de préparation (**Figure 17**), pour atteindre un total de 40 patients. L'objectif de cette cohorte, appelée "Supergut", est de couvrir la plus grande hétérogénéité de profils possibles pour améliorer notre compréhension des processus

immunitaires dans les MICI. La cohorte est constituée de 26 patients atteints de MC, 7 atteints de RCH et 7 contrôles qui correspondent à des patients n'ayant pas développé de MICI mais qui ont nécessité une opération chirurgicale au niveau de l'intestin pour une autre raison (**Annexe 2**). Parmi ces patients, nous disposons pour certains de deux échantillons, l'un provenant des muqueuses inflammées, l'autre de parties non inflammées pour servir de contrôle interne. Grâce à ces échantillons non inflammés, on espère identifier les profils cellulaires spécifiquement recrutés en cas d'inflammation mais aussi identifier des potentielles populations cellulaires présentes à l'homéostasie qu'on ne retrouverait pas dans les échantillons de patients qui ne sont pas atteints de MICI.

Nous avons aussi séquencé des patients à Nantes, pour lesquels une fraction enrichie en MNP a été analysée, en plus des cellules totales de la lamina propria, avec pour objectif d'obtenir plus d'informations sur les populations de MNP. La stratégie de tri cellulaire pour capturer les MNP a été de trier parmi les cellules vivantes celles marquées par le CD45<sup>+</sup>, marqueur des leucocytes mais négatives pour CD326 (EpCAM), pour exclure les cellules épithéliales, CD66b pour exclure les granulocytes et CD3, CD19 pour exclure les LT et LB. Au sein de ces cellules, celles exprimant HLA-DR ont été triées (**Figure 18**).



**Figure 18. Stratégie de tri par cytométrie en flux.**

Des biopsies prélevées sur des pièces de résections chirurgicales d'iléons inflammatoires de patients atteints de maladie de Crohn ont été préparées pour générer des suspensions unicellulaires, de cellules de la lamina propria (A) à partir desquelles ont été triées les cellules totales, excluant les débris et cellules épithéliales (B), et les MNP (C).

En comparant les proportions de cellules au sein des MNP entre les échantillons triés et non triés (**Figure 19A**), on a retrouvé une grande similarité, qui s'est aussi vu au niveau transcriptomique avec des valeurs d'expressions comparables. Le tri cellulaire et l'enrichissement des fractions de MNP n'a donc pas altéré la distribution ni les profils d'expression des gènes sur nos cellules d'intérêt. De cette manière, nous avons un jeu de données plus riche de ces populations de MNP, là où pour les autres patients séquencés, ces cellules ne représentaient en moyenne que 10 à 20 % du total des cellules capturées.

La précédente étude de notre équipe (Jerome C. Martin et al. 2019) avait déjà identifié une ségrégation des macrophages inflammatoires autour de deux programmes anti-corrélés, l'un contenant de nombreux gènes de monocytes classiques du sang, et l'autre exprimant des gènes de macrophages matures.

Pour explorer ces données et faire les analyses primaires, j'ai utilisé l'outil scDissector (<https://github.com/effiken/scDissector>) développé lors de la précédente étude de notre équipe. Cet outil permet d'inclure les différents patients et de grouper les cellules en clusters pour pouvoir projeter ensuite les cellules de nouveaux patients sur un modèle prédéfini. Le contrôle qualité est effectué par un premier tri des cellules ne gardant que celles dont le nombre d'UMI est compris entre 800 et 25000, le pourcentage de gènes mitochondriaux inférieur à 25% et de gènes épithéliaux d'1%. L'intégration et la comparaison des patients dans cet outil reposent sur 2 principes. Le premier est une correction du bruit de fond, qui correspond à une contamination par des transcrits d'autres cellules détectées artificiellement dans la cellule observée. La corrélation entre le profil transcriptomique de l'échantillon, le nombre d'UMI de la cellule et son niveau de bruit de fond permet de déterminer la probabilité d'observer chaque gène en tant que bruit de fond et de corriger l'expression de chaque cellule avec cette information, de manière indépendante par échantillon (Paul et al. 2015). Le second principe repose sur un sous-échantillonnage du nombre d'UMI par cellule (par exemple 2000 UMIs par cellule), afin d'avoir une distribution semblable entre cellules sans passer par une normalisation classique.

À partir de la cohorte plus large que j'ai analysée, et notamment grâce aux deux échantillons enrichis en MNP, j'ai pu définir plus finement la composition en macrophages, en les décrivant en 4 populations (**Figure 19B**). On y trouve deux groupes de macrophages, avec d'un côté des cellules exprimant une signature "immature", plus proche des monocytes circulants du sang, avec *VCAN*, *S100A8*, *S100A9*, et d'un autre côté, des cellules exprimant une signature de macrophages "matures", avec les chaînes du système du complément (*C1QA*, *C1QB* ...) pour la clairance apoptotique ou *MAFB*, le facteur de transcription qui régule la différenciation des macrophages vers un programme anti-inflammatoire (**Figure 19D**). Parmi les macrophages inflammatoires, deux profils principaux se dégagent, les Macs1 qui sont présents dans tous les patients, et les Macs2, enrichis dans les patients ayant un score GIMATS élevé (**Figure 19B**). Ces Macs2 expriment fortement *TNF* et *IL1B*, et pourraient être à l'origine de l'accumulation excessive de  $TNF\alpha$  dans les tissus (Friedrich, Pohin, et Powrie 2019). Pour les macrophages matures, les Macs4 sont majoritaires dans les échantillons sains ou non inflammés mais deviennent très minoritaires dans les cas les plus sévères de MICI. Les Macs3 par contre sont une nouvelle population spécifique à un petit sous-groupe de patients GIMATS<sup>high</sup>, et sont caractérisés par un enrichissement en gènes de métalloprotéases *MMP9* et *MMP12*. Ces métalloprotéases n'étaient par le passé que décrites pour leur rôle de dégradation de la matrice extra-cellulaire (MEC) et donc pour un rôle de remodelage et de maintien des tissus (Dufour et al. 2018), mais elles ont aussi été décrites comme pro-fibrotique dans certains organes (H. Zhao et al. 2013). La comparaison du transcriptome des cellules Macs3 comparées aux autres macrophages à l'aide du package DESeq2 (Love, Huber, et Anders 2014) a montré un groupe de gènes différentiels spécifiques

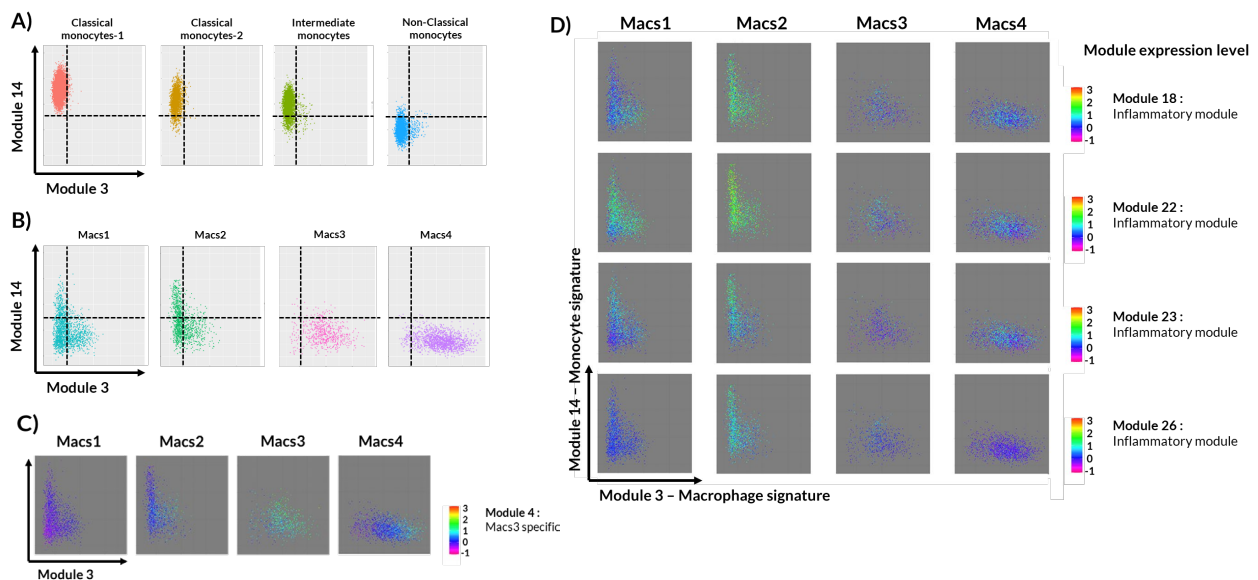




qu'en ne prenant que les macrophages issus des tissus, il est difficile d'interpréter les différences transcriptomiques liées uniquement à l'infiltration des monocytes dans le tissu, et celles liées à l'environnement inflammatoire et les interactions avec les cellules cibles. On a donc dans nos PBMC le moyen d'établir un profil de monocyte que l'on comparera aux profils de macrophages. À partir de là, j'ai sélectionné le sous-ensemble comprenant les MNP et les monocytes et j'ai cherché à identifier des modules de gènes pour décrire les programmes en jeu. On définit les modules de gènes comme des groupes de gènes fortement co-exprimés, en supposant que leurs similitudes de profil ne sont pas dues au hasard mais proviennent d'une régulation commune ou d'un processus biologique commun. Ces modules peuvent donc ressortir comme étant spécifiques à une lignée cellulaire ou à un cluster, mais de façon plus intéressante, ils permettent aussi d'identifier des voies et programmes cellulaires partagés entre différentes cellules. Dans mon analyse, j'ai d'abord sélectionné les gènes les plus variables pour réduire le poids des calculs et m'affranchir des gènes peu exprimés ou peu variables qui formeraient des modules incohérents biologiquement. Cette sélection repose sur la relation entre la moyenne d'expression (log-transformée pour réduire l'impact des valeurs élevées) et la variance des gènes, le tout suivant une régression locale dite de LOESS (locally estimated scatterplot smoothing) pour s'adapter à l'augmentation naturelle de la variance pour les gènes les plus exprimés. De cette manière j'identifie les gènes les plus variables de mon jeu de donnée et potentiellement pertinents dans la caractérisation des processus biologiques. J'ai ensuite calculé le score de corrélation (de Pearson) d'expression gène-gène entre tous les patients et appliqué un clustering hiérarchique pour regrouper les gènes les plus corrélés entre eux. On détermine ensuite un nombre de modules pour savoir où couper le dendrogramme hiérarchique, en essayant d'obtenir un nombre de modules suffisamment haut pour éviter de grouper des gènes qui n'auraient pas de rapports entre eux, mais pas trop élevé non plus pour éviter les modules ne contenant qu'un ou deux gènes. Pour cette analyse j'ai sélectionné 374 gènes variables distribués en 30 modules (**Annexe 1**). Parmi ces modules, deux modules signatures sont ressortis. Le Module 3 est caractérisé comme une signature de macrophages matures car il comprend les gènes de chaînes du complément (*C1QA*, *C1QB*, *C1QC*), *FOLR2* qui est un marqueur de macrophages résidents (Samaniego et al. 2020), *STAB1* un marqueur de résidence acquis durant l'organogenèse (Mass et al. 2016) ou encore *MRC1* (CD206). À l'opposé, on trouve le Module 14 qui lui contient des gènes fortement liés à une signature de monocytes classiques circulants, avec le marqueur de lignée *CD14*, le gène *LYZ* du lysozyme et les gènes de fixation du calcium *S100A8* et *S100A9* connues pour être abondantes dans les monocytes. J'ai ensuite projeté les cellules en utilisant ces deux signatures comme axes pour décrire les macrophages et monocytes (**Figure 20 A, B**). Parmi les monocytes du sang, on confirme que la signature mature (Module 3) n'est pas exprimée, tandis que le niveau de signature de monocyte (Module 14) l'est, et surtout enrichi dans les monocytes classiques. Chez les macrophages de l'intestin en revanche, l'expression des signatures diffère dans chaque cluster. La signature de macrophage est faiblement exprimée dans les Macs1 et Macs2, et anti-corrélée à la signature de monocyte. Les Macs3 et Macs4 sont eux enrichis en signature de

macrophage mais n'expriment pas la signature monocyte. Il semble donc que nos populations de macrophages inflammatoires sont un mélange de cellules fraîchement infiltrées et de macrophages immatures. En validation supplémentaire, les cellules dendritiques capturées sont négatives aux deux signatures.

Pour identifier ensuite les modules intéressants dans l'organisation des MNP, le score d'expression de chaque module est projeté en gardant la distribution selon les deux signatures précédentes (**Figure 20 C, D**). Le module 4 est caractérisé par une expression quasi exclusive au cluster de Macs 3 et comprend les gènes précédemment identifiés comme spécifiques à cette population, avec les métalloprotéases *MMP9*, *MMP12* associées au remodelage de la matrice extracellulaire et à la fibrose, les gènes associés au métabolisme des lipides (*APOE*, *APOC1*) et à la réponse au stress oxydatif (*APOE*). Ce module contient aussi les gènes de cathepsine B, C, D et L, qui sont des protéases retrouvées dans les lysosomes. Ces enzymes dégradent les protéines de manière non spécifiques et leur surexpression a déjà été décrite dans les macrophages des patients MICI (Menzel et al. 2006) avec pour conséquence une dégradation accrue des tissus riches en élastine. À terme ces macrophages participeraient par cette voie aux atteintes des muqueuses en condition pathologique chronique. Ces Macs 3 semblent donc partager à la fois des fonctions bénéfiques pour les tissus et d'autres plus pro-inflammatoires, qui pourraient être liés à une stimulation chronique excessive qui déréguleraient leurs fonctions (Vogel et al. 2022).



**Figure 20. Projection des sous-types de MNP de la cohorte SuperGut**

A) Projection des 4 sous-types de monocytes identifiés dans les échantillons de PBMC inclus dans la cohorte SuperGut. B) Projection des 4 sous-types de macrophages identifiés dans les échantillons de tissus inclus dans la cohorte SuperGut. C) À partir de la projection précédente, le score du module 4 est marqué par un gradient de couleur, bleu=module non exprimé, rouge=module fortement enrichi. D) Distribution des scores des 4 modules inflammatoires à travers les sous-types de macrophages.

Quatre modules (18, 22, 23, 26, voir **Table 2**) sont ressortis enrichis dans les deux clusters de macrophages inflammatoires, révélant des programmes inflammatoires exprimés différemment selon les profils cellulaires.

On observe que la signature inflammatoire classiquement associée aux macrophages inflammatoires dans leur ensemble est ici découpée parmi les modules et présentent des profils d'expression différents. Les chimiokines et *TNF* sont exprimés par le module 18, qui est spécifique à notre cluster de Macs2, le gène de l'IL-1 $\beta$  est exprimé dans le module 22 qui est enrichi dans la majorité des macrophages inflammatoires. L'expression de *NLRP3* dans le module 23 révèle une potentielle activation de l'inflammasome localisée dans les macrophages inflammatoires dont la signature monocyte est la plus forte, et donc probablement infiltrés récemment. Enfin, dans le module 26, on retrouve *CXCL1*, *IL6*, et *IL23*, qui sont exprimés dans les Macs 2 avec un score de macrophage mature le plus faible. On a donc par cette analyse une stratégie pour explorer plus finement l'expression des programmes moléculaires qui va plus loin que les définitions phénotypiques classiques. Cela souligne aussi que la caractérisation des cellules nécessite des outils plus adaptés à cette granularité, et c'est l'une des raisons pour laquelle j'ai utilisé la méthode des Metacells par la suite.

Gene composition of the 4 inflammatory modules	
Mod18	ABL2, PTGS2, PLEK, CCL20, NFKBIZ, CXCL8, CXCL3, CXCL2, NFKB1, SERPINB9, TNF, TNFAIP3, AC058791.1, INSIG1, NR4A3, BCL2A1, LITAF, KDM6B, CCL3, CCL4, CCL3L3, CCL4L2, ICAM1, MIR155HG
Mod22	G0S2, IL1B, EREG, ELL2, IER3, VEGFA, SLC16A10, SOD2, TIMP1, SRGN, MMP19, PHLDA1, NFKBIA, THBS1, PLAUR
Mod23	NLRP3, SLC11A1, NAMPT, SAT1, SLC25A37, C11orf96, NEAT1, SLC2A3, YBX3, VMP1, CEBPB
Mod26	IL1A, IL1RN, TNFAIP6, PTX3, CXCL1, SLC39A8, TNIP3, MIR3945HG, IL6, INHBA, GK, HSPA5, CD44, PPIF, IL23A, C15orf48, AQP9, SERPINB2
Gene composition of the Macs 3-specific module	
Mod4	FUCA1, FCER1G, FCGR3A, GLUL, GPNMB, CTSS, CTSL, CTSD, MS4A7, CTSC, MMP12, CD63, RNASE1, LGALS2, MT2A, MT1G, MT1X, VMO1, CCL18, MMP9, CFD, ACP5, APOE, APOC1, HMOX1

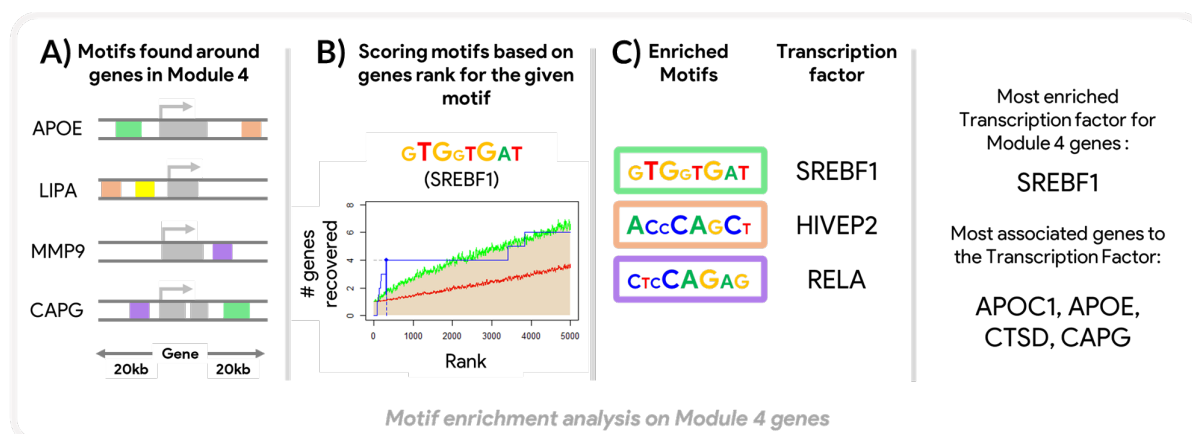
**Table 2. Liste des gènes impliqués dans les modules d'intérêt.**

Les différents programmes moléculaires identifiés sur la base de corrélations d'expression, suggèrent que les gènes les définissant pourraient participer de mécanismes de régulation similaires, partageant l'implication commune de facteurs de transcription (FT). En l'absence de données épigénétiques sur notre jeu de données, il est néanmoins possible d'utiliser des outils permettant d'estimer les FT les plus probablement impliqués. C'est notamment le cas du package RcisTarget (Aibar et al. 2017), qui utilise une liste de gènes qu'on lui fournit pour identifier les motifs de fixation des FT sur-représentés. La recherche de motifs s'effectue autour du site d'initiation de la transcription (TSS) à l'aide d'un des deux répertoires de RcisTarget : celui proximal des motifs retrouvés jusqu'à une distance en paire de bases de 500bp en amont du TSS des gènes, et celui plus distal couvrant les 20kb autour du TSS (soit 10kb de chaque côté). En croisant la liste de gène et tous les motifs retrouvés proches des TSS de ces gènes, on identifie les motifs les plus enrichis pour le module, que

l'on annote ensuite avec les FT qui se fixent à ces motifs (**Figure 21**). Une dernière étape va réattribuer à chaque FT identifié les gènes cibles candidats parmi ceux de la liste de gènes dans le cas où ils ont un site de fixation à ce FT.

Les FT associés à la signature monocyte (Module 14) signatures sont *CEBPA* et *CEBPB*, deux FT importants dans le développement des lignées myéloïdes et pour la signature macrophage (Module 3), c'est le FT *SPI1* (PU.1) qui est le plus enrichi. PU.1 est un FT exprimé durant la différenciation en macrophage (Medzhitov et Horng 2009). Dans le module 18, les FT les plus enrichis correspondent à trois gènes codants des sous-unités de NF- $\kappa$ B : *REL* (Rel-c), *RELA* (p65) et *NFKB1* (p50/p105). La voie NF- $\kappa$ B est importante dans l'activité immunitaire des macrophages et c-Rel participe fortement à la transcription d'*IL12B* (IL-12p40) (Dorrington et Fraser 2019), ce qui est cohérent avec le fait de retrouver cette signature dans les Macs2 inflammatoires enrichis dans les patients GIMATS<sup>high</sup>, sachant que les patients Crohn ont une sur-expression d'IL-12p40.

Associés au Module 22, on retrouve *CEBPB* et *CEBPA*, similairement à la signature de monocyte, et étant donné que le Module 22 est exprimé dans la majorité des Macs1 et 2, on a probablement là une voie assez générale d'induction de l'inflammation. Enfin, le module 4 spécifique aux macrophages Macs 3 présente un enrichissement de *SREBF1* (**Figure 21**), codant pour SREBP1 qui est effectivement décrit comme régulateur de gènes liés au métabolisme des lipides et du cholestérol et pourrait jouer un rôle dans la résolution via la production d'acides gras anti-inflammatoires (Oishi et al. 2017)



**Figure 21. Résumé de l'analyse d'enrichissement de motifs avec RcisTarget**

A) Sélection des motifs de fixation des FT en amont et aval des TSS de gènes appartenant au module d'intérêt. B) Calcul du score d'enrichissement de chaque motif. C) Classement des motifs les plus enrichis et de leur FT associé

Une fois ces premiers résultats acquis, l'objectif suivant a été d'augmenter l'échelle de l'étude, pour chercher à intégrer le plus grand nombre de cas possibles et couvrir une plus grande diversité de cas. Cependant la quantité de données a été un facteur limitant dans le choix de la méthode puisque l'outil scDissector ne permet pas de charger un trop grand nombre de cellules en simultanément.

J'ai alors utilisé la méthode des Metacells (Ben-Kiki et al. 2022, 2) comme décrit précédemment pour réduire la charge de calcul et surtout tenter d'obtenir des profils transcriptomiques plus précis. Étant donné que l'on repart de données brutes non corrigées et sans sous-échantillonnage, le batch effect observé entre tous les patients inclus ne permet pas de conduire des analyses directement.

Pour réduire les différences liées entre les différents patients et les différents laboratoires, il a été nécessaire d'identifier les cellules contaminantes ainsi que les gènes induisant la plus grande variabilité. Cette première étude se restreint à l'inclusion de patients des centres de Nantes et de Mount Sinaï à notre disposition. Comme dans une analyse en scRNA-seq classique, j'ai d'abord exclu les cellules basées sur le pourcentage de gènes mitochondriaux, car au-dessus de 20% de gènes mitochondriaux, on estime que la qualité ne sera pas suffisante ou que les transcrits captés proviennent de cellules en apoptose. On exclut aussi les cellules exprimant plus de 10% de transcrits de gènes d'hémoglobine (*HBB*, *HBA1*, *HBA2*) qui sont soit issus d'une hématie, soit d'un doublet entre un érythrocyte et une autre cellule. Les cellules épithéliales ont ensuite été identifiées et exclues puisque l'analyse se concentre sur les cellules stromales et immunitaires des tissus intestinaux, et que la digestion des cellules endommage fortement les cellules épithéliales. Pour identifier les cellules épithéliales, la liste de gènes utilisée provient de Martin et al. (2019), et les cellules exprimant plus de 10% de ces gènes sont exclues de l'analyse. Liste de gènes épithéliaux : *PLA2G2A*, *CLCA1*, *REG4*, *S100A14*, *ITLN1*, *ELF3*, *PIGR*, *EPCAM*, *REG1B*, *REG1A*, *REG3A*, *FABP1*, *RBP2*, *SST*, *FABP2*, *SPINK1*, *FABP6*, *AGR2*, *AGR3*, *CLDN3*, *CLDN4*, *DEFA6*, *DEFA5*, *SPINK4*, *ALDOB*, *LCN2*, *MUC2*, *KRT8*, *KRT18*, *TSPAN8*, *OLFM4*, *GPX2*, *IFI27*, *PHGR1*, *MT1G*, *CLDN7*, *KRT19*, *FXD3*, *LGALS4*, *FCGBP*, *TFF3*, *TFF1*.

Suite à la sélection des cellules d'intérêt, l'objet obtenu montre un effet de batch important puisque les cellules des patients se distribuent dans des clusters de metacells indépendants entre les deux centres d'étude au lieu d'être regroupées ensembles (**Figure 22A**). Il est donc nécessaire d'identifier les variations transcriptomiques liées aux différences techniques et biologiques non pertinentes entre les échantillons dans notre étude. Ces variations peuvent être ignorées lors de l'intégration des cellules pour former les metacells, en fournissant une liste de gènes à ignorer à l'étape du calcul des graphes. Ce processus de nettoyage permettra d'améliorer la qualité des metacells sans pour autant perdre l'information transcriptomique pour les étapes suivantes.

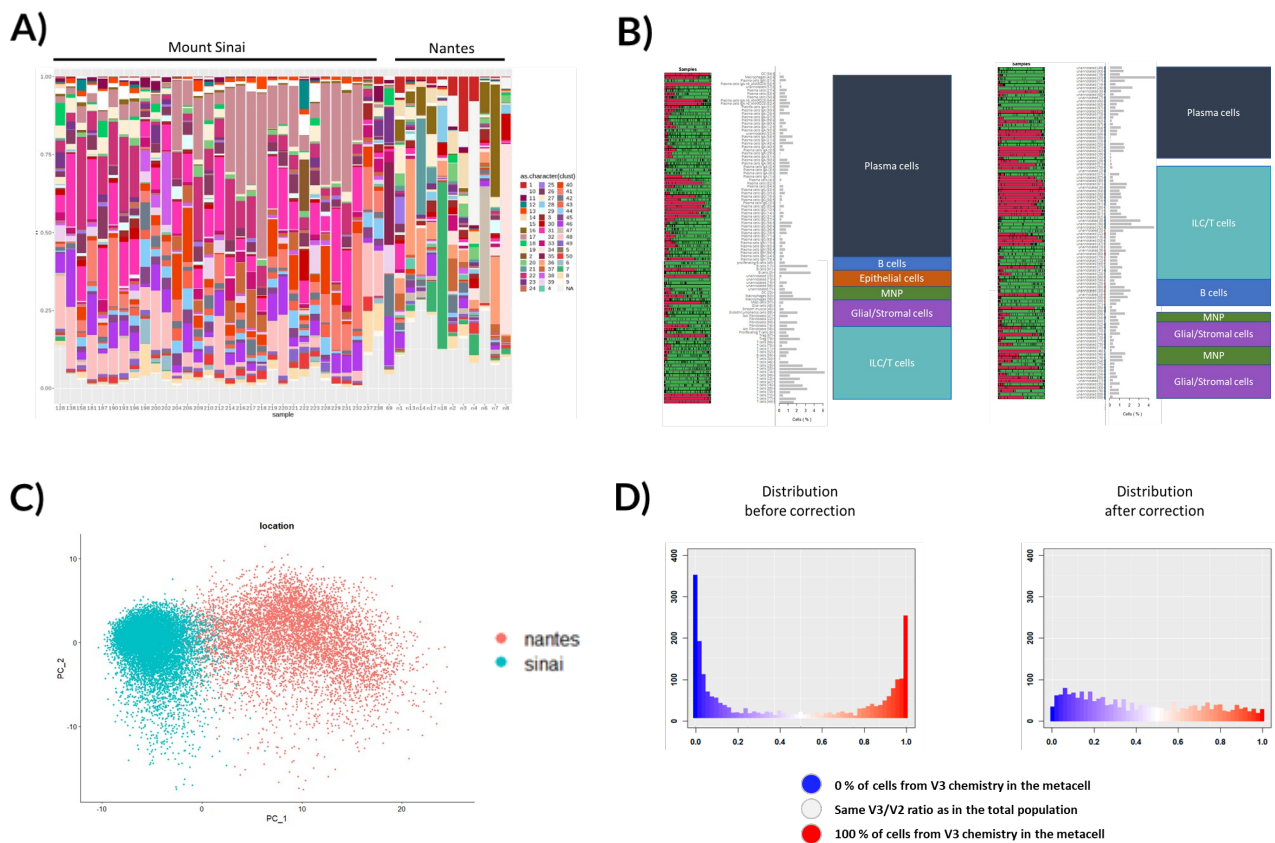
Parmi les gènes que je détecte comme étant bruyants ou spécifiques aux patients, et donc causant un biais d'origine biologique mais problématique pour l'intégration des différents patients, il y a en premier lieu *MALAT1* et *NEAT1*, deux ARNs longs non codants (ARNInc) abondamment exprimés dans les cellules qui bien que déjà décrits dans des pathologies comme le cancer (X. Zhang, Hamblin, et Yin 2017, 1) sont ici beaucoup trop variables entre patients pour être intéressants à garder dans la construction des metacells. Le gène *XIST* est aussi exclu de la construction puisque c'est un ARNInc associé à l'inactivation de l'un des deux chromosomes X chez les femmes, et donc exclusivement exprimé chez nos patientes femmes. J'ai aussi exclu de la construction des metacells le gène *JCHAIN* (exprimé par les plasmocytes) et les



gènes HLA de classe I et II. En plus de ces gènes on ajoute la liste comprenant *DUSP10*, *CSRNP1*, *IRF1*, *HSP90AB1*, *DNAJA1*, *TUBB4B*, *DDIT4*, *RGCC*, *SOCS1*, *UBB*, *PMAIP1*, *SERTAD1*, *MYADM*, *HSPA6*, *HSPE1*, *HSPA1A*, *HSPA1B*, *BAG3*, *HSPH1*, *HSP90AA1*, *DNAJB1*, *PIK3R3*, qui sont des gènes précédemment identifiés comme “broyants”, liés au stress et inégalement capturés entre patients (Jerome C. Martin et al. 2019). Avec cette sélection, j’ai construit un nouveau modèle de Metacell, et recalculé la distribution en 100 clusters. Néanmoins, la distribution des metacells dans cet objet se trouvait fortement enrichie en clusters correspondants aux plasmocytes (plus de 50%) alors qu’ils représentent moins d’un quart du nombre de metacells totaux. Après analyse des gènes différentiels, il est apparu que les gènes codant les chaînes variable et constantes des Ig introduisaient une forte variabilité, ils ont donc été exclus pour retrouver une distribution plus représentative du nombre de metacells dans chaque lignées (**Figure 22B**). Malgré une distribution corrigée des types cellulaires, les cellules provenant des 2 laboratoires ne s’intègrent pas ensemble pour former les metacells, et ce biais est le plus fort puisqu’il ressort en tant que 1ère composante principale dans la PCA avec une séparation claire des cellules venant du centre de Nantes et de celles du Mount Sinai (**Figure 22C**).

Après ce constat, j’ai intégré des patients du centre de Mount Sinai qui avaient été préparés avec le même kit V3 de 10X Genomics et exploré l’expression des gènes en séparant les patients selon l’origine et le kit utilisé. Le résultat montre une différence d’expression uniquement due à la version du kit utilisé puisque les quelques patients V3 de Mount Sinai affichaient des profils d’expression similaires aux patients de Nantes fait en V3. Cependant, malgré ce biais, l’objet metacell créé après toutes mes corrections montre une distribution plus homogène des cellules au sein des metacells. L’objet final du modèle Supergut en Metacells a permis d’intégrer 189401 cellules provenant de 43 patients en 1685 metacells représentatives de l’hétérogénéité cellulaire des tissus intestinaux. Au regard de la distribution des metacells, à l’origine, on observe que plus de 600 de ces metacells étaient constitués uniquement de cellules V2 (n=357) ou de cellules V3 (n=254), tandis qu’après l’affinage du calcul des metacells, moins de 70 metacells le sont (**Figure 22D**). Toutefois, la différence de distribution des gènes capturés selon la version du kit expérimental utilisé est une variable qu’il sera nécessaire de corriger lors de l’analyse secondaire pour garantir des comparaisons équilibrées.

Une fois le modèle Supergut en Metacells établi, je l’ai enrichi en intégrant un jeu de données publique (Friedrich, Pohin, et Powrie 2019) comprenant 5 échantillons contrôles, 8 échantillons de tissus inflammés en RCH, et 4 échantillons de tissu en zone non inflammée en RCH. Avec l’intégration de nouveaux patients séquencés en local, le modèle actuel contient les échantillons de 64 patients : 7 du centre nantais, 40 du centre de Mount Sinai et 17 de données publiques.

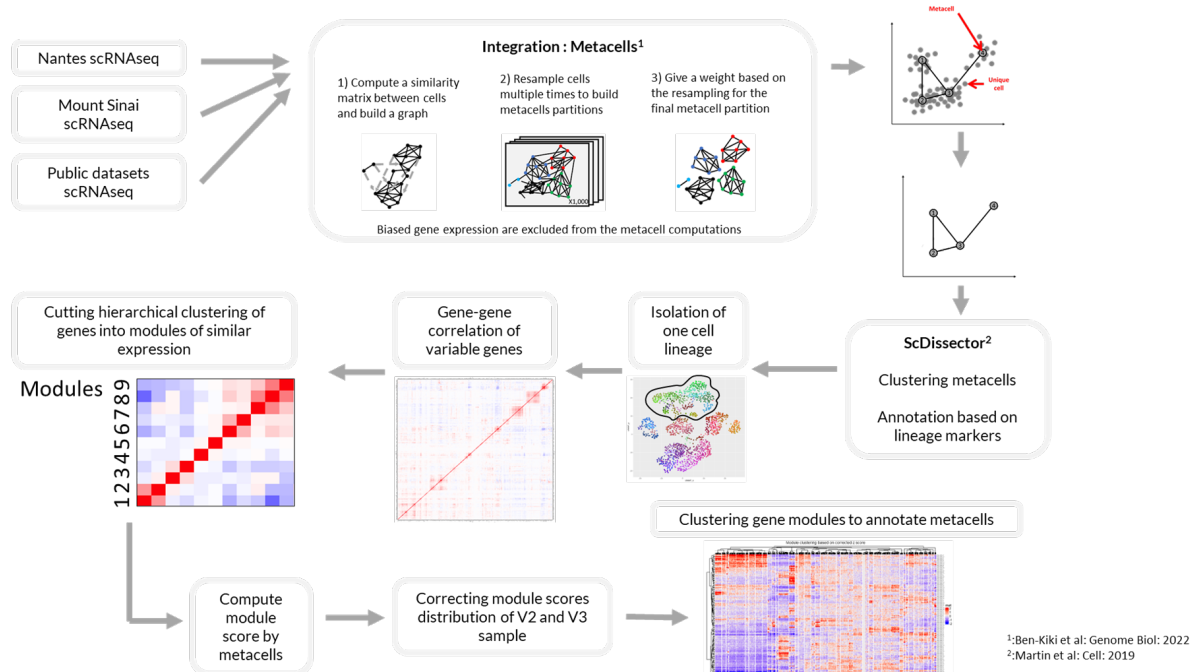


**Figure 22. Identification des biais influençant la construction des metacells.**

A) Distribution des cellules entre les clusters de metacells selon leur échantillon d'origine. Basé sur le premier modèle de Metacells sans exclusion de gènes. B) Proportions de chaque lignée cellulaire parmi les 100 clusters identifiés par Metacells avant (1) et après (2) exclusion des gènes liés aux Ig. La barre de gauche indique la proportion de patients venant des centres de Nantes ou New York, le graphique central représente le nombre de cellules par cluster et le groupement à droite résume les lignées présentes dans ces clusters C) PCA plot des cellules annotées comme LT CD8<sup>+</sup> dans le dataset supergut. D) Distribution cumulée du nombre de metacells selon la proportion de cellules provenant d'une expérience avec le kit V3 qui les constituent, avant la correction (gauche), et après correction (droite).

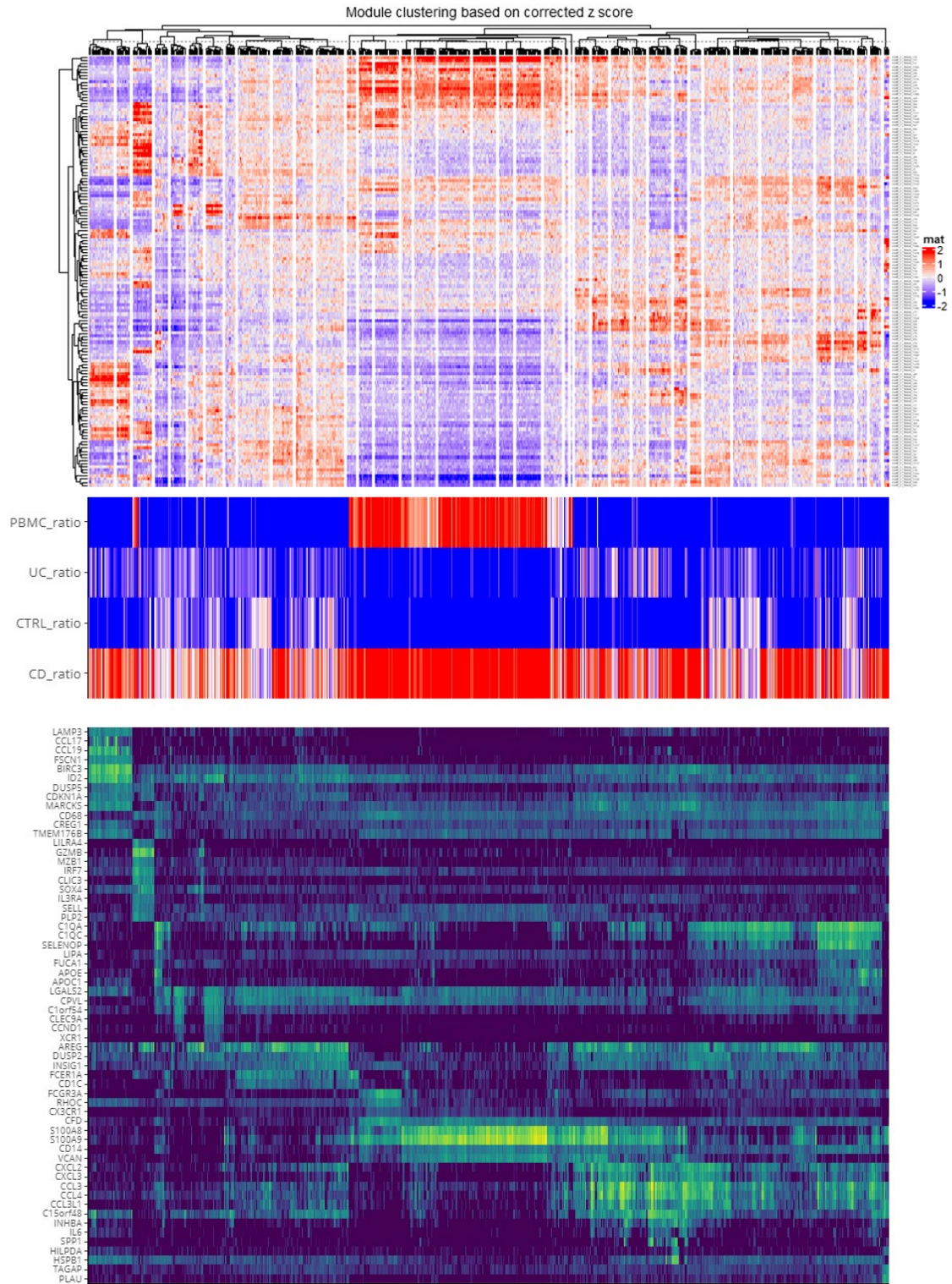
La construction du modèle désormais appelé "Ultragut" par la méthode Metacell suit le même déroulé que le modèle précédent et sera amené à s'enrichir avec des nouveaux patients (**Figure 23**). L'objet Ultragut a été subdivisé pour garder d'un côté les cellules stromales et de l'autre les MNP qui sont nos deux populations d'intérêt. Sur le groupe des metacells de MNP, j'ai effectué une recherche de 150 modules basée sur l'expression des metacells (la liste est en **Annexe 1**). Parmi ces modules on retrouve nos deux modules signature (le module 14 est la signature de macrophage mature, le module 24 est la signature de monocyte).





**Figure 23. Sommaire du pipeline d'analyse final appliqué aux cohortes de MICI**

La caractérisation de ce modèle "Ultragut" est encore en cours, mais cette analyse m'a permis de coder mon propre outil d'exploration des données de Metacell pour faciliter la comparaison des modules et de leur distribution en s'affranchissant d'une définition des clusters trop rigides (**Figure 24**).



**Figure 24. Heatmap de l'enrichissement relatif des modules par metacells**

A) Dans cette heatmap, chaque colonne représente un metacell de MNP et chaque ligne représente l'expression relative d'un des 150 modules de gènes. Les metacells sont triés selon leur profil d'enrichissement en module par clustering hiérarchique. B) Ratio de cellules provenant d'échantillons de PBMC par metacell suivi du ratio selon le phénotype. Rouge = 100% de cellules constituant le metacell, Bleu = 0% C) Heatmap d'expression de marqueurs spécifiques aux différentes sous-populations de MNP.

En parallèle, j'ai aussi regardé l'association du module GIMATS et la réponse au traitement anti-IL12/23p40 à partir de données de microarray sur des biopsies de colon dans des patients RCH. Les précédents travaux de mon équipe s'étaient intéressés au rôle des cytokines IL-17A et IL-22 dans la rectocolite. Ces deux cytokines ont un rôle protecteur de la barrière intestinale à l'homéostasie en induisant l'expression de peptides anti-microbiens, en stimulant la prolifération de cellules épithéliales et par la production des protéines des jonctions serrées (Valeri et Raffatellu 2016). Suite à la reconnaissance de pathogènes, les DC et macrophages sécrètent l'IL-23 qui va induire l'immunité de type 3 : les cellules Th17 et les ILC3 ont un récepteur à l'IL-23 (IL-23R) et vont sécréter de l'IL-17A et IL-22 en réponse.

Cependant, dans les MICI il a aussi été démontré que la voie IL-23/IL-17 pouvait entraîner l'augmentation de la sécrétion de cytokine pro-inflammatoires (Schmitt, Neurath, et Atreya 2021) et IL-22 seul ou en combinaison avec l'IL-17A, participe aussi à l'inflammation en induisant un stress du réticulum endoplasmique des cellules épithéliales et leur apoptose (Powell et al. 2020). Bien que le blocage de l'IL23 ou de son récepteur aient montré une rémission dans une partie de patients atteints de MICI (Feagan et al. 2016; Sands et al. 2017; Sands, Sandborn, et al. 2019), le blocage seul de la voie IL17 aggrave la sévérité de la maladie (Hueber et al. 2012; Targan et al. 2016) et l'effet du blocage de l'IL-22 seul n'a pas encore été montré cliniquement. On notera aussi que durant les MICI, un inhibiteur de l'IL-22, l'IL-22BP (pour Binding protein), est enrichi dans la muqueuse intestinale et diminue les fonctions protectrices de l'IL-22 (J. C. Martin et al. 2016).

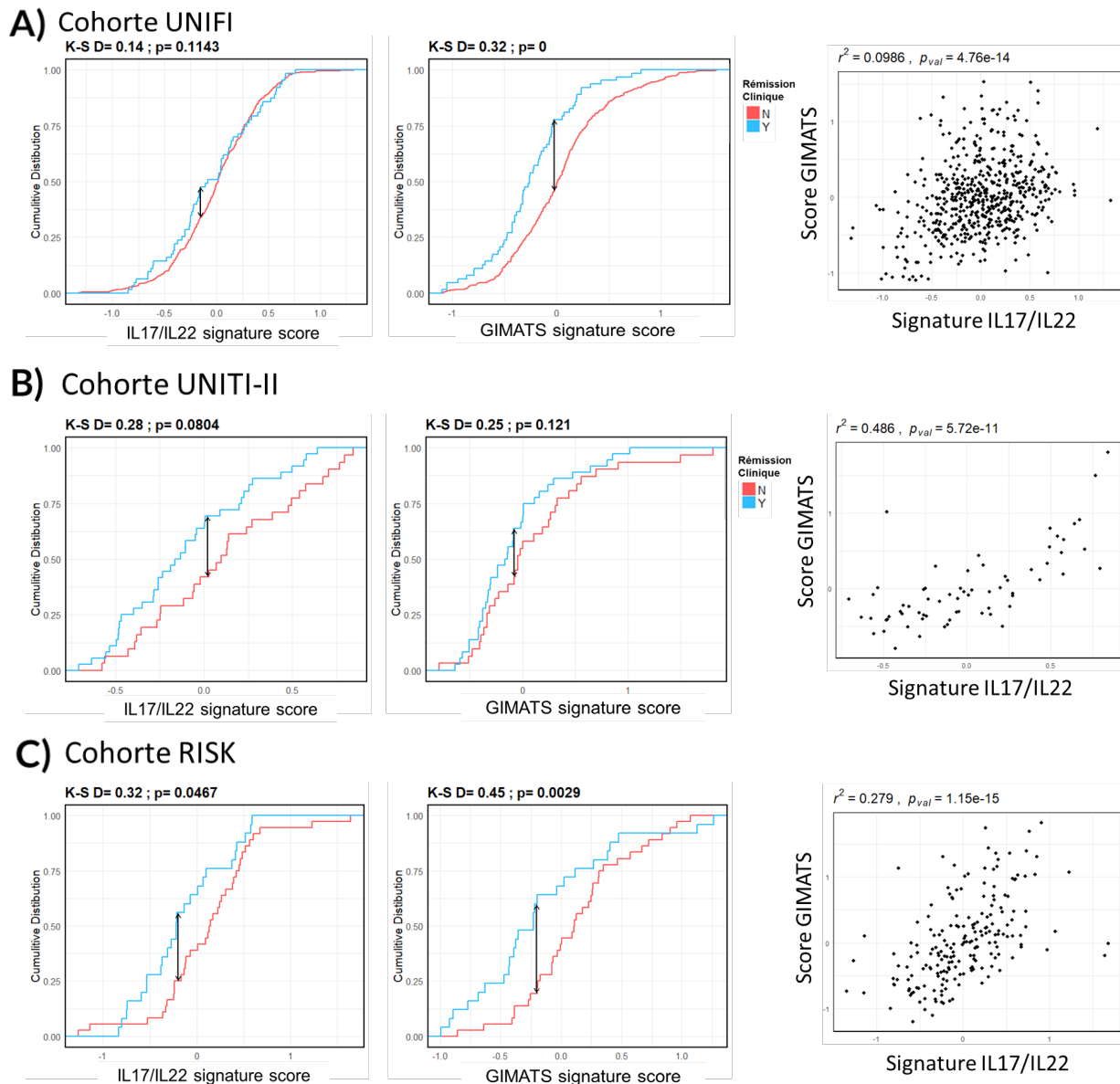
Pour tenter de suivre l'action de l'IL-17 et l'IL-22 dans l'intestin, mon équipe a utilisé des biopsies de colon pour observer les différences induites par la stimulation par IL-17 seul, IL-22 seul ou la combinaison de l'IL-17 et IL-22 (**Figure 25 A**). L'expression des gènes a ensuite été observée en RNA-seq. On a observé que les cytokines seules n'induisent que peu de différences transcriptomiques (3 et 31 gènes différentiels pour la stimulation par IL-22 et IL-17 respectivement), tandis que la combinaison des deux induit une synergie et nous permet d'établir une signature IL-17/IL-22 de 131 gènes (**Figure 25 B**). Les gènes induits dans cette signature sont des gènes participant au fonctionnement de la barrière épithéliale (induction de peptides anti-microbiens, chimiokines, métabolisme oxydatif...).

Pour voir l'effet de cette signature dans les patients atteints de MICI, j'ai utilisé les données de microarray de la cohorte UNIFI (Argmann et al. 2022) contenant le séquençage de biopsies de colon de 550 patients divisés en 18 volontaires sains, 186 patients RCH ayant reçu un placebo et 364 patients RCH ayant reçu de l'ustekinumab. L'étude de l'expression des gènes de la signature IL17/IL22 montre une stratification des patients, avec une partie des patients qui est enrichie dans cette signature (**Figure 25 C**) mais ne semble pas être enrichie dans un sous-type de patients particuliers.





Pour vérifier ça, j'ai comparé les distributions des patients répondeurs ou non en fonction de la signature IL17/IL22 (**Figure 26 A**). Le test ne relève pas de différence dans la distribution, et donc pas d'association entre cette signature et la réponse au traitement.



**Figure 26. Association des signatures IL17/IL22 et GIMATS dans les cohortes RISK, UNITI-II et UNIFI.**

Distribution cumulée du score IL17/22 (gauche) ou GIMATS (centre) dans les répondeurs et non répondeurs. Corrélation entre le score GIMATS et le score IL17/IL22 (droite)

A) Cohorte UNIFI : Les non-répondeurs ont un score GIMATS plus élevé. Corrélation des deux scores :  $r^2=0.009$

A) Cohorte UNITI-II : Les non-répondeurs ont un score IL17/IL22 et GIMATS plus élevé. Corrélation des deux scores :  $r^2=0.486$

A) Cohorte RISK : Les non-répondeurs ont un score IL17/IL22 et GIMATS plus élevé. Corrélation des deux scores :  $r^2=0.279$

Cependant avec le score de gènes enrichis dans la signature GIMATS<sup>high</sup>, on observe une différence (Kolmogorov-Smirnov D = 0.35; p < 0.0001), le test est significatif donc les patients en rémission et ceux qui ne répondent pas suivent une distribution différente, la réponse GIMATS est associée à la résistance à l'ustekinumab dans la RCH. J'ai étendu mon analyse à deux cohortes de patients Crohn à atteinte iléale, la cohorte RISK (Haberman et al. 2014) testant l'effet de l'anti-TNF, et la cohorte UNIT-II (Feagan et al. 2016) testant l'effet de l'ustekinumab. J'ai pu à nouveau observer que les gènes de la signature IL17/IL22 étaient bien corrélés entre eux et étaient enrichis différemment entre patients. Dans la cohorte RISK, les signatures IL17/IL22 et GIMATS sont peu corrélées ( $r^2=0.279$ ) mais toutes deux ressortent significativement associées à une résistance au traitement anti-TNF (**Figure 26 C**). Dans la cohorte UNIT-II, les deux signatures sont corrélées ( $r^2=0.475$ ) mais seule la signature GIMATS est associée à la résistance à l'ustekinumab (**Figure 26 B**).

On a donc deux signatures corrélées dans la maladie de Crohn, mais pas dans la RCH, et leurs différences d'association avec une résistance aux traitements montre un fonctionnement distinct entre l'inflammation chronique iléale et celle du colon.

### 3) Projets annexes

Durant ma thèse j'ai pu mettre à disposition mes compétences d'analyse de données de séquençage à haut débit au profit de divers projets.

Projet Abidi (Abidi et al. 2020, **Annexe 4**): J'ai effectué les analyses primaires et secondaires sur les données de scRNA-seq d'ILC de rat afin d'identifier les populations présentes et les marqueurs d'intérêt pour valider les résultats expérimentaux

Projet Roquilly : J'ai participé à l'analyse primaire de scRNA-seq et la caractérisation des populations de cellules immunitaires présentes dans les poumons de souris sur plusieurs jours suivant une septicémie

Projet Chapelle : J'ai participé à l'analyse primaire des échantillons de cancers gastriques séquencés en scRNA-seq, en plus d'un rôle de conseil sur la suite de l'analyse bio-informatique.

Projet Fantou : (Fantou et al. 2022) : J'ai participé aux analyses de RNA-seq et de statistiques sur les données cliniques. Le projet aboutira à un second papier, qui est en cours de rédaction et pour lequel j'ai montré les résultats sur la signature IL17/IL22 et son association à la MC et la RCH dans différentes cohortes et son association aux traitements.

# Discussion

L'arrivée de la première machine d'encapsulation à haut débit pour produire des expériences de scRNA-seq capables d'atteindre des milliers de cellules concordait avec les débuts de mon projet. J'ai été l'un des tout premiers utilisateurs à Nantes à appliquer cette technique pour en faire bénéficier un projet médical concret et je me suis donc formé en continu pendant ces trois ans sur les bonnes pratiques d'analyse. Cette technique est récente et évolue en continu, j'ai donc suivi en parallèle de mes recherches l'évolution des méthodes, qu'elles couvrent la préparation des cellules, le contrôle qualité, l'identification des cellules, ou des analyses secondaires comme la recherche de différences entre échantillons et la façon d'intégrer ses données. J'ai appliqué dans mes travaux mes propres solutions ou recherché les outils les plus adaptés aux différents usages du scRNA-seq pour tenter de répondre au mieux aux questions biologiques de mes deux projets principaux.

## Partie Transplantation

Dans la transplantation rénale, la survenue d'un rejet chronique est souvent associée à une prise en charge trop tardive qui ne permet plus de ralentir son évolution, que ça soit dû à l'espacement du suivi sur le long terme ou à cause d'un rejet infraclinique qui n'a pas été détecté par le suivi au long terme. Le rejet infraclinique n'a pas de répercussions sur la fonction rénale (les niveaux de créatinine sont stables) mais une dégradation des tissus peut être observée en histologie. Dans ces cas, le traitement ne peut pas être adapté à temps pour améliorer la survie du greffon. Les rejets de greffe sont associés à une infiltration monocytaire et lymphocytaire et nous avons donc choisi une approche exploratoire des cellules immunitaires circulantes pour tenter de déterminer si les signaux à l'origine de cette infiltration étaient détectables dans le sang. Dans mon travail sur une cohorte rétrospective de patients transplantés, c'est la recherche d'une signature dans le sang à des fins de prévention qui m'a conduit à identifier le module de gènes ALARM associé au rejet et par extension, à l'inflammation des tissus, aussi retrouvée dans diverses maladies immunitaires.

Pour caractériser une population cellulaire, il est crucial de définir l'hétérogénéité cellulaire présente et dans notre cas, nous cherchons des différences de distribution des types cellulaires ou de profils transcriptomiques au sein du compartiment des PBMC inter- et intra-patient, au cours du temps. Cependant, il faut garder à l'esprit qu'il existe déjà une certaine variabilité au niveau basal, que ça soit des différences inter-individuelles par leur historique immunitaire ou même pour un même individu au cours du temps (Kashima et al. 2022). J'ai alors mis en place un "pipeline" d'analyse commençant par la préparation des échantillons et l'optimisation de l'expérience de scRNA-seq. À la suite du séquençage, j'effectue les différentes étapes de démultiplexage et de contrôle qualité pour garantir la fiabilité des données, puis l'intégration des données et l'annotation précédemment manuelle et désormais semi-supervisée pour avoir une annotation des cellules consistante entre les échantillons.

Et enfin j'ai appliqué différentes méthodes pour décortiquer les profils transcriptomiques.

Dans le premier groupe de patients que nous avons analysé avant la survenue du COVID-19 pour la mise au point de notre étude (et qui n'est pas montré dans le papier), l'analyse de l'hétérogénéité cellulaire est ce qui nous a empêché de les intégrer. En effet, la fraction des monocytes était quasiment absente dans nos données de scRNA-seq et en conflit avec les numérations cellulaires cliniques des mêmes patients. Cela souligne d'ailleurs l'une des problématiques soulevées avec la technique : comment peut-on être sûr que nos observations correspondent à la réalité ? De nombreuses étapes peuvent en effet altérer le résultat : la dissociation dans le cas de tissus, la décongélation, et dans notre cas la technique d'élimination des cellules mortes. Nous sommes donc passé à un tri cellulaire en cytométrie de flux pour l'élimination des cellules mortes des échantillons de notre deuxième cohorte, et nous avons pu montrer en comparant nos résultats aux numérations cellulaires cliniques que de cette manière l'hétérogénéité complète des PBMCs était capturée en scRNA-seq. L'étude de cet atlas des cellules immunitaires circulantes au cours du temps et à travers les conditions, en incluant des échantillons d'ABMR, TCMR et de patients stables, nous a permis de décrire plus en détail la composition cellulaire des différents patients. Nous avons montré que les variations de proportions cellulaires entre conditions et au cours du temps étaient minimales et ne permettent pas d'associer une réorganisation cellulaire à la progression du rejet dans les PBMCs de patients transplantés. Pour mieux comprendre l'organisation moléculaire, nous avons analysé les variations d'expression entre conditions et au cours du temps. Plusieurs gènes différentiels sont sortis de l'analyse et parmi les plus différentiels le gène *CD69*, un marqueur d'activation précoce des lymphocytes T et de rétention dans les tissus (Ziegler, Ramsdell, et Alderson 1994; Werfel, Boeker, et Kapp 1997; J.-E. Turner et al. 2018), que l'on retrouve enrichi à l'émergence des rejets et qui diminue avec le temps. Ce marqueur nous intéresse puisqu'il a déjà été décrit lors d'inflammations chroniques (Afeltra et al. 1993) et enrichi dans le cas de rejet de greffe (Schowengerdt et al. 2000; Posselt et al. 2003). Pour mieux comprendre les voies moléculaires associées, nous avons effectué une analyse de module en parallèle sur les 3 groupes d'expériences qui ont révélé 9 modules communs dont le module ALARM. Parmi ces 9 modules, 5 d'entre eux sont exclusifs à un type cellulaire, et 3 autres sont exprimés dans des types cellulaires proches (sous-types de lymphocytes T ou sous-types de monocytes). Cette spécificité de programmes d'identité semble liée au fait que parmi la diversité de cellules trouvée dans les PBMCs, les gènes coexprimés dans un type cellulaire spécifique sont plus conservés et donc plus facilement détectés en analyse de modules. Le dernier module n'est pas restreint à un type cellulaire mais a été retrouvé dans la plupart des PBMCs et majoritairement exprimé dans les cellules T, NK et B. Certains des modules spécifiques à des types cellulaires sont toutefois enrichis au cours du temps dont deux modules de monocytes, et le module de NK, révélant une potentielle activation de ces cellules. Le module ALARM est cependant le seul module que l'on peut définir comme module "d'activité" en opposition aux modules "d'identité" (Kotliar et al. 2019). Le module ALARM contient le gène *CD69*



mais aussi un enrichissement en gènes codant pour des facteurs de transcription, comme les gènes *JUN* et *FOS* qui constituent des sous unités du facteur de transcription AP-1 ainsi que *REL*, une sous unité de NFκB. Or il a déjà été décrit que *CD69* est sous le contrôle de AP-1 et NFκB (López-Cabrera et al. 1995; Castellanos et al. 1997). L'augmentation de la protéine CD69 en surface des lymphocytes T a précédemment été montrée dans les cas de rejets aigu de cœur (Schowengerdt et al. 2000) et de reins (Posselt et al. 2003), et suppose un rôle dans l'activation des LT impliqués dans le rejet et leur rétention dans les tissus. Le module ALARM contient en plus le gène codant pour le récepteur à chimiokines CXCR4, exprimé par les LT et les monocytes, qui a été décrit dans les lésions ischémiques comme participant à la migration des leucocytes vers le tissu inflammé via la chimiokine CXCL12, son seul ligand. De plus, un taux plus élevé de CXCL12 a aussi été observé dans les biopsies en cas de rejets de rein chroniques (Hoffmann et al. 2006) et sachant qu'un gradient de CXCL12 participe à la chimiotaxie des cellules exprimant le CXCR4 (Britton, Poznansky, et Reeves 2021), CXCL12 participerait donc à l'attraction des cellules immunitaires dans les tissus inflammés du greffon. L'analyse des ontologies fonctionnelles à partir des gènes du module montrent que son activité est structurée autour de l'expression de gènes, en partie influencée par l'enrichissement en facteurs de transcriptions, avec des fonctions de transcription, de traitement de l'ARN et d'ubiquitination.

L'hypothèse d'un recrutement de ces cellules activées vers les tissus du greffon est étayée par les données de tissus humain et du modèle de rejet de greffe chez le porc que nous avons analysé. L'analyse transcriptomique des gènes du module ALARM ont démontré un enrichissement significatif dans les tissus de greffons rejetés comparés aux greffons stables de notre modèle de rejet aigu chez le porc. Il apparaît aussi une augmentation des monocytes et une diminution des lymphocytes dans le sang, ainsi qu'une forte diminution du module ALARM dans les cellules circulantes restantes après le jour 2. Associé à l'infiltration des tissus par des lymphocytes que l'on a pu observer sur les biopsies lors du rejet, ces résultats semblent suggérer que les cellules exprimant le module ALARM sont préférentiellement recrutées dans les tissus inflammés.

En support des données des PBMC sur le modèle de rejet chez le porc, nous avons cherché à récupérer les cellules immunitaires présentes dans les fragments de rein prélevés le jour du rejet, avec un fragment du rein sain et un fragment du rein transplanté en cours de rejet aigu. Cependant, la rareté des cellules immunitaires dans un tissu sain, et le processus de digestion des tissus et d'isolement des cellules d'intérêt ne m'a pas permis de réaliser une expérience de scRNA-seq comparative entre tissus sains et en rejet, malgré les tentatives d'optimisation de la méthode. Nous avons alors utilisé deux jeux de données publics (Reeve et al., s. d.; Callemeyn et al. 2020) de microarray sur des biopsies de patients transplantés du rein, encourageant un rejet ou non. Nous avons montré un enrichissement de notre module ALARM, notamment d'une partie des gènes le constituant dont *CD69* et *CXCR4* décrits précédemment, dans les échantillons de TCMR et ABMR issus de ces jeux de

données. Ces résultats corroborent l'hypothèse d'une infiltration préférentielle de cellules enrichies en ce module. Afin d'identifier des pistes de signaux participant à ce recrutement cellulaire, nous avons conduit une analyse des ligands exprimés dans les tissus inflammés et des récepteurs exprimés par les PBMCs, à partir de laquelle nous avons identifié une forte association du couple CXCL12-CXCR4. CXCL12 semble donc jouer un rôle similaire à son rôle de homing décrit dans la moelle osseuse (Moll et Ransohoff 2010; Yellowley 2013) et dans les tissus cancéreux (Burger et Kipps 2006) et permettrait le recrutement spécifique des cellules CXCR4+ et par extension, enrichies en module ALARM.

En partant de l'observation d'un enrichissement du module ALARM dans les cellules circulantes en cas de rejet, nous avons émis l'hypothèse que ce module pourrait être retrouvé dans d'autres pathologies impliquant les PBMCs. Pour vérifier cette hypothèse, nous avons analysé deux jeux de données. Le premier contient des patients sains stimulés avec de l'endotoxine (aussi connue sous le nom de LPS pour lipopolysaccharides) en intra-veineuse pour simuler un sepsis et provoquer un syndrome de réponse inflammatoire systémique (SIRS) et le second comporte des patients atteints ou non du COVID, à différents niveaux de sévérité. La stimulation par le LPS induit une diminution rapide (en 90 min) du taux de cellules enrichies au module ALARM dans le sang. Dans la cohorte de patients COVID, le module ALARM était diminué de manière corrélée à la sévérité. Il est intéressant de noter que nos résultats montrent une exception chez les patients asymptomatiques pour qui le module est globalement augmenté, mais dont l'enrichissement n'est en réalité associé qu'aux lymphocytes puisque le module est diminué dans les monocytes. Ces résultats montrent que le module de gènes que l'on a identifié comme enrichi dans les cellules immunitaires circulantes participant à la réponse immunitaire durant le rejet de greffe de rein est aussi retrouvé dans différentes conditions inflammatoires. La proportion de cellules enrichies en module ALARM dans le sang varierait avec la condition et la sévérité de la pathologie.

## Partie MICI

Dans mon second projet, nous nous sommes intéressés à la réorganisation cellulaire de la lamina propria à la suite d'une réponse inflammatoire dans les MICI. L'identification de la signature GIMATS dans les précédents travaux de mon équipe avait permis une première stratification des patients en deux groupes. L'enrichissement en signature GIMATS est associée à une résistance accrue aux traitements aux anti-TNF, même dans des patients d'apparence clinique homogène. Cependant cette stratification à haute résolution ne couvre qu'un nombre réduit de patients et l'intégration de nouveaux patients m'a permis de couvrir de plus nombreux cas de MICI.

Des échantillons additionnels d'iléon de patients atteints de la MC ayant des degrés divers d'atteinte endoscopique ont été inclus, ainsi que des échantillons de colon venant de patients avec une RCH puisque la pathogénie est proche, bien qu'il reste important de distinguer les différences qui mènent à l'un ou l'autre. Les MNP ont été décrits dans l'étude précédente de notre équipe comme acteurs importants de

l'organisation cellulaire en réponse aux infections, nous avons donc réalisé un enrichissement de ces populations et effectué le scRNA-seq sur ces échantillons. Par comparaison avec des échantillons appariés non enrichis, nous avons montré que l'enrichissement en MNP n'impactait pas la distribution des cellules capturées, ni leurs profils d'expression, et permet donc d'étudier avec plus de précision les programmes moléculaires qui expliquent leur hétérogénéité.

L'implémentation de l'algorithme Metacells au problème des MICI a nécessité une longue période de mise en place et de réglage des différents paramètres pour s'assurer une intégration des données optimisée. La difficulté la plus importante que j'ai rencontré dans l'analyse de ces données est due aux biais techniques et biologiques. À travers des étapes d'affinage successives, j'ai été capable d'identifier les différents obstacles à l'intégration de nos données. Les variations inter-individuelles des gènes liés aux immunoglobulines et aux chaînes du CMH induisaient une forte variabilité des profils de plasmocytes qui impacte la sélection des metacells et des clusters. La population observée de cellules immunitaires et de cellules épithéliales est un panel large de profils transcriptomiques, et il a fallu réduire le poids des plasmocytes pour avoir une représentativité équilibrée des différentes populations. J'ai ensuite observé qu'un biais majeur entre les échantillons des deux laboratoires se rapportait en fait à une différence dans la chimie des réactifs utilisés pour la capture et l'amplification des transcrits en scRNA-seq. En effet, la distribution des gènes diffère, avec un taux de capture de transcrits et un nombre de gènes détectés plus élevés en v3 que dans la version précédente (Yamawaki et al. 2021). J'ai pu répondre à certaines limites de la technique de scRNA-seq en utilisant des outils adaptés. L'application d'un calcul de metacells à partir de notre jeu de données apporte une sensibilité accrue dans la définition des programmes moléculaires tout en réduisant les contraintes liées à l'importante quantité de données. La résolution permise par le scRNA-seq et l'intégration en metacells permet de découper la réponse inflammatoire des macrophages, là où les études précédentes se contentaient de définir une réponse unique comprenant les différentes cytokines inflammatoires comme TNF $\alpha$ , IL-1 $\beta$ , IL-6, IL-12 et IL-23 (Bain et Mowat 2014). J'ai pu montrer que l'expression de ces cytokines peut être regroupée en plusieurs modules de gènes co-exprimés dont les profils d'expression ne se chevauchent que partiellement : certains ne sont exprimés que dans les monocytes fraîchement infiltrés tandis que d'autres sont retrouvés uniformément parmi les macrophages inflammatoires.

En condition normale, les macrophages résidents ont un rôle de phagocytes et de maintien de l'homéostasie. Cependant, lors d'une inflammation, d'autres macrophages entrent en jeu, des macrophages dérivés des monocytes, qui sont recrutés du sang. Une partie de ces macrophages, que l'on avait appelé Macs2 dans la première analyse, est spécifique aux patients pour qui on retrouve une activation stromale associée. On retrouve en effet dans les mêmes patients les cellules endothéliales ACKR1+ et les Fibroblastes inflammatoires (**Annexe 3**). Cette présence complémentaire est cohérente avec le module inflammatoire spécifiquement exprimé par les Macs2 (le module 18) qui contient les gènes codant les chimiokines CXCL2,

CXCL3 et CXCL8, qui sont des chimiokines décrites dans le recrutement des neutrophiles, et que les protéines d'ACKR1 des cellules endothéliales présentent CXCL2 pour faciliter la transmigration des neutrophiles vers la lamina propria (Girbl et al. 2018).

La description détaillée des MNP a mis en évidence une population rare que l'on retrouve enrichi chez une faible fraction des patients GIMATS<sup>high</sup>, que l'on identifie par "Macs3". Leur profil transcriptomique suggère une fonction qui semble être à double tranchant. Ces cellules expriment un programme pro-résolutif impliqué dans la réparation des lésions de la barrière, et dans la phagocytose à l'homéostasie mais en condition pathologique chronique, la surexpression de métalloprotéase semble en contrepartie favoriser la dégradation de la matrice extra cellulaire et provoquerait une fibrose irréversible des tissus, qui est la complication principale dans la maladie de Crohn. Il serait pertinent de trouver comment pousser ces cellules à garder un profil pro-résolutif et si leur induction dans d'autres cas de MICI permettrait d'assurer une meilleure résolution de l'inflammation.

Les macrophages résidents de la muqueuse intestinale proviennent en premier lieu de précurseurs embryonnaires qui colonisent le tissu et se maintiennent à l'âge adulte (Ginhoux et Williams 2016; De Schepper et al. 2018), puis à l'âge adulte, les monocytes circulants sont recrutés pour garder un nombre constant de macrophages dans les tissus (Bain et al. 2014; Ginhoux et Williams 2016). C'est ce qu'on semble voir ici avec les macrophages immatures, fraîchement infiltrés, et un spectre de différenciation vers des profils plus matures, comme Macs 3 qui en plus exprime un programme phagocytaire fort, et de nombreux gènes impliqués dans le remodelage tissulaire. Cela laisse penser que les macrophages inflammatoires persistants chez les patients réfractaires aux traitements de base peuvent jouer un rôle dans la fibrose par l'expression de MMP qui participent au remodelage de la matrice extra cellulaire, la fibrose étant une des complications fréquentes des MICIs. Il serait aussi intéressant d'analyser si ces cellules communiquent avec les fibroblastes et si elles encourageant la fibrose par ce moyen aussi. De plus, l'analyse en modules a permis de révéler une grande diversité de programmes inflammatoires exprimés différemment au sein des macrophages inflammatoires, et traduit les différences de réponses des patients. On s'éloigne de la polarisation simpliste M1/M2 des macrophages qui était la norme il y a quelques années mais les avancées récentes, notamment la description plus précise des programmes cellulaires grâce au scRNA-seq, a permis d'établir un spectre plus large de macrophages caractéristiques comme vu ici.

À partir de ces signatures je pourrais tester dans la cohorte qui est maintenant large et diversifiée leur association avec les différents paramètres cliniques, qu'ils soient accessoires à notre étude (âge, sexe, origine, pathologies additionnelles, ...) ou d'intérêt comme la sévérité de la maladie ou l'efficacité des traitements promulgués. Ce premier pas dans la description des communications cellulaires des MNP devra s'accompagner d'une étude des cellules stromales pour comprendre les communications potentielles qui ont lieu dans un même sous-groupe de patient. L'étude des MNP qui semblent jouer un rôle central dans l'organisation de la réponse

immunitaire, pourrait soulever de nouvelles pistes pour traiter les patients chez qui les traitements qui ciblent en majorité les réponses lymphocytaires ont échoué. De plus, une meilleure compréhension des mécanismes moléculaires en jeu permettrait de mieux adapter le traitement en fonction du profil du patient.

En perspective sur le plus long terme, nous souhaitons appliquer une méthode de transcriptomique spatiale afin d'identifier la localisation de chaque cellule et mieux caractériser les communications cellulaires dans la muqueuse. Pour tester nos hypothèses sur les programmes identifiés et les facteurs de transcription, nous envisageons de faire une analyse en MultiOme (10X Genomics) pour capturer en parallèle l'information transcriptomique et épigénétique. La caractérisation détaillée des MNP devrait aussi nous permettre d'identifier des marqueurs pour le tri en cytométrie, ainsi que des cibles moléculaires pour la différenciation et l'induction de signaux *in vitro* avec pour objectif de prendre des monocytes de patients Crohn qui semblent incapables d'effectuer leur rôle de nettoyage et de résolution et essayer de leur induire ces programmes à nouveau.

# Bibliographie

- Abidi, Ahmed, Thomas Laurent, Gaëlle Bériou, Laurence Bouchet-Delbos, Cynthia Fourgeux, Cédric Louvet, Raja Triki-Marrakchi, Jeremie Poschmann, Régis Josien, et Jérôme Martin. 2020. « Characterization of Rat ILCs Reveals ILC2 as the Dominant Intestinal Subset ». *Frontiers in Immunology* 11. <https://doi.org/10.3389/fimmu.2020.00255>.
- Adami, J, H Gäbel, B Lindelöf, K Ekström, B Rydh, B Glimelius, A Ekblom, H-O Adami, et F Granath. 2003. « Cancer risk following organ transplantation: a nationwide cohort study in Sweden ». *British Journal of Cancer* 89 (7): 1221-27. <https://doi.org/10.1038/sj.bjc.6601219>.
- Afeltra, A., M. Galeazzi, G. M. Ferri, A. Amoroso, O. De Pità, F. Porzio, et L. Bonomo. 1993. « Expression of CD69 Antigen on Synovial Fluid T Cells in Patients with Rheumatoid Arthritis and Other Chronic Synovitis ». *Annals of the Rheumatic Diseases* 52 (6): 457-60. <https://doi.org/10.1136/ard.52.6.457>.
- Aibar, Sara, Carmen Bravo González-Blas, Thomas Moerman, Vân Anh Huynh-Thu, Hana Imrichova, Gert Hulselmans, Florian Rambow, et al. 2017. « SCENIC: Single-Cell Regulatory Network Inference and Clustering ». *Nature Methods* 14 (11): 1083-86. <https://doi.org/10.1038/nmeth.4463>.
- Alatab, Sudabeh, Sadaf G. Sepanlou, Kevin Ikuta, Homayoon Vahedi, Catherine Bisignano, Saeid Safiri, Anahita Sadeghi, et al. 2020. « The Global, Regional, and National Burden of Inflammatory Bowel Disease in 195 Countries and Territories, 1990–2017: A Systematic Analysis for the Global Burden of Disease Study 2017 ». *The Lancet Gastroenterology & Hepatology* 5 (1): 17-30. [https://doi.org/10.1016/S2468-1253\(19\)30333-4](https://doi.org/10.1016/S2468-1253(19)30333-4).
- Alberts, Bruce, Dennis Bray, Karen Hopkin, Alexander D. Johnson, Julian Lewis, Martin Raff, Keith Roberts, et Peter Walter. 2015. *Essential Cell Biology*. 4<sup>e</sup> éd. New York: W.W. Norton & Company. <https://doi.org/10.1201/9781315815015>.
- Allez, Matthieu, Konstantinos Karmiris, Edouard Louis, Gert Van Assche, Shomron Ben-Horin, Amir Klein, Janneke Van der Woude, et al. 2010. « Report of the ECCO Pathogenesis Workshop on Anti-TNF Therapy Failures in Inflammatory Bowel Diseases: Definitions, Frequency and Pharmacological Aspects ». *Journal of Crohn's & Colitis* 4 (4): 355-66. <https://doi.org/10.1016/j.crohns.2010.04.004>.
- Al-Sadi, Rana, Shuhong Guo, Dongmei Ye, Manmeet Rawat, et Thomas Y. Ma. 2016. « TNF- $\alpha$  Modulation of Intestinal Tight Junction Permeability Is Mediated by NIK/IKK- $\alpha$  Axis Activation of the Canonical NF-KB Pathway ». *The American Journal of Pathology* 186 (5): 1151-65. <https://doi.org/10.1016/j.ajpath.2015.12.016>.
- Angus, Hamish C. K., A. Grant Butt, Michael Schultz, et Roslyn A. Kemp. 2020. « Intestinal Organoids as a Tool for Inflammatory Bowel Disease Research ». *Frontiers in Medicine* 6 (janvier): 334. <https://doi.org/10.3389/fmed.2019.00334>.
- Argmann, Carmen, Ruixue Hou, Ryan C. Ungaro, Haritz Irizar, Zainab Al-Taie, Ruiqi Huang, Roman Kosoy, et al. 2022. « Biopsy and Blood-Based Molecular Biomarker of Inflammation in IBD ». *Gut*, septembre. <https://doi.org/10.1136/gutjnl-2021-326451>.
- Arijs, Ingrid, Roel Quintens, Leentje Van Lommel, Kristel Van Steen, Gert De Hertogh, Katleen Lemaire, Anica Schraenen, et al. 2010. « Predictive Value of

- Epithelial Gene Expression Profiles for Response to Infliximab in Crohn's Disease ». *Inflammatory Bowel Diseases* 16 (12): 2090-98.  
<https://doi.org/10.1002/ibd.21301>.
- Atreya, Raja, Markus F. Neurath, et Britta Siegmund. 2020. « Personalizing Treatment in IBD: Hype or Reality in 2020? Can We Predict Response to Anti-TNF? » *Frontiers in Medicine* 7 (septembre): 517.  
<https://doi.org/10.3389/fmed.2020.00517>.
- Atreya, Raja, Michael Zimmer, Brigitte Bartsch, Maximilian J. Waldner, Imke Atreya, Helmut Neumann, Kai Hildner, et al. 2011. « Antibodies against Tumor Necrosis Factor (TNF) Induce T-Cell Apoptosis in Patients with Inflammatory Bowel Diseases via TNF Receptor 2 and Intestinal CD14<sup>+</sup> Macrophages ». *Gastroenterology* 141 (6): 2026-38.  
<https://doi.org/10.1053/j.gastro.2011.08.032>.
- Bain, Calum C., Alberto Bravo-Blas, Charlotte L. Scott, Elisa Gomez Perdiguero, Frederic Geissmann, Sandrine Henri, Bernard Malissen, Lisa C. Osborne, David Artis, et Allan Mcl Mowat. 2014. « Constant Replenishment from Circulating Monocytes Maintains the Macrophage Pool in the Intestine of Adult Mice ». *Nature Immunology* 15 (10): 929-37.  
<https://doi.org/10.1038/ni.2967>.
- Bain, Calum C., et Allan Mcl Mowat. 2014. « Macrophages in intestinal homeostasis and inflammation ». *Immunological Reviews* 260 (1): 102-17.  
<https://doi.org/10.1111/imr.12192>.
- Baker, Shawn C., Steven R. Bauer, Richard P. Beyer, James D. Brenton, Bud Bromley, John Burrill, Helen Causton, et al. 2005. « The External RNA Controls Consortium: A Progress Report ». *Nature Methods* 2 (10): 731-34.  
<https://doi.org/10.1038/nmeth1005-731>.
- Baran, Yael, Akhiad Bercovich, Arnau Sebe-Pedros, Yaniv Lubling, Amir Giladi, Elad Chomsky, Zohar Meir, Michael Hoichman, Aviezer Lifshitz, et Amos Tanay. 2019. « MetaCell: analysis of single-cell RNA-seq data using K-nn graph partitions ». *Genome Biology* 20 (1): 206. <https://doi.org/10.1186/s13059-019-1812-2>.
- Baumgart, Daniel C., et Catherine Le Berre. 2021. « Newer Biologic and Small-Molecule Therapies for Inflammatory Bowel Disease ». *New England Journal of Medicine* 385 (14): 1302-15. <https://doi.org/10.1056/NEJMra1907607>.
- Ben-Kiki, Oren, Akhiad Bercovich, Aviezer Lifshitz, et Amos Tanay. 2022. « Metacell-2: a divide-and-conquer metacell algorithm for scalable scRNA-seq analysis ». *Genome Biology* 23 (1): 100. <https://doi.org/10.1186/s13059-022-02667-1>.
- Berlin, C., E. L. Berg, M. J. Briskin, D. P. Andrew, P. J. Kilshaw, B. Holzmann, I. L. Weissman, A. Hamann, et E. C. Butcher. 1993. « Alpha 4 Beta 7 Integrin Mediates Lymphocyte Binding to the Mucosal Vascular Addressin MAdCAM-1 ». *Cell* 74 (1): 185-95. [https://doi.org/10.1016/0092-8674\(93\)90305-a](https://doi.org/10.1016/0092-8674(93)90305-a).
- Bernink, Jochem H., Lisette Krabbendam, Kristine Germar, Esther de Jong, Konrad Gronke, Michael Kofoed-Nielsen, J. Marius Munneke, et al. 2015. « Interleukin-12 and -23 Control Plasticity of CD127<sup>+</sup> Group 1 and Group 3 Innate Lymphoid Cells in the Intestinal Lamina Propria ». *Immunity* 43 (1): 146-60. <https://doi.org/10.1016/j.immuni.2015.06.019>.
- Bloom, Jesse D. 2018. « Estimating the Frequency of Multiplets in Single-Cell RNA Sequencing from Cell-Mixing Experiments ». *PeerJ* 6: e5578.  
<https://doi.org/10.7717/peerj.5578>.
- Briskin, M., D. Winsor-Hines, A. Shyjan, N. Cochran, S. Bloom, J. Wilson, L. M. McEvoy, et al. 1997. « Human Mucosal Addressin Cell Adhesion Molecule-1

- Is Preferentially Expressed in Intestinal Tract and Associated Lymphoid Tissue ». *The American Journal of Pathology* 151 (1): 97-110.
- Britton, C., M. C. Poznansky, et P. Reeves. 2021. « Polyfunctionality of the CXCR4/CXCL12 Axis in Health and Disease: Implications for Therapeutic Interventions in Cancer and Immune-Mediated Diseases ». *FASEB Journal: Official Publication of the Federation of American Societies for Experimental Biology* 35 (4): e21260. <https://doi.org/10.1096/fj.202001273R>.
- Bunte, Kübra, et Thomas Beikler. 2019. « Th17 Cells and the IL-23/IL-17 Axis in the Pathogenesis of Periodontitis and Immune-Mediated Inflammatory Diseases ». *International Journal of Molecular Sciences* 20 (14): E3394. <https://doi.org/10.3390/ijms20143394>.
- Burger, Jan A., et Thomas J. Kipps. 2006. « CXCR4: A Key Receptor in the Crosstalk between Tumor Cells and Their Microenvironment ». *Blood* 107 (5): 1761-67. <https://doi.org/10.1182/blood-2005-08-3182>.
- Burisch, Johan, Gediminas Kiudelis, Limas Kupcinskis, Hendrika Adriana Linda Kievit, Karina Winther Andersen, Vibeke Andersen, Riina Salupere, et al. 2019. « Natural Disease Course of Crohn's Disease during the First 5 Years after Diagnosis in a European Population-Based Inception Cohort: An Epi-IBD Study ». *Gut* 68 (3): 423-33. <https://doi.org/10.1136/gutjnl-2017-315568>.
- Butler, Andrew, Paul Hoffman, Peter Smibert, Efthymia Papalexi, et Rahul Satija. 2018. « Integrating Single-Cell Transcriptomic Data across Different Conditions, Technologies, and Species ». *Nature Biotechnology* 36 (5): 411-20. <https://doi.org/10.1038/nbt.4096>.
- Callemeyn, Jasper, Evelyne Lerut, Henriette de Loor, Ingrid Arijs, Olivier Thauinat, Alice Koenig, Vannary Meas-Yedid, et al. 2020. « Transcriptional Changes in Kidney Allografts with Histology of Antibody-Mediated Rejection without Anti-HLA Donor-Specific Antibodies ». *Journal of the American Society of Nephrology* 31 (9): 2168-83. <https://doi.org/10.1681/ASN.2020030306>.
- Cao, Junyue, Jonathan S. Packer, Vijay Ramani, Darren A. Cusanovich, Chau Huynh, Riza Daza, Xiaojie Qiu, et al. 2017. « Comprehensive single-cell transcriptional profiling of a multicellular organism ». *Science* 357 (6352): 661-67. <https://doi.org/10.1126/science.aam8940>.
- Cao, Yirui, Yue Qiu, Guowei Tu, et Cheng Yang. 2020. « Single-cell RNA Sequencing in Immunology ». *Current Genomics* 21 (8): 564-75. <https://doi.org/10.2174/1389202921999201020203249>.
- Caruso, Roberta, Bernard C. Lo, et Gabriel Núñez. 2020. « Host–Microbiota Interactions in Inflammatory Bowel Disease ». *Nature Reviews Immunology* 20 (7): 411-26. <https://doi.org/10.1038/s41577-019-0268-7>.
- Casamassimi, Amelia, Antonio Federico, Monica Rienzo, Sabrina Esposito, et Alfredo Ciccodicola. 2017. « Transcriptome Profiling in Human Diseases: New Advances and Perspectives ». *International Journal of Molecular Sciences* 18 (8): E1652. <https://doi.org/10.3390/ijms18081652>.
- Castellanos, M. C., C. Muñoz, M. C. Montoya, E. Lara-Pezzi, M. López-Cabrera, et M. O. de Landázuri. 1997. « Expression of the Leukocyte Early Activation Antigen CD69 Is Regulated by the Transcription Factor AP-1 ». *Journal of Immunology (Baltimore, Md.: 1950)* 159 (11): 5463-73.
- Castro-Dopico, Tomas, et Menna R. Clatworthy. s. d. « Mucosal IgG in inflammatory bowel disease – a question of (sub)class? ». *Gut Microbes* 12 (1): 1651596. <https://doi.org/10.1080/19490976.2019.1651596>.
- Cavanaugh, J. et IBD International Genetics Consortium. 2001. « International Collaboration Provides Convincing Linkage Replication in Complex Disease through Analysis of a Large Pooled Data Set: Crohn Disease and



- Chromosome 16 ». *American Journal of Human Genetics* 68 (5): 1165-71.  
<https://doi.org/10.1086/320119>.
- Cazzato, Gerardo, Anna Colagrande, Valeria Andriola, Teresa Lettini, Sebastiano Cicco, Pragnell Mary Victoria Candance, Leonardo Resta, Leonardo Vincenti, et Giuseppe Ingravallo. 2021. « Histological Hallmarks of Mucosal Healing in Inflammatory Bowel Diseases in the Era of Monoclonal Antibodies Therapy: New Insights and Perspectives ». *Diagnostics* 11 (9): 1570.  
<https://doi.org/10.3390/diagnostics11091570>.
- Cella, Marina, Anja Fuchs, William Vermi, Fabio Facchetti, Karel Otero, Jochen K. M. Lennerz, Jason M. Doherty, Jason C. Mills, et Marco Colonna. 2009. « A Human Natural Killer Cell Subset Provides an Innate Source of IL-22 for Mucosal Immunity ». *Nature* 457 (7230): 722-25.  
<https://doi.org/10.1038/nature07537>.
- Chang, John T. 2020. « Pathophysiology of Inflammatory Bowel Diseases ». *The New England Journal of Medicine* 383 (27): 2652-64.  
<https://doi.org/10.1056/NEJMra2002697>.
- Chen, Haide, Fang Ye, et Guoji Guo. 2019. « Revolutionizing immunology with single-cell RNA sequencing ». *Cellular and Molecular Immunology* 16 (3): 242-49. <https://doi.org/10.1038/s41423-019-0214-4>.
- Childers, Ryan E., Swathi Eluri, Christine Vazquez, Rayna Matsuno Weise, Theodore M. Bayless, et Susan Hutfless. 2014. « Family History of Inflammatory Bowel Disease among Patients with Ulcerative Colitis: A Systematic Review and Meta-Analysis ». *Journal of Crohn's & Colitis* 8 (11): 1480-97. <https://doi.org/10.1016/j.crohns.2014.05.008>.
- Cho, Judy H., et Clara Abraham. 2007. « Inflammatory Bowel Disease Genetics: Nod2 ». *Annual Review of Medicine* 58: 401-16.  
<https://doi.org/10.1146/annurev.med.58.061705.145024>.
- Clevers, Hans. 2013. « The Intestinal Crypt, A Prototype Stem Cell Compartment ». *Cell* 154 (2): 274-84. <https://doi.org/10.1016/j.cell.2013.07.004>.
- Cleynen, Isabelle, Gabrielle Boucher, Luke Jostins, L. Philip Schumm, Sebastian Zeissig, Tariq Ahmad, Vibeke Andersen, et al. 2016. « Inherited Determinants of Crohn's Disease and Ulcerative Colitis Phenotypes: A Genetic Association Study ». *Lancet (London, England)* 387 (10014): 156-67.  
[https://doi.org/10.1016/S0140-6736\(15\)00465-1](https://doi.org/10.1016/S0140-6736(15)00465-1).
- Coillard, Alice, et Elodie Segura. 2019. « In vivo Differentiation of Human Monocytes ». *Frontiers in Immunology* 10 (août): 1907.  
<https://doi.org/10.3389/fimmu.2019.01907>.
- Colombel, J. F., B. Grandbastien, C. Gower-Rousseau, S. Plegat, J. P. Evrard, J. L. Dupas, J. P. Gendre, et al. 1996. « Clinical Characteristics of Crohn's Disease in 72 Families ». *Gastroenterology* 111 (3): 604-7.  
<https://doi.org/10.1053/gast.1996.v111.pm8780563>.
- Dall, Aaron, et Sundaram Hariharan. 2008. « BK Virus Nephritis after Renal Transplantation ». *Clinical Journal of the American Society of Nephrology : CJASN* 3 (Suppl 2): S68-75. <https://doi.org/10.2215/CJN.02770707>.
- Dantas, S. R. P. E., R. H. Kuboyama, M. Mazzali, et M. L. Moretti. 2006. « Nosocomial Infections in Renal Transplant Patients: Risk Factors and Treatment Implications Associated with Urinary Tract and Surgical Site Infections ». *The Journal of Hospital Infection* 63 (2): 117-23.  
<https://doi.org/10.1016/j.jhin.2005.10.018>.
- Dashti-Khavidaki, Simin, Reza Saidi, et Hong Lu. 2021. « Current status of glucocorticoid usage in solid organ transplantation ». *World Journal of*

- Transplantation* 11 (11): 443-65. <https://doi.org/10.5500/wjt.v11.i11.443>.
- De Schepper, Sebastiaan, Simon Verheijden, Javier Aguilera-Lizarraga, Maria Francesca Viola, Werend Boesmans, Nathalie Stakenborg, Iryna Voytyuk, et al. 2018. « Self-Maintaining Gut Macrophages Are Essential for Intestinal Homeostasis ». *Cell* 175 (2): 400-415.e13. <https://doi.org/10.1016/j.cell.2018.07.048>.
- De Voogd, Floris AE, Richard B Geary, Christopher J Mulder, et Andrew S Day. 2016. « Osteoprotegerin: A Novel Biomarker for Inflammatory Bowel Disease and Gastrointestinal Carcinoma ». *Journal of Gastroenterology and Hepatology* 31 (8): 1386-92. <https://doi.org/10.1111/jgh.13324>.
- Dicksved, Johan, Jonas Halfvarson, Magnus Rosenquist, Gunnar Järnerot, Curt Tysk, Juha Apajalahti, Lars Engstrand, et Janet K. Jansson. 2008. « Molecular Analysis of the Gut Microbiota of Identical Twins with Crohn's Disease ». *The ISME Journal* 2 (7): 716-27. <https://doi.org/10.1038/ismej.2008.37>.
- Dobin, Alexander, Carrie A. Davis, Felix Schlesinger, Jorg Drenkow, Chris Zaleski, Sonali Jha, Philippe Batut, Mark Chaisson, et Thomas R. Gingeras. 2013. « STAR: Ultrafast Universal RNA-Seq Aligner ». *Bioinformatics (Oxford, England)* 29 (1): 15-21. <https://doi.org/10.1093/bioinformatics/bts635>.
- Dominguez-Bello, Maria Gloria, Martin J. Blaser, Ruth E. Ley, et Rob Knight. 2011. « Development of the Human Gastrointestinal Microbiota and Insights from High-Throughput Sequencing ». *Gastroenterology* 140 (6): 1713-19. <https://doi.org/10.1053/j.gastro.2011.02.011>.
- Dorrington, Michael G., et Iain D. C. Fraser. 2019. « NF-κB Signaling in Macrophages: Dynamics, Crosstalk, and Signal Integration ». *Frontiers in Immunology* 10. <https://www.frontiersin.org/articles/10.3389/fimmu.2019.00705>.
- Dorschner, P., L. M. McElroy, et M.g. Ison. 2014. « Nosocomial Infections within the First Month of Solid Organ Transplantation ». *Transplant Infectious Disease* 16 (2): 171-87. <https://doi.org/10.1111/tid.12203>.
- Duerr, Richard H., Kent D. Taylor, Steven R. Brant, John D. Rioux, Mark S. Silverberg, Mark J. Daly, A. Hillary Steinhart, et al. 2006. « A Genome-Wide Association Study Identifies IL23R as an Inflammatory Bowel Disease Gene ». *Science* 314 (5804): 1461-63. <https://doi.org/10.1126/science.1135245>.
- Dufour, Antoine, Caroline L. Bellac, Ulrich Eckhard, Nestor Solis, Theo Klein, Reinhild Kappelhoff, Nikolaus Fortelny, et al. 2018. « C-Terminal Truncation of IFN-γ Inhibits Proinflammatory Macrophage Responses and Is Deficient in Autoimmune Disease ». *Nature Communications* 9 (1): 2416. <https://doi.org/10.1038/s41467-018-04717-4>.
- Dumoutier, L., D. Lejeune, D. Colau, et J. C. Renauld. 2001. « Cloning and Characterization of IL-22 Binding Protein, a Natural Antagonist of IL-10-Related T Cell-Derived Inducible Factor/IL-22 ». *Journal of Immunology (Baltimore, Md.: 1950)* 166 (12): 7090-95. <https://doi.org/10.4049/jimmunol.166.12.7090>.
- Dumoutier, L., J. Louahed, et J. C. Renauld. 2000. « Cloning and Characterization of IL-10-Related T Cell-Derived Inducible Factor (IL-TIF), a Novel Cytokine Structurally Related to IL-10 and Inducible by IL-9 ». *Journal of Immunology (Baltimore, Md.: 1950)* 164 (4): 1814-19. <https://doi.org/10.4049/jimmunol.164.4.1814>.
- Dunham, Ian, Anshul Kundaje, Shelley F. Aldred, Patrick J. Collins, Carrie A. Davis, Francis Doyle, Charles B. Epstein, et al. 2012. « An Integrated Encyclopedia

- of DNA Elements in the Human Genome ». *Nature* 489 (7414): 57-74. <https://doi.org/10.1038/nature11247>.
- Durrbach, A., J. M. Pestana, T. Pearson, F. Vincenti, V. D. Garcia, J. Campistol, M. del Carmen Rial, et al. 2010. « A Phase III Study of Belatacept versus Cyclosporine in Kidney Transplants from Extended Criteria Donors (BENEFIT-EXT Study) ». *American Journal of Transplantation: Official Journal of the American Society of Transplantation and the American Society of Transplant Surgeons* 10 (3): 547-57. <https://doi.org/10.1111/j.1600-6143.2010.03016.x>.
- Eddy, Justin L., Karen Krukowski, Linda Janusek, et Herbert L. Mathews. 2014. « Glucocorticoids regulate natural killer cell function epigenetically ». *Cellular immunology* 290 (1): 120-30. <https://doi.org/10.1016/j.cellimm.2014.05.013>.
- Eder, Piotr, Michal Michalak, Katarzyna Katulska, Liliana Lykowska-Szuber, Iwona Krela-Kazmierczak, Kamila Stawczyk-Eder, Katarzyna Klimczak, Aleksandra Szymczak, et Krzysztof Linke. 2015. « Magnetic resonance enterographic predictors of one-year outcome in ileal and ileocolonic Crohn's disease treated with anti-tumor necrosis factor antibodies ». *Scientific Reports* 5 (mai): 10223. <https://doi.org/10.1038/srep10223>.
- Elsen, Lieke W. J. van den, Johan Garssen, Remy Burcelin, et Valerie Verhasselt. 2019. « Shaping the Gut Microbiota by Breastfeeding: The Gateway to Allergy Prevention? » *Frontiers in Pediatrics* 7. <https://www.frontiersin.org/articles/10.3389/fped.2019.00047>.
- External RNA Controls Consortium. 2005. « Proposed Methods for Testing and Selecting the ERCC External RNA Controls ». *BMC Genomics* 6 (novembre): 150. <https://doi.org/10.1186/1471-2164-6-150>.
- Famulski, K. S., G. Einecke, B. Sis, M. Mengel, L. G. Hidalgo, B. Kaplan, et P. F. Halloran. 2010. « Defining the Canonical Form of T-Cell-Mediated Rejection in Human Kidney Transplants ». *American Journal of Transplantation: Official Journal of the American Society of Transplantation and the American Society of Transplant Surgeons* 10 (4): 810-20. <https://doi.org/10.1111/j.1600-6143.2009.03007.x>.
- Fantou, Aurélie, Eric Lagrue, Thomas Laurent, Laurence Delbos, Stéphanie Blandin, Anne Jarry, Gaëlle Beriou, et al. 2022. « IL-22BP Production Is Heterogeneously Distributed in Crohn's Disease ». *Frontiers in Immunology* 13: 1034570. <https://doi.org/10.3389/fimmu.2022.1034570>.
- Feagan, Brian G., William J. Sandborn, Christopher Gasink, Douglas Jacobstein, Yinghua Lang, Joshua R. Friedman, Marion A. Blank, Jewel Johanns, Long-Long Gao, Ye Miao, Omoniyi J. Adedokun, Bruce E. Sands, Stephen B. Hanauer, Severine Vermeire, Stephan Targan, Subrata Ghosh, Willem J. de Villiers, Jean-Frédéric Colombel, Zsolt Tulassay, Ursula Seidler, Bruce A. Salzberg, Pierre Desreumaux, Scott D. Lee, Edward V. Loftus, Levinus A. Dieleman, Seymour Katz, Paul Rutgeerts, et al. 2016. « Ustekinumab as Induction and Maintenance Therapy for Crohn's Disease ». *The New England Journal of Medicine* 375 (20): 1946-60. <https://doi.org/10.1056/NEJMoa1602773>.
- . 2016. « Ustekinumab as Induction and Maintenance Therapy for Crohn's Disease ». *New England Journal of Medicine* 375 (20): 1946-60. <https://doi.org/10.1056/NEJMoa1602773>.
- Flier, Laurens G. van der, et Hans Clevers. 2009. « Stem Cells, Self-Renewal, and Differentiation in the Intestinal Epithelium ». *Annual Review of Physiology* 71: 241-60. <https://doi.org/10.1146/annurev.physiol.010908.163145>.
- Fodor, S. P., J. L. Read, M. C. Pirrung, L. Stryer, A. T. Lu, et D. Solas. 1991. « Light-

- Directed, Spatially Addressable Parallel Chemical Synthesis ». *Science (New York, N. Y.)* 251 (4995): 767-73. <https://doi.org/10.1126/science.1990438>.
- Friedrich, Matthias, Mathilde Pohin, et Fiona Powrie. 2019. « Cytokine Networks in the Pathophysiology of Inflammatory Bowel Disease ». *Immunity* 50 (4): 992-1006. <https://doi.org/10.1016/j.immuni.2019.03.017>.
- Fritz, Teresa, Lukas Niederreiter, Timon Adolph, Richard S Blumberg, et Arthur Kaser. 2011. « Crohn's disease: NOD2, autophagy and ER stress converge ». *Gut* 60 (11): 1580-88. <https://doi.org/10.1136/gut.2009.206466>.
- Frolkis, Alexandra, Levinus A Dieleman, Herman W Barkema, Remo Panaccione, Subrata Ghosh, Richard N Fedorak, Karen Madsen, et Gilaad G Kaplan. 2013. « Environment and the inflammatory bowel diseases ». *Canadian Journal of Gastroenterology* 27 (3): e18-24.
- Fuior, Elena V., et Anca V. Gafencu. 2019. « Apolipoprotein C1: Its Pleiotropic Effects in Lipid Metabolism and Beyond ». *International Journal of Molecular Sciences* 20 (23): 5939. <https://doi.org/10.3390/ijms20235939>.
- Fukunaga, Shuhei, Kotaro Kuwaki, Keiichi Mitsuyama, Hidetoshi Takedatsu, Shinichiro Yoshioka, Hiroshi Yamasaki, Ryosuke Yamauchi, et al. 2018. « Detection of Calprotectin in Inflammatory Bowel Disease: Fecal and Serum Levels and Immunohistochemical Localization ». *International Journal of Molecular Medicine* 41 (1): 107-18. <https://doi.org/10.3892/ijmm.2017.3244>.
- Garrett, Giorgia L., Paul D. Blanc, John Boscardin, Amanda Abramson Lloyd, Rehana L. Ahmed, Tiffany Anthony, Kristin Bibee, et al. 2017. « Incidence of and Risk Factors for Skin Cancer in Organ Transplant Recipients in the United States ». *JAMA Dermatology* 153 (3): 296-303. <https://doi.org/10.1001/jamadermatol.2016.4920>.
- Gerbe, François, Emmanuelle Sidot, Danielle J. Smyth, Makoto Ohmoto, Ichiro Matsumoto, Valérie Dardalhon, Pierre Cesses, et al. 2016. « Intestinal Epithelial Tuft Cells Initiate Type 2 Mucosal Immunity to Helminth Parasites ». *Nature* 529 (7585): 226-30. <https://doi.org/10.1038/nature16527>.
- Geremia, Alessandra, Carolina V. Arancibia-Cárcamo, Myles P. P. Fleming, Nigel Rust, Baljit Singh, Neil J. Mortensen, Simon P. L. Travis, et Fiona Powrie. 2011. « IL-23-Responsive Innate Lymphoid Cells Are Increased in Inflammatory Bowel Disease ». *The Journal of Experimental Medicine* 208 (6): 1127-33. <https://doi.org/10.1084/jem.20101712>.
- Ginhoux, Florent, et Martin Williams. 2016. « Tissue-Resident Macrophage Ontogeny and Homeostasis ». *Immunity* 44 (3): 439-49. <https://doi.org/10.1016/j.immuni.2016.02.024>.
- Girbl, Tamara, Tchern Lenn, Lorena Perez, Loïc Rolas, Anna Barkaway, Aude Thiriou, Carlos Del Fresno, et al. 2018. « Distinct Compartmentalization of the Chemokines CXCL1 and CXCL2 and the Atypical Receptor ACKR1 Determine Discrete Stages of Neutrophil Diapedesis ». *Immunity* 49 (6): 1062-1076.e6. <https://doi.org/10.1016/j.immuni.2018.09.018>.
- Goodman, Wendy A., Ian P. Erkkila, et Theresa T. Pizarro. 2020. « Sex Matters: Impact on Pathogenesis, Presentation and Treatment of Inflammatory Bowel Disease ». *Nature Reviews Gastroenterology & Hepatology* 17 (12): 740-54. <https://doi.org/10.1038/s41575-020-0354-0>.
- Gorfu, G., J. Rivera-Nieves, et K. Ley. 2009. « Role of Beta7 Integrins in Intestinal Lymphocyte Homing and Retention ». *Current Molecular Medicine* 9 (7): 836-50. <https://doi.org/10.2174/156652409789105525>.
- Graham, Daniel B., et Ramnik J. Xavier. 2020. « Pathway Paradigms Revealed from the Genetics of Inflammatory Bowel Disease ». *Nature* 578 (7796): 527-39.

- <https://doi.org/10.1038/s41586-020-2025-2>.
- Greenlee-Wacker, Mallary C. 2016. « Clearance of apoptotic neutrophils and resolution of inflammation ». *Immunological reviews* 273 (1): 357-70. <https://doi.org/10.1111/imr.12453>.
- Grün, Dominic, Anna Lyubimova, Lennart Kester, Kay Wiebrands, Onur Basak, Nobuo Sasaki, Hans Clevers, et Alexander van Oudenaarden. 2015. « Single-Cell Messenger RNA Sequencing Reveals Rare Intestinal Cell Types ». *Nature* 525 (7568): 251-55. <https://doi.org/10.1038/nature14966>.
- Günther, Claudia, Eva Martini, Nadine Wittkopf, Kerstin Amann, Benno Weigmann, Helmut Neumann, Maximilian J. Waldner, et al. 2011. « Caspase-8 Regulates TNF- $\alpha$ -Induced Epithelial Necroptosis and Terminal Ileitis ». *Nature* 477 (7364): 335-39. <https://doi.org/10.1038/nature10400>.
- Günther, Claudia, Helmut Neumann, Markus F. Neurath, et Christoph Becker. 2013. « Apoptosis, Necrosis and Necroptosis: Cell Death Regulation in the Intestinal Epithelium ». *Gut* 62 (7): 1062-71. <https://doi.org/10.1136/gutjnl-2011-301364>.
- Haas, M., A. Loupy, C. Lefaucheur, C. Roufosse, D. Glotz, D. Seron, B. J. Nankivell, et al. 2018. « The Banff 2017 Kidney Meeting Report: Revised diagnostic criteria for chronic active T cell-mediated rejection, antibody-mediated rejection, and prospects for integrative endpoints for next-generation clinical trials ». *American Journal of Transplantation* 18 (2): 293-307. <https://doi.org/10.1111/ajt.14625>.
- Haberman, Yael, Timothy L. Tickle, Phillip J. Dexheimer, Mi-Ok Kim, Dora Tang, Rebekah Karns, Robert N. Baldassano, et al. 2014. « Pediatric Crohn Disease Patients Exhibit Specific Ileal Transcriptome and Microbiome Signature ». *The Journal of Clinical Investigation* 124 (8): 3617-33. <https://doi.org/10.1172/JCI75436>.
- Habtezion, Aida, Linh P. Nguyen, Husein Hadeiba, et Eugene C. Butcher. 2016. « Leukocyte Trafficking to the Small Intestine and Colon ». *Gastroenterology* 150 (2): 340-54. <https://doi.org/10.1053/j.gastro.2015.10.046>.
- Haghverdi, Laleh, Aaron T. L. Lun, Michael D. Morgan, et John C. Marioni. 2018. « Batch effects in single-cell RNA sequencing data are corrected by matching mutual nearest neighbours ». *Nature biotechnology* 36 (5): 421-27. <https://doi.org/10.1038/nbt.4091>.
- Hampsey, Michael. 1998. « Molecular Genetics of the RNA Polymerase II General Transcriptional Machinery ». *Microbiology and Molecular Biology Reviews* 62 (2): 465-503.
- Haque, Ashraful, Jessica Engel, Sarah A. Teichmann, et Tapio Lönnberg. 2017. « A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications ». *Genome Medicine* 9 (1): 75. <https://doi.org/10.1186/s13073-017-0467-4>.
- Heldal, Kristian, Anders Hartmann, Diana C. Grootendorst, Dinanda J. de Jager, Torbjørn Leivestad, Aksel Foss, et Karsten Midtvedt. 2010. « Benefit of Kidney Transplantation beyond 70 Years of Age ». *Nephrology, Dialysis, Transplantation: Official Publication of the European Dialysis and Transplant Association - European Renal Association* 25 (5): 1680-87. <https://doi.org/10.1093/ndt/gfp681>.
- Hendel, Sebastian Kjærgaard, Lauge Kellermann, Annika Hausmann, Niels Bindslev, Kim Bak Jensen, et Ole Haagen Nielsen. 2022. « Tuft Cells and Their Role in Intestinal Diseases ». *Frontiers in Immunology* 13: 822867. <https://doi.org/10.3389/fimmu.2022.822867>.
- Hirsch, Hans H., et David R. Snyderman. 2005. « BK Virus: Opportunity Makes a

- Pathogen ». *Clinical Infectious Diseases* 41 (3): 354-60.  
<https://doi.org/10.1086/431488>.
- Hoffmann, U., B. Banas, B. Krüger, M. Banas, T. Bergler, C. Böger, M. Kammerl, et al. 2006. « SDF-1 Expression Is Elevated in Chronic Human Renal Allograft Rejection ». *Clinical Transplantation* 20 (6): 712-18.  
<https://doi.org/10.1111/j.1399-0012.2006.00540.x>.
- Hon, Chung-Chau, Jordan A. Ramiłowski, Jayson Harshbarger, Nicolas Bertin, Owen J. L. Rackham, Julian Gough, Elena Denisenko, et al. 2017. « An atlas of human long non-coding RNAs with accurate 5' ends ». *Nature* 543 (7644): 199-204. <https://doi.org/10.1038/nature21374>.
- Hotelling, H. 1933. « Analysis of a complex of statistical variables into principal components ». *Journal of Educational Psychology* 24: 417-41.  
<https://doi.org/10.1037/h0071325>.
- Huang, Hailiang, Ming Fang, Luke Jostins, Maša Umićević Mirkov, Gabrielle Boucher, Carl A. Anderson, Vibeke Andersen, et al. 2017. « Fine-Mapping Inflammatory Bowel Disease Loci to Single-Variant Resolution ». *Nature* 547 (7662): 173-78. <https://doi.org/10.1038/nature22969>.
- Huang, Zhaohao, Binyao Chen, Xiuxing Liu, He Li, Lihui Xie, Yuehan Gao, Runping Duan, et al. 2021. « Effects of sex and aging on the immune cell landscape as assessed by single-cell transcriptomic analysis ». *Proceedings of the National Academy of Sciences* 118 (33): e2023216118.  
<https://doi.org/10.1073/pnas.2023216118>.
- Hueber, Wolfgang, Bruce E. Sands, Steve Lewitzky, Marc Vandemeulebroecke, Walter Reinisch, Peter D. R. Higgins, Jan Wehkamp, et al. 2012. « Secukinumab, a Human Anti-IL-17A Monoclonal Antibody, for Moderate to Severe Crohn's Disease: Unexpected Results of a Randomised, Double-Blind Placebo-Controlled Trial ». *Gut* 61 (12): 1693-1700.  
<https://doi.org/10.1136/gutjnl-2011-301668>.
- Hugot, Jean-Pierre, Mathias Chamaillard, Habib Zouali, Suzanne Lesage, Jean-Pierre Cézard, Jacques Belaiche, Sven Almer, et al. 2001. « Association of NOD2 Leucine-Rich Repeat Variants with Susceptibility to Crohn's Disease ». *Nature* 411 (6837): 599-603. <https://doi.org/10.1038/35079107>.
- Hwang, John M, et Madhulika G Varma. 2008. « Surgery for inflammatory bowel disease ». *World Journal of Gastroenterology: WJG* 14 (17): 2678-90.  
<https://doi.org/10.3748/wjg.14.2678>.
- Ip, W. K. Eddie, Namiko Hoshi, Dror S. Shouval, Scott Snapper, et Ruslan Medzhitov. 2017. « Anti-Inflammatory Effect of IL-10 Mediated by Metabolic Reprogramming of Macrophages ». *Science (New York, N. Y.)* 356 (6337): 513-19. <https://doi.org/10.1126/science.aal3535>.
- Islam, Saiful, Amit Zeisel, Simon Joost, Gioele La Manno, Pawel Zajac, Maria Kasper, Peter Lönnerberg, et Sten Linnarsson. 2014. « Quantitative Single-Cell RNA-Seq with Unique Molecular Identifiers ». *Nature Methods* 11 (2): 163-66. <https://doi.org/10.1038/nmeth.2772>.
- Jolliffe, Ian T., et Jorge Cadima. 2016. « Principal Component Analysis: A Review and Recent Developments ». *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences* 374 (2065): 20150202.  
<https://doi.org/10.1098/rsta.2015.0202>.
- Jostins, Luke, Stephan Ripke, Rinse K Weersma, Richard H Duerr, Dermot P McGovern, Ken Y Hui, James C Lee, et al. 2012. « Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease ». *Nature* 491 (7422): 119-24. <https://doi.org/10.1038/nature11582>.

- Jou, W. Min, G. Haegeman, M. Ysebaert, et W. Fiers. 1972. « Nucleotide Sequence of the Gene Coding for the Bacteriophage MS2 Coat Protein ». *Nature* 237 (5350): 82-88. <https://doi.org/10.1038/237082a0>.
- Kabat, Agnieszka M., Johanna Pott, et Kevin J. Maloy. 2016. « The Mucosal Immune System and Its Regulation by Autophagy ». *Frontiers in Immunology* 7. <https://www.frontiersin.org/articles/10.3389/fimmu.2016.00240>.
- Kamada, Nobuhiko, Sang-Uk Seo, Grace Y. Chen, et Gabriel Núñez. 2013. « Role of the Gut Microbiota in Immunity and Inflammatory Disease ». *Nature Reviews Immunology* 13 (5): 321-35. <https://doi.org/10.1038/nri3430>.
- Kang, Hyun Min, Meena Subramaniam, Sasha Targ, Michelle Nguyen, Lenka Maliskova, Elizabeth McCarthy, Eunice Wan, et al. 2018. « Multiplexed Droplet Single-Cell RNA-Sequencing Using Natural Genetic Variation ». *Nature Biotechnology* 36 (1): 89-94. <https://doi.org/10.1038/nbt.4042>.
- Kanodia, Parna, Gurvinder Kaur, Poonam Coshic, Kabita Chatterjee, Teresa Neeman, Anna George, Satyajit Rath, Vineeta Bal, et Savit B. Prabhu. 2019. « Characterization of Biological Variation of Peripheral Blood Immune Cytome in an Indian Cohort ». *Scientific Reports* 9 (1): 14735. <https://doi.org/10.1038/s41598-019-51294-7>.
- Kashima, Yukie, Keiya Kaneko, Patrick Reteng, Nina Yoshitake, Lucky Ronald Runtuwene, Sato Nagasawa, Masaya Onishi, et al. 2022. « Intensive Single-Cell Analysis Reveals Immune-Cell Diversity among Healthy Individuals ». *Life Science Alliance* 5 (7). <https://doi.org/10.26508/lsa.202201398>.
- Kawai, Taro, et Shizuo Akira. 2009. « The roles of TLRs, RLRs and NLRs in pathogen recognition ». *International Immunology* 21 (4): 317-37. <https://doi.org/10.1093/intimm/dxp017>.
- Kenny, Eimear E., Itsik Pe'er, Amir Karban, Laurie Ozelius, Adele A. Mitchell, Sok Meng Ng, Monica Erazo, et al. 2012. « A Genome-Wide Scan of Ashkenazi Jewish Crohn's Disease Suggests Novel Susceptibility Loci ». *PLOS Genetics* 8 (3): e1002559. <https://doi.org/10.1371/journal.pgen.1002559>.
- Kimura, Akihiro, et Tadimitsu Kishimoto. 2010. « IL-6: Regulator of Treg/Th17 Balance ». *European Journal of Immunology* 40 (7): 1830-35. <https://doi.org/10.1002/eji.201040391>.
- Kissmeyer-Nielsen, F., S. Olsen, V. P. Petersen, et O. Fjeldborg. 1966. « Hyperacute Rejection of Kidney Allografts, Associated with Pre-Existing Humoral Antibodies against Donor Cells ». *Lancet (London, England)* 2 (7465): 662-65. [https://doi.org/10.1016/s0140-6736\(66\)92829-7](https://doi.org/10.1016/s0140-6736(66)92829-7).
- Klein, Allon M., Linas Mazutis, Ilke Akartuna, Naren Tallapragada, Adrian Veres, Victor Li, Leonid Peshkin, David A. Weitz, et Marc W. Kirschner. 2015. « Droplet Barcoding for Single-Cell Transcriptomics Applied to Embryonic Stem Cells ». *Cell* 161 (5): 1187-1201. <https://doi.org/10.1016/j.cell.2015.04.044>.
- Kobayashi, Taku, Britta Siegmund, Catherine Le Berre, Shu Chen Wei, Marc Ferrante, Bo Shen, Charles N. Bernstein, Silvio Danese, Laurent Peyrin-Biroulet, et Toshifumi Hibi. 2020. « Ulcerative Colitis ». *Nature Reviews Disease Primers* 6 (1): 1-20. <https://doi.org/10.1038/s41572-020-0205-x>.
- Koh, Ara, Filipe De Vadder, Petia Kovatcheva-Datchary, et Fredrik Bäckhed. 2016. « From Dietary Fiber to Host Physiology: Short-Chain Fatty Acids as Key Bacterial Metabolites ». *Cell* 165 (6): 1332-45. <https://doi.org/10.1016/j.cell.2016.05.041>.
- Kotliar, Dylan, Adrian Veres, M Aurel Nagy, Shervin Tabrizi, Eran Hodis, Douglas A Melton, et Pardis C Sabeti. 2019. « Identifying gene expression programs of cell-type identity and cellular activity with single-cell RNA-Seq ». Édité par

- Alfonso Valencia, Naama Barkai, Elisabetta Mereu, et Berthold Göttgens. *eLife* 8 (juillet): e43803. <https://doi.org/10.7554/eLife.43803>.
- Kuipers, Jack, Katharina Jahn, et Niko Beerwinkel. 2017. « Advances in Understanding Tumour Evolution through Single-Cell Sequencing ». *Biochimica Et Biophysica Acta. Reviews on Cancer* 1867 (2): 127-38. <https://doi.org/10.1016/j.bbcan.2017.02.001>.
- Lakshmikanth, Tadeally, Sayyed Auwn Muhammad, Axel Olin, Yang Chen, Jaromir Mikes, Linn Fagerberg, Anders Gummesson, Göran Bergström, Mathias Uhlen, et Petter Brodin. 2020. « Human Immune System Variation during 1 Year ». *Cell Reports* 32 (3): 107923. <https://doi.org/10.1016/j.celrep.2020.107923>.
- Lange, Katrina M. de, Loukas Moutsianas, James C. Lee, Christopher A. Lamb, Yang Luo, Nicholas A. Kennedy, Luke Jostins, et al. 2017. « Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease ». *Nature genetics* 49 (2): 256-61. <https://doi.org/10.1038/ng.3760>.
- Le Berre, Catherine, Laurent Peyrin-Biroulet, et SPIRIT-IOIBD study group. 2021. « Selecting End Points for Disease-Modification Trials in Inflammatory Bowel Disease: The SPIRIT Consensus From the IOIBD ». *Gastroenterology* 160 (5): 1452-1460.e21. <https://doi.org/10.1053/j.gastro.2020.10.065>.
- Levine, Jonathan S., et Robert Burakoff. 2011. « Extraintestinal Manifestations of Inflammatory Bowel Disease ». *Gastroenterology & Hepatology* 7 (4): 235-41.
- Li, Jing, et Changning Liu. 2019. « Coding or Noncoding, the Converging Concepts of RNAs ». *Frontiers in Genetics* 10. <https://www.frontiersin.org/articles/10.3389/fgene.2019.00496>.
- Liang, Spencer C., Xiang-Yang Tan, Deborah P. Luxenberg, Riyez Karim, Kyriaki Dunussi-Joannopoulos, Mary Collins, et Lynette A. Fouser. 2006. « Interleukin (IL)-22 and IL-17 are coexpressed by Th17 cells and cooperatively enhance expression of antimicrobial peptides ». *The Journal of Experimental Medicine* 203 (10): 2271-79. <https://doi.org/10.1084/jem.20061308>.
- Liebowitz, D. 1998. « Epstein-Barr Virus and a Cellular Signaling Pathway in Lymphomas from Immunosuppressed Patients ». *The New England Journal of Medicine* 338 (20): 1413-21. <https://doi.org/10.1056/NEJM199805143382003>.
- Lim, Mary Ann, Jatinder Kohli, et Roy D. Bloom. 2017. « Immunosuppression for Kidney Transplantation: Where Are We Now and Where Are We Going? ». *Transplantation Reviews (Orlando, Fla.)* 31 (1): 10-17. <https://doi.org/10.1016/j.trre.2016.10.006>.
- Lim, Wai H., Germaine Wong, Helen L. Pilmore, Stephen P. McDonald, et Steven J. Chadban. 2017. « Long-Term Outcomes of Kidney Transplantation in People with Type 2 Diabetes: A Population Cohort Study ». *The Lancet Diabetes & Endocrinology* 5 (1): 26-33. [https://doi.org/10.1016/S2213-8587\(16\)30317-5](https://doi.org/10.1016/S2213-8587(16)30317-5).
- Lim, Wee-Chian, Yongjun Wang, John K MacDonald, et Stephen Hanauer. 2016. « Aminosalicylates for induction of remission or response in Crohn's disease ». *The Cochrane Database of Systematic Reviews* 2016 (7): CD008870. <https://doi.org/10.1002/14651858.CD008870.pub2>.
- Lim, Wei-Wen, Benjamin Ng, Anissa Widjaja, Chen Xie, Liping Su, Nicole Ko, Sze-Yun Lim, et al. 2020. « Transgenic Interleukin 11 Expression Causes Cross-Tissue Fibro-Inflammation and an Inflammatory Bowel Phenotype in Mice ». *PLoS One* 15 (1): e0227505. <https://doi.org/10.1371/journal.pone.0227505>.
- Liu, Jimmy Z, Suzanne van Sommeren, Hailiang Huang, Siew C Ng, Rudi Alberts, Atsushi Takahashi, Stephan Ripke, et al. 2015. « Association analyses



- identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations ». *Nature genetics* 47 (9): 979-86. <https://doi.org/10.1038/ng.3359>.
- López-Cabrera, Manuel, Eduardo Muñoz, M. Valle Blázquez, Maria A. Ursa, Ana G. Santis, et Francisco Sánchez-Madrid. 1995. « Transcriptional Regulation of the Gene Encoding the Human C-Type Lectin Leukocyte Receptor AIM/CD69 and Functional Characterization of Its Tumor Necrosis Factor- $\alpha$ -Responsive Elements (\*) ». *Journal of Biological Chemistry* 270 (37): 21545-51. <https://doi.org/10.1074/jbc.270.37.21545>.
- Loupy, A., M. Haas, K. Solez, L. Racusen, D. Glotz, D. Seron, B. J. Nankivell, et al. 2017. « The Banff 2015 Kidney Meeting Report: Current Challenges in Rejection Classification and Prospects for Adopting Molecular Pathology ». *American Journal of Transplantation* 17 (1): 28-41. <https://doi.org/10.1111/ajt.14107>.
- Loupy, Alexandre, Michael Mengel, et Mark Haas. 2022. « Thirty Years of the International Banff Classification for Allograft Pathology: The Past, Present, and Future of Kidney Transplant Diagnostics ». *Kidney International* 101 (4): 678-91. <https://doi.org/10.1016/j.kint.2021.11.028>.
- Love, Michael I., Wolfgang Huber, et Simon Anders. 2014. « Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2 ». *Genome Biology* 15 (12): 550. <https://doi.org/10.1186/s13059-014-0550-8>.
- Lowe, Rohan, Neil Shirley, Mark Bleackley, Stephen Dolan, et Thomas Shafee. 2017. « Transcriptomics Technologies ». *PLoS Computational Biology* 13 (5): e1005457. <https://doi.org/10.1371/journal.pcbi.1005457>.
- Luecken, Malte D., et Fabian J. Theis. 2019. « Current Best Practices in Single-Cell RNA-Seq Analysis: A Tutorial ». *Molecular Systems Biology* 15 (6): e8746. <https://doi.org/10.15252/msb.20188746>.
- Lun, Aaron T. L., Fernando J. Calero-Nieto, Liora Haim-Vilmovsky, Berthold Göttgens, et John C. Marioni. 2017. « Assessing the Reliability of Spike-in Normalization for Analyses of Single-Cell RNA Sequencing Data ». *Genome Research* 27 (11): 1795-1806. <https://doi.org/10.1101/gr.222877.117>.
- Maaten, Laurens van der. 2014. « Accelerating t-SNE using Tree-Based Algorithms ». *Journal of Machine Learning Research* 15 (93): 3221-45.
- Maaten, Laurens van der, et Geoffrey Hinton. 2008. « Visualizing data using t-SNE ». *Journal of Machine Learning Research* 9 (novembre): 2579-2605.
- MacDonald, T. T., P. Hutchings, M. Y. Choy, S. Murch, et A. Cooke. 1990. « Tumour Necrosis Factor-Alpha and Interferon-Gamma Production Measured at the Single Cell Level in Normal and Inflamed Human Intestine ». *Clinical and Experimental Immunology* 81 (2): 301-5. <https://doi.org/10.1111/j.1365-2249.1990.tb03334.x>.
- Macosko, Evan Z., Anindita Basu, Rahul Satija, James Nemesh, Karthik Shekhar, Melissa Goldman, Itay Tirosh, et al. 2015. « Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets ». *Cell* 161 (5): 1202-14. <https://doi.org/10.1016/j.cell.2015.05.002>.
- Mager, Riccardo, Giulia Roda, Mohammad Khaled Shalaby, et Stefania Vetrano. 2021. « Fibrotic Strictures in Crohn's Disease: Mechanisms and Predictive Factors ». *Current Drug Targets* 22 (2): 241-51. <https://doi.org/10.2174/1389450121666201020160803>.
- Marinov, Georgi K., Brian A. Williams, Ken McCue, Gary P. Schroth, Jason Gertz, Richard M. Myers, et Barbara J. Wold. 2014. « From single-cell to cell-pool transcriptomes: Stochasticity in gene expression and RNA splicing ». *Genome*

- Research* 24 (3): 496-510. <https://doi.org/10.1101/gr.161034.113>.
- Márquez, Eladio J., Cheng-han Chung, Radu Marches, Robert J. Rossi, Djamel Nehar-Belaid, Alper Eroglu, David J. Mellert, George A. Kuchel, Jacques Banchereau, et Duygu Ucar. 2020. « Sexual-Dimorphism in Human Immune System Aging ». *Nature Communications* 11 (1): 751. <https://doi.org/10.1038/s41467-020-14396-9>.
- Martin, J. C., G. Bériou, M. Heslan, C. Bossard, A. Jarry, A. Abidi, P. Hulin, et al. 2016. « IL-22BP Is Produced by Eosinophils in Human Gut and Blocks IL-22 Protective Actions during Colitis ». *Mucosal Immunology* 9 (2): 539-49. <https://doi.org/10.1038/mi.2015.83>.
- Martin, Jerome C., Christie Chang, Gilles Boschetti, Ryan Ungaro, Mamta Giri, John A. Grout, Kyle Gettler, et al. 2019. « Single-Cell Analysis of Crohn's Disease Lesions Identifies a Pathogenic Cellular Module Associated with Resistance to Anti-TNF Therapy ». *Cell* 178 (6): 1493-1508.e20. <https://doi.org/10.1016/j.cell.2019.08.008>.
- Mass, Elvira, Ivan Ballesteros, Matthias Farlik, Florian Halbritter, Patrick Günther, Lucile Crozet, Christian E. Jacome-Galarza, et al. 2016. « Specification of tissue-resident macrophages during organogenesis ». *Science (New York, N.Y.)* 353 (6304): aaf4238. <https://doi.org/10.1126/science.aaf4238>.
- McDole, Jeremiah R., Leroy W. Wheeler, Keely G. McDonald, Baomei Wang, Vjollca Konjufca, Kathryn A. Knoop, Rodney D. Newberry, et Mark J. Miller. 2012. « Goblet Cells Deliver Luminal Antigen to CD103+ Dendritic Cells in the Small Intestine ». *Nature* 483 (7389): 345-49. <https://doi.org/10.1038/nature10863>.
- McInnes, Leland, John Healy, et James Melville. 2018. « UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction ». *arXiv:1802.03426 [cs, stat]*, février. <http://arxiv.org/abs/1802.03426>.
- Medzhitov, Ruslan, et Tiffany Horng. 2009. « Transcriptional Control of the Inflammatory Response ». *Nature Reviews Immunology* 9 (10): 692-703. <https://doi.org/10.1038/nri2634>.
- Menzel, K., M. Hausmann, F. Obermeier, K. Schreiter, N. Dunger, F. Bataille, W. Falk, J. Scholmerich, H. Herfarth, et G. Rogler. 2006. « Cathepsins B, L and D in Inflammatory Bowel Disease Macrophages and Potential Therapeutic Effects of Cathepsin Inhibition in Vivo ». *Clinical and Experimental Immunology* 146 (1): 169-80. <https://doi.org/10.1111/j.1365-2249.2006.03188.x>.
- Mereu, Elisabetta, Atefeh Lafzi, Catia Moutinho, Christoph Ziegenhain, Davis J. McCarthy, Adrián Álvarez-Varela, Eduard Batlle, et al. 2020. « Benchmarking Single-Cell RNA-Sequencing Protocols for Cell Atlas Projects ». *Nature Biotechnology* 38 (6): 747-55. <https://doi.org/10.1038/s41587-020-0469-4>.
- Mizoguchi, Atsushi, Arisa Yano, Hidetomo Himuro, Yui Ezaki, Takayuki Sadanaga, et Emiko Mizoguchi. 2018. « Clinical Importance of IL-22 Cascade in IBD ». *Journal of Gastroenterology* 53 (4): 465-74. <https://doi.org/10.1007/s00535-017-1401-7>.
- Moll, Natalia M., et Richard M. Ransohoff. 2010. « CXCL12 and CXCR4 in Bone Marrow Physiology ». *Expert Review of Hematology* 3 (3): 315-22. <https://doi.org/10.1586/ehm.10.16>.
- Molodecky, Natalie A., Ing Shian Soon, Doreen M. Rabi, William A. Ghali, Mollie Ferris, Greg Chernoff, Eric I. Benchimol, et al. 2012. « Increasing Incidence and Prevalence of the Inflammatory Bowel Diseases With Time, Based on Systematic Review ». *Gastroenterology* 142 (1): 46-54.e42. <https://doi.org/10.1053/j.gastro.2011.10.001>.
- Montgomery, Robert A., Jeffrey M. Stern, Bonnie E. Lonze, Vasishta S. Tatapudi,

- Massimo Mangiola, Ming Wu, Elaina Weldon, et al. 2022. « Results of Two Cases of Pig-to-Human Kidney Xenotransplantation ». *The New England Journal of Medicine* 386 (20): 1889-98.  
<https://doi.org/10.1056/NEJMoa2120238>.
- Moschen, Alexander R., Herbert Tilg, et Tim Raine. 2019. « IL-12, IL-23 and IL-17 in IBD: Immunobiology and Therapeutic Targeting ». *Nature Reviews Gastroenterology & Hepatology* 16 (3): 185-96.  
<https://doi.org/10.1038/s41575-018-0084-8>.
- Murch, S. H., C. P. Braegger, J. A. Walker-Smith, et T. T. MacDonald. 1993. « Location of Tumour Necrosis Factor Alpha by Immunohistochemistry in Chronic Inflammatory Bowel Disease ». *Gut* 34 (12): 1705-9.  
<https://doi.org/10.1136/gut.34.12.1705>.
- Nagalakshmi, Marehalli L., Anne Rascle, Sandra Zurawski, Satish Menon, et Rene de Waal Malefyt. 2004. « Interleukin-22 Activates STAT3 and Induces IL-10 by Colon Epithelial Cells ». *International Immunopharmacology* 4 (5): 679-91.  
<https://doi.org/10.1016/j.intimp.2004.01.008>.
- Nagano, Yuji, Kikuji Itoh, et Kenya Honda. 2012. « The Induction of Treg Cells by Gut-Indigenous Clostridium ». *Current Opinion in Immunology, Host pathogens / Immune senescence*, 24 (4): 392-97.  
<https://doi.org/10.1016/j.coi.2012.05.007>.
- Nemati, Eghlim, Behzad Einollahi, Mahboob Lesan Pezeshki, Vahid Porfarziani, et Mohamad Reza Fattahi. 2014. « Does Kidney Transplantation With Deceased or Living Donor Affect Graft Survival? ». *Nephro-urology Monthly* 6 (4): e12182. <https://doi.org/10.5812/numonthly.12182>.
- Neurath, Markus F. 2014. « Cytokines in Inflammatory Bowel Disease ». *Nature Reviews. Immunology* 14 (5): 329-42. <https://doi.org/10.1038/nri3661>.
- Nie, Kai, Kejia Ma, Weiwei Luo, Zhaohua Shen, Zhenyu Yang, Mengwei Xiao, Ting Tong, Yuanyuan Yang, et Xiaoyan Wang. 2021. « Roseburia intestinalis: A Beneficial Gut Organism From the Discoveries in Genus and Species ». *Frontiers in Cellular and Infection Microbiology* 11.  
<https://www.frontiersin.org/articles/10.3389/fcimb.2021.757718>.
- Nood, Els van, Anne Vrieze, Max Nieuwdorp, Susana Fuentes, Erwin G. Zoetendal, Willem M. de Vos, Caroline E. Visser, et al. 2013. « Duodenal Infusion of Donor Feces for Recurrent Clostridium difficile ». *New England Journal of Medicine* 368 (5): 407-15. <https://doi.org/10.1056/NEJMoa1205037>.
- Oellerich, Michael, Karen Sherwood, Paul Keown, Ekkehard Schütz, Julia Beck, Johannes Stegbauer, Lars Christian Rump, et Philip D. Walson. 2021. « Liquid Biopsies: Donor-Derived Cell-Free DNA for the Detection of Kidney Allograft Injury ». *Nature Reviews Nephrology* 17 (9): 591-603.  
<https://doi.org/10.1038/s41581-021-00428-0>.
- Oishi, Yumiko, Nathanael J. Spann, Verena M. Link, Evan D. Muse, Tobias Strid, Chantle Edillor, Matthew J. Kolar, et al. 2017. « SREBP1 contributes to resolution of pro-inflammatory TLR4 signaling by reprogramming fatty acid metabolism ». *Cell metabolism* 25 (2): 412-27.  
<https://doi.org/10.1016/j.cmet.2016.11.009>.
- Okai, Natsuki, Tomohiro Watanabe, Kosuke Minaga, Ken Kamata, Hajime Honjo, et Masatoshi Kudo. 2022. « Alterations of autophagic and innate immune responses by the Crohn's disease-associated ATG16L1 mutation ». *World Journal of Gastroenterology* 28 (26): 3063-70.  
<https://doi.org/10.3748/wjg.v28.i26.3063>.
- Opelz, Gerhard, et Bernd Döhler. 2007. « Effect of Human Leukocyte Antigen

- Compatibility on Kidney Graft Survival: Comparative Analysis of Two Decades ». *Transplantation* 84 (2): 137-43.  
<https://doi.org/10.1097/01.tp.0000269725.74189.b9>.
- Oppmann, Birgit, Robin Lesley, Bianca Blom, Jackie C. Timans, Yuming Xu, Brisdell Hunte, Felix Vega, et al. 2000. « Novel P19 Protein Engages IL-12p40 to Form a Cytokine, IL-23, with Biological Activities Similar as Well as Distinct from IL-12 ». *Immunity* 13 (5): 715-25. [https://doi.org/10.1016/S1074-7613\(00\)00070-4](https://doi.org/10.1016/S1074-7613(00)00070-4).
- O'Sullivan, Shane, John F. Gilmer, et Carlos Medina. 2015. « Matrix Metalloproteinases in Inflammatory Bowel Disease: An Update ». *Mediators of Inflammation* 2015: 964131. <https://doi.org/10.1155/2015/964131>.
- Paul, Franziska, Ya'ara Arkin, Amir Giladi, Diego Adhemar Jaitin, Ephraim Kenigsberg, Hadas Keren-Shaul, Deborah Winter, et al. 2015. « Transcriptional Heterogeneity and Lineage Commitment in Myeloid Progenitors ». *Cell* 163 (7): 1663-77.  
<https://doi.org/10.1016/j.cell.2015.11.013>.
- Pavlidis, Polychronis, Anastasia Tsakmaki, Eirini Pantazi, Katherine Li, Domenico Cozzetto, Jonathan Digby- Bell, Feifei Yang, et al. 2022. « Interleukin-22 Regulates Neutrophil Recruitment in Ulcerative Colitis and Is Associated with Resistance to Ustekinumab Therapy ». *Nature Communications* 13 (1): 5820.  
<https://doi.org/10.1038/s41467-022-33331-8>.
- Peeters, T., et G. Vantrappen. 1975. « The Paneth Cell: A Source of Intestinal Lysozyme. » *Gut* 16 (7): 553-58. <https://doi.org/10.1136/gut.16.7.553>.
- Pera, A., P. Bellando, D. Caldera, V. Ponti, M. Astegiano, C. Barletti, E. David, A. Arrigoni, G. Rocca, et G. Verme. 1987. « Colonoscopy in Inflammatory Bowel Disease. Diagnostic Accuracy and Proposal of an Endoscopic Score ». *Gastroenterology* 92 (1): 181-85.
- Peral de Castro, Celia, Sarah A. Jones, Clíona Ní Cheallaigh, Claire A. Hearnden, Laura Williams, Jan Winter, Ed C. Lavelle, Kingston H. G. Mills, et James Harris. 2012. « Autophagy Regulates IL-23 Secretion and Innate T Cell Responses through Effects on IL-1 Secretion ». *Journal of Immunology (Baltimore, Md.: 1950)* 189 (8): 4144-53.  
<https://doi.org/10.4049/jimmunol.1201946>.
- Peyrin-Biroulet, L., W. Sandborn, B. E. Sands, W. Reinisch, W. Bemelman, R. V. Bryant, G. D'Haens, et al. 2015. « Selecting Therapeutic Targets in Inflammatory Bowel Disease (STRIDE): Determining Therapeutic Goals for Treat-to-Target ». *Official Journal of the American College of Gastroenterology | ACG* 110 (9): 1324-38.  
<https://doi.org/10.1038/ajg.2015.233>.
- Picarella, D., P. Hurlbut, J. Rottman, X. Shi, E. Butcher, et D. J. Ringler. 1997. « Monoclonal Antibodies Specific for Beta 7 Integrin and Mucosal Addressin Cell Adhesion Molecule-1 (MAdCAM-1) Reduce Inflammation in the Colon of Scid Mice Reconstituted with CD45RBhigh CD4+ T Cells ». *Journal of Immunology (Baltimore, Md.: 1950)* 158 (5): 2099-2106.
- Piovani, Daniele, Silvio Danese, Laurent Peyrin-Biroulet, Georgios K. Nikolopoulos, Theodore Lytras, et Stefanos Bonovas. 2019. « Environmental Risk Factors for Inflammatory Bowel Diseases: An Umbrella Review of Meta-Analyses ». *Gastroenterology* 157 (3): 647-659.e4.  
<https://doi.org/10.1053/j.gastro.2019.04.016>.
- Poggio, Emilio D., Joshua J. Augustine, Susana Arrigain, Daniel C. Brennan, et Jesse D. Schold. 2021. « Long-Term Kidney Transplant Graft Survival—Making Progress When Most Needed ». *American Journal of Transplantation*

- 21 (8): 2824-32. <https://doi.org/10.1111/ajt.16463>.
- Polito, J. M., B. Childs, E. D. Mellits, A. Z. Tokayer, M. L. Harris, et T. M. Bayless. 1996. « Crohn's Disease: Influence of Age at Diagnosis on Site and Clinical Type of Disease ». *Gastroenterology* 111 (3): 580-86. <https://doi.org/10.1053/gast.1996.v111.pm8780560>.
- Posselt, Andrew M., Flavio Vincenti, Melanie Bedolli, Marianne Lantz, John P. Roberts, et Ryutaro Hirose. 2003. « CD69 Expression on Peripheral CD8 T Cells Correlates with Acute Rejection in Renal Transplant Recipients ». *Transplantation* 76 (1): 190-95. <https://doi.org/10.1097/01.TP.0000073614.29680.A8>.
- Powell, Nick, Eirini Pantazi, Polychronis Pavlidis, Anastasia Tsakmaki, Katherine Li, Feifei Yang, Aimee Parker, et al. 2020. « Interleukin-22 Orchestrates a Pathological Endoplasmic Reticulum Stress Response Transcriptional Programme in Colonic Epithelial Cells ». *Gut* 69 (3): 578-90. <https://doi.org/10.1136/gutjnl-2019-318483>.
- Pruenster, Monika, Liesbeth Mudde, Paula Bombosi, Svetla Dimitrova, Marion Zsak, Jim Middleton, Ann Richmond, et al. 2009. « The Duffy Antigen Receptor for Chemokines Transports Chemokines and Supports Their Promigratory Activity ». *Nature Immunology* 10 (1): 101-8. <https://doi.org/10.1038/ni.1675>.
- Quévrain, E., M. A. Maubert, C. Michon, F. Chain, R. Marquant, J. Tailhades, S. Miquel, et al. 2016. « Identification of an Anti-Inflammatory Protein from Faecalibacterium Prausnitzii, a Commensal Bacterium Deficient in Crohn's Disease ». *Gut* 65 (3): 415-25. <https://doi.org/10.1136/gutjnl-2014-307649>.
- Rama, Inés, et Josep M. Grinyó. 2010. « Malignancy after Renal Transplantation: The Role of Immunosuppression ». *Nature Reviews. Nephrology* 6 (9): 511-19. <https://doi.org/10.1038/nrneph.2010.102>.
- Reeve, Jeff, Georg A. Böhmig, Farsad Eskandary, Gunilla Einecke, Carmen Lefaucheur, Alexandre Loupy, et Philip F. Halloran. s. d. « Assessing rejection-related disease in kidney transplant biopsies based on archetypal analysis of molecular phenotypes ». *JCI Insight* 2 (12): e94197. <https://doi.org/10.1172/jci.insight.94197>.
- Regev, Aviv, Sarah A. Teichmann, Eric S. Lander, Ido Amit, Christophe Benoist, Ewan Birney, Bernd Bodenmiller, et al. 2017. « The Human Cell Atlas ». *ELife* 6 (décembre): e27041. <https://doi.org/10.7554/eLife.27041>.
- Robinson, James, Dominic J Barker, Xenia Georgiou, Michael A Cooper, Paul Flicek, et Steven G E Marsh. 2020. « IPD-IMGT/HLA Database ». *Nucleic Acids Research* 48 (D1): D948-55. <https://doi.org/10.1093/nar/gkz950>.
- Rogler, Gerhard. 2020. « Efficacy of JAK Inhibitors in Crohn's Disease ». *Journal of Crohn's & Colitis* 14 (Supplement\_2): S746-54. <https://doi.org/10.1093/ecco-jcc/jjz186>.
- Rohlke, Faith, et Neil Stollman. 2012. « Fecal microbiota transplantation in relapsing Clostridium difficile infection ». *Therapeutic Advances in Gastroenterology* 5 (6): 403-20. <https://doi.org/10.1177/1756283X12453637>.
- Rosenberg, Alexander B., Charles M. Roco, Richard A. Muscat, Anna Kuchina, Paul Sample, Zizhen Yao, Lucas Gray, et al. 2018. « SPLiT-seq reveals cell types and lineages in the developing brain and spinal cord ». *Science (New York, N.Y.)* 360 (6385): 176-82. <https://doi.org/10.1126/science.aam8999>.
- Rosenberg, Laura, Garrett O. Lawlor, Talia Zenlea, Jeffrey D. Goldsmith, Anne Gifford, Kenneth R. Falchuk, Jacqueline L. Wolf, Adam S. Cheifetz, Simon C. Robson, et Alan C. Moss. 2013. « Predictors of Endoscopic Inflammation in Patients with Ulcerative Colitis in Clinical Remission ». *Inflammatory Bowel*

- Diseases* 19 (4): 779-84. <https://doi.org/10.1097/MIB.0b013e3182802b0e>.
- Rosenfeld, Nitzan, Jonathan W. Young, Uri Alon, Peter S. Swain, et Michael B. Elowitz. 2005. « Gene Regulation at the Single-Cell Level ». *Science* 307 (5717): 1962-65. <https://doi.org/10.1126/science.1106914>.
- Roufosse, Candice, Naomi Simmonds, Marian Clahsen-van Groningen, Mark Haas, Kammi J. Henriksen, Catherine Horsfield, Alexandre Loupy, et al. 2018. « A 2018 Reference Guide to the Banff Classification of Renal Allograft Pathology ». *Transplantation* 102 (11): 1795-1814. <https://doi.org/10.1097/TP.0000000000002366>.
- Rozenblatt-Rosen, Orit, Aviv Regev, Philipp Oberdoerffer, Tal Nawy, Anna Hupalowska, Jennifer E. Rood, Orr Ashenberg, et al. 2020. « The Human Tumor Atlas Network: Charting Tumor Transitions across Space and Time at Single-Cell Resolution ». *Cell* 181 (2): 236-49. <https://doi.org/10.1016/j.cell.2020.03.053>.
- Rutgeerts, Paul, William J. Sandborn, Brian G. Feagan, Walter Reinisch, Allan Olson, Jewel Johanns, Suzanne Travers, et al. 2005. « Infliximab for Induction and Maintenance Therapy for Ulcerative Colitis ». *New England Journal of Medicine* 353 (23): 2462-76. <https://doi.org/10.1056/NEJMoa050516>.
- Samaniego, Rafael, Ángeles Domínguez-Soto, Manohar Ratnam, Takami Matsuyama, Paloma Sánchez-Mateos, Ángel L. Corbí, et Amaya Puig-Kröger. 2020. « Folate Receptor  $\beta$  (FR $\beta$ ) Expression in Tissue-Resident and Tumor-Associated Macrophages Associates with and Depends on the Expression of PU.1 ». *Cells* 9 (6): 1445. <https://doi.org/10.3390/cells9061445>.
- Sands, Bruce E., Jingjing Chen, Brian G. Feagan, Mark Penney, William A. Rees, Silvio Danese, Peter D. R. Higgins, et al. 2017. « Efficacy and Safety of MEDI2070, an Antibody Against Interleukin 23, in Patients With Moderate to Severe Crohn's Disease: A Phase 2a Study ». *Gastroenterology* 153 (1): 77-86.e6. <https://doi.org/10.1053/j.gastro.2017.03.049>.
- Sands, Bruce E., Laurent Peyrin-Biroulet, Edward V. Loftus, Silvio Danese, Jean-Frédéric Colombel, Murat Törüner, Laimas Jonaitis, et al. 2019. « Vedolizumab versus Adalimumab for Moderate-to-Severe Ulcerative Colitis ». *New England Journal of Medicine* 381 (13): 1215-26. <https://doi.org/10.1056/NEJMoa1905725>.
- Sands, Bruce E., William J. Sandborn, Remo Panaccione, Christopher D. O'Brien, Hongyan Zhang, Jewel Johanns, Omoniyi J. Adedokun, et al. 2019. « Ustekinumab as Induction and Maintenance Therapy for Ulcerative Colitis ». *The New England Journal of Medicine* 381 (13): 1201-14. <https://doi.org/10.1056/NEJMoa1900750>.
- Sasaki, Makoto, D. Ostanin, J. W. Elrod, T. Oshima, P. Jordan, M. Itoh, T. Joh, A. Minagar, et J. S. Alexander. 2003. « TNF-Alpha -Induced Endothelial Cell Adhesion Molecule Expression Is Cytochrome P-450 Monooxygenase Dependent ». *American Journal of Physiology. Cell Physiology* 284 (2): C422-428. <https://doi.org/10.1152/ajpcell.00271.2002>.
- Satsangi, J, M S Silverberg, S Vermeire, et J-F Colombel. 2006. « The Montreal classification of inflammatory bowel disease: controversies, consensus, and implications ». *Gut* 55 (6): 749-53. <https://doi.org/10.1136/gut.2005.082909>.
- Schafer, Sebastian, Sivakumar Viswanathan, Anissa A. Widjaja, Wei-Wen Lim, Aida Moreno-Moral, Daniel M. DeLaughter, Benjamin Ng, et al. 2017. « IL-11 Is a Crucial Determinant of Cardiovascular Fibrosis ». *Nature* 552 (7683): 110-15. <https://doi.org/10.1038/nature24676>.
- Schena, Mark, Dari Shalon, Ronald W. Davis, et Patrick O. Brown. 1995.

- « Quantitative Monitoring of Gene Expression Patterns with a Complementary DNA Microarray ». *Science* 270 (5235): 467-70.  
<https://doi.org/10.1126/science.270.5235.467>.
- Schleier, Lena, Maximilian Wiendl, Karin Heidbreder, Marie-Theres Binder, Raja Atreya, Timo Rath, Emily Becker, et al. 2020. « Non-Classical Monocyte Homing to the Gut via A4 $\beta$ 7 Integrin Mediates Macrophage-Dependent Intestinal Wound Healing ». *Gut* 69 (2): 252-63. <https://doi.org/10.1136/gutjnl-2018-316772>.
- Schmitt, Heike, Ulrike Billmeier, Walburga Dieterich, Timo Rath, Sophia Sonnewald, Stephen Reid, Simon Hirschmann, et al. 2019. « Expansion of IL-23 receptor bearing TNFR2+ T cells is associated with molecular resistance to anti-TNF therapy in Crohn's disease ». *Gut* 68 (5): 814-28.  
<https://doi.org/10.1136/gutjnl-2017-315671>.
- Schmitt, Heike, Markus F. Neurath, et Raja Atreya. 2021. « Role of the IL23/IL17 Pathway in Crohn's Disease ». *Frontiers in Immunology* 12 (mars): 622934.  
<https://doi.org/10.3389/fimmu.2021.622934>.
- Schowengerdt, K. O., F. J. Fricker, K. S. Bahjat, et S. T. Kuntz. 2000. « Increased Expression of the Lymphocyte Early Activation Marker CD69 in Peripheral Blood Correlates with Histologic Evidence of Cardiac Allograft Rejection ». *Transplantation* 69 (10): 2102-7. <https://doi.org/10.1097/00007890-200005270-00023>.
- Schraufstatter, Ingrid U, Ming Zhao, Sophia K Khaldoyanidi, et Richard G DiScipio. 2012. « The chemokine CCL18 causes maturation of cultured monocytes to macrophages in the M2 spectrum ». *Immunology* 135 (4): 287-98.  
<https://doi.org/10.1111/j.1365-2567.2011.03541.x>.
- Sedano, J, R Swamy, K Jain, et S Gupta. 2015. « Brunner's gland hamartoma of the duodenum ». *Annals of The Royal College of Surgeons of England* 97 (5): e70-72. <https://doi.org/10.1308/003588415X14181254790040>.
- Segerstolpe, Åsa, Athanasia Palasantza, Pernilla Eliasson, Eva-Marie Andersson, Anne-Christine Andréasson, Xiaoyan Sun, Simone Picelli, et al. 2016. « Single-Cell Transcriptome Profiling of Human Pancreatic Islets in Health and Type 2 Diabetes ». *Cell Metabolism* 24 (4): 593-607.  
<https://doi.org/10.1016/j.cmet.2016.08.020>.
- Sellon, Rance K., Susan Tonkonogy, Michael Schultz, Levinus A. Dieleman, Wetonia Grenther, Ed Balish, Donna M. Rennick, et R. Balfour Sartor. 1998. « Resident Enteric Bacteria Are Necessary for Development of Spontaneous Colitis and Immune System Activation in Interleukin-10-Deficient Mice ». *Infection and Immunity* 66 (11): 5224-31.
- Selsted, M. E., S. I. Miller, A. H. Henschen, et A. J. Ouellette. 1992. « Enteric Defensins: Antibiotic Peptide Components of Intestinal Host Defense ». *The Journal of Cell Biology* 118 (4): 929-36. <https://doi.org/10.1083/jcb.118.4.929>.
- Shalek, Alex K., et Mikael Benson. 2017. « Single-cell analyses to tailor treatments ». *Science Translational Medicine* 9 (408): eaan4730.  
<https://doi.org/10.1126/scitranslmed.aan4730>.
- Shen, Zhaohua, Weiwei Luo, Bei Tan, Kai Nie, Minzi Deng, Shuai Wu, Mengwei Xiao, et al. 2022. « Roseburia Intestinalis Stimulates TLR5-Dependent Intestinal Immunity against Crohn's Disease ». *EBioMedicine* 85 (novembre).  
<https://doi.org/10.1016/j.ebiom.2022.104285>.
- Sherman, Michael, Donald N. Tsynman, Albert Kim, Jyoti Arora, Timothy Pietras, Susan Messing, Lydia St Hilaire, et al. 2014. « Sustained Improvement in Health-Related Quality of Life Measures in Patients with Inflammatory Bowel Disease Receiving Prolonged Anti-Tumor Necrosis Factor Therapy ». *Journal*

- of *Digestive Diseases* 15 (4): 174-79. <https://doi.org/10.1111/1751-2980.12125>.
- Shivaswamy, Vijay, Brian Boerner, et Jennifer Larsen. 2016. « Post-Transplant Diabetes Mellitus: Causes, Treatment, and Impact on Outcomes ». *Endocrine Reviews* 37 (1): 37-61. <https://doi.org/10.1210/er.2015-1084>.
- Silverberg, Mark S., Jack Satsangi, Tariq Ahmad, Ian D. R. Arnott, Charles N. Bernstein, Steven R. Brant, Renzo Caprilli, et al. 2005. « Toward an Integrated Clinical, Molecular and Serological Classification of Inflammatory Bowel Disease: Report of a Working Party of the 2005 Montreal World Congress of Gastroenterology ». *Canadian Journal of Gastroenterology = Journal Canadien De Gastroenterologie* 19 Suppl A (septembre): 5A-36A. <https://doi.org/10.1155/2005/269076>.
- Sitruk, Lola, Cécile Couchoud, Maryvonne Hourmant, Philippe Tuppin, Marie-Alice Macher, et Camille Legeai. 2018. « Description des traitements immunosuppresseurs d'entretien post-greffe rénale à travers le système national de l'Assurance maladie ». *Néphrologie & Thérapeutique* 14 (7): 523-30. <https://doi.org/10.1016/j.nephro.2018.03.004>.
- Slack, Frank J., et Arul M. Chinnaiyan. 2019. « The Role of Non-coding RNAs in Oncology ». *Cell* 179 (5): 1033-55. <https://doi.org/10.1016/j.cell.2019.10.017>.
- Smillie, Christopher S., Moshe Biton, Jose Ordovas-Montanes, Keri M. Sullivan, Grace Burgin, Daniel B. Graham, Rebecca H. Herbst, et al. 2019. « Intra- and Inter-Cellular Rewiring of the Human Colon during Ulcerative Colitis ». *Cell* 178 (3): 714-730.e22. <https://doi.org/10.1016/j.cell.2019.06.029>.
- Sokol, H., P. Seksik, J. P. Furet, O. Firmesse, I. Nion-Larmurier, L. Beaugerie, J. Cosnes, G. Corthier, P. Marteau, et J. Doré. 2009. « Low Counts of Faecalibacterium Prausnitzii in Colitis Microbiota ». *Inflammatory Bowel Diseases* 15 (8): 1183-89. <https://doi.org/10.1002/ibd.20903>.
- Souza, Heitor S. P. de, et Claudio Fiocchi. 2016. « Immunopathogenesis of IBD: Current State of the Art ». *Nature Reviews Gastroenterology & Hepatology* 13 (1): 13-27. <https://doi.org/10.1038/nrgastro.2015.186>.
- Stoeckius, Marlon, Shiwei Zheng, Brian Houck-Loomis, Stephanie Hao, Bertrand Z. Yeung, William M. Mauck, Peter Smibert, et Rahul Satija. 2018. « Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics ». *Genome Biology* 19 (1): 224. <https://doi.org/10.1186/s13059-018-1603-1>.
- Stuart, Tim, Andrew Butler, Paul Hoffman, Christoph Hafemeister, Efthymia Papalexi, William M. Mauck, Yuhan Hao, Marlon Stoeckius, Peter Smibert, et Rahul Satija. 2019. « Comprehensive integration of single-cell data ». *Cell* 177 (7): 1888-1902.e21. <https://doi.org/10.1016/j.cell.2019.05.031>.
- Svensson, Valentine, Roser Vento-Tormo, et Sarah A. Teichmann. 2018. « Exponential Scaling of Single-Cell RNA-Seq in the Past Decade ». *Nature Protocols* 13 (4): 599-604. <https://doi.org/10.1038/nprot.2017.149>.
- Tait Wojno, Elia D., et David Artis. 2012. « Innate Lymphoid Cells: Balancing Immunity, Inflammation, and Tissue Repair in the Intestine ». *Cell Host & Microbe* 12 (4): 445-57. <https://doi.org/10.1016/j.chom.2012.10.003>.
- Tan, Tze Guan, Esen Sefik, Naama Geva-Zatorsky, Lindsay Kua, Debdut Naskar, Fei Teng, Lesley Pisman, et al. 2016. « Identifying Species of Symbiont Bacteria from the Human Gut That, Alone, Can Induce Intestinal Th17 Cells in Mice ». *Proceedings of the National Academy of Sciences of the United States of America* 113 (50): E8141-50. <https://doi.org/10.1073/pnas.1617460113>.
- Tang, Fuchou, Catalin Barbacioru, Ellen Nordman, Nanlan Xu, Vladimir I Bashkirov,



- Kaiqin Lao, et M. Azim Surani. 2010. « RNA-Seq analysis to capture the transcriptome landscape of a single cell ». *Nature protocols* 5 (3): 10.1038/nprot.2009.236. <https://doi.org/10.1038/nprot.2009.236>.
- Tang, Fuchou, Catalin Barbacioru, Yangzhou Wang, Ellen Nordman, Clarence Lee, Nanlan Xu, Xiaohui Wang, et al. 2009. « MRNA-Seq Whole-Transcriptome Analysis of a Single Cell ». *Nature Methods* 6 (5): 377-82. <https://doi.org/10.1038/nmeth.1315>.
- Targan, Stephan R., Brian Feagan, Severine Vermeire, Remo Panaccione, Gil Y. Melmed, Carol Landers, Dalin Li, et al. 2016. « A Randomized, Double-Blind, Placebo-Controlled Phase 2 Study of Brodalumab in Patients With Moderate-to-Severe Crohn's Disease ». *Official Journal of the American College of Gastroenterology | ACG* 111 (11): 1599-1607. <https://doi.org/10.1038/ajg.2016.298>.
- Targan, Stephan R., Stephen B. Hanauer, Sander J.H. van Deventer, Lloyd Mayer, Daniel H. Present, Tanja Braakman, Kimberly L. DeWoody, Thomas F. Schaible, et Paul J. Rutgeerts. 1997. « A Short-Term Study of Chimeric Monoclonal Antibody cA2 to Tumor Necrosis Factor  $\alpha$  for Crohn's Disease ». *New England Journal of Medicine* 337 (15): 1029-36. <https://doi.org/10.1056/NEJM199710093371502>.
- Torres, Joana, Saurabh Mehandru, Jean-Frédéric Colombel, et Laurent Peyrin-Biroulet. 2017. « Crohn's Disease ». *The Lancet* 389 (10080): 1741-55. [https://doi.org/10.1016/S0140-6736\(16\)31711-1](https://doi.org/10.1016/S0140-6736(16)31711-1).
- Travassos, Leonardo H., Leticia A. M. Carneiro, Mahendrasingh Ramjeet, Seamus Hussey, Yun-Gi Kim, João G. Magalhães, Linda Yuan, et al. 2010. « Nod1 and Nod2 Direct Autophagy by Recruiting ATG16L1 to the Plasma Membrane at the Site of Bacterial Entry ». *Nature Immunology* 11 (1): 55-62. <https://doi.org/10.1038/ni.1823>.
- Tremaroli, Valentina, et Fredrik Bäckhed. 2012. « Functional Interactions between the Gut Microbiota and Host Metabolism ». *Nature* 489 (7415): 242-49. <https://doi.org/10.1038/nature11552>.
- Trinchieri, Giorgio. 1995. « Interleukin-12: A Proinflammatory Cytokine with Immunoregulatory Functions that Bridge Innate Resistance and Antigen-Specific Adaptive Immunity ». *Annual Review of Immunology* 13 (1): 251-76. <https://doi.org/10.1146/annurev.iy.13.040195.001343>.
- Turner, Dan, Amanda Ricciuto, Ayanna Lewis, Ferdinando D'Amico, Jasbir Dhaliwal, Anne M. Griffiths, Dominik Bettenworth, et al. 2021. « STRIDE-II: An Update on the Selecting Therapeutic Targets in Inflammatory Bowel Disease (STRIDE) Initiative of the International Organization for the Study of IBD (IOIBD): Determining Therapeutic Goals for Treat-to-Target Strategies in IBD ». *Gastroenterology* 160 (5): 1570-83. <https://doi.org/10.1053/j.gastro.2020.12.031>.
- Turner, Jan-Eric, Martina Becker, Hans-Willi Mittrücker, et Ulf Panzer. 2018. « Tissue-Resident Lymphocytes in the Kidney ». *Journal of the American Society of Nephrology : JASN* 29 (2): 389-99. <https://doi.org/10.1681/ASN.2017060599>.
- Ungaro, Ryan, Saurabh Mehandru, Patrick B Allen, Laurent Peyrin-Biroulet, et Jean-Frédéric Colombel. 2017. « Ulcerative colitis ». *Lancet (London, England)* 389 (10080): 1756-70. [https://doi.org/10.1016/S0140-6736\(16\)32126-2](https://doi.org/10.1016/S0140-6736(16)32126-2).
- Uo, Michihide, Tadakazu Hisamatsu, Jun Miyoshi, Daiki Kaito, Kazuaki Yoneno, Mina T. Kitazume, Maiko Mori, et al. 2013. « Mucosal CXCR4+ IgG Plasma Cells Contribute to the Pathogenesis of Human Ulcerative Colitis through

- FcγR-Mediated CD14 Macrophage Activation ». *Gut* 62 (12): 1734-44. <https://doi.org/10.1136/gutjnl-2012-303063>.
- Uzzan, Mathieu, Jerome C. Martin, Luka Mesin, Alexandra E. Livanos, Tomas Castro-Dopico, Ruiqi Huang, Francesca Petralia, et al. 2022. « Ulcerative Colitis Is Characterized by a Plasmablast-Skewed Humoral Response Associated with Disease Activity ». *Nature Medicine* 28 (4): 766-79. <https://doi.org/10.1038/s41591-022-01680-y>.
- Vaishnava, Shipra, Miwako Yamamoto, Kari M. Severson, Kelly A. Ruhn, Xiaofei Yu, Omry Koren, Ruth Ley, Edward K. Wakeland, et Lora V. Hooper. 2011. « The Antibacterial Lectin RegIIIγ Promotes the Spatial Segregation of Microbiota and Host in the Intestine ». *Science (New York, N.Y.)* 334 (6053): 255-58. <https://doi.org/10.1126/science.1209791>.
- Valdes, Ana M., Jens Walter, Eran Segal, et Tim D. Spector. 2018. « Role of the Gut Microbiota in Nutrition and Health ». *BMJ* 361 (juin): k2179. <https://doi.org/10.1136/bmj.k2179>.
- Valeri, Maria, et Manuela Raffatellu. 2016. « Cytokines IL-17 and IL-22 in the host response to infection ». *Pathogens and Disease* 74 (9): ftw111. <https://doi.org/10.1093/femspd/ftw111>.
- Van Assche, G., Séverine Vermeire, et Paul Rutgeerts. 2010. « The Potential for Disease Modification in Crohn's Disease ». *Nature Reviews Gastroenterology & Hepatology* 7 (2): 79-85. <https://doi.org/10.1038/nrgastro.2009.220>.
- Vavricka, Stephan R., Lionel Brun, Pierluigi Ballabeni, Valérie Pittet, Bettina Mareike Prinz Vavricka, Jonas Zeitz, Gerhard Rogler, et Alain M. Schoepfer. 2011. « Frequency and Risk Factors for Extraintestinal Manifestations in the Swiss Inflammatory Bowel Disease Cohort ». *The American Journal of Gastroenterology* 106 (1): 110-19. <https://doi.org/10.1038/ajg.2010.343>.
- Verstockt, Bram, Sare Verstockt, Brecht Creyns, Sophie Tops, Gert Van Assche, Ann Gils, Jan L. Ceuppens, Séverine Vermeire, Marc Ferrante, et Christine Breynaert. 2019. « Mucosal IL13RA2 expression predicts nonresponse to anti-TNF therapy in Crohn's disease ». *Alimentary Pharmacology & Therapeutics* 49 (5): 572-81. <https://doi.org/10.1111/apt.15126>.
- Viladomiu, Monica, Charles Kivolowitz, Ahmed Abdulhamid, Belgin Dogan, Daniel Victorio, Jim G. Castellanos, Viola Woo, et al. 2017. « IgA-coated E. coli enriched in Crohn's Disease Spondyloarthritis Promote Th17-dependent Inflammation ». *Science translational medicine* 9 (376): eaaf9655. <https://doi.org/10.1126/scitranslmed.aaf9655>.
- Villani, Alexandra-Chloé, Rahul Satija, Gary Reynolds, Siranush Sarkizova, Karthik Shekhar, James Fletcher, Morgane Griesbeck, et al. 2017. « Single-Cell RNA-Seq Reveals New Types of Human Blood Dendritic Cells, Monocytes, and Progenitors ». *Science (New York, N.Y.)* 356 (6335): eaah4573. <https://doi.org/10.1126/science.aah4573>.
- Vincenti, Flavio, Lionel Rostaing, Joseph Grinyo, Kim Rice, Steven Steinberg, Luis Gaité, Marie-Christine Moal, et al. 2016. « Belatacept and Long-Term Outcomes in Kidney Transplantation ». *New England Journal of Medicine* 374 (4): 333-43. <https://doi.org/10.1056/NEJMoa1506027>.
- Vivier, Eric, David Artis, Marco Colonna, Andreas Diefenbach, James P. Di Santo, Gérard Eberl, Shigeo Koyasu, et al. 2018. « Innate Lymphoid Cells: 10 Years On ». *Cell* 174 (5): 1054-66. <https://doi.org/10.1016/j.cell.2018.07.017>.
- Vlachos, Christoforos, Georgios Gaitanis, Konstantinos H. Katsanos, Dimitrios K. Christodoulou, Epameinondas Tsianos, et Ioannis D. Bassukas. 2016. « Psoriasis and Inflammatory Bowel Disease: Links and Risks ». *Psoriasis (Auckland, N.Z.)* 6: 73. <https://doi.org/10.2147/PTT.S85194>.

- Vogel, Andrea, Julia Stefanie Brunner, Alexander Hajto, Omar Sharif, et Gernot Schabbauer. 2022. « Lipid Scavenging Macrophages and Inflammation ». *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids* 1867 (1): 159066. <https://doi.org/10.1016/j.bbaliip.2021.159066>.
- Waldner, Maximilian J., et Markus F. Neurath. 2014. « Master Regulator of Intestinal Disease: IL-6 in Chronic Inflammation and Cancer Development ». *Seminars in Immunology* 26 (1): 75-79. <https://doi.org/10.1016/j.smim.2013.12.003>.
- Wang, Dingzhi, Raymond N. DuBois, et Ann Richmond. 2009. « The role of chemokines in intestinal inflammation and cancer ». *Current opinion in pharmacology* 9 (6): 688-96. <https://doi.org/10.1016/j.coph.2009.08.003>.
- Wang, Zhong, Mark Gerstein, et Michael Snyder. 2009. « RNA-Seq: A Revolutionary Tool for Transcriptomics ». *Nature Reviews Genetics* 10 (1): 57-63. <https://doi.org/10.1038/nrg2484>.
- Wehkamp, J., J. Harder, M. Weichenthal, M. Schwab, E. Schäffeler, M. Schlee, K. R. Herrlinger, et al. 2004. « NOD2 (CARD15) Mutations in Crohn's Disease Are Associated with Diminished Mucosal  $\alpha$ -Defensin Expression ». *Gut* 53 (11): 1658-64. <https://doi.org/10.1136/gut.2003.032805>.
- Wéra, Odile, Patrizio Lancellotti, et Cécile Oury. 2016. « The Dual Role of Neutrophils in Inflammatory Bowel Diseases ». *Journal of Clinical Medicine* 5 (12): 118. <https://doi.org/10.3390/jcm5120118>.
- Werfel, T., M. Boeker, et A. Kapp. 1997. « Rapid Expression of the CD69 Antigen on T Cells and Natural Killer Cells upon Antigenic Stimulation of Peripheral Blood Mononuclear Cell Suspensions ». *Allergy* 52 (4): 465-69. <https://doi.org/10.1111/j.1398-9995.1997.tb01031.x>.
- West, C. E., M. C. Jenmalm, et S. L. Prescott. 2015. « The Gut Microbiota and Its Role in the Development of Allergic Disease: A Wider Perspective ». *Clinical and Experimental Allergy: Journal of the British Society for Allergy and Clinical Immunology* 45 (1): 43-53. <https://doi.org/10.1111/cea.12332>.
- West, Nathaniel R., Ahmed N. Hegazy, Benjamin M. J. Owens, Samuel J. Bullers, Bryan Linggi, Sofia Buonocore, Margherita Coccia, et al. 2017. « Oncostatin M Drives Intestinal Inflammation and Predicts Response to Tumor Necrosis Factor-Neutralizing Therapy in Patients with Inflammatory Bowel Disease ». *Nature Medicine* 23 (5): 579-89. <https://doi.org/10.1038/nm.4307>.
- Wolfe, R. A., V. B. Ashby, E. L. Milford, A. O. Ojo, R. E. Ettenger, L. Y. Agodoa, P. J. Held, et F. K. Port. 1999. « Comparison of Mortality in All Patients on Dialysis, Patients on Dialysis Awaiting Transplantation, and Recipients of a First Cadaveric Transplant ». *The New England Journal of Medicine* 341 (23): 1725-30. <https://doi.org/10.1056/NEJM199912023412303>.
- Wolk, Kerstin, Stefanie Kunz, Khusru Asadullah, et Robert Sabat. 2002. « Cutting Edge: Immune Cells as Sources and Targets of the IL-10 Family Members? ». *Journal of Immunology (Baltimore, Md.: 1950)* 168 (11): 5397-5402. <https://doi.org/10.4049/jimmunol.168.11.5397>.
- Wolk, Kerstin, Stefanie Kunz, Ellen Witte, Markus Friedrich, Khusru Asadullah, et Robert Sabat. 2004. « IL-22 Increases the Innate Immunity of Tissues ». *Immunity* 21 (2): 241-54. <https://doi.org/10.1016/j.immuni.2004.07.007>.
- Yamawaki, Tracy M., Daniel R. Lu, Daniel C. Ellwanger, Dev Bhatt, Paolo Manzanillo, Vanessa Arias, Hong Zhou, et al. 2021. « Systematic comparison of high-throughput single-cell RNA-seq methods for immune cell profiling ». *BMC Genomics* 22 (1): 66. <https://doi.org/10.1186/s12864-020-07358-4>.
- Yan, Jun, Jing Yu, Ke Liu, Yijia Liu, Changchun Mao, et Wenda Gao. 2021. « The Pathogenic Roles of IL-22 in Colitis: Its Transcription Regulation by Musculin

- in T Helper Subsets and Innate Lymphoid Cells ». *Frontiers in Immunology* 12. <https://www.frontiersin.org/articles/10.3389/fimmu.2021.758730>.
- Yellowley, Clare. 2013. « CXCL12/CXCR4 signaling and other recruitment and homing pathways in fracture repair ». *BoneKEy Reports 2* (mars): 300. <https://doi.org/10.1038/bonekey.2013.34>.
- Zhang, Xuejing, Milton H. Hamblin, et Ke-Jie Yin. 2017. « The long noncoding RNA Malat1: Its physiological and pathophysiological functions ». *RNA Biology* 14 (12): 1705-14. <https://doi.org/10.1080/15476286.2017.1358347>.
- Zhang, Yang, Li Yang, et Ling-Ling Chen. 2014. « Life without A Tail: New Formats of Long Noncoding RNAs ». *The International Journal of Biochemistry & Cell Biology* 54 (septembre): 338-49. <https://doi.org/10.1016/j.biocel.2013.10.009>.
- Zhao, Hong, Yanting Dong, Xinrui Tian, Thian Kui Tan, Zhuola Liu, Ye Zhao, Yun Zhang, David Ch Harris, et Guoping Zheng. 2013. « Matrix Metalloproteinases Contribute to Kidney Fibrosis in Chronic Kidney Diseases ». *World Journal of Nephrology* 2 (3): 84-89. <https://doi.org/10.5527/wjn.v2.i3.84>.
- Zhao, Mirabella, Lóránt Gönczi, Peter L Lakatos, et Johan Burisch. 2021. « The Burden of Inflammatory Bowel Disease in Europe in 2020 ». *Journal of Crohn's and Colitis* 15 (9): 1573-87. <https://doi.org/10.1093/ecco-jcc/jjab029>.
- Zheng, G. X., J. M. Terry, P. Belgrader, P. Ryvkin, Z. W. Bent, R. Wilson, S. B. Ziraldo, T. D. Wheeler, G. P. McDermott, et J. Zhu. 2017. « Massively parallel digital transcriptional profiling of single cells ». *Nat Commun* 8. <https://doi.org/10.1038/ncomms14049>.
- Zhou, Guang Xi, et Zhan Ju Liu. 2017. « Potential Roles of Neutrophils in Regulating Intestinal Mucosal Inflammation of Inflammatory Bowel Disease ». *Journal of Digestive Diseases* 18 (9): 495-503. <https://doi.org/10.1111/1751-2980.12540>.
- Zhu, Sibó, Tao Qing, Yuanting Zheng, Li Jin, et Leming Shi. 2017. « Advances in single-cell RNA sequencing and its applications in cancer research ». *Oncotarget* 8 (32): 53763-79. <https://doi.org/10.18632/oncotarget.17893>.
- Ziegler, Steven F., Fred Ramsdell, et Mark R. Alderson. 1994. « The activation antigen CD69 ». *Stem Cells* 12 (5): 456-65. <https://doi.org/10.1002/stem.5530120502>.

# Annexes

## Annexe 1 : Liste des 150 modules identifiés dans le modèle supergut

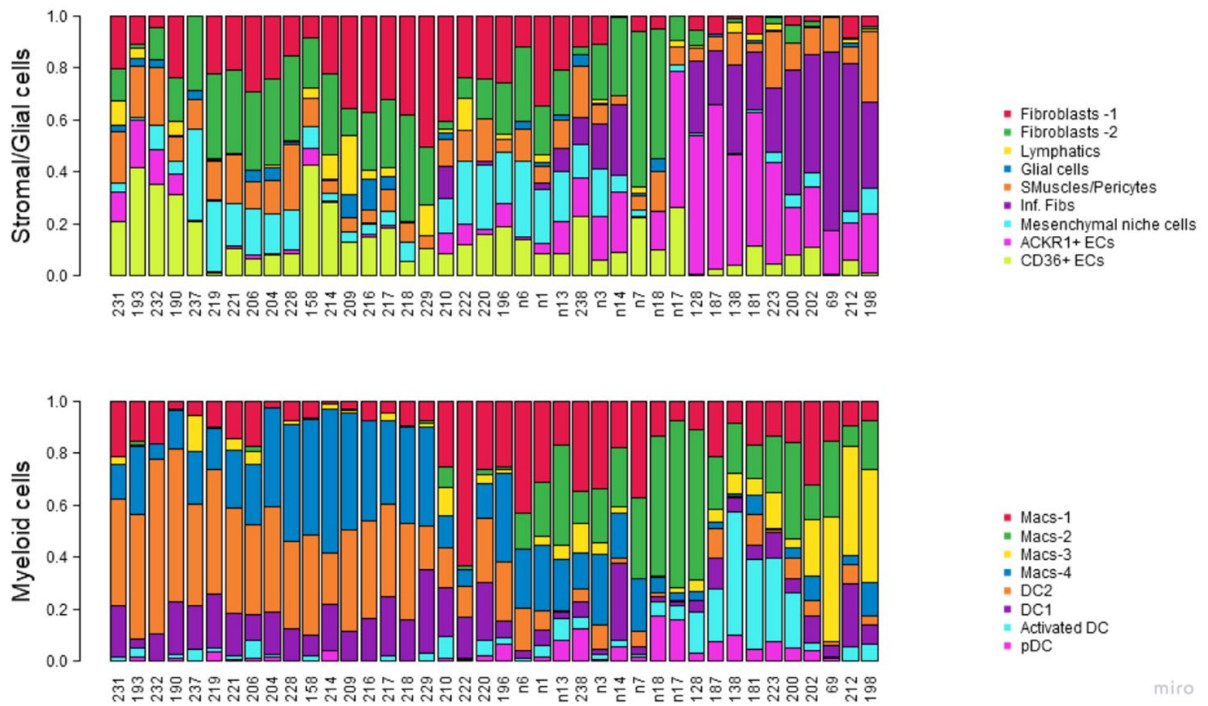
Mod_1	AL645608.7,AL139246.5,TNFRSF14-AS1,PLPP3,CSF1,ADAMTSL4-AS1,DYF,TET3,GBE1,TBC1D7,TAGAP,AL356417.3,AP521,ZNF316,SLC25A13,AC007032.1,TMEM140,N5MAF,AC079209.1,SLC45A4,LINCO2649,ZNF438,PLAU,STX3,AC023157.3,RESF1,VP518,CLDN7,AC099489.1,DNAH17,FAM210A,EPG5,GDF15,CCNT2,LINCO2154
Mod_2	HE54,GBP5,GCLM,IVNS1,ABF1,IF1,BC16,DDX60L,SGTB,ERBIN,AC034199.1,EDN1,STEAP4,EGR3,NB,N,RALGDS,PP1F,NP5A,CAAMK2,CTC2,N4B,P1,VP9S1D,SCARF1,CPD,TRIM25,AC0515912.3,PGS1,TNFSF14,FPR2,BX640514.2,HELZ2,LIMK2,APOL6
Mod_3	TNFRSF18,TNFRSF4,RAB13,SLAMF9,TGFA,CKX19,CKX10,CKX11,AL021978.1,ETV7,TPST1,CAV1,IL2RA,ARID5B,ANKRD22,CLEC6A,CYP27B1,CSRP2,BATF,IL27,VMY1,ALOX15B,C17orf58,RBF0X3,FSTL3,TNFRSF6B,MIR503HG
Mod_4	MIB2,DFFA,EPHA2,ATP13A2,LINCO1226,TUT4,LINCO2812,PPM11,PKFB2,OPN3,ZBTB18,SMC6,AC007381.1,BCL11A,CDBL11,SCN9A,CHPF,ITM2C,HIGD1A,MANF,SIDT1,EPHB1,CCDC5Q,OCIA2,STAR4D-AS1,MZB1,CXXC5,TXNDC5,NOTCH4,PTCRA,KLC4,PTK7,RUNX2,AL096865.1,CARD11,CDC47,NM8E,ABP1,MDF1,CREB3L2,PNO,C,AC023590.1,ZFAT,SIT1,TPM2,TEM8B,UGCG,AL138756.1,SUSD1,STRBP,ST6GALN,AC4,CLIC3,FUT7,LRR,C26B,LNK,CCDC186,JRF7,ECT,EP58L2,NUCB2,CYB561A3,VEGFB,PP1R14B,PPP1R14B-AS1,TM75F2,GRAMD18,ST3GAL4,CLEC4C,PCED1B-AS1,ATG101,ZDHHC17,C12orf75,ALOX5AP,PCDH9,GAS6,GZMB,PLEKH1D,AC004846.1,TCL1A,CYP46A1,PLD4,FAM330A,SNRPN,SNHG14,MAP1A,MYO1E,SCAMPS,CIB2,TNFRSF17,GG2,SEPTIN1,C,CCDC189,LRC36,SMPD3,SERPINF1,TRAF4,ABHD15,PHB,TEX2,MAP2,K6,SM,IMS,SETBP1,TCF4,TMIGD2,CD320,AC008764.7,NIBAN3,FUZ,SPB,LILRA4,LAMP5,ASIP,SLA2,TOX2,LIME1,APP,YPEL1,DELR3,GPM6B,MAEGA2,CXCR3
Mod_5	MMP23B,FLNB,SEL1L3,SOX4,CD2AP,TSPAN13,BLK,POLB,SEPHS1,NUMT2B-AS1,TP53I11,NECTIN1,SLC15A4,EGLN3,MNAT1,SLC12A3,NEK8,LDRAD4,CHAF1A,MYBL2,IL3RA
Mod_6	NADK,CUX1,ARID3A,CYTH4
Mod_7	RNF207,TNFSF4,HIST1H2AC,HIST1H2BJ,HIST1H3H,GN11,CD5,LINCO1146,TOB1,TRPM4
Mod_8	GRP153,H6PD,MTFR,LINCO1358,LOR,TNNT2,HSD11B1-AS1,EDARDD,SLC5A6,TARB2A,TTN-AS1,ZBTB38,PDGFRA,HOPX,ARHGAP10,LINCO2362,SLC22A23,ADTRP,UBD,AFDN,EPDR1,TFPI2,BX25923.2,AL162390.1,DNM1,DEPP1,ARHGAP22,ANK3,TSPAN15,ADAM12,DD2,PA,C51,TLCD5,WNT5B,CDK2,LINCO2384,NA2S5,TRAFD1,BICDL1,ABCC4,IL21R,CC117,C,DYL2,ENO3,SPEC11,MAP3K14,SNX11,ALPK2,ZNF260,MIA,AJ009632.2,APOBEC3H,CB7X,JADE3,ATP11C
Mod_9	ACOT7,SGO1,SMIM14,SEPTIN11,ZNF827,TIFAB,REPIN1,PSIP1,LINCO2397
Mod_10	KLHL21,UBR4,RLF,SEC22B,MAP4K4,CCDC93,RVBP,LPCAT1,AC010359.3,SEC24A,CREBRF,CPEB4,RABGEF1,CD82,SLC43A3,MAML2,RAB21,NU58,LINCO0641,DENN4A,FLCN,MLL2E,SPAG9,MAFG,ADNP2
Mod_11	TNFRSF9,NIBAN1,NABP1,ARA2P,ANXA3,SYNPO2,STRIP2,CD274,SLC44A1,PALM2-AKAP2,RAB8B,TRIP10,ADORA2A
Mod_12	ERRF1,PDNP,TM45F1,SUCNR1,CXCL1,CXCL5,MYO10,AKAP12,AC009630.4,CDKN2B,MMP10,MMP1,CASP5,AC015871.7,AC106028.4,DNAAF1,CCL2,CCL8,KCNK4,TIAM1
Mod_13	GRP157,MARCKS11,GADD45A,SIPR1,LAO1,NUAK2,MBOAT2,LV75,LINCO1857,MREG,RAMP1,ST3GAL6,CD200,POGLUT1,MGLL,SIAH2,LAM P3,MAP3K13,LJMCH1,GRSF1,CCNG2,TEMEM150C,DAPP1,IL15,ANKRD33B,KIF2A,ST8IIA4,SYNPO,ANXA6,TREM1,LPLEKH1,FSCN1,ACHE,TEMEM176A,ACO1,NUB1,PSD3,SLCO5A1,ZBTB10,PCDD1L2,CCL19,NM1R1,RASFA4,H2AFY2,SWAP70,PRR4,RAB30,BCL2L114,ARNTL2,TUBA1C,SOCS2,KDM2B,TBC1D4,SFPL2A,IL32,TVP23A,NDE1,AC026401.3,CCL22,CCR7,L3M,BT4,TNFRSF11A,HMSD,CYB5A,EBI3,CD70,NCCR1,PVR,RELB,INSM1,ADORA2A-AS1,UBP1,U62317.1,CRF2,RAB9A
Mod_14	RBP7,AGTRAP,CD4,MTMR11,S100A9,S100A12,S100A8,MN,DM,AR1,AC096667.1,E1F4E3,FBX15,BST1,TEMEM154,VN2,CREB5,CD36,SMARCD3,TMEM71,MEGF9,NU2P14,FCN1,KCN E3,PLB D1,MGST1,NFE2L1A4H,RNASE2,RAB27A,TGAM,CE1,ASGR1,RET,N,AC011472.2,RAB3D,TEMEM91,VTM1,TNNT1,IL17R,AL034397.3
Mod_15	CENPF,CLSPN,CD20,CKS1B,ASPM,UB2T,CENPF,RRM2,SGO2,SMC4,BAN K3,CENPE,MAD2L1,CCNA2,CENPU,CENPK,CCN B1,DHFR,PTTG1,HM MR,GMN,HIST1H4C,KIFC1,CENPW,FBXO5,RA3,ATAO2,SMC2,WDR34,ZWINT,CDK1,DNAIC9,KIF11,HELLS,MK167,CCDC34,M5441,H2AFX,RAD51AP1,TEMEM106C,TEX30,CDKN3,NUSAP1,WDR76,CCN B2,PCLAF,FANCI,PRC1,SHCB8,C,CT1.1,URK8,TOP2A,BCR1,T,K1,BIRC5,TYMS,UHRF1,ASF1B,CD79A,PCNA,TPX2,UBE2C,VPRB3,CENPM,GTSE1
Mod_16	AL021155.5,OFDFL2,AC245014.3,AC239799.2,AC009404.1,ZDBF2,KZF2,FBXW11,TRIM73,PALD1,ZNF891,ARID3B,LINCO2207,SDR4E2,TMEM107,CTC11,GLO N5,SPAG4,Z93241.1,AL031595.3
Mod_17	TNFRSF8,FCGR3A,SHISA5,MAPKAPK3,HK3,GRK6,LFNG,PILRA,ZNF703,CDH23,TCF7L2,CDKN1C,AC009559.1,BIN2,SHISA1,CALML4,CD79B,PECAM1,AC064805.1,SLC9A3R1,ADGRE1,RRAS,LILRA1,PLA2G2,SMIM25
Mod_18	PRDM2,ETV3,KDMSB,GLS,ANTXR2,HMGCS1,MALT1,PTPN1,AL021707.6
Mod_19	AL450998.2,AC253572.2,SORL1,NUDT4,WDFY2,LINCO2256,RLNB,AC007952.4
Mod_20	PAD12,CK2,CD2,LAX1,LINCO1871,IBH,CDBA,LINCO1943,CHN1,CD28,CTLA4,PAK3,TRAT1,TIGIT,SPIN K2,CXCL13,SPARCL1,HPGD,AC114977.1,CSF2,CFB,FOXPA,TRGC2,TRB C1,TRBC2,RARRES2,C,LU,T,CM,SPOCK2,CAVIN3,CD3E,C,D9D,C,D3G,C,D27,AC092821.3,KLRB1,IFNG,CLU105,TRAC,BCL11B,IGHG,IGHM,RASGRP1,RORAN,MB,TPSB2,CCL5,AC025048.4,C,D7,MIAT,ITM2A,SH2D1A
Mod_21	PAD14,F5,AL445524.1,CYP1B1,VCAN,CD14,VNN1,PTGER2,CRISPLD2,PLCB1,CYBB
Mod_22	PLA2G2D,CTSK,FCGR2B,CHI3L1,HSD11B1,FN1,SLC9B2,SPARC,LPL,CA2,MMP12,BHLHE41,RND1,AC020656.1,TEMEM258,CCL13,CCL18,SI,GLEC15,HAMP,MMP9
Mod_23	ECE1,B3GNT2,TNIP1,ADAM19,OPTN,DENND5A,AL62717.1.2,FURIN,SNB2,SLC5A3
Mod_24	C1QA,C1QC,C1QB,FUC1,SDC3,TM1GB3,ADORA3,SHE,EPHX1,ERTK,TMEM37,SLC40A1,DIH1,SRGAP3,ITGA9,STAB1,FRMD4B,TGB5,ABCC5,PDGFC,ANKH,SELEN OF,CHO2,ME1,ADGRG6,PDK4,CTT,NBP2,MSR1,GFR4,ADAMDEC1,WVW1,ENG,FRMD4A,CXCL12,LIPA,AP2A2,LVYE1,MS444A,FOLR2,SLCO2B1,CD163L1,A2M1,GF1,CMKLR1,SLC7A8,SLC38A6,SPRED1,GATM,HEXA,MAF,P,LXDC1,LGALS3BP,CD209,PLD3,AXLL1L1R8,PLTP,GRP34,VSIG4
Mod_25	ID3,RAB42,LMN A,B,3GN7,RB P1,GPX3,TREM2,GNM B,FABP5,FABP4,ATP6V0D2,SCD,CD59,CD9,TSC22D1,RNASE1,OTOA,NUPR1,FAM20A,ACPF5,AC092069.1,APOE,APOC1,APOC2,HSD17B14,TCFAL9,BEX3,AC004596.3
Mod_26	CLIC4,GBP3,DYRK3,PGAP1,SGPP2,SLC7A11,ELOVL7,PIM1,TEMEM217,AL355304.1,AC002456.1,KDM7A-DT,MFHAS1,AC083837.1,LRP12,TJP2,MSANTD3,C11orf96,CLCF1,AC005280.1,ABHD17C,HS3T3B1,AC005224.2,STR2,RB8P,MIR155HG,ZFX,STK26
Mod_27	RUNX3,LIMNB1,RALA,SNHG15
Mod_28	STMN1,PARP1,MCM7,KIF20B,TNNT2,CDK2A2,PEBP1,SNRN P25,NME4
Mod_29	PAFAH2,CDCC18-AS1,BIN1,ZDHHC4,GGH,PTGDS,AL135925.1,SVIP,TEMEM109,P2R Y6,GSEC,ST14,PCED1B,TEMEM19,MLEC,APEX1,KCTD5,SPNS3,AC005726.5,HOXB2,CUEDC1,MIF4GD,SEC11C,PLVAP,MAP4K1,TASP1,B,CAS4,BIK,CXorf21
Mod_30	UBXN11,IFI6,IFI44,QPCT,CSTA,CHST13,FAM200B,S100Z,ZNF467,NACC2,NRGN,AC139768.1,HAL,CARS2,RNASE4,SLC244A,PSTPIP1,MSR1,SU1T1A1,NLRP12,ME1,MAP3K7C,L
Mod_31	SYT11,C1orf162,ARHGAP30,RGS18,TRAF3IP3,CNIH4,AC009950.1,AC1147067.1,PDLM2,HACD4,HEX,CAHLM2,C11orf21,TSPAN3B,RIN1,TBCL1D3C,TEC,C,RPPE1,TPP3,ZDHHC1,DPEP1,CAMF2,CD300C,RAB37,CD300LF,S1PR4,ICAM3,LRR C25,POU2F2,C20orf27,RGS19,AATBC,CFP,NHSL2
Mod_32	FGR,NCF2,RXRA,DENN D10,CASP1,CLEC12A,PYGL,SERPINA1,OSCAR,LILRA5
Mod_33	PTAFR,FCGR2A,ADM,CLEC4E,MLKL,PRNP
Mod_34	AL360012.1,ADCY3,RANBP2,PPP1R3B,RALGAPA1,PLC G2,USP36,TNFSF9,TSIX
Mod_35	CCDC28B,TST1D,CAFE,NAV1,CEP68,GCSAM,VOPP1,TSPAN33,PLGRKT,GSN,PTGES2,SIN HCAF,RHOF,ARPP19,M,MD,CLPTM1,TOM M34,ASMTL P2RY10,CLIC2
Mod_36	RNF198,MTF1,GBP2,FAM102B,SLAMF1,LAMB3,RHOU,AC079753.2,IL1RN,MAP3K20,AGF1,EAF1,TEMEM158,GK5,SGMS2,TNIP3,MIR3945HG,LUCAT1,ERGIC1,PPARD,C2109,DSE,UPP1,KDM7A,TNFRSF10D,GPAT4,TRPS1,TNFSF15,ZBTB43,NIBAN2,LIPN,SLC2A3,GPR C5A,EMP1,GRP84,OASL,FLT1,ACOD1,DNAJC3-DT,PNF,AC025580.2,ANPEP,C,HSY1,MIR22HG,LRG1,RIN2,ACSL4
Mod_37	SH3D21,AC092803.1,AC104695.4,AC079905.3,SA7B1-AS1,GRIN3,PHACTR1,AL109615.2,AL355881.1,AL499604.1,TEMEM273,ATG2A,DDIT3,CYSLTR2,SETDB2,AC004817.3,VA5H1,APOB,AC009093.2,HIC1,AC091271.1,MR C2,GZF1,CASS4,LINCO0528,AL118516.1,SCML1
Mod_38	LSM10,CCDC12,CCND3,SH2D3C,CAT,RASGRP2,SPN,AC004687.1,CDKN2D
Mod_39	CSFR3,CCR1,AC072022.2,S100P,HERC5,VN N3,DDX5B,IFI T3,LINCO1001,BATF2,AP5B1,GN G2,PLEKH2,Z,XAF1,MCOLN1,FFAR2,RALGAPA2,KCN J15
Mod_40	ZC3H12A,MFSD2A,NOTCH2NLC,NBPF19,SLAMF7,ABL2,XMO,FOSL2,IL1R1,SH3BP5,SKIL,RASGEF1B,TE2,RARGEF2,ZFYVE16,TNFAI3P,HIVEP2,PURB,AC016831.7,AC016831.1,AZIN1-AS1,ABCA1,ARL5B,ZMIZ1,NCR3LG1,AC124798.1,FNDC3A,SAV1,MAPK6,KDM6B,CYB5D1,DOT1L,ICAM1,SRC
Mod_41	R,RAGC,RHOH,CYSTM1,ZNF292,TRIB1,HEL8
Mod_42	SLC2A1,MARCB9,ST3GAL2,CSF2RB
Mod_43	ARMH1,CCR2,ARHGAP24,EMB3,CCDC69,C9orf139,C16orf54,SYNGR1
Mod_44	PLK3,MXD1,IRAK2,CCR L2,FNDC3B,B3GNT5,JLIRAP,CXCL8,TLR2,ACSL1,ARHGAP26,TREM1,VEGFA,PNPLA8,DOCK4,SLC25A37,CSGALNACT2,ITPRIP,SLC11A2,JRAK3,SUSD6,AQ,P9,MEFV,DCUN1D3,TGAX,KCNJ2,FCAR,ETS2,TOM1,GK,LONRF3
Mod_45	BTD19,NOTCH2NLAJL10,TNFAIP6,SESTD1,CCL20,PTX3,UBTD2,AGPAT4,CCDC71L,LY6K,NCS1,FOSL1,ZC3H12C,GALNT6,IL23A,HAP2,PHLDA1,GJ82,MPHOSP6,SPAG5,SLC16A6,AC015802.6,SERPIN2,SCN1B,HAS1,AL121899.4,PLCXD1,MIR22HG
Mod_46	GRP7,ODC1,AFB3,P2RY14,P3H2,DANCR,DCK,NREP,FZD3,SNHG7,PAUX,DHTKD1,SLC38A1,MYL6B,C12orf45,JDH3A,ADGRG5,CBFA2T3,ACAP1,GN G7,LINCO1684,OFD1,SLC9A7
Mod_47	LRP8,CEP350,CER6,CALCRLL,RFTN1,MAP4,CBLB,LRCH3,CEP135,MOB1B,TM2A2,ENPP4,WAKMAR2,AUTS2,ZEB1,SH3PKX2A,SMNDC1,DUSP5,UVRAG,RAB5B,ATP6VOA2,N4BP2L1,LRK1,ZC3H7A,CHD3,AC008105.3,DSG2,FAM118A
Mod_48	TACSTD2,VCAM1,GSTM4,LINCO1357,TSPAN2,C1orf54,AIM2,C1orf21,RAB7B,FKB P1B,PLCD1,XCR1,CLNK,CAMK2D,DB N1,TMEM14A,ENPP1,AC005162.3,TOX,LRRC1,SHNTN1,CCND1,CADM1,CLEC9A,CLEC1A,PAWR,SNX22,LINCO2206,KIF16B
Mod_49	NFIA,WLS,NPL,KLHDC8B,F13A1,TSPAN4,ACP2,TIMP2,DNASE2
Mod_50	AK4,CXCL3,CXC L2,SPF1,SLC39A8,LINCO0243,DLL1,TWISTN B,TG88,IL6,GSDME,TRGC1,INHBA,CLEC5A,HEY1,SDC2,PHLDA2,BCAT1,LINCO0346,C15orf48,CEMIP,KIFC3,CCL3,CCL4,CCL3L1,CCL4L2,CSF3,TGM2,GGT5,AL031590.1
Mod_51	PDE4B,PKFB3,ZNF267,SOCS3,MAFF
Mod_52	DNAJB4,RGS16,HK2,PDK1,TN11,CDCP1,ALAS1,ANKRD37,ARRDC3,HSPA11,ZFAN D2A,HSPB1,SERPINE1,HILPDA,DEFB1,CTNNAL1,ERSL1,P4HA1,BNIP3,SERPINH1,CHORDC1,FKBP4,P2RX7,ERO1A,HIF1A-AS3,SNACP1,HSPA2,DNAJA4,STARDS,RRAD,SPAG5-AS1,C17orf67,SH2D3A
Mod_53	AC103591.3,CEP170,EML4,KDM3A,KANS1L,FEM1C,NDRG1,AL354733.3,PCNXN1,IMJ06,PLEKHG2,RUNX1,Z93930.2,PIGA
Mod_54	IFI44L,IPOR2,OAS2,TNFSF13
Mod_55	MCOLN2,ANKRD36C,ANKRD36,BCL2L11,TMEM39A,TENT4A,CLINT1,AC058791.1,PDE7A,ATF7IP,GRASP,CDK17,FBRS1L,BRC2A,ARHGFE7,ICOSLG,PIM2
Mod_56	GBP1,GOS2,NEURL3,IGHG2,IGHG1,O,SM,GMK-IT1,AC112496.1
Mod_57	GBP4,RASGRP3,CFLAR-AS1,TPRA1,NFKBIE,PRDM1,STX4,TGIF1,BCL2
Mod_58	RNBP1L,RGS13,IL18R1,SH3BP4,H1FX,SPON2,STAP1,HIST1H1C,HIST1H3D,HIST1H1D,KCN K17,TM6D132,MYB,AKR1C3,SEC61A2,FX YD2,CDH1,CDK5R1,U BE25,AL118508.1
Mod_59	F3ZNF697,PDE4D,IL6R,PTGS2,NLRP3,LI1A,LI1B,PLD1,ATP13A3,LINCO0884,EREG,PRDM8,EL2L2,MIR3142HG,SERPINBP1,JEK3,TNF,SLC16A10,AC007384.1,LINCO2605,NRIP3,AMPD3,AL133330.1,EHD1,OLR1,TGA5,DRAM1,P2RX4,SIP1A1L1,IL4R,AC248329.4,SPHK1,PLAUR,SDC4,B4GALT5,SLCO4A1

Mod_60	DPVD,C9orf72,TLE4,TLR4,CHST15,LRP1,DUSP6,ATP2B1-AS1,AC005280.2,SECTM1,KCNE1
Mod_61	GSTM3,COL13A1,IGFBP5,HRH1,HS3ST1,FAM13A,CXCL14,MYL1,NEU1,CD24,COL1A2,CALD1,RNF122,TP53INP1,C15,LIU,DCN,MMP2,CCL11,COL1A1,RCAN1,COL18A1,COL6A2,A4GALT,DDX3Y
Mod_62	RHOC,RABGAP1L,PRADC1,SRD5A3,FAM153CP,TMEM134,UNC119,NAC42,AP2A1,MAGED2,EIF1A1Y
Mod_63	TF2,STK17A,AC016405.1,SQLE,ETS1,SLC41A2,TUBB6,ATF5,IL2RG
Mod_64	ACPE,RSAD2,CXCR2,PROK2,AC09343.3,AC084871.1,AL095590.2,AC083862.3,IFIT2,ERLIN1,IFITM1,AC136475.5,LINC02422,HCAR2,HCAR3,KATNB1,ZNF429,ZBP1
Mod_65	NBPF14,XRN1,RB,IC1,ASPH,PHF20L1,PIK3AP1,SPATA13,MAP2K3,MX2
Mod_66	FCGR1A,TFEC,PLIN2,C3AR1,RAB20,YPEL2
Mod_67	HIST2H2A4,KMT2E-AS1,SYNE2,EIF4A3
Mod_68	C1orf56,PSENE2,PCD13,EGOT,ZNF366,ACOT13,BAG2,RCL1,NR6A1,NET1,MDK,PLEKHA5,TMEM120B,FLT3,PNMA1,RASGRF1,AC015871.1,MYH11,VAC14,IZUMO4
Mod_69	TNFAIP8L2,CAMK1,P2RY13,SCIMP,ABI3,CD300A,CD33
Mod_70	TCHH,IL6R-AS1,DUSP2,PTGER4,HOMER1,AC008771.1,LINC01010,INSG1,FGFR1L,PKN,ATF7IP2,SLC26A11,GLA
Mod_71	S100A13,CD84,FMNL2,BMP2K,PLAGL1,GADD45G,AL162414.1,EMP2,IGFBP4
Mod_72	ARHGFE2,ZNF281,JMY,ST3GAL1,KCNO10T1,TMCC3,SNN,NPIP84,TTYH2,JAG1
Mod_73	CD1D,LY86,RAB32,CLEC4A
Mod_74	AL138899.1,AC092809.2,CNRIPI,CLEC4F,RND3,ETV5,RSG12,CRHB,LTCA5,AKR1B1,GF1B,MACORIS,RHOBTB1,EGR2,ZNF503,IL18,SCN4B,GRIP1,STON2,QPRT,CLEC10A,ACE,AC124283.1,FCER2,CLEC11A,H1FO,TTYT14
Mod_75	CD1A,CD1C,CD1B,CD1E,FCER1A,CD207,CACNA2D3,GPAT3,ENHO,C19orf33
Mod_76	SLAMF8,PAPOLG,HESI,ZNF141,RHOBTB3,GM2A,GLC,SASH1,FAM20C,TTYH3,AHR,KLF9,TMEM236,MRC1,IL18BP,CD163,CPM,SD5,AC109630.1,SLC38A7,ZFXH3,C5AR2,FPR3,SIGLEC1,ITSN1
Mod_77	PEA15,TUBB2A,C6orf223,AC007991.2,MSC-AS1,MSC,FSD1L,SLC1A2,ISG20,LYPD3,PTGIR,RIPOR3
Mod_78	LY9,TARBP1,LINC00342,AC092683.1,ARL4C,GNL3,DIPK2A,FAM53C,MAN1A1,DNAJB9,CLN8,BCL7A,EIF2AK4,DDX27,USP11,CLCN5
Mod_79	ALS92295.5,AL136987.1,CACNB4,AC006230.1,ZDHHC14,ARL4A,HOXA5,GAL3ST4,DNAJC5B,ENPP2,AC067930.8,ADIRF,CH25H,PRAP1,APOC3,APOA1,LOH12CR2,PKP2,STX2,RAMP2,FOXRED2
Mod_80	OLFML2B,RGL1,MI7F,SLC1A3,C2,PLA2G7,SMPD13A,GIMAP5,SGP1,KCNMA1,CD81,NR1H3,SERPING1,TCN2
Mod_81	CREG1,TMEM176B
Mod_82	ATP1B1,MIR181A1,HG,HAVCR2,NRARP,AVP1,METTL15
Mod_83	SLC19A2,SLC4A7,TRC,SH3PKD2B,DAPK1,ITPR2,SRGAP1,LACC1,FMN1,TRPM2
Mod_84	SELL,PLP2
Mod_85	RGS1,TPD52,PTMS
Mod_86	AL136454.1,BATF3,ARF5,ZDHHC12,CLN51A,HSB2D,APOBEC3G,UPK3A
Mod_87	DENN1B,DST,ARHGAP5,MYEF2,DMD
Mod_88	PPP1R15B,TP53BP2,ARL8B,TSC22D2,STARD4,AGO2,ZFC3H1,NCOR2,GPR132,NFATS,MED13,SMAD7,PDE4A,BCL3
Mod_89	RAB29,ST3GAL5,TBC1D8,COMMD2,GTFE2,GRHR,CBWD5,SESN3,LINC00173,SPG11,TRADD,ZNF296,RASSF2
Mod_90	SRGAP2,ANKRD28,SATB1,AREG,SPRY1,GFOD1,REV3L,SYTL3,HIPK2,UBASH3B,CNCD2,INSR
Mod_91	MAPKAPK2,OGT
Mod_92	CR1,LRK2,ARHGFE4O,CTMT2
Mod_93	TRAF5,LPP,NSUN7,TRIO,CRYBG1,FAM53B,GOLGA8A,GOLGA8B,WDR59,CCDC144A,AC005332.4,AC100786.1
Mod_94	ATF3,SLC8A1,NRA42,EPB4112,SGK1,ADAM28,ZNF331
Mod_95	VASH2,LPIN1,INPP4A,LINC01215,RUBCN,CENPC,GAB1,MCC,IRF4,CDYL,MDN1,AH1,GLCC1,ENTPD7,FADS1,FC HSD2,CEP126,HYOU1,CORO1C,ZNF10,AL355075.4,SLC25A29,FAM214A,RASD1,C BX4,CTDP1,SLC35E1,MED26,ZNF506,ACTN4,RRBP1,PMEPAL1,TSPYL2
Mod_96	CAPN2,TUBA4A,LINC01506,IFIT1,IFIT5,C15orf39,OLIG1
Mod_97	LYST,BLVRA,MTSS1,CD42EP4
Mod_98	GPR137B,SPRED2,NR2P,CD80,IL7R,DUSP4,IDO1,IDO2,SLC6A12,HAPLN3,NECTIN2,IL4I1,EBF4,CST7,CD40
Mod_99	ERO1B,CHML,SLC20A1,GOLGB1,GN12,LINC-PINT,FMNL3,BANP,ERN1
Mod_100	LINC00299,RAPGEF4,PLK2,ZNF165,IL22RA2,LSMEM1,GEM,ACY3,CD69,AC005480.1,IGHA2,IGHG4,IGHA1,IGHG3,AC018362.1,PIF1,ALDOC,CEACAM1,MYL9,S100B,TNFRSF13C
Mod_101	RHOB,SNX18,TENTSA,CITED2,ABCC3,CEBPA,HMOX1,S3BP1
Mod_102	PTRHD1,MZT2A,HMGA1,PON2,CKS2,SUSD3,PDLLIM1,JDH2,DCTPP1,NME1
Mod_103	CRIM1,NGLY1,ZNR F2,KANK1,SNHG9,GN15
Mod_104	EIF2AK2,DDX60,ERV3-1,SAMD9,SAMD9L,OAS3,LINC00921
Mod_105	CD42EP3,SLC11A1,HPSE,FOLR3,CLEC4D,KAP4,CD300E,MCEMP1,FPR1,IL1LRB3,CD93,APOBEC3A
Mod_106	RHOQ,ITGA5,ZSWIM6,GSAP,CEMIP2,OTUD1,PAPSS2,CD42EP2,MMP19,DMXL2,SLC43A2,EMILIN2,EPB4113,ADGR2
Mod_107	PEL1,ATXN1,RNF144B,DENN3D,SLC15A3
Mod_108	AAK1,SLC12A7,IL6ST,SERINC5,KIF13B,DAPK1-T1,NSUN6,GAS2L3,SLC8B1,AC135050.3,PDGFB
Mod_109	GNLY,CD8B,CAVIN2,FGFBP2,IGFBP7,PF4,PPBP,GZMK,GZMA,GMPR,MYC,OLFMI,PRKCQ-AS1,PRF1,CTSW,ACRBP,KLRF1,KLRD1,KLRK1,IGFBP6,TRDC,GZMH,CBK,MTAK,NKG7,IDI1,TUBB1,GP1B,IL2RB
Mod_110	ARID5A,KBTD8,NR4A3,CREM,NR4A1,GPR183
Mod_111	LINC02611,DAB2,ARHGAP18,NRP1,LGMN,FN3K
Mod_112	IL1R2,HBEGF,NEDD9,GABARAPL1,PSTPI2,AC015819.2,SNAI1
Mod_113	GCC2,AP153,G3BP2,RAB11FIP1,UFP2,GABPB1-AS1,C11TA,CBX6,SEPTIN6
Mod_114	MIR4435-2HG,SYNJ2,ADAM8,VDR,ACTN1
Mod_115	MARCO,CYP27A1,NR G1,FBP1,CATSPER1,ASGR2
Mod_116	HNMT,CD302,PID1,GASK1B,EGR1,AOAH,ALDH1A1,KCTD12,IGSF6
Mod_117	DPH4,BTLA,TLR10,PLPPI,HDAC9,CPNE3,PTK2,WDFY4,RUBCNL,DAPK2,SERPINF2
Mod_118	GPR155,CSF1R,FGL2,MS4A7,LPAR6,ADAP2,SIGLEC10,MAFB,NC F4
Mod_119	FPI1,BHLHE40,EDEM1,GOLGA4,ALCAM,N4BP2,HIVEP1,SNX9,LINC00513,AC016831.5,ZC3HAV1,ZHX2,RAPGEF1,FAM107B,ETV6,JNTS6,NPIP5,ABCA5,NFKBID
Mod_120	STAT4,BASP1,EZH2,NFKB2,ANKLE2,STAT5A,GRAMD1A,BTG3,PIM3
Mod_121	AC114760.2,INPP5D,NHLRC3,MIS18BP1,AL627171.1,ILF3-DT,LINC00685
Mod_122	SPATS2,LNASE1L3,PRMT9,ELMO1,ENTPD1,ITGB7,NFATC2
Mod_123	GPBAR1,CX3CR1,TNFSF10,LINC02432,GIMAP7,GIMAP4,GIMAP2,GIMAP1,DOK2,SLC7A7,CFD,PRAM1,C19orf38,LYL1,STX10,LILRA2
Mod_124	ACSL3,PDE4D,PTK2B,TUT7,NFIL3,FGD4,VPS37B,THBS1,AC037198.1,SNHG25,METRN1,C5AR1
Mod_125	AC104809.2,CHST2,MXD3,SCGB3A1,PPP1R17,PTP4A3,PHF19,NEURL1,TAGLN,MEG3,LINC02345,TIPIN,MSLN,DEDD2,PPM1N,USP18,RGL4
Mod_126	IQSEC1,PAG1,GLT1D1,MX1
Mod_127	NUP210,LIMD1,CYFIP2,PIK3CG,UNC93B1,SELPLG,DHR57,P2RX1,RNF126,SLC44A2,ADA
Mod_128	ABHD5,SAP30,SLA,ZBTB16,IRS2,JD2,SIGLEC12
Mod_129	LZTFL1,ATP8A1,CASC15,HIST1H2BD,DDAH2,PLAAT3,AC026369.3,C12orf57,PRH1,DLEU1,PPP1R14A,A1B3G
Mod_130	LINC00877,REC8,SEPHS2
Mod_131	ROBO1,GPR171,KIT,CCR6,EHF,LAG3,DBNDD2
Mod_132	FILIP1L,USP53,PALLD,SESN1,IFRD1,WASHC2A,CHKA,PER1
Mod_133	TIPARP,FNIP2,AGAP3,SFMBT2,PHLDB1,AC090617.10,SEMA6B,LDLR,NRIP1
Mod_134	KLHL6,AF4,PPP1R10,BTAF1,JOSS1
Mod_135	MAGEF1,PARM1,C,CSER1,PKIB,CDH17,CCDC26,TCTN3,NDRG2,NUBPL,ASB2,TMEM97,LGALS2
Mod_136	FAM43A,HIST1H1E,HIST1H2BG,AL121944.1,LTB,ALS91115.1,PGM2L1,FKBP11,GIHCG,ATP10A,GAMT,SIGLEC6,BEX2,FHL1,USP9Y
Mod_137	AC024560.5,FYTTD1,PPM1K,DAAM1,SEMA7A,PPP1R16B
Mod_138	KL F3,GPR65
Mod_139	SCARB2,TGFB1,UTRN,LAIR1
Mod_140	PLAC8,MRPL36,GAPT,MEF2C,NHP2,RNF5,SDHAF3,DCPS,SIVA1,UBL7,CARHSP1,ITGAE,SMARCB1,CYSLTR1
Mod_141	MGST2,OTU,LIN,L,MPEG1,RNASE6
Mod_142	CCDC152,TNFRSF21,LINC00996,2DHHHC21,PRXL2A,LILRB4
Mod_143	N AIP,CLEC7A,KLF2,PIK3IP1
Mod_144	LHFP2,SMIM3,IGF2R,B4GALT1,SLC7A5,NPC1
Mod_145	ADR B2,SLC2A6,GC H1,ICAM4,LILRB2,LILRB1
Mod_146	RNF145,CD83,CDKN1A,RIK2,TRAF1,BIRC2,USP12,CYTH1,SAMSN1
Mod_147	CPVL,EVL,KCNK6,FAM110A
Mod_148	MYO1G,CHST7
Mod_149	CTSL,MMP14,PMP22
Mod_150	NDS2,ELK3,TRMT6,THBD

## Annexe 2 - Metadata de la cohorte SuperGut

Sample	Disease	Origin	Tissue	Status	10X chemistry	Location	subset
69	CD	Surgical resection	ILEUM	Involved	V1	mtsinai	total cells
128	CD	Surgical resection	ILEUM	Involved	V2	mtsinai	total cells
138	CD	Surgical resection	ILEUM	Involved	V2	mtsinai	total cells
158	CD	Surgical resection	ILEUM	Involved	V2	mtsinai	total cells
181	CD	Surgical resection	ILEUM	Involved	V2	mtsinai	total cells
187	CD	Surgical resection	ILEUM	Involved	V2	mtsinai	total cells
190	CD	Surgical resection	ILEUM	Involved	V2	mtsinai	total cells
193	CD	Surgical resection	ILEUM	Involved	V2	mtsinai	total cells
196	CD	Surgical resection	ILEUM	Involved	V2	mtsinai	total cells
198	UC	Biopsies	COLON	Involved	V2	mtsinai	total cells
200	UC	Biopsies	COLON	Involved	V2	mtsinai	total cells
202	UC	Biopsies	COLON	Involved	V2	mtsinai	total cells
204	CTRL	Biopsies	COLON	CONTROL	V2	mtsinai	total cells
206	CTRL	Biopsies	COLON	CONTROL	V2	mtsinai	total cells
209	CD	Surgical resection	ILEUM	Involved	V2	mtsinai	total cells
210	CD	Biopsies	ILEUM	Involved	V2	mtsinai	total cells
212	CD	Biopsies	COLON	Involved	V2	mtsinai	total cells
214	UC	Biopsies	ILEUM	Uninvolved	V2	mtsinai	total cells
216	UC	Biopsies	ILEUM	Uninvolved	V2	mtsinai	total cells
217	UC	Biopsies	ILEUM	Uninvolved	V2	mtsinai	total cells
218	CTRL	Biopsies	ILEUM	CONTROL	V2	mtsinai	total cells
219	CTRL	Biopsies	COLON	CONTROL	V2	mtsinai	total cells
220	CTRL	Biopsies	ILEUM	CONTROL	V2	mtsinai	total cells
221	CTRL	Biopsies	COLON	CONTROL	V2	mtsinai	total cells
222	CD	Biopsies	COLON	Involved	V2	mtsinai	total cells
223	CD	Biopsies	ILEUM	Involved	V2	mtsinai	total cells
228	CD	Biopsies	COLON	Involved	V2	mtsinai	total cells
229	CD	Biopsies	ILEUM	Involved	V2	mtsinai	total cells
231	CD	Biopsies	ILEUM	Involved	V2	mtsinai	total cells
232	CD	Biopsies	COLON	Involved	V2	mtsinai	total cells
237	CTRL	Biopsies	COLON	CONTROL	V3	mtsinai	total cells
238	UC	Biopsies	COLON	Involved	V3	mtsinai	total cells
n1	CD	Surgical resection	ILEUM	Involved	V3	nantes	total cells
n2	CD	Surgical resection	ILEUM	Involved	V3	nantes	mnp_stromal
n3	CD	Surgical resection	ILEUM	Involved	V3	nantes	total cells
n4	CD	Surgical resection	ILEUM	Involved	V3	nantes	mnp
n6	CD	Surgical resection	ILEUM	Involved	V3	nantes	total cells
n7	CD	Surgical resection	ILEUM	Involved	V3	nantes	total cells
n13	CD	Biopsies	ILEUM	Involved	V3	nantes	total cells
n14	CD	Surgical resection	ILEUM	Involved	V3	nantes	total cells
n17	CD	Surgical resection	ILEUM	Involved	V3	nantes	total cells
n18	CD	Surgical resection	ILEUM	Involved	V3	nantes	mnp

## Annexe 3 - Distribution des compartiments de cellules myéloïdes et stromales dans la cohorte SuperGut



## Annexe 4 - Characterization of Rat ILCs Reveals ILC2 as the Dominant Intestinal Subset

(Front Immunol. 2020 Feb 19;11:255. doi: 10.3389/fimmu.2020.00255)

Ahmed Abidi, Thomas Laurent, Gaëlle Bériou, Laurence Bouchet-Delbos, Cynthia Fourgeux, Cédric Louvet, Raja Triki-Marrakchi, Jeremie Poschmann, Régis Josien, Jérôme Martin



**Titre :** Développement de nouvelles approches pour l'analyse des interactions cellulaires dans les maladies immunitaires

**Mots clés :** Immunologie - single cell RNAseq - Bioinformatique - Greffe de rein - Maladie de Crohn

**Résumé :** Le développement des technologies de séquençage a ouvert la voie vers la caractérisation des mécanismes immuno-pathologiques en atteignant la précision des cellules uniques. L'application de ce type d'analyses transcriptomiques à des échantillons de cellules circulantes provenant de receveur de greffe de rein nous a permis de suivre de manière longitudinale les dynamiques post-transplantation des différentes populations immunitaires. Nous avons mis en avant l'émergence d'une signature d'activation précoce préférentiellement exprimée par les cellules immunitaires des patients qui subissent un rejet de greffe par la suite. Cette signature semble indiquer un recrutement favorisé de lymphocytes T et monocytes via l'axe CXCR4-CXCL12 qui participerait à la détérioration des tissus menant au rejet.

Dans le second projet, l'application du séquençage en cellule unique a permis d'étudier les programmes moléculaires en jeu au sein des tissus d'iléons de patients atteints de la maladie de Crohn. Par l'adaptation de l'algorithme de Metacell combinée à une analyse de modules de gènes sur notre jeu de données, il a été possible d'identifier avec précision une hétérogénéité de profils au sein des macrophages avec la présence de signatures inflammatoires et pro-fibrotiques enrichies chez les patients ne répondant pas au traitement clinique d'anti-TNFa actuel. Ces résultats proposent de nouvelles pistes de compréhension pour expliquer les différences de réponses de patients et permettent d'envisager de nouveaux traitements plus adaptés pour cibler les programmes moléculaires actives chez les patients non-répondeurs.

**Title:** Development of new approaches to the analysis of cellular interactions in immune diseases

**Keywords:** Immunology - single cell RNAseq - Bioinformatics - Kidney graft - Crohn's disease

**Abstract:** The development of sequencing technologies has paved the way for the characterisation of immuno-pathological mechanisms by achieving single-cell precision. The application of this type of transcriptomic analysis to samples of circulating cells from kidney transplant recipients has enabled us to follow the post-transplantation dynamics of multiple immune populations longitudinally. We highlighted the emergence of an early activation signature preferentially expressed by the immune cells of patients who subsequently undergo graft rejection. This signature seems to indicate a preferential recruitment of T lymphocytes and monocytes via the CXCR4-CXCL12 axis, which is thought to contribute to tissue damage leading to rejection.

In the second project, the application of single-cell sequencing allowed us to study the molecular programs at play within the ileum tissues of Crohn's disease patients. By adapting the Metacell algorithm, combined with gene module analysis on our dataset, it was possible to accurately identify heterogeneous profiles within macrophages, with the presence of enriched inflammatory and pro-fibrotic signatures in patients not responding to current clinical anti-TNFa treatment. These results offer new ways of understanding the differences in patient responses and allow us to envisage new treatments that are better adapted to targeting the molecular programs active in non-responder patients.