HAL
open science

Caractérisation des premières étapes de différenciation des cellules hématopoïétiques à l'échelle de la cellule unique

Alice Moussy

▶ **To cite this version:**

Alice Moussy. Caractérisation des premières étapes de différenciation des cellules hématopoïétiques à l'échelle de la cellule unique. Biologie cellulaire. Université Paris sciences et lettres, 2017. Français. NNT : 2017PSLEP029 . tel-02106673

HAL Id: tel-02106673

https://theses.hal.science/tel-02106673v1

Submitted on 23 Apr 2019

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

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

THÈSE DE DOCTORAT

de l'Université de recherche Paris Sciences et Lettres
PSL Research University

Préparée à l'École Pratique des Hautes Études

Caractérisation des premières étapes de différenciation des cellules hématopoïétiques à l'échelle de la cellule unique

École doctorale de l'EPHE – ED 472

Spécialité : Systèmes intégrés, environnement et biodiversité

COMPOSITION DU JURY :

M. András PALDI
Ecole Pratique des Hautes Etudes
Directeur de thèse

M. Olivier GANDRILLON
CNRS
Président du jury

M. François DELHOMMEAU
Hôpital Saint-Antoine AP-HP
Rapporteur

M. Thierry JAFFREDO
CNRS
Rapporteur

Mme Leïla PERIE
Institut Curie
Membre du jury

Soutenue par :
Alice MOUSSY
le 31 octobre 2017

Dirigée par :
András PALDI
Daniel STOCKHOLM

École Pratique des Hautes Études | PSL RESEARCH UNIVERSITY PARIS

REMERCIEMENTS

Je tiens tout d'abord à remercier les membres du jury d'avoir accepté d'évaluer mon travail de thèse. Je remercie le Dr Olivier Gandrillon pour avoir accepté de présider mon jury de thèse. Je remercie le Dr Thierry Jaffredo et le Pr François Delhommeau d'avoir accepté et pris de leur temps pour évaluer ce manuscrit et de m'avoir fait part de leurs commentaires. Je remercie également le Dr Leïla Perié d'avoir accepté de faire partie des membres du jury de ma thèse.

Je remercie sincèrement le Pr András Páldi, mon directeur de thèse, de m'avoir encadrée ces 3 années. Je te remercie pour la confiance que tu m'as accordée et de m'avoir apporté cette nouvelle vision, plus philosophique de la biologie. Je te remercie pour tes nombreux conseils sur les présentations orales comme écrites qui m'ont permis de progresser au fil des années. Tu as toujours su trouver les mots pendant les périodes difficiles pour me remonter le moral. Tu as été un directeur de thèse présent, à l'écoute et sur qui j'ai pu compter pendant ces années. Je te remercie encore pour ton aide précieuse sur le manuscrit, c'est en grande partie grâce à toi si j'ai pu soutenir ma thèse en 3 ans, merci beaucoup !

Je remercie également le Dr Daniel Stockholm, mon co-encadrant de thèse, pour m'avoir épaulée pendant toutes ces années. Grâce à toi, j'ai pu faire mon stage de fin d'études à Généthon et faire une thèse dans l'équipe EPHE. Je te remercie aussi de m'avoir fait confiance, de m'avoir écoutée, de m'avoir initiée au logiciel R et d'avoir toujours été là quand j'en avais besoin. Je te remercie pour ta bonne humeur, tes « stockholmites », tes expériences toujours plus surprenantes les unes que les autres qui m'ont permis de décompresser et de bien rigoler. Je garderai un très bon souvenir de ma thèse (mon foie aussi s'en souviendra pour tous ces bonbons que tu m'as apportés au bureau !)

Je remercie le Dr Anne Galy, directrice de l'INSERM_UMR_S951, de m'avoir accueillie au sein de son unité, à Généthon.

Cette thèse n'aurait pas été la même sans le Dr Jérémie Cosette, le post-doc de l'équipe Imagerie-Cytométrie de Généthon. Jérémie, je te remercie pour tous tes conseils tant sur le plan expérimental que sur les présentations orales (M. Expert en Power Point). Je te remercie beaucoup pour ton aide et tes corrections sur mon manuscrit. Ensemble, nous avons réalisé 3

papiers en 3 ans et je crois que cela n'aurait été pas possible sans toi ! Tu m'as aussi aidée à m'épanouir en tant que doctorante. Ta mémoire et tes connaissances sur tout et n'importe quoi font que tu es une personne surprenante que je ne pourrai pas oublier ! Merci pour tous ces moments de rigolade, toutes tes blagues qui créent une bonne entente au sein de Généthon.

Je remercie également les stagiaires de mon équipe, Romuald Parmentier et Cindy Da Silva, que j'ai pu encadrer et qui ont contribué massivement à ma thèse. Merci pour votre aide sur les analyses d'images, je sais que ça a été pénible et très long, sans vous, je n'aurais pas pu avoir autant de données ! Cindy, ma petite portugaise du bureau, on s'est connues au tout début de ma thèse et tu es restée encore pendant 1 an et demi ; merci d'avoir été là, ta joie et ton enthousiasme sont communicatifs, surtout ne change pas ! Romuald, le geek qui me fait me sentir vieille, mais grâce à toi, je revois ma culture web ! Merci encore pour tes analyses d'images et tes scripts R, maintenant c'est à ton tour d'être thésard et si je peux te donner un conseil : éclate toi !

Je tiens aussi à remercier le Dr Guillaume Corre, bio-informaticien de Généthon, pour son aide précieuse dans les analyses multiparamétriques. Merci Guillaume d'avoir toujours pris le temps pour trouver la solution (en 2 min) alors que moi, ça faisait déjà 3h que je me prenais la tête avec R ! C'est toujours avec plaisir que tu acceptais de m'aider et pour cela je t'en remercie.

Je remercie également, Peggy Sanatine et Simon Jimenez, membres de l'équipe Imagerie-Cytométrie. Merci beaucoup pour votre aide en cytométrie, je n'aurais jamais pu avoir d'aussi jolis résultats en cellules uniques sans vous ☺. Votre rigueur et votre disponibilité m'ont permis d'en arriver là aujourd'hui. Et je tiens à te remercier tout particulièrement, Peggy, d'avoir été aussi présente personnellement, de m'avoir rassurée et conseillée.

Je remercie aussi, Fanny Onodi, qui a fait son stage au même moment que moi et qui a réalisé la partie enrichissement des cellules CD4+ dans mon premier projet de thèse. Merci beaucoup pour ton aide et aussi pour tous ces bons moments que nous avons passés ensemble avec Julie et fils, les nombreux afterworks au New Village et surtout on sait toutes les 2 que le plus mignon de tous c'est pfft pfft ;).

Je tiens également à remercier le Dr Thi My Anh Neildez, pour ses nombreux conseils sur les LT et sur les expériences en général. TMA, je te remercie tout particulièrement pour avoir été là quand ça n'allait pas, pour m'avoir toujours soutenue et pour avoir pris de ton temps pour me soutenir. Je sais que je peux venir de voir à n'importe quel moment et que tu seras toujours là pour moi !

Bien sûr, il y a aussi des personnes qui n'ont pas forcément contribué directement à mon travail de thèse mais qui ont contribué fortement au bon déroulement de ma thèse et je tiens à les remercier eux aussi !

Tout d'abord, il y a l'équipe IMBI, ma première famille, avec les « anciens » d'un côté avec Armelle, Catherine, David, Khalil, Maxime, Mirella, Sophie, Sylvie et Roseline qui m'ont accueillie et intégrée rapidement dans leur équipe. Merci à tous pour votre enthousiasme, votre bonne humeur et pour tous les bons moments passés ensemble comme les goûters IMBI ou les parties de bowling après le boulot. Maxime, merci d'avoir débriefé avec moi chaque mardi midi les épisodes de l'ADP et pour tes nombreuses blagues (sur les blondes) ; tu m'as mise tout de suite à l'aise et je t'en remercie ! Mirella, merci pour tes nombreux conseils sur les cellules CD34+ et pour les fous rires lors des pauses café. Sylvie, merci d'avoir toujours été là, ton sourire et ta joie m'ont manqué cette dernière année. Et Papa Khalil, le doyen de l'équipe, je suis contente de t'avoir connu avant ton départ à la retraite. Merci pour tes nombreux conseils, tu représentes la sagesse, tu es mon modèle ☺.

Et de l'autre côté, il y a les jeunes avec Julie, Saliha, Youna, Laetitia et Gaëlle avec qui nous nous sommes toutes très vite entendues et soutenues. Merci pour tous les moments passés ensemble à faire les mots fléchés et à rigoler. Youna, la petite thésarde de l'équipe et l'experte en mots fléchés, tu es toujours là quand on a besoin d'aide et toujours de bonne humeur. Je te souhaite bon courage pour la fin de ta thèse surtout pour supporter Maxime encore 2ans... ☺. Même s'il ne fait pas partie officiellement de l'équipe IMBI, je suis obligée de faire un petit mot ici, à Antoine, le squatteur du grand L1 et de notre équipe pour m'avoir fait prendre des kilos à cause des bonbons qu'il ramenait tous les jours. D'ailleurs je ne sais pas si je dois vraiment te remercier... en tout cas, tu m'auras bien fait rire ces 2 années !

Et il y a ma seconde famille, ma famille d'adoption, avec Céline, Sandrine et Greg. Merci à vous pour vos conseils sur les cellules CD34+, sur le processing, je suis (trop) souvent venue vous déranger, j'espère que vous ne m'en voulez pas trop ;). Céline et Sandrine, merci pour

toutes ces discussions sur tout et sur rien, votre bonne humeur et votre sourire, vous avez toujours été présentes pour moi, merci pour tout !

Il y a aussi notre groupe du midi avec Céline, Guillaume, Jérémy, Julie et Saliha sur qui je pouvais compter et avec qui je savais que j'allais passer un bon moment de détente.

Je tiens aussi à remercier Maguy et Stéphane, pour avoir toujours trouvé des réponses à mes questions. Vous avez contribué à la bonne organisation de ma thèse.

Et enfin, il y a mes collègues de la team Rose, Sophie, Saliha et Julie. Sophie, quand j'ai vu pour la première fois ton bureau, je me suis tout de suite dit qu'on allait bien s'entendre toutes les 2 : tout ce rose !! j'ai essayé de te battre mais là je crois que tu es à un haut niveau ☺. Et je te l'ai déjà dit plusieurs fois mais je te le redis encore, tu es un modèle pour moi. Tu es une experte en bio mol et j'espère un jour être autant calée que toi ! Merci pour tes nombreux conseils, tu as toujours pris le temps de m'expliquer les choses, de me montrer comment cela fonctionnait et ça, dès le premier jour de ma thèse ! Saliha, ma copine de thèse, qu'est-ce qu'on a pu papoter toutes les deux, heureusement, tu n'étais pas loin de mon bureau, ça m'a évité de trop marcher ! Tu as été plus qu'une simple collègue pour moi, tu as été et es toujours une amie, une grande sœur, qui a toujours été là pour me conseiller tant au niveau manip que dans la vie perso. J'ai adoré ces années avec toi, tous ces fous rire le midi, les pauses café, à EMBO, les afterworks... Tu es drôle, généreuse et bosseuse, cette dernière année de thèse sans toi était différente, tu m'as beaucoup manqué :'(. Mais sache qu'aujourd'hui je suis super fière de toi, tu as enfin réalisé ton rêve : vivre aux States ! Julie, ma copine de thèse, on a commencé la thèse ensemble et on va la finir ensemble !!! Quand je suis arrivée à Généthon, tu as été la première personne à être venue vers moi et à me présenter à tout le monde. Tu savais que c'était difficile d'être une nouvelle, de s'intégrer et tu as été là dès le début pour moi .Je ne te remercierai jamais pour cela. Tu as toujours pris ton temps pour m'aider avec les présentations, tu es douée pour cela ☺. Je sais que je peux compter sur toi tout le temps. Tu as été là pour mon déménagement, pour aller chercher mon canapé à But...Tu te souviens de la fois où j'étais en galère de RER et que t'es venue me chercher ? Heureusement que tu étais là ! Tu es généreuse et altruiste, surtout ne change pas. Torpille a beaucoup de chance de t'avoir ☺.

Il y a aussi mes amies, mes copines de l'UTC, Elodie, Lucie, Murielle, Nathalie et Sophie, je sais que je n'ai pas forcément été présente pendant ces 3 années mais je sais que vous le

comprenez. Vous êtes des vraies amies. Merci d'être venues me soutenir et m'écouter blablater pendant 2h. Il y a aussi mes copines d'enfance, Aurélie et Lucie, qui m'ont toujours soutenue. Vos encouragements, tout au long de la thèse, m'ont fait du bien !

Je tiens aussi à faire un petit mot à la famille de mon chéri, à ses parents, Gildo et Deolinda, à sa sœur Sandrine et son mari Quentin et à sa nièce et son neveu, Cassandra et Anthony. Merci encore pour ces fous rires surtout sur le titre de ma thèse : et non je ne fais pas de la « poésie » ☺. Et Cassandra, ça y est je suis une adulte, je ne vais plus aller à l'école ;).

Bien sûr cette thèse n'aurait pas pu bien se dérouler sans le soutien de ma famille, en particulier sans mes parents qui ont toujours pris le temps de m'écouter, de m'encourager et d'essayer de comprendre ce que je faisais. Même dans les moments les plus difficiles, vous avez toujours cru en moi et vous m'avez toujours poussée vers le haut. Ce titre de Doctorat, je le dois aussi et surtout grâce à vous. Maman, merci d'avoir pris le temps de corriger mes fautes d'orthographe dans mon manuscrit. Quand même, je t'ai fait lire plus de 100 pages sur des termes que tu ne comprenais rien. Je t'en suis reconnaissante, je sais que tu seras toujours là pour moi. Je vous dis encore mille fois merci. Je ne vous le dis pas souvent mais je vous aime fort, Papa et Maman.

J'ai beaucoup de chance car vous n'étiez pas les seules à m'avoir soutenue, il y a aussi ma grand-mère, ma mamie de Corse, ma sœur, Amélie, mon frère, Adrien, mon beau-frère et ma belle-sœur, Axel et Marion, ma tante, Chantal et mon parrain et sa femme, Gilles et Momo qui ont toujours cru en moi et qui m'ont apporté toute l'aide qu'ils pouvaient me fournir. Et même si vous êtes encore trop jeunes pour comprendre ce qu'a fait Tatalice, grâce à vous, mes neveux et nièce, Alex, Raphaël, Charlotte et Angel, j'ai su/dû faire des breaks le WE. Vous avez, vous aussi, participé quelque part à la bonne réussite de Tata ☺.

Et enfin, je ne pouvais pas finir ces remerciements sans faire un petit mot sur toi mon Doudou. Tu m'as toujours épaulée, soutenue et conseillée dans les moments de stress et de doute. C'est toi qui m'as supportée pendant ces 2 années, surtout que ce n'était pas les plus faciles pour moi. Tu as su me changer les idées quand il le fallait et je sais que je peux compter sur toi. Merci à toi.

TABLE DES MATIERES

LISTE DES ABREVIATIONS **5**

LISTE DES ILLUSTRATIONS **7**

INTRODUCTION **9**

PARTIE 1 LES CELLULES HEMATOPOÏETIQUES 11

I. Caractéristiques des cellules souches hématopoïétiques 14

I.A Découverte des cellules souches hématopoïétiques 14

I.B Ontogenèse de l'hématopoïèse 15

I.B.1 Hématopoïèse primitive 16

I.B.2 Hématopoïèse définitive 17

I.C Caractérisation des cellules souches hématopoïétiques humaines 19

I.C.1 Enrichissement des CSH 19

I.C.2 Mécanismes moléculaires et cellulaires des CSH humaines 20

I.C.3 La cellule souche : une entité ou un état ? 22

II. Différents types de cellules sanguines 24

II.A Erythrocytes 24

II.B Thrombocytes 25

II.C Leucocytes 27

II.C.1 Monocytes 27

II.C.2 Lymphocytes 28

II.C.3 Granulocytes 30

III. Différenciation des cellules souches hématopoïétiques : l'hématopoïèse 34

III.A Vision historique de l'hématopoïèse : modèle déterministe 34

III.A.1 Modèle classique de l'hématopoïèse : premier modèle explicatif 34

III.A.2 Autres modèles déterministes : modèle classique revisité 37

III.B Vision dynamique de l'hématopoïèse : modèle stochastique 40

III.B.1 Paysage épigénétique de Waddington 41

III.B.2 Technologie de cellules uniques : la fin du modèle déterministe 45

IV. Applications thérapeutiques des cellules souches hématopoïétiques 48

IV.A Différentes sources de cellules souches hématopoïéiques 48

IV.A.1 CSH de la moelle osseuse 49

IV.A.2 CSH du sang périphérique 50

IV.A.3 CSH du sang de cordon ombilical 50

IV.B Thérapie génique 52
IV.B.1 Deux stratégies de thérapie génique 53
IV.B.2 Apport du transgène 55
IV.B.3 Limites de la thérapie génique..... 57
IV.C Utilisation des cellules CD4+ et des cellules CD34+ dans ma thèse 58

PARTIE 2 IMPORTANCE DES ETUDES REALISEES A L'ECHELLE DE LA CELLULE UNIQUE 59

I. Hétérogénéité des cellules souches hématopoïétiques 63

II. Etude des voies de différenciation 65

III. Choix du destin cellulaire 67

IV. Réseaux de régulation 70

V. Hématopoïèse développementale.....72

PARTIE 3 DIFFERENTES TECHNIQUES DE CAPTURES ET D'ANALYSES DE CELLULES HEMATOPOÏETIQUES UNIQUES 75

I. Méthodes de capture de cellules uniques dans une population cellulaire.....77
I.A Trieur de cellules par cytométrie en flux (FACS) 77
I.B Microfluidiques 78
I.B.1 Piège hydrodynamique..... 79
I.B.2 Isolement par gouttes 81
I.C Microgrilles 84

II. Etude transcriptomique à l'échelle de la cellule unique 86
II.A Mesure quantitative de l'expression de multiples gènes : la RT-qPCR 86
II.A.1 Principales étapes de la RT-qPCR à l'échelle de la cellule unique 87
II.A.2 SYBR Green I..... 88
II.A.3 Sonde TaqMan..... 89
II.B Etude des ARN totaux d'une cellule unique : le séquençage 91
II.B.1 Méthode de Tang 91
II.B.2 Smart-seq et Smart-seq 2..... 92
II.B.3 Quartz-seq 94
II.B.4 Expression cellulaire par amplification linéaire et séquençage (Cel-seq)..... 95
II.B.5 Cellule unique marquée lors de la reverse transcription (STRT-seq)..... 96
II.B.6 ARNm-seq quantitative à l'échelle de la cellule unique..... 97

III. Etude protéomique à l'échelle de la cellule unique.....98

IV. Analyse des données multiparamétriques.....102

IV.A Inconvénients de la méthode « gating » manuel en cytométrie 102
IV.B Différentes méthodes d'analyses multiparamétriques 104
IV.B.1 Réduction multidimensionnelle en 2D : ACP et t-SNE 104
IV.B.2 Approche par regroupement : SPADE et Citrus 106
IV.B.3 Détection de la trajectoire développementale 108
V. Lien entre phénotype et morphologie des cellules individuelles 111
V.A Images instantanées de cellules uniques avec marqueurs phénotypiques 111
V.B Suivi temporel de la dynamique phénotypique des cellules uniques in vitro : l'imagerie en time lapse 114
OBJECTIFS ET STRATEGIES 117
RESULTATS 121
PARTIE 1 MODELE SIMPLE DE DIFFERENCIATION : LA DIFFERENCIATION DES LYMPHOCYTES T EN LYMPHOCYTES T REGULATEURS 123
PARTIE 2 MODELE PLUS COMPLEXE DE DIFFERENCIATION : LES CELLULES CD34+ HUMAINES ISSUES DE SANG DE CORDON 145
I. Etude de la dynamique morphologique des cellules CD34+ issues de sang de cordon à l'échelle de la cellule unique 147
II. Etude cinétique des changements phénotypiques des cellules CD34+ humaines issues de sang de cordon à l'échelle de la cellule unique 167
DISCUSSION ET PERSPECTIVES 203
I. Discussion 205
I.A Hétérogénéité 205
I.B Choix du destin cellulaire 207
I.C Proposition d'un modèle de différenciation stochastique 209
II. Perspectives 212
BIBLIOGRAPHIE 217

LISTE DES ABREVIATIONS

2D ou 3D	2 ou 3 Dimensions
A	
AAV	Virus Adéno-Associés
ACP	Analyse en Composantes Principales
ADN	Acide Désoxyribonucléique
ADNc	Acide Désoxyribonucléique complémentaire
AGM	Aorte-Gonade-Mésonéphros
ARN	Acide Ribonucléique
ARNm	Acide Ribonucléique messager
C	
CD	Cluster de Différenciation
CFI	Cytométrie en Flux en Images
CFU-S	Colony Forming Unit in Spleen
CGD	Granulomatose Septique Chronique
CLP	Progéniteur Commun Lymphoïde
CMLP	Progéniteur Commun Myélo-Lymphoïde
CMP	Progéniteur Commun Myéloïde
CSE	Cellule Souche Embryonnaire
CSH	Cellule Souche Hématopoïétique
CSPi	Cellule Souche Pluripotente induite
CyTOF	Cytometry by Time-Of-Flight
F	
FACS	Flow Activated Cell Sorting
FISH	Hybridation Fluorescente In Situ
FSC	Foward Scatter
G	
G-CSF	Facteur de Stimulation des Colonies de Granulocytes
GFP	Green Fluorescent Protein
GM-CSF	Facteur de Stimulation de la Colonie Macrophage-Granulocyte
GMP	Progéniteur Granulocyte-Monocyte
H	
HDAC	Histone Déacétylase
I	
IMU	Identificateur Moléculaire Unique

L	
LB	Lymphocyte B
LMPP	Progéniteur Multipotent biaisé Lymphoïde
LT	Lymphocyte T
LT_H	Lymphocyte T Helper
LTreg	Lymphocyte T régulateur
M	
M-CSF	Facteur de Stimulation des Colonies de Macrophages
MEP	Progéniteur Erythroïde Mégacaryocyte
MO	Moelle Osseuse
MPP	Progéniteur Multipotent
N	
NK	Natural Killer
P	
PCR	Polymerase Chain Reaction
PDMS	Polydiméthylsiloxane
qPCR	Polymerase Chain Reaction quantitative
R	
RT	Réverse Transcription
S	
SCID	Immunodéficiences Combinés Sévères
SDF-1	Stromal-cell Derived Factor 1
Smart-seq	Switching Mechanism At the 5’ end of RNA Template sequencing
SPADE	Spanning-tree Progression Analysis of Density-normalized Event
SSC	Side Scatter
T	
TIV	Transcription In Vitro
TL	Time Lapse
TPO	Thrombopoïétine
t-SNE	t-distributed Stochastic Neighbor Embedding
V	
VPA	Acide Valproïque
W	
WAS	Syndrome de Wiskott Aldrich

LISTE DES ILLUSTRATIONS

FIGURES

Figure 1 : Rates de souris irradiées 10 jours après injection de $6x10^4$ cellules nucléées. 15
Figure 2 : Ontogenèse des cellules hématopoïétiques. 18
Figure 3 : Erythrocytes. 24
Figure 4 : Thrombocytes en microscopie électronique à balayage. 26
Figure 5 : Cellules dendritiques et macrophages en microscopie électronique à balayage. 28
Figure 6 : Lymphocytes B et T en microscopie électronique à balayage. 30
Figure 7 : Les trois types de granulocytes par coloration au May Grünwald Giemsa. 32
Figure 8 : Premier modèle « classique » de l'hématopoïèse. 36
Figure 9 : Différents modèles déterministes au cours du temps 39
Figure 10 : Modèle stochastique et « paysage épigénétique » de Waddington. 41
Figure 11 : Différents types de bifurcation. 45
Figure 12 : Modèles redéfinis de Perié et al. (2016) et de Notta et al. (2016). 46
Figure 13 : Différents essais cliniques de thérapie génique. 53
Figure 14 : Stratégies in vivo et ex vivo de la thérapie génique. 54
Figure 15 : Différents vecteurs viraux et non viraux utilisés dans les essais cliniques de thérapie génique. 56
Figure 16 : Etudes faites à l'échelle de la cellule unique révèlent l'hétérogénéité populationnelle. 62
Figure 17 : Diverses applications réalisées à l'échelle de la cellule unique. 64
Figure 18 : Différents choix du destin cellulaire. 67
Figure 19 : Méthodes de capture des CSH uniques. 77
Figure 20 : Puce avec des coupelles en forme de U. 80
Figure 21 : Système hydrodynamique avec des canaux pour isoler des cellules uniques. 81
Figure 22 : Conception modulaire des composants microfluidiques. 82
Figure 23 : Illustration de la préparation des ARNm-seq de cellule unique par la technique de Drop-seq. 83
Figure 24 : Schéma explicatif du mode de déroulement pour isoler des cellules grâce à la microgrille. 84
Figure 25 : Déroulement de la RT-qPCR à l'échelle de la cellule unique. 88
Figure 26 : Deux méthodes d'analyse quantitative de la PCR. 90

Figure 27 : Séquençage des ARN totaux par la méthode de Tang. 92
Figure 28 : Séquençage des ARN totaux par la méthode de Smart-seq 2. 93
Figure 29 : Séquençage des ARN totaux par la méthode de Quartz-seq. 94
Figure 30 : Séquençage des ARN totaux par la méthode de Cel-seq. 95
Figure 31 : Séquençage des ARN totaux par la méthode de STRT-seq. 96
Figure 32 : Séquençage des ARN totaux par la méthode des ARNm-seq quantitative. 97
Figure 33 : Analyse protéique dans une cellule individuelle. 98
Figure 34 : Représentation schématique d'un cytomètre en flux. 99
Figure 35 : Exemple de réduction multidimensionnelle en 2D ou 3D : ACP et t-SNE. 106
Figure 36 : Exemple de représentation des données multiparamétriques avec la méthode SPADE et citrus. 108
Figure 37 : Description de l'algorithme de Wanderlust. 110
Figure 38 : Images obtenues avec un cytomètre en flux en images (ImageStream®). 113
Figure 39 : Comparaison entre des images prises à des instants fixes et des images prises en continu. 115
Figure 40 : Comparaison de la prolifération des LTreg entre 21% et 5% d'O_2. 125
Figure 41 : Exemple de bifurcation aléatoire. 206
Figure 42 : Analyses instantanées vs dynamique phénotypique. 208
Figure 43 : Proposition d'un modèle stochastique de différenciation. 210

TABLEAUX

Tableau 1 : Comparaison des différentes sources de CSH. 51
Tableau 2 : Différentes méthodes d'analyses multiparamétriques. 103
Tableau 3 : Récapitulatif des différentes technologies de cellule unique et de leur résolution. 111

INTRODUCTION

PARTIE 1

Les cellules hématopoïétiques

Les cellules souches hématopoïétiques (CSH), du grec haímatos pour sang et poíêsis pour création, sont des cellules souches adultes multipotentes, capables de donner toutes les cellules du sang et de s'auto-renouveler. On considère traditionnellement que la plupart des CSH se trouvent dans un état de quiescence ou dormance dans des niches qui les maintiennent souches et répondent à des signaux lorsque l'équilibre des cellules sanguines ou des CSH est perturbé par des stimuli intrinsèques ou extrinsèques. Les CSH, de par leur potentiel multipotent, sont largement étudiées en biologie surtout dans le domaine thérapeutique, dans les traitements des maladies du sang (syndrome de Wiskott-Aldrich (WAS), granulomatose septique chronique (CGD), etc.), dans les maladies auto-immunes (anémie hémolytique auto-immune, Purpura thrombopénique idiopathique, etc.) et dans les maladies héréditaires du sang (drépanocytose, thalassémie, etc.). Bien que les CSH soient étudiées depuis de nombreuses années, notre compréhension sur le maintien de leur état souche et sur sa différenciation ne sont pas encore claires.

Dans cette première partie, je vais tout d'abord développer un paragraphe sur les caractéristiques des CSH, avec leur découverte, l'ontogenèse et leur phénotype, ensuite un paragraphe sur les différents types de cellules sanguines et leur fonction suivi des différents modèles de l'hématopoïèse qui existent et enfin un résumé bref sur les différentes applications thérapeutiques des CSH.

I. Caractéristiques des cellules souches hématopoïétiques

Dans le cadre de mon sujet de thèse, il me paraît essentiel de connaître la découverte des CSH, leur ontogenèse et leur phénotype pour étudier les premières étapes de différenciation des CSH. C'est pourquoi, ces trois points seront développés dans ce paragraphe.

I.A Découverte des cellules souches hématopoïétiques

En 1896, Artur Pappenheim a été le premier à utiliser le terme de cellules souches pour décrire une cellule précurseur capable de générer à la fois des globules rouges et des globules blancs (Pappenheim, 1896; Ramalho-Santos and Willenbring, 2007). Ce concept était une construction intellectuelle relativement peu importante jusqu'à l'avènement des armes nucléaires pendant la Seconde Guerre mondiale (Sykes and Scadden, 2013). Dans ce contexte, la possibilité d'avoir des cellules souches comme une contre-mesure contre les conséquences dévastatrices des rayonnements sur la production de cellules sanguines a considérablement augmenté l'intérêt à déterminer leur existence. A cette époque, des débats persistaient encore sur l'existence d'une cellule souche, avec l'hypothèse que ces cellules étaient capables de se régénérer, une notion largement guidée par des observations sur les cellules tumorales.

La découverte des CSH par Till and McCulloch au Canada en 1961 a permis d'éclairer les études de biologie de cellules souches dans le système hématopoïétique incluant les connexions avec les autres cellules souches adultes à travers les concepts de base de la différenciation, de la multipotence et de l'auto-renouvellement (Till and McCulloch, 1961). Au préalable, des études avaient déjà démontré que l'injection de cellules de la moelle osseuse (MO) dans des souris irradiées les maintenait en vie, ce qui a prouvé le potentiel de transplantation de ces cellules (Ford et al., 1956; Jacobson et al., 1951). Till et McCulloch ont constaté que l'injection d'un petit nombre de cellules de MO entraînait la formation de nodules hématopoïétiques sept-huit jours après injection à l'intérieur de la rate des souris irradiées (Figure 1). Plus tard, des études ont démontré que ces colonies spléniques, appelées CFU-S (Colony Forming Unit in Spleen), étaient d'origine clonale souvent composées de précurseurs de granulocytes/mégacaryocytes et érythroïdes et que leur transplantation chez des souris secondaires donnait à nouveau des colonies spléniques (Curry et al., 1967; Lewis and Trobaugh, 1964; Siminovitch et al., 1963; Wu et al., 1967). Au départ, les scientifiques pensaient que les CFU-S provenaient de la différenciation des CSH mais il a été démontré plus tard que les CFU-S dérivaient plutôt de cellules progénitrices engagées (Schofield, 1978).

La découverte de Till et McCulloch a permis d'entreprendre de nouvelles investigations sur ces CSH comme sa caractérisation fonctionnelle, sa purification, sa culture et d'autres recherches en général sur les cellules souches.

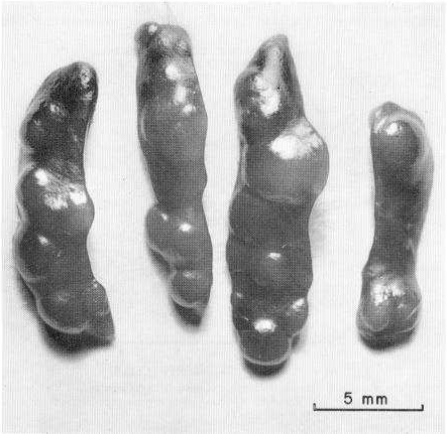


Figure 1 : Rates de souris irradiées 10 jours après injection de 6x10^4 cellules nucléées.
Expérience de Till et McCulloch, 1961. Les nodules sur lesquels repose l'analyse sont facilement visibles.

Ces études ont été un point de repère pour l'étude de l'hématopoïèse et ont fourni la base scientifique pour la poursuite des efforts cliniques dans la transplantation de MO. La démonstration de l'existence des CSH a établi ces cellules comme des « prototypes » de cellules souches : des cellules capables de s'auto-renouveler presque indéfiniment et de créer des cellules spécialisées (multi-lignage). Finalement, ces études ont servi de modèle général pour étudier la différenciation et pour définir expérimentalement les cellules souches dans d'autres tissus comme la peau, le système nerveux central, etc. D'ailleurs, notre compréhension sur le choix du destin des CSH reste encore peu claire. Selon moi, pour comprendre au mieux le destin de ces cellules, il faut savoir d'où elles proviennent. C'est pourquoi le prochain paragraphe portera sur l'ontogenèse de l'hématopoïèse afin d'avoir tous les « outils » nécessaires pour connaître le passé de ces cellules afin de mieux comprendre leur destin.

I.B Ontogenèse de l'hématopoïèse

Le système hématopoïétique est l'un des premiers tissus complexes à se former dans l'embryon chez les mammifères (Dzierzak and Speck, 2008). L'ontogenèse du système hématopoïétique est caractérisée par deux grandes étapes : la première, l'hématopoïèse primitive, génère des populations cellulaires hématopoïétiques transitoires (sauf les microglies) et la deuxième étape, l'hématopoïèse définitive, arrive plus tard dans le développement des CSH multipotentes définitives (Julien et al., 2016).

I.B.1 Hématopoïèse primitive

La première émergence hématopoïétique, appelée hématopoïèse primitive, commence à la troisième semaine du développement humain, avant l'établissement du flux sanguin, dans un tissu extra-embryonnaire composé d'endoderme viscéral et de mésoderme, le sac vitellin (Figure 2) (Tavian et al., 1999). Initialement, ce sont les expériences de Moore et Metcalf, dans les années 70, qui ont montré que les cellules hématopoïétiques, se trouvant dans le sac vitellin, pouvaient générer des progéniteurs hématopoïétiques uniquement des lignées érythroïdes et myéloïdes (Moore and Metcalf, 1970).

Chez les vertébrés, les cellules mésodermiques du sac vitellin forment des agrégats cellulaires qui sont à l'origine à la fois du système vasculaire et de l'hématopoïèse extra-embryonnaire. Les cellules périphériques acquièrent des marqueurs endothéliaux et les cellules centrales disparaissent pour créer la lumière des premiers vaisseaux sanguins. Certaines cellules mésodermiques restent encore attachées à l'endothélium vasculaire nouvellement formé et constituent les îlots sanguins à l'origine de l'hématopoïèse extra-embryonnaire (Péault et al., 1988; Sabin, 1920). Ces observations ont suggéré un précurseur mésodermique commun des CSH et des cellules endothéliales, appelé hémangioblaste (Figure 2) (Choi et al., 1998; Fehling et al., 2003). Des études faites chez la souris au stade précoce ont montré que des hémangioblastes présomptifs pouvaient exprimer à la fois le marqueur mésodermique Brachyury et le récepteur 2 au VEGF (VEGFR2) dans la région postérieure de la ligne primitive (Huber et al., 2004).
Les agrégats mésodermiques apparaissent dans le sac vitellin humain à environ 16 jours de gestation et les îlots sanguins, observés à partir du 19ème jour, produisent des érythrocytes, des macrophages et des mégacaryocytes primitifs essentiellement à courte durée (Bloom and Bartelmez, 1940; Luckett, 1978). La séparation du sac vitellin avec l'embryon au 19ème jour, juste avant le début de la circulation sanguine, a montré que la pré-circulation du sac vitellin était limitée à la lignée érythro-myéloïde et qu'il manquait le potentiel souche lympho-hématopoïétique (surtout des lymphocytes T) (Tavian et al., 2001). Ces découvertes ont suggéré que l'hématopoïèse chez les mammifères adultes aurait plutôt une origine intra-embryonnaire, en accord avec les précédentes observations faites sur les embryons aviaires (Dieterlen-Lièvre et al., 1976).

I.B.2 Hématopoïèse définitive

La seconde étape de l'hématopoïèse, appelée l'hématopoïèse définitive, apparait plus tard dans le développement des vertébrés et a lieu dans l'embryon plus particulièrement dans la région aorte-gonade-mésonéphros (AGM), dans le foie fœtal, dans le placenta et dans les grands vaisseaux vitellins et ombilicaux (Figure 2). La mise en évidence de l'existence de cellules sanguines dans l'embryon humain provient de tests de colonies in vitro réalisés avec des cellules CD34+ (CD pour cluster de différenciation), marqueur connu pour identifier les progéniteurs hématopoïétiques immatures humain, purifiées à partir soit du sac vitellin, soit des tissus intra-embryonnaires (Huyhn et al., 1995).

L'origine de l'hématopoïèse définitive proviendrait directement de l'endothélium vasculaire spécialisé appelé endothélium hémogénique dans la région AGM. La première étude expérimentale prouvant l'existence de cet endothélium transitoire provient d'embryons de poulet, dans lesquels les cellules endothéliales ont été marquées avant l'émergence des CSH (Jaffredo et al., 1998). Comme le marqueur a été détecté par la suite dans les cellules hématopoïétiques, Jaffredo et ses collègues en ont conclu que ces cellules provenaient des cellules endothéliales initialement marquées. Des résultats similaires ont été obtenus chez les embryons de souris, en utilisant la même technique de traçage (Sugiyama et al., 2003). En utilisant une approche transgénique, l'équipe de Zovein et al., ont montré que l'endothélium était capable de créer des CSH (Zovein et al., 2008). Les cellules provenant de l'endothélium migrent vers le foie fœtal, puis vers la MO. Elles ont des capacités d'expansion, d'auto-renouvellement et de différenciation hématopoïétique multilignée. D'autres études ont montré que le rôle du gène Runx1 était essentiel dans les cellules endothéliales pour former des progéniteurs hématopoïétiques et des CSH à partir du système vasculaire (Chen et al., 2009). D'autres études faites sur l'embryon humain ont aussi montré que les cellules endothéliales CD34+CD45- prises dans la région AGM et cultivées en présence d'une couche stromale se différenciaient en cellules hématopoïétiques (Oberlin et al., 2002). Cela a confirmé le concept de l'endothélium hémogénique dans les artères intra-embryonnaires humaines à l'origine du système hématopoïétique embryonnaire.

Le rôle crucial de l'AGM dans la formation de cellules hématopoïétique définitives a suggéré que cet organe avait la capacité autonome de générer des progéniteurs hématopoïétiques. De plus, l'émergence autonome d'une lignée progénitrice myélo-lymphoïde dans l'embryon a été trouvée dans le mésoderme splanchnopleure (feuillet interne de la portion ventrale), dérivé de l'aorte tandis que des progéniteurs plus restreints ont été générés dans le sac vitellin (Tavian et al., 2001).

Même si la principale source de l'hématopoïèse embryonnaire est considérée comme étant située dans l'AGM incluant les vaisseaux vitellins et ombilicaux, il est possible que la population produite dans ces régions soit insuffisante pour tout l'embryon. En effet, la population rare néoformée colonise ensuite le foie fœtal pour le développement des CSH, une quantité suffisante de cellules est nécessaire pour qu'elles puissent continuer de se produire et de se développer dans le thymus fœtal, dans la rate et dans la MO durant toute la vie post-natale. Récemment, le placenta, un organe extra-embryonnaire a été considéré comme un autre organe hématopoïétique (Gekas et al., 2005; Ottersbach and Dzierzak, 2005; Robin et al., 2009). Cela serait dû à sa physiologie contenant des vaisseaux sanguins hautement vascularisés, des cytokines et des facteurs de croissance présents dans l'environnement riche pour l'hématopoïèse et le développement (Cross, 2005). Cependant, la génération de novo des CSH dans cet organe n'a pas été prouvée expérimentalement et le placenta reste encore un organe hématopoïétique de transition avec une fonction unique comme une niche spécialisée pour supporter le développement des CSH exogènes.

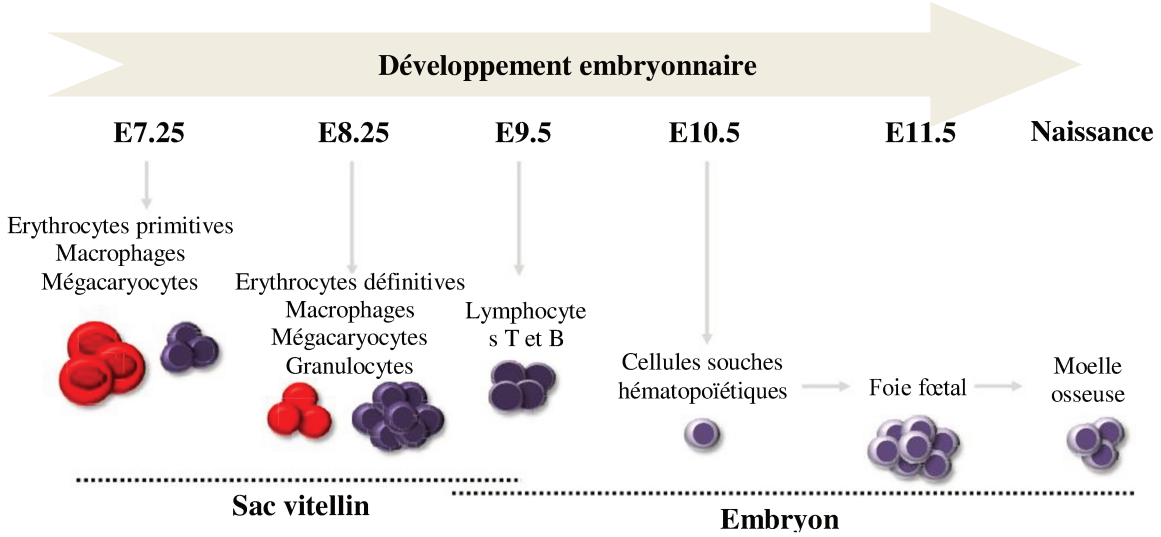


Figure 2 : Ontogenèse des cellules hématopoïétiques.

Représentation schématique du développement hématopoïétique au cours de l'embryogenèse. Schéma adapté de Lacaud and Kouskoff, 2017.

La production des CSH durant le développement est une étape clé dans la création d'un organisme vivant et implique de multiples interactions de régulation durant l'embryogénèse. Il est maintenant relativement accepté que les CSH naissent dans l'embryon à travers les cellules endothéliales hémogéniques.

I.C Caractérisation des cellules souches hématopoïétiques humaines

Malgré sa complexité, le sang est probablement le système de développement le mieux compris. Cliniquement, la transplantation des CSH représente le traitement de régénération le plus développé bien que leur caractérisation soit relativement récente. Le concept de la cellule souche et la vision hiérarchique de l'hématopoïèse suggèrent implicitement que les CSH soient très rares dans l'organisme. Afin de pouvoir les étudier ou les utiliser pour la transplantation, un effort considérable a été investi dans la recherche de marqueurs spécifiques pour pouvoir enrichir la population en CSH (Doulatov et al., 2012; Wang et al., 1997). Pour être définie comme une cellule souche, elle doit être capable de s'auto-renouveler et de se différencier dans tous les types cellulaires qui composent le tissu.

I.C.1 Enrichissement des CSH

L'enrichissement des CSH a nécessité une détection simultanée de plusieurs marqueurs de surface cellulaires. Le CD34 a été le premier marqueur trouvé pour enrichir les CSH et les progéniteurs humains (Civin et al., 1984). Le CD34 identifie aussi des progéniteurs plus différenciés ce qui a incité les scientifiques à chercher d'autres marqueurs membranaires. En utilisant le modèle SCID humain (immunodéficiences combinés sévères), modèle de souris humanisée, Baum et al. ont identifié la protéine CD90 aussi nommée Thy1 comme un marqueur de cellules souches (Baum et al., 1992; Péault et al., 1993; Weissman et al., 2001). En effet, ils ont démontré in vitro et in vivo que les cellules $CD34^{+}CD90^{+}Lin^{-}$ généraient tous les types cellulaires des progéniteurs lymphoïdes et myéloïdes contrairement aux cellules $CD34^{+}CD90^{-}Lin^{-}$. Les marqueurs Lin comprennent des marqueurs spécifiques des cellules T, B, NK (Natural Killer) et myélo-érythroïde.

Un enrichissement supplémentaire des CSH humaines dans la population $CD34^{+}$ a été obtenu grâce à l'expression différentielle du marqueur de surface CD38. Bien que la plupart des cellules $CD34^{+}$ expriment aussi CD38, il a été démontré que les cellules $CD34^{+}CD90^{+}$ et CD38 faibles/négatives pouvaient se différencier à la fois en cellules lymphoïdes et myéloïdes (Huang and Terstappen, 1994; Miller et al., 1999). D'autres études ont introduit le marqueur de surface CD45RA comme marqueur des progéniteurs plus différenciés qui négativement enrichit les CSH (Lansdorp et al., 1990).

Toutes ces données permettraient de conclure que les cellules sélectionnées sur les marqueurs $CD34^{+}CD38^{-}CD90^{+}CD45RA^{-}Lin^{-}$ enrichiraient les CSH humaines alors que les cellules $CD34^{+}CD38^{-}CD90^{-}CD45RA^{-}Lin^{-}$ enrichiraient la population en cellules progénitrices multipotentes (MPP) avec une capacité d'auto-renouvellement incomplète (Majeti et al., 2007). Récemment, Notta et al. auraient identifié un nouveau marqueur spécifique des CSH dans la population cellulaire $CD34^{+}CD38^{-}CD90^{+}CD45RA^{-}Lin^{-}$: le CD49f (intégrine α6) (Notta et al., 2011). En résumant tous les résultats obtenus depuis ces dernières années, il est possible aujourd'hui de dire que les CSH humaines seraient enrichies dans la population cellulaire $CD34^{+}CD38^{-}CD90^{+}CD45RA^{-}CD49f^{+}Lin^{-}$.

I.C.2 Mécanismes moléculaires et cellulaires des CSH humaines

Au cours de ma thèse, j'ai étudié (entre autres) le transcriptome des cellules CD34+ au cours du temps. Je vais donc brièvement présenter quelques gènes que j'ai utilisés pour analyser ces cellules (gènes qui joueraient un rôle dans la différenciation des CSH).

I.C.2.a Régulations moléculaires des CSH

Après la division cellulaire d'une cellule souche, le potentiel multipotent des cellules filles néoformées peut être conservé dans les deux cellules filles, dans l'une des deux ou aucune des deux. Till et al. ont été les premiers à décrire ces résultats en terme de mécanisme stochastique (Till et al., 1964). Les mécanismes moléculaires laisseraient suggérer qu'ils régulent l'équilibre entre l'auto-renouvellement et la différenciation. Ces choix dans le devenir cellulaire seraient associés à des changements dans l'expression des gènes, conduits par des facteurs de transcription. Ils seraient accompagnés, et souvent précédés par des changements épigénétiques (Bernstein et al., 2006). La comparaison transcriptionnelle des CSH $CD49f^{+}$ et des MPP $CD49f^{-}CD90^{-}$ a fourni des informations plus précises sur l'état des cellules souches. En effet, plusieurs facteurs de transcription comme les gènes ID3, SOX18 et NFIB, seraient associés à l'état souche des cellules alors que MYC et IL1RL1 seraient sur-régulés lors de la différenciation en MPP (Notta et al., 2011).
L'extinction du gène BMI1 dans les cellules CD34+ entrainerait une perte du potentiel clonal alors qu'une surexpression de ce gène augmenterait le potentiel multipotent des cellules (Rizo et al., 2009, 2008). En effet, la transplantation en série dans des souris « Non-obese diabetic-SCID » (NOD-SCID) a montré l'effet de BMI1 dans l'auto-renouvellement.

Voie de signalisation de SDF-1 (ou CXCL12)/CXCR4

La voie de signalisation de SDF-1 (stromal-cell derived factor 1) / CXCR4 jouerait un rôle dans le maintien des CSH et régulerait également leur fixation dans les niches (Heissig et al., 2002). L'inactivation ou la suppression de CXCR4 chez la souris a entraîné une réduction des CSH et une hyper-prolifération (Nie et al., 2008; Sugiyama et al., 2006). De plus, Tzeng et al. ont confirmé le rôle de SDF-1 dans la maintenance des CSH en utilisant une souris déficiente en SDF-1 (Tzeng et al., 2011).

Voie de signalisation c-Mpl/TPO

Le récepteur c-Mpl et son ligand, la thrombopoïétine (TPO), sont connus pour réguler la mégacaryopoïèse (Broudy and Kaushansky, 1995). C-Mpl est exprimé principalement à la surface des CSH et à un niveau moins important chez les progéniteurs mégacaryocytes, les mégacaryocytes et les plaquettes (Debili et al., 1995). Cette voie de signalisation est impliquée dans la maintenance des CSH post-natales à l'état d'équilibre (Qian et al., 2007; Yoshihara et al., 2007). Il a été démontré que les CSH à long terme exprimant le récepteur c-Mpl se trouvaient dans un état quiescent dans la MO adulte et étaient en étroite association avec les cellules ostéoblastes productrices de la TPO (Yoshihara et al., 2007). TPO a donc une fonction importante dans la maintenance et l'auto-renouvellement des CSH.

Voie de signalisation de Notch

La voie de signalisation de Notch jouerait aussi un rôle essentiel dans la prolifération, la différenciation et la prise de décision du destin cellulaire des CSH (Artavanis-Tsakonas et al., 1999; Chotinantakul and Leeanansaksiri, 2012). Quatre récepteurs (Notch 1 à 4) et cinq ligands (Jagged 1 et 2 et Delta 1, 3 et 4) ont été identifiés dans la voie de signalisation de Notch chez les mammifères. Les cellules exprimant des ligands Notch pourraient maintenir ou améliorer l'auto-renouvellement des CSH dans la culture (Ohishi et al., 2002; Varnum-Finney et al., 2011). En effet, certaines recherches ont démontré in vitro et in vivo une diminution de la différenciation des CSH suite à l'interaction des récepteurs Notch et de ses ligands (Stier et al., 2002; Varnum-Finney et al., 2003). Ces études confirment le rôle de la signalisation de Notch médiée dans l'hématopoïèse et la maintenance des CSH. Néanmoins, certaines études ont montré que la voie de signalisation de Notch ne serait pas importante pour l'auto-renouvellement et la maintenance des CSH (Maillard et al., 2008; Mancini et al., 2005). L'inactivation de Notch-1 et Jagged-1 dans les progéniteurs de la MO et les cellules stromales de la MO n'aurait pas altéré le maintien et la reconstitution des CSH (Mancini et al., 2005).

I.C.2.b Propriétés cellulaires des CSH

Les CSH de la MO de souris sont connues pour se trouver dans un état de dormance, de quiescence (Lord et al., 1976). Etant donné que l'existence de CSH humaines quiescentes ne peut être évaluée directement, des mesures dites « statiques » par cytométrie en flux ont permis de montrer que la grande majorité des CSH humaines se trouvait en phase G0 du cycle cellulaire (Passegué et al., 2005). Des études qui s'appuient sur des paramètres mesurés dans des cellules différenciées, telles que la longueur moyenne des télomères ou les taux de changement des phénotypes des chromosome X maternels / paternels avec l'âge, ont estimé que les CSH humaines se divisaient tous les 175 à 350 jours (Catlin et al., 2011; Shepherd et al., 2004). Ainsi, cela supposerait que la division cellulaire peu fréquente serait un mécanisme de protection limitant l'accumulation de mutations d'acide désoxyribonucléique (ADN) dus aux réplications et aux stress oxydatifs.

Mohrin et ses collègues ont montré chez la souris que les CSH possédaient des mécanismes intrinsèques uniques qui assureraient leur survie en réponse à l'irradiation ionisante (Mohrin et al., 2010). Cela inclueraient l'expression de gènes « pro-survis » et une forte activation de p53 suite à la réponse de l'ADN endommagé. Ils ont aussi montré que les CSH quiescentes utilisaient différents types de mécanismes de réparation de l'ADN comme la jonction d'extrémités non homologues (NHEJ de l'anglais non-homologous end-joining) qui permet de réparer des lésions provocant des cassures double brin. Tous ces mécanismes moléculaires et protéiques sembleraient maintenir un équilibre entre l'auto-renouvellement et la différenciation en cellules sanguines.

I.C.3 La cellule souche : une entité ou un état ?

Les cellules souches, comme nous avons vu précédemment, sont généralement définies en rétrospective par des analyses fonctionnelles qui testent les propriétés cellulaires. Par exemple, dans le système hématopoïétique, nous avons vu que plusieurs marqueurs étaient utilisés pour définir une population qui sera ensuite isolée et injectée chez des animaux examinés plusieurs mois plus tard. Cependant, même avec des techniques de séparation de cellules souches les plus strictes, la population de CSH résultante purifiée reste hétérogène (Jobin et al., 2015; Levsky et al., 2002). Des études récentes sur la MO prise entièrement, non triée, ont montré que les CSH greffées à long terme se trouvaient dans un cycle cellulaire actif et changeaient continuellement de phénotype (Goldberg et al., 2014).

Par conséquent, ces cellules ne peuvent pas être purifiées par ces approches dépendant de l'expression stable de l'épitope de surface car les marqueurs de surface changent continuellement aussi. Ces cellules cycliques sont éliminées avec les méthodes de purification actuelle des cellules souches (Quesenberry et al., 2014). Par exemple, l'analyse de transplantation à une seule cellule a révélé que les cellules qui présentaient des capacités de repopulation les plus élevées (marqueur hoechst) et le phénotype $CD34^{-}cKit^{+}Sca\text{-}1^{+}Lin^{-}$ (marqueurs de surface des CSH de souris) présentaient une prolifération très forte et une capacité de différenciation multipotente (Matsuzaki et al., 2004). 96% des souris irradiées ont reçu efficacement la cellule purifiée. Au premier coup d'œil, cela semblerait bien prouver qu'il s'agirait d'une population pure de cellules souches. Cependant, il existe d'autres cellules qui ne s'ensemencent pas bien dans la MO mais qui présentent un phénotype clair de cellules souches lors de la greffe dans la MO (Mazurier et al., 2003). Par conséquent, la définition des cellules souches selon les propriétés de purification exclurait d'autres cellules qui possèderaient des phénotypes évidents fonctionnels de cellules souches et qui pourraient être aussi importantes ou abondantes.

On pourrait voir la cellule souche plutôt comme un multi-état transitoire et réversible et non comme une entité définie par des marqueurs spécifiques (Blau et al., 2001; Zipori, 2004). D'ailleurs, chez les amphibiens et les plantes, cette réversibilité existe et est pleinement acceptée alors que la différenciation chez les mammifères est décrite en générale comme étant unidirectionnelle et non réversible (Tanaka, 2003; Weigel and Jürgens, 2002). On peut alors se poser la question : pourquoi cette réversibilité n'« existe » pas chez les mammifères ? Cela serait-il dû à une mauvaise compréhension et définition d'une cellule souche ?

Malgré cela, la recherche qui définit des caractéristiques comme la capacité d'auto-renouvellement, l'engagement de lignage, les niches de MO et l'état prolifératif des CSH continue de se concentrer principalement sur cette petite sous-population de cellules purifiées (Quesenberry et al., 2015). Bien que les CSH soient largement étudiées, le choix du destin cellulaire entre soit rester dans un état multipotent, soit s'engager dans une voie de différenciation est encore mal compris.
Avant d'aborder la question sur les différents modèles d'hématopoïèse, je vais développer un paragraphe sur les différents types de cellules sanguines pour lesquelles la CSH se différencie avec leurs caractéristiques et leurs fonctions.

II. Différents types de cellules sanguines

Les CSH, comme dit précédemment, sont des cellules multipotentes capables de donner toutes les cellules du sang. Afin de mieux comprendre mes travaux qui portent sur le devenir des lymphocytes T (LT) en lymphocytes T régulateurs (LTreg) et sur les cellules $CD34^+$, une brève description des différents types de cellules sanguines et de leur fonction sera faite. Les cellules du sang se trouvent en suspension dans le plasma, dans le sang et ont une durée de vie limitée ; c'est pourquoi il y a un perpétuel équilibre entre leur production, l'hématopoïèse, et leur destruction. Il existe trois grandes catégories de cellules sanguines caractérisées selon leur fonction et leur phénotype : les érythrocytes (globules rouges), les thrombocytes (plaquettes) et les leucocytes (globules blancs).

II.A Erythrocytes

Les érythrocytes ou les globules rouges ont pour rôle de transporter l'oxygène aux différents tissus du corps humain et de rapporter le gaz carbonique aux alvéoles pulmonaires. Ce ne sont pas des cellules à proprement parler car ils sont dépourvus de noyau et ne peuvent donc pas se reproduire ; ils sont dits anucléés. Ils représentent plus de 90% des cellules du sang et ont une forme biconcave, aplati au centre (Figure 3.A). Le transport de l'oxygène et du gaz carbonique se fait par l'intermédiaire d'une hémoprotéine appelée hémoglobine (Figure 3.B). Cette hémoglobine est formée de globine, protéine associée à quatre groupements hèmes qui contiennent chacun un atome de fer entouré d'un anneau organique appelé porphyrine.

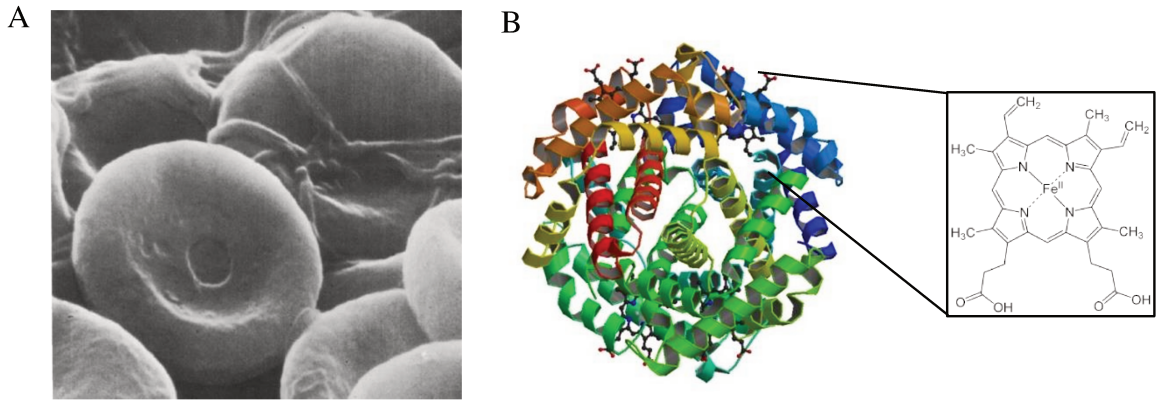


Figure 3 : Erythrocytes.

A. Globules rouges au microscope électronique à balayage (x13 500) (Source : McDonald and Hayes, 1969).

B. Représentation hétéro-tétramère de l'hémoglobine avec un grossissement de la structure de l'hème. (Source : rcsb.org).

En utilisant des effets synergiques de l'hémoglobine, de l'anhydrase carbonique et des échanges d'anion de la protéine « bande 3 », les érythrocytes jouent un rôle important dans le système et constituent un transporteur idéal pour délivrer l'oxygène aux tissus en fonction de leur activité métabolique (Hamasaki and Yamamoto, 2000).

Les globules rouges, comparés aux autres cellules, ne contiennent pas d'organites intracellulaires comme les mitochondries, les ribosomes ou les réticulums endoplasmiques granuleux et sont incapables de synthétiser des acides nucléiques, des protéines ou des lipides. La glycolyse est l'unique voie métabolique dans ces éléments figurés pour produire de l'énergie. Les érythrocytes sont assez flexibles et peuvent ainsi pénétrer rapidement à travers les vaisseaux capillaires (de l'ordre d'une cellule par milliseconde pour passer à travers les vaisseaux capillaires pulmonaires) pour apporter l'oxygène dans tout le corps (Wagner, 1977; Wieth et al., 1982). Les érythrocytes ont une durée de vie de 120 jours et sont programmés pour mourir. Le signal de mort cellulaire provient probablement des altérations de leur membrane soit par des dommages accumulés conduisant à une rigidité de leur membrane, soit par l'exposition d'antigènes qui lient les immunoglobulines, ce qui les rend accessible aux macrophages (Landaw, 1988; Turrini et al., 1991).

II.B Thrombocytes

Les thrombocytes ou les plaquettes sanguines représentent 4 à 7% des cellules du sang et sont les plus petits éléments figurés du sang. Comme les érythrocytes, ils ne contiennent pas de noyau et sont présents en très grand quantité dans le système sanguin (entre 150 et $400x10^9$ plaquettes/L) (Klinger and Jelkmann, 2002). Les plaquettes sanguines sont des fragments de cellules mégacaryocytes issues de la MO (Figure 4).

Ils jouent un rôle essentiel dans les processus de coagulation, d'hémostase, d'inflammation et de réparation des tissus (Klinger and Jelkmann, 2002). Le taux de production des plaquettes est augmenté en réponse à des épisodes de thrombopénies ($<150x10^9$ plaquettes/L) et supprimé suite à l'induction de thrombocytoses ($>400x10^9$ plaquettes/L). Ils ont une durée de vie très courte (comprise entre 7 à 10 jours) et se dégradent dans la rate.

Dans le cas de l'hémostase, les thrombocytes sont les premiers éléments à aller sur le site endommagé de la paroi vasculaire et dans le tissu environnant. Ils s'attachent alors à des fibres de collagènes ou à d'autres composants de la matrice sous-endothéliale qui sont normalement cachés (Rossaint and Zarbock, 2015). La formation de liaisons stables entre les récepteurs d'adhésion plaquettaires et le collagène induit l'activation des plaquettes entrainant ainsi leurs agrégations grâce aux faisceaux de microtubules et à la libération de cytokines et de médiateurs pro-coagulant hautement actif comme la thrombine ou la prostaglandine.

Ce phénomène provoque la formation d'un thrombus qui incorpore également des érythrocytes et des leucocytes (Gawaz et al., 2005). Cependant, la lésion tissulaire n'entraine pas que la perte de sang mais implique aussi une entrée possible d'agents pathogènes exogènes comme les bactéries ou les virus, dans les tissus plus profonds, permettant ainsi leur circulation, ce qui laisse l'organisme vulnérable aux infections locales et/ou systémiques. Les plaquettes peuvent donc soit stimuler, soit inhiber les fonctions des autres cellules dans l'inflammation et favoriser la migration trans-endothéliale des leucocytes vers le tissu environnant.

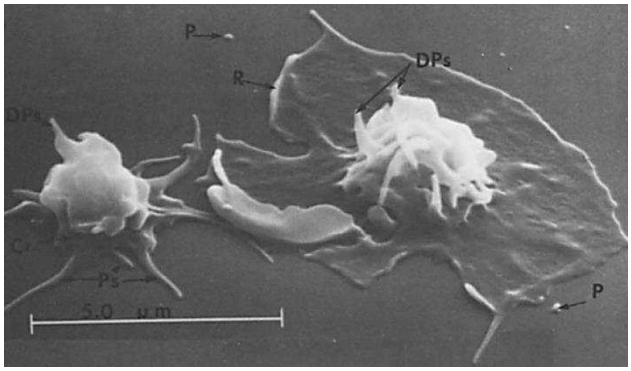


Figure 4 : Thrombocytes en microscopie électronique à balayage.

Plaquette à gauche : stade de propagation précoce avec pseudopodes attachés (Ps) et pseudopodes dorsaux non attachés (DP). Plaquette à droite : plaquette avec hyalomère partiellement étendues avant la rétraction des pseudopodes dorsaux (DP). Des particules exocytées (P) sont visibles au voisinage de ces plaquettes. R : ondulation. (Source : Allen et al., 1979).

Le cytoplasme des plaquettes est très riche en organites (mitochondries, appareil de Golgi, cytosquelettes), en inclusion (glycogène) et contient de nombreuses granulations (α-granules, granulations denses et granules de lysosome) (Rendu and Brohard-Bohn, 2001). Les granulations α contiennent des facteurs de coagulation (fibrinogène), de la fibronectine (protéine adhésive) et des facteurs de croissance. Les granulations denses sont reliées à la membrane et à la surface de la cellule et contiennent du calcium, de l'adénosine diphosphate et de la sérotonine qui active la vasoconstriction. Les plaquettes sont aussi équipées d'une large gamme de récepteurs à leur surface et des molécules d'adhésion.

Les plaquettes sanguines jouent donc un rôle important dans les défenses antimicrobiennes et induisent l'inflammation et la réparation des tissus en plus de leur participation dans l'hémostase (Yeaman, 1997).

II.C Leucocytes

Les leucocytes ou les globules blancs sont les cellules sanguines les moins nombreuses (0,1 à 0,2% des cellules du sang). Contrairement aux érythrocytes et aux thrombocytes, ils sont caractérisés par une taille plus grande et par la présence d'un noyau. Ce sont des cellules produites dans la MO qui jouent un rôle dans le système immunitaire en participant à la protection contre les agressions d'organismes extérieurs et en faisant la discrimination entre le « soi » et le « non soi ». Suite à l'introduction d'un agent infectieux dans le corps tel que des virus, des bactéries, des champignons, etc., le système immunitaire se met en route. L'immunité innée est d'abord déclenchée, ce qui entraine une réponse immédiate et non spécifique de l'antigène. Elle se caractérise par une phagocytose du pathogène qui entraine une inflammation et une activation du système du complément. Ensuite, vient l'immunité adaptative qui est caractérisée par une réponse plus lente et spécifique de l'antigène ; elle met en place la mémoire immunitaire à long terme.

Les leucocytes circulent dans le sang et dans le système lymphatique (rate, thymus, ganglion lymphatique, etc.) et sont recrutés lors des lésions tissulaires et des infections. Il existe trois types de globules blancs : les monocytes, les lymphocytes et les granulocytes distinguables selon leurs caractéristiques fonctionnelles et physiques.

II.C.1 Monocytes

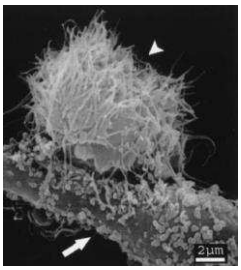
Les monocytes sont les plus grosses cellules qui ne prolifèrent pas et qui circulent dans le sang, dans la MO et dans la rate (Swirski et al., 2009). Ils représentent 5 à 10% des globules blancs totaux chez l'adulte et ont une morphologie hétérogène circulaire ou ovale (Pardali and Waltenberger, 2012). Ils appartiennent à la famille des phagocytes mononucléés et représentent les cellules immunitaires innées effectrices équipées de récepteurs à chimiokines et d'adhésion qui servent d'intermédiaire à la migration du sang aux tissus pendant l'infection (Geissmann et al., 2010). Les monocytes produisent des cytokines inflammatoires et peuvent absorber les cellules et les molécules toxiques. Ils peuvent se différencier soit en macrophages, soit en cellules dendritiques durant l'inflammation (Figure 5). La migration vers les tissus et la différenciation en cellules dendritiques inflammatoires et macrophages sont probablement déterminées par le milieu inflammatoire et les récepteurs de reconnaissance des motifs associés aux pathogènes (Serbina et al., 2008).

II.C.1.a Cellules dendritiques

les cellules dendritiques sont des cellules présentatrices d'antigène qui ont deux fonctions principales (Figure 5.A) (Banchereau and Steinman, 1998). La première est le déclenchement de la réponse immunitaire adaptive : une fois avoir phagocyté une particule antigénique, les cellules dendritiques migrent vers les organes lymphoïdes et présentent l'antigène du « non soi » aux lymphocytes T naïfs. La deuxième fonction est le maintien de la tolérance du « soi » dans la rate ou le thymus en présentant des antigènes du « soi » aux lymphocytes T.

II.C.1.b Macrophages

Les macrophages résident dans les tissus lymphoïdes et non lymphoïdes et ont pour rôle de phagocyter les débris cellulaires et les agents pathogènes (Figure 5.B). Ils sont aussi capables de se comporter comme des cellules présentatrices d'antigène.

A. 

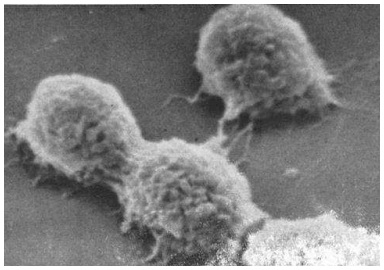
B. 

Figure 5 : Cellules dendritiques et macrophages en microscopie électronique à balayage.
A. Cellule dendritique (tête de la flèche) accolée à une cellule endothéliale (flèche). (Source : Van Nguyen et al., 2004). B. Macrophages posés sur une surface en verre (x 4 000). (Source : Carr and Carr, 1970)

Les monocytes jouent un rôle essentiel dans l'élimination des bactéries, des virus, des substances toxiques et dans l'éradication des cellules apoptotiques et nécrotiques. Ils jouent aussi un rôle important dans l'angiogenèse (formation de vaisseaux sanguins), dans l'artériogenèse et dans la réparation des tissus lésés (Bergmann et al., 2006; Shi and Pamer, 2011).

II.C.2 Lymphocytes

Les lymphocytes sont des cellules mononucléées de petites tailles qui ont un rôle majeur dans le système immunitaire adaptatif (Figure 6.A). Ils représentent 30 à 50% des globules blancs circulant dans le sang (Mestas and Hughes, 2004). Il existe deux grandes lignées lymphocytaires différentes : les lymphocytes T (LT) et les lymphocytes B (LB). Un troisième groupe de lymphocytes, les Natural Killer (NK), ressemblant aux LT, existe mais ne sera pas développé ici.

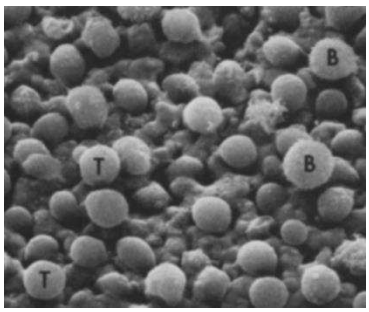
II.C.2.a Lymphocytes B

Les LB naissent, prolifèrent et se différencient dans la MO. Ils sont responsables de l'immunité humorale c'est-à-dire une immunité adaptative liée à la production d'anticorps (Figure 6.B). Les LB naïfs une fois matures quittent la MO vers la circulation sanguine puis lymphatique. Ils expriment à leur surface différents récepteurs immunoglobulines reconnaissant spécifiquement les épitopes des antigènes présentés par les cellules présentatrices d'antigène (macrophages, cellules dendritiques) (LeBien and Tedder, 2008). Si l'immunoglobuline a reconnu spécifiquement l'antigène, les LB sont alors activés, prolifèrent et se transforment soit en plasmocytes capables de fabriquer des anticorps dirigés contre l'antigène activateur, soit en LB mémoires circulant indéfiniment dans l'organisme pour réagir plus rapidement lors d'une deuxième exposition à l'antigène complémentaire.

II.C.2.b Lymphocytes T

Les LT, eux, se différencient dans le thymus d'où le « T » du lymphocyte T (Figure 6.C). Tous les LT matures expriment à leur surface membranaire le récepteur protéique CD3 ; c'est un marqueur « pan T ». Ce sont des cellules responsables à la fois d'immunité humorale et cellulaire c'est-à-dire des cellules infectées par des virus sont détruites par un mécanisme complexe (Liu et al., 2016). Ils existent plusieurs LT matures :

- Les LT cytotoxiques exprimant le CD8 à leur membrane et détruisant les cellules infectées. Après avoir reconnu la cellule étrangère, les LT cytotoxiques libèrent une substance chimique, appelée perforine, qui vient s'insérer dans la membrane plasmique de l'hôte et la perfore ce qui entraine une explosion de la cellule.
- Les LT auxiliaire ou helper (LT_H) exprimant le CD4 et activant les cellules de la réaction immunitaire (LT CD8+, LB). LT_H les plus classiques sont :
    - Les LT $CD4^+$ T_H1 qui s'orientent vers une réponse cytotoxique en participant à l'activation des LT $CD8^+$;
    - Les LT $CD4^+$ T_H2 qui s'orientent vers une réponse plus humorale en activant les LB ;
    - Les LT $CD4^+$ T_H17 qui protègent l'organisme contre les bactéries et les champignons ;
- Les LT suppresseurs ou régulateurs (LTreg) ayant à leur surface le CD4 et le CD25 et exprimant spécifiquement Foxp3 (Khattri et al., 2003; Sakaguchi et al., 2001). Ils ont la propriété d'inhiber la prolifération d'autres LT ou de les supprimer. Ils sont nécessaires au maintien de la tolérance immunitaire.

A. 

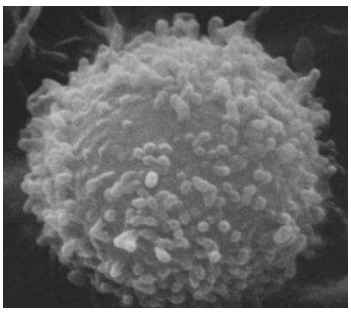
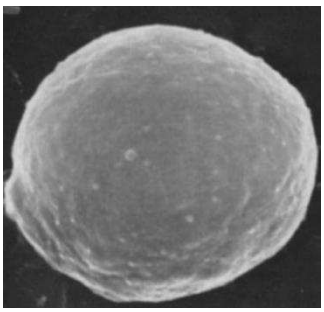
B. 
C. 

Figure 6 : Lymphocytes B et T en microscopie électronique à balayage.

A. Mélange de lymphocytes T et B (LT et LB) (x2 250). La majorité des LT ont une surface lisse. La majorité des LB ont un petit nombre de villosités à leur surface. B. Grossissement d'un LB (x18 000). C. Grossissement d'un LT (x14 000). (Source : Polliack et al., 1973).

II.C.3 Granulocytes

Les granulocytes, appelés aussi cellules polynucléaires, sont des cellules qui ont un noyau plurilobé dont les lobes sont reliés ensembles par la chromatine et qui sont caractérisés par des granules dans leur cytoplasme. Il existe trois types de granulocytes aux fonctions différentes selon le contenu de leurs granules : les neutrophiles, les éosinophiles et les basophiles (Figure 7). Ils jouent un rôle important dans l'inflammation à la fois dans l'immuno-régulation et dans le nettoyage des agents pathogènes. Ils se différencient dans la MO, puis circulent dans le sang et restent en phase G0 du cycle cellulaire pendant la durée de leur vie plutôt courte (quelques jours) (Geering et al., 2013). Ce sont des globules blancs dits « non spécifiques » car ils ne sont pas dirigés contre un seul antigène.

II.C.3.a Neutrophiles

50 à 70% des globules blancs circulants sont des neutrophiles (Mestas and Hughes, 2004). Leur noyau est généralement trilobé mais il peut varier de deux à cinq lobes et leur cytoplasme renferme des granules contenant des protéines pro-inflammatoires et des vésicules sécrétoires (Figure 7.A). En l'absence de stimuli extracellulaire, les neutrophiles circulants subissent une apoptose spontanée dans les 1 à 5 jours après leur formation (Pillay et al., 2010).

Les neutrophiles font partis de la famille des phagocytes et peuvent éliminer les agents pathogènes par de multiples moyens intra et/ou extracellulaires (Kolaczkowska and Kubes, 2013). Lorsque les neutrophiles rencontrent des microorganismes, ils les phagocytent et les tuent grâce à des protéines antibactériennes contenues dans ces granules.

Ces protéines antibactériennes peuvent aussi être libérées dans le milieu extracellulaire, agissant ainsi sur des agents pathogènes extracellulaires.

Les neutrophiles hautement activés peuvent aussi éliminer les microorganismes extracellulaires en libérant des « pièges extracellulaires neutrophiles ». Ces pièges sont constitués d'ADN auquel sont attachées des histones, des protéines antibactériennes et des enzymes (comme l'élastane). Ce piège immobilise les agents pathogènes, ce qui les empêche de se propager et facilite la phagocytose des microorganismes piégés. Ils tueraient également directement les agents pathogènes au moyen d'histones et de protéases antimicrobiennes (Papayannopoulos and Zychlinsky, 2009).

II.C.3.b Eosinophiles

Ces cellules représentent 1 à 3% des globules blancs dans le sang, ont un noyau bilobé et se trouvent à l'état immature dans la MO (Figure 7.B). Les éosinophiles matures sont libérés dans le sang périphérique où ils y restent très peu de temps (demi-vie de 18h) (Steinbach et al., 1979). Une fois activée par des stimuli inflammatoires, ils migrent vers les tissus périphériques où ils peuvent vivre pendant plusieurs semaines. Comme les neutrophiles, les éosinophiles ont un noyau bilobé et contiennent dans leur cytoplasme de nombreux granules volumineux. Ces granules sont constitués de plusieurs protéines, d'enzymes, de cytokines et de facteurs de croissances (Rosenberg et al., 2013).

La fonction principale de l'éosinophile est de détruire les parasites de l'organisme, sans les phagocyter : c'est la dégranulation c'est-à-dire la libération de leur contenu granulé protéique dans l'espace extracellulaire (Rosenberg et al., 2013). Il existe trois mécanismes de dégranulation :

- La dégranulation cytolytique, mécanisme par lequel une attaque pathogène entraîne une extrusion complète du contenu (protéines cationiques) des éosinophiles.
- La dégranulation progressive ou fragmentaire (Dvorak et al., 1994). Des morceaux du contenu granulé des éosinophiles sont libérés par bourgeonnement de petites vésicules en réponse à un stimulus donné, ce qui leur permet de rester en vie. C'est le mode de dégranulation le plus fréquent.
- L'exocytose. La membrane des granules fusionne avec la membrane plasmique de l'éosinophile et libère le contenu granulé vers l'extérieur.

II.C.3.c Basophiles

Les basophiles sont les cellules les moins nombreuses parmi tous les granulocytes (< à 1 %) ; c'est l'une des raisons pour laquelle leur rôle est peu connu (Chirumbolo, 2012). De tous les globules blancs, ils ont la durée de demi-vie la plus longue dans le sang (62h) (Steinbach et al., 1979). Comme les neutrophiles et les éosinophiles, ils ont un noyau plurilobé et contiennent de multiples granules dans leur cytoplasme (Figure 7.C). Ils ont un rôle essentiel dans les réactions allergiques et dans le système immunitaire (Karasuyama et al., 2011). Lorsqu'ils rencontrent un allergène ou sont déclenchés par des stimuli, les basophiles s'activent et provoquent des réactions d'hypersensibilité larguant des substances vaso-actives dans le microenvironnement (Chirumbolo, 2012). Ils utilisent le même mécanisme que les éosinophiles, la dégranulation, libérant ainsi le contenu des granules renfermant des histamines (vaso-dilatateur), des protéoglycanes (ex : héparine) et des enzymes. Cette dégranulation due à l'allergène serait donc responsable des manifestations allergiques.

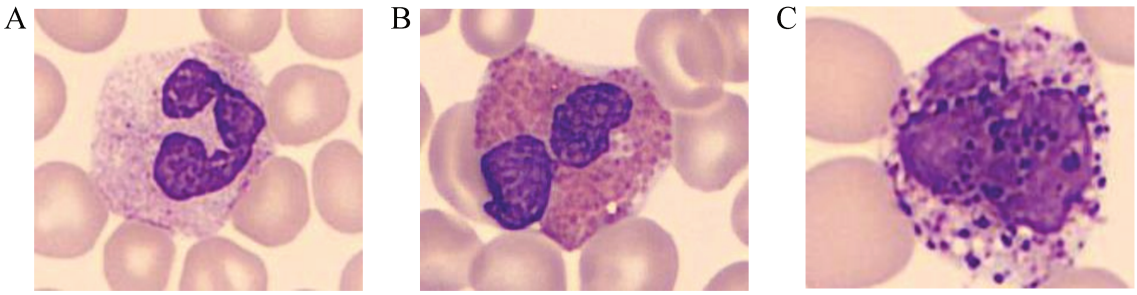


Figure 7 : Les trois types de granulocytes par coloration au May Grünwald Giemsa.

A. Neutrophile avec un noyau trilobé et peu de granules. B. Eosinophile avec un noyau bilobé et dont les granules sont colorés en rouge dus à l'éosine contenu dedans. C. Basophile avec un noyau multilobé avec des granules colorés en bleu. (Source des images : hematocell.fr).

Toutes ces cellules sanguines proviennent donc d'un long processus de différenciation d'un seul type de cellule : la CSH. Malgré les nombreuses recherches sur la compréhension du devenir des CSH, notre connaissance sur les mécanismes de décision du développement cellulaire reste insuffisante. En effet, il n'existe pas à l'heure actuelle un seul modèle de différenciation reconnu par tous les scientifiques qui expliquerait la différenciation de ces cellules en ces différents types cellulaires (Kawamoto and Katsura, 2009).

Deux modèles de différenciation totalement opposés sont proposés dans la littérature :

- une modèle purement déterministe où finalement la cellule est déjà prédéterminée à se diriger dans une voie de différenciation et une fois engagée, elle ne peut pas revenir en arrière ;
- un modèle stochastique où la cellule, due aux fluctuations aléatoires internes et externes, a toutes les probabilités de se différencier en n'importe quels types cellulaires voire même de pouvoir changer de voie au cours du processus.

Dans le cadre de mon sujet de thèse, ces deux modèles seront détaillés dans le prochain paragraphe afin de comprendre la prise de décision dans le choix du modèle dans la suite du travail.

III. Différenciation des cellules souches hématopoïétiques : l'hématopoïèse

Les CSH ont la capacité à s'auto-renouveler et à se différencier en toutes les cellules du sang. L'engagement de ces cellules dans la différenciation, l'auto-renouvellement, la prolifération ou la mort cellulaire est une question largement abordée en biologie depuis des décennies. Le débat conceptuel tourne autour de la nature des cellules souches : s'agit-il d'un type cellulaire ou d'un état transitoire ? La réponse à cette question a des conséquences pratiques et théoriques notamment sur les explications du déroulement de l'hématopoïèse. Plusieurs modèles sont proposés pour expliquer l'hématopoïèse ; ils peuvent principalement être classés en deux groupes : le modèle déterministe où la cellule est « prédéterminée » à s'engager dans une voie et le modèle stochastique où la cellule a le choix de se différencier dans n'importe quelle voie avec une certaine probabilité.

III.A Vision historique de l'hématopoïèse : modèle déterministe

Le modèle déterministe est un modèle où le destin de la CSH est défini à l'avance. La différenciation procède à travers des étapes intermédiaires, prédéterminées où il n'est pas possible de revenir à son état initial. En effet, une fois engagée pour devenir un type cellulaire, la cellule ne peut pas changer en cours de route ou revenir à son état d'origine pour changer de chemin. Plusieurs modèles existent dans la littérature et ont évolués au cours du temps. La plupart des études sur l'hématopoïèse ont été réalisées chez la souris pour des raisons évidentes d'accessibilité et de manipulations in vivo. Bien que le sang humain soit différent de celui de la souris, l'hypothèse générale sur la hiérarchie de l'hématopoïèse humaine est qu'elle serait compatible avec les modèles proposés chez la souris. Dans le prochain paragraphe, l'évolution de ces différents scénarios sera développée.

III.A.1 Modèle classique de l'hématopoïèse : premier modèle explicatif

Pour rappel, il existe deux grands groupes fonctionnels dans le sang : des cellules qui ont un rôle dans le système immunitaire, les leucocytes (les monocytes, les lymphocytes et les granulocytes) et les autres cellules qui n'ont aucun rôle dans l'immunité cellulaire : les érythrocytes et les thrombocytes. Une première explication simple de l'hématopoïèse serait qu'elle est basée sur les fonctionnalités des cellules. Cependant, historiquement, les cellules sanguines sont catégorisées différemment, principalement basées sur des observations histologiques (Kawamoto and Katsura, 2009).

L'observation en microscopie a été l'outil le plus élémentaire pour étudier l'hématologie depuis plus de 100 ans. En effet, de nombreuses cellules sanguines ont été décrites sur leurs

caractéristiques morphologiques comme les érythrocytes, les granulocytes, les macrophages et les lymphocytes (Bloom et al., 1975). Comme dit précédemment, les granulocytes peuvent être classifiés en 3 sous-groupes avec les neutrophiles, les éosinophiles et les basophiles et les monocytes comme des précurseurs des macrophages et des cellules dendritiques. Les granulocytes et les monocytes font parties d'une même lignée appelée lignée myéloïde car ces cellules se différencient dans la MO comparées aux lymphocytes qui appartiennent à la lignée lymphoïde. Malgré que les LB se différencient aussi dans la MO, les morphologistes, à l'époque, ne l'ont pas reconnu comme telles car les cellules B se ressemblent indépendamment de leur stade de développement.

Bien que les globules rouges, les plaquettes, les monocytes et les granulocytes se différencient au même endroit dans la MO, les morphologistes avaient observé que leurs précurseurs étaient différents en termes de formes et de propriétés de coloration. C'est pour toutes ces raisons que les hématologues ont placé la différenciation des érythrocytes, des plaquettes et des cellules de la lignée myéloïde sous un même précurseur commun, appelé progéniteur commun érythro-myéloïde ou plus simplement progéniteur commun myéloïde (CMP) (Figure 8) (Akashi et al., 2000).

Contrairement aux autres cellules immunitaires, les LB et les LT, même s'ils ont des fonctions différentes, répondent de manière spécifique à l'antigène. C'est pourquoi les chercheurs ont séparé ces cellules des autres dans l'hématopoïèse et ont pensé que leur différenciation était issue d'un même progéniteur commun autre que le CMP : le progéniteur commun lymphoïde (CLP) (Kondo et al., 1997).

En résumé, dans ce premier modèle classique, l'hématopoïèse est décrite comme une hiérarchie cellulaire maintenue par des CSH auto-renouvelables qui résident au sommet de la structure pyramidale des cellules différenciées (Till and McCulloch, 1980). Une fois qu'elles s'engagent dans la différenciation, elles génèrent soit des CMP, soit des CLP qui produisent des cellules myéloïdes ou érythroïdes et des cellules B ou T, respectivement.

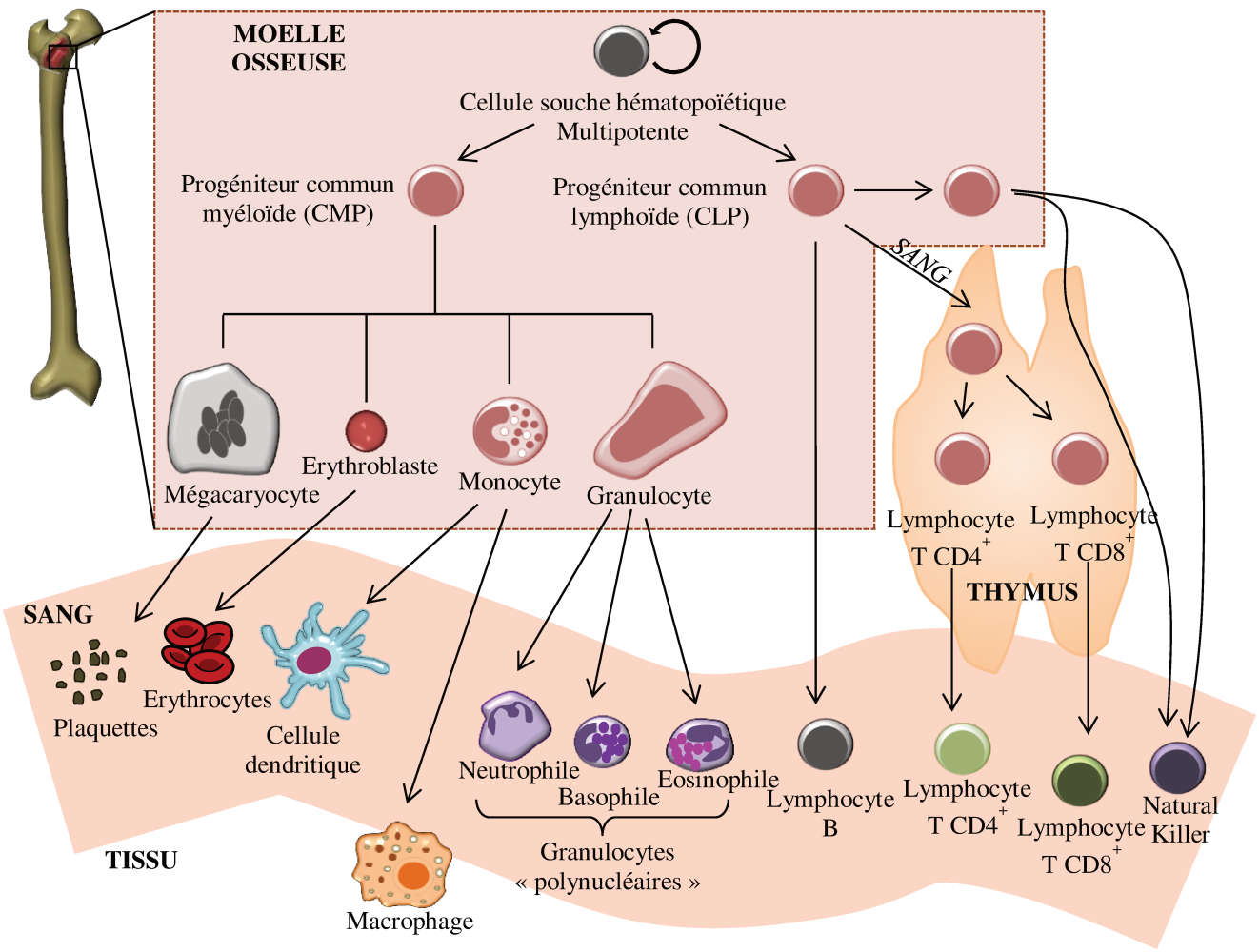


Figure 8 : Premier modèle « classique » de l'hématopoïèse.

Les CSH multipotentes génèrent deux types de cellules progénitrices : les progéniteurs communs myéloïdes (CMP) et les progéniteurs communs lymphoïdes (CLP). Les érythrocytes, les plaquettes, les monocytes et les granulocytes proviennent des CMP et les lymphocytes T, B et NK proviennent des CLP. Schéma adapté de Fisson and Vendomèle, 2017.

L'isolement des progéniteurs de souris engagés dans les lignées myéloïdes et lymphoïdes (CMP et CLP, respectivement) en utilisant la cytométrie en flux a conduit à la formulation du premier modèle « classique » complet de l'hématopoïèse (Figure 9.A) (Akashi et al., 2000; Doulatov et al., 2012; Kondo et al., 1997).

La première hypothèse de ce modèle est que la perte de la capacité d'auto-renouvellement des CSH pendant la différenciation précède l'engagement de la lignée. L'état transitoire d'une population, appelée MPP, avant l'engagement dans une lignée serait à l'origine de cette perte (Morrison et al., 1997).

La seconde hypothèse est qu'après la première séparation des cellules myéloïdes et lymphoïdes, les CMP donnent naissance à :

- des progéniteurs granulocytes-monocytes (GMP), qui s'engagent dans le devenir des monocytes et des granulocytes ;
- des progéniteurs érythroïdes-mégacaryocytes (MEP), qui ne produisent que des cellules érythroïdes et des mégacaryocytes.

Du côté lymphoïde, les CLP donnent naissance à des précurseurs de cellules B et les premiers progéniteurs thymiques engagés dans les lignées T et NK.

III.A.2 Autres modèles déterministes : modèle classique revisité

Depuis quelques décennies, plusieurs modèles déterministes autres que le modèle classique ont été proposés dans la littérature. Une grande partie de ces modèles va être décrite dans ce paragraphe.

En 1985, l'équipe de Brown a proposé un modèle, appelé « détermination séquentielle », où les MPP modifient leur potentiel pour créer des cellules érythroïdes puis myéloïdes, puis des LB et enfin des LT (Figure 9.B). Au cours de la différenciation, des progéniteurs bi-potentiels sont produits entre chacune des étapes (Brown et al., 2007, 1985).

En 1996, l'équipe de Singh a proposé un autre modèle où les MPP donnent soit des cellules érythroïdes, soit des progéniteurs communs myélo-lymphoïdes (CMLP) qui se différencient ensuite en cellules myéloïdes ou en CLP (Figure 9.C) (Singh, 1996). Ce modèle est basé sur les résultats obtenus chez des souris déficientes pour certains facteurs de transcription :

- chez les souris PU.1-/-, les cellules érythroïdes et les mégacaryocytes ont été formés normalement, mais les cellules myéloïdes, les LT et les LB se sont trouvés défectueux (Scott et al., 1994).
- chez les souris ikaros-/-, gène qui code pour une famille de facteurs de transcription précoces pour les cellules hématopoïétiques et les lymphocytes, seuls les LT et LB n'étaient pas présents (Georgopoulos et al., 1994). Néanmoins, des études ultérieures ont révélé que le développement des LT dans ces cas n'était pas tellement affecté (Spain et al., 1999; Wang et al., 1996). Cependant, il est à noter que le manque de cellules à lignages multiples dans les souris knock-out ne prouve pas nécessairement la présence de progéniteurs communs pour eux.

En 1997, grâce à des tests in vitro, Kawamoto et al. ont catalogué méticuleusement la production de lignées progénitrices de foie fœtal de souris (Kawamoto et al., 1997). Ils ont

trouvé que le devenir des cellules érythroïdes, des LB et des LT était presque toujours associé à un potentiel myéloïde (Kawamoto et al., 1999). Les auteurs ont donc proposé un nouveau modèle, appelé « modèle basé-myéloïde », dans lequel le devenir des lymphoïdes et des myéloïdes reste couplé, au lieu de se diviser au début de la différenciation (Kawamoto et al., 2010). Dans ce modèle, les CSH génèrent des CMLP et des CMP et le potentiel myéloïde persiste dans les branches LB et LT même après la divergence de ces lignées (Figure 9.D). Ce modèle a été confirmé en 2005 par l'isolement de progéniteurs multipotents biaisés lymphoïdes (LMPP) dans des souris adultes (Figure 9.E) (Adolfsson et al., 2005). En effet, dans ce modèle, appelé « modèle composite », les CSH peuvent donner :

- soit des LMPP qui donnent ensuite des CLP ou des GMP ;
- soit des CMP qui donnent des GMP ou des MEP.

les LMPP sont donc décrits comme des progéniteurs avec peu ou pas de potentiels érythroïdes-mégacaryocytes (Månsson et al., 2007).

En 2013, l'équipe de Görgens a revisité le modèle proposé par Adolfsson et ses collègues sur le critère de l'expression de la protéine membranaire CD133 (Görgens et al., 2013b). Ils ont montré que les granulocytes éosinophiles et basophiles comme les monocytes, les érythrocytes et les mégacaryocytes étaient issus d'un même CMP alors que les granulocytes neutrophiles appartenaient à la branche des LMPP (Figure 9.F) (Görgens et al., 2013a).

Finalement, au cours de ces dernières années, il n'existe pas un modèle explicatif de l'hématopoïèse sur lequel les scientifiques seraient tous d'accord. On ne sait pas vraiment comment la CSH fait son choix entre se différencier ou garder ses capacités multipotentes. On peut même se poser la question comment les CSH peuvent-elles s'engager dans une voie plutôt qu'une autre alors qu'elles sont toutes soumises aux mêmes stimuli extérieurs, qu'elles sont toutes dans le même environnement. De plus, une difficulté dans ce modèle déterministe est de délimiter les différentes catégories de progéniteurs comme les MPP, CLP, CMP, etc. Cette difficulté peut être surmontée en attribuant une part aléatoire définie grâce à des probabilités dans le devenir de ces cellules : c'est ce que propose le modèle stochastique.

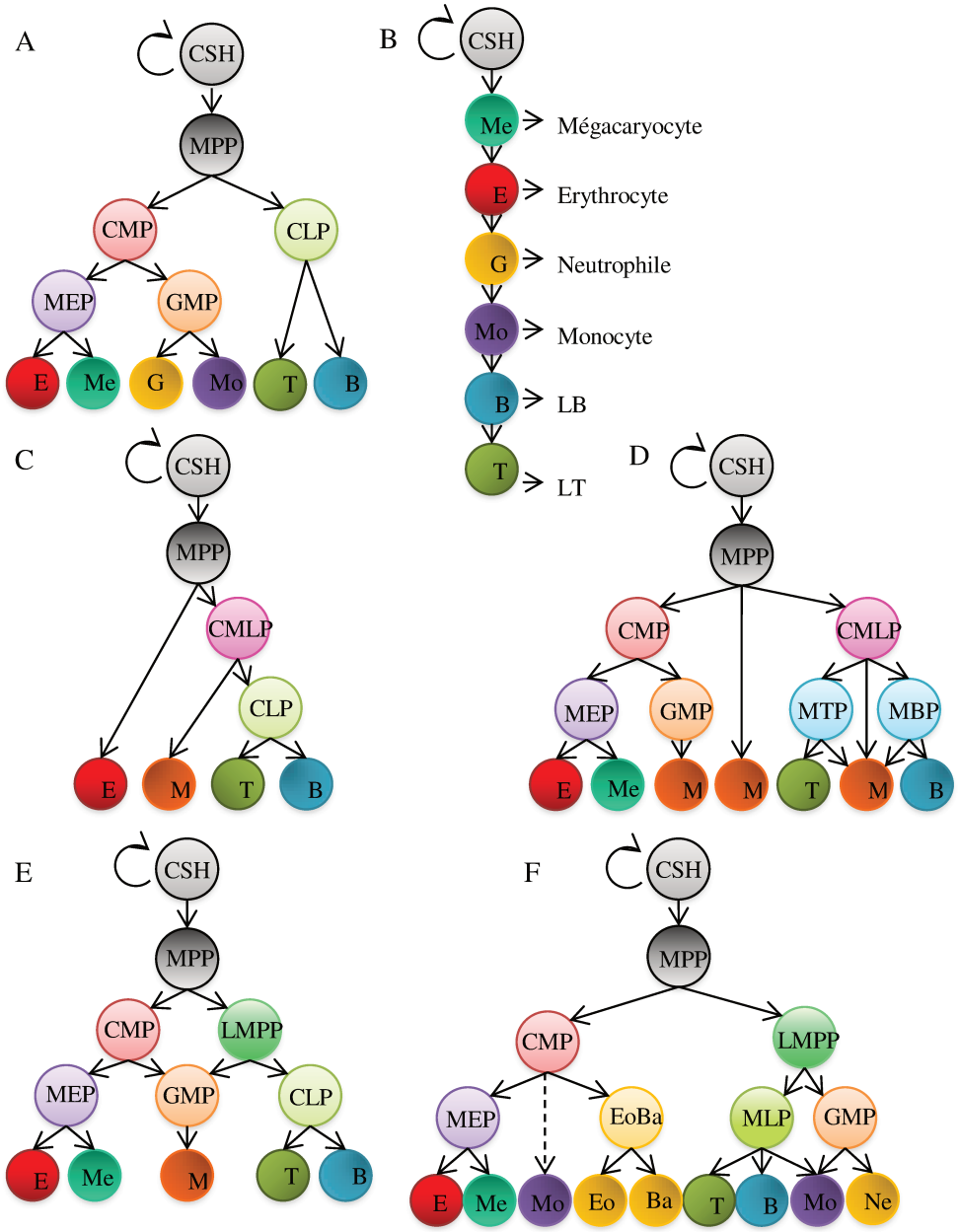


Figure 9 : Différents modèles déterministes au cours du temps

A. Modèle classique. B. Modèle séquentiel de Brown et al. (1985) basé sur le potentiel des progéniteurs qui change au cours de la différenciation. C. Modèle proposé par Singh et al. (1996) basé sur les études de facteurs de transcription dans les cellules hématopoïétiques. D. Modèle basé-myéloïde de Kawamoto et al. (1997) où les lymphoïdes et les myéloïdes restent associés dans la différenciation. E. Modèle composite d'Adolfsson et al. (2005) où les LMPP produisent des lymphocytes et des myéloïdes mais pas d'E ou Me. F. Modèle proposé par Görgens et al. (2013) basé sur la séparation des Ne, des Ba et Eo dans la différenciation.

Abréviations : CSH : cellules souche hématopoïétique. MPP : progéniteur multipotent. CMP : progéniteur commun myéloïde. CLP : progéniteur commun lymphoïde. MEP : progéniteur érythroïde-mégacaryocyte. GMP : progéniteur granulocyte-monocyte. CMLP : progéniteur commun myélo-lymphoïde. LMPP : progéniteur multipotent biaisé lymphoïde. MTP : progéniteur myéloïde-LT. MBP : progéniteur myéloïde-LB. MLP : progéniteur multi-lymphoïde. E : érythroïde. Me : mégacaryocyte. G : granulocyte. Mo : monocyte. M : myéloïde. T : LT. B : LB. Eo : éosinophile. Ba : basophile. Ne : neutrophile.

III.B Vision dynamique de l'hématopoïèse : modèle stochastique

Bien que les modèles déterministes existent depuis quelques décennies, le premier modèle explicatif de l'hématopoïèse est un modèle stochastique proposé par Till et McCulloch dans les années 60 (Till et al., 1964). Dans leur modèle, les cellules ont deux devenirs possibles :

- Soit elles se différencient et quittent le compartiment prolifératif ;
- Soit elles subissent des divisions symétriques formant deux colonies de cellules.

Une probabilité est attribuée à chaque devenir. En tirant des nombres aléatoires pour déterminer le sort de chaque cellule, Till et McCulloch ont calculé la diversité des populations de cellules souches sur plusieurs générations. La génération de colonies apparaît comme un processus bien défini, même si les décisions individuelles du devenir des cellules sont aléatoires. Finalement, il est possible de dire, comme en physique avec l'entropie, que l'ordre vient du désordre (Paldi, 2003). Plus tard, d'autres équipes ont aussi proposé ce modèle stochastique pour expliquer l'hématopoïèse (Abkowitz et al., 1996; Ogawa et al., 1983). Dans ces modèles, le paradigme du processus de ramification est utilisé à son plus simple, avec des cellules représentées comme des individus indépendants, se séparant au hasard et pouvant interagir avec un nombre limité d'entités plus petites (Figure 10.A). Cette part aléatoire dans le devenir de la cellule peut aussi être perçue comme une difficulté car, contrairement au modèle déterministe, le modèle stochastique ne propose pas de mécanismes moléculaires exacts pour la différenciation.

Un autre concept actuel est celui de la variabilité non génétique en tant que « substrat » pour la sélection naturelle. De nombreuses tentatives ont été faites pour comprendre le lien entre l'expression des gènes et la prise de décision du devenir cellulaire dans les organismes en développement. L'un des concepts les plus utilisés est la notion de « paysage épigénétique » de Conrad Waddington (Waddington, 1957).

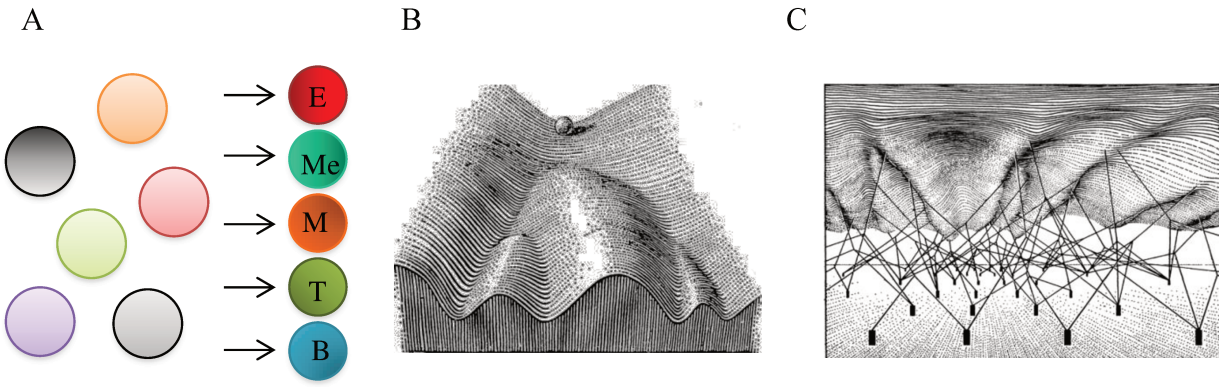


Figure 10 : Modèle stochastique et « paysage épigénétique » de Waddington.

A. Modèle stochastique proposé par Ogawa et al. (1983) où les ramifications ont toutes disparu. B. « Paysage épigénétique » de Waddington. Une cellule, représentée par une bille, est au sommet d'une colline et descend dans le paysage à travers une série de points de ramification qui représentent des événements décisionnels. C. La forme du paysage selon Waddington est déterminée par l'activité des gènes, représentés ici par des clous sous les collines et les vallées (Source images : Waddington, 1957).

III.B.1 Paysage épigénétique de Waddington

Dans le paysage de Waddington, une cellule, imagée par une bille, est au sommet d'une colline et suit les chemins existants dans le paysage, entraînée par une force gravitationnelle qui l'amène à l'un des nombreux destins possibles représentés comme des vallées (Figure 10.B). Selon Waddington, sa forme est déterminée par l'activité des gènes (Figure 10.C). En principe, les décisions concernant le destin sont stochastiques, donc, s'il y a de nombreuses cellules qui roulent dans le paysage, elles choisissent leur voie indépendamment des autres. Une fois qu'une cellule prend une décision, elle ne peut pas rebrousser chemin ce qui représente une diminution du potentiel cellulaire et une restriction du destin (Moris et al., 2016). Bien que Waddington soit conscient de la théorie des systèmes dynamiques, il a utilisé le « paysage épigénétique » plus comme une métaphore pour le développement, sans être explicite sur les détails. Un fondement formel de ce paysage a été suggéré en premier par Kauffman en 1969 : il a formalisé le paysage en identifiant les vallées avec des attracteurs dans les réseaux d'interaction génique (Kauffman, 1969, 1993). Plus tard, Huang et ses collègues ont utilisé et développé davantage cette notion dans un cadre de systèmes dynamiques multidimensionnels (Huang, 2012; Huang et al., 2005; Mojtahedi et al., 2016). Dans son concept, les profils d'expression génique sont projetés dans un espace à n-dimension, avec des champs vectoriels où les états stables correspondent à des attracteurs et où la différenciation est le passage d'un attracteur à un autre.

Selon la métaphore du paysage de Waddington, les cellules ont finalement deux moyens pour changer d'état :

- soit par des fluctuations stochastiques intrinsèques ;
- soit par des changements extrinsèques au système qui modifient les paramètres du réseau génique et modifient la géométrie du paysage (position, forme ou taille des éléments du paysage) ou la topologie (le nombre d'éléments dans le paysage) (Huang et al., 2007; Verd et al., 2014).

III.B.1.a Fluctuations stochastiques intrinsèques

Les fluctuations stochastiques, appelées aussi « bruit », sont des caractéristiques intrinsèques des systèmes biologiques comme la transcription, avec des variabilités aléatoires dans l'expression des gènes qui peuvent jouer un rôle biologique très important dans la différenciation. Cette idée est apparue très tôt, mais seulement récemment avec la technologie de cellules uniques, elle a pu être testée expérimentalement (Kupiec, 1983).

A l'origine des fluctuations stochastiques dans la cellule est un faible nombre de molécules de chaque type. Par exemple, les facteurs de transcription provoquent un engagement sporadique de l'acide ribonucléique (ARN) polymérase et conduisent à une synthèse discontinue de l'ARN messager (ARNm). Cela entraîne des écarts par rapport à l'état stable de l'attracteur (Eldar and Elowitz, 2010; Raj and van Oudenaarden, 2008). Habituellement, ces perturbations sont assez faibles pour que la force de l'attracteur ramène le système à son état stable d'origine, mais dans certaines conditions, les cellules peuvent traverser une barrière dans le paysage qui divise les bassins attracteurs, provoquant un événement spontané de changement d'état (Enver et al., 2009). En biologie, ces évènements peuvent être dus à des fluctuations stochastiques, dues aux différents niveaux de facteurs de transcription sur un seuil qui s'avère suffisant pour promouvoir les changements d'identité cellulaire (Schröter et al., 2015). Aujourd'hui, on considère plutôt la fluctuation des réseaux de gènes comme inducteur de la transition que la fluctuation des gènes individuels. Une architecture de réseaux, y compris les interactions, peut être modélisée mathématiquement dont les solutions indiqueront si un gène particulier est actif ou non dans un ensemble particulier de conditions. Les réseaux de gènes représentent des interactions régulatrices entre les gènes individuels fournissant des informations sur les mécanismes moléculaires sous-jacents des décisions et des processus de développement (Enver et al., 2009).

Un exemple de module de réseaux de gènes est l'interaction entre les facteurs de transcription GATA-1 et PU.1 dans une cellule CMP. Ces deux facteurs de transcription s'inhibent mutuellement et établissent ainsi une situation de décision pour la cellule progénitrice entre les destins érythroïdes-mégacaryocytes ou myéloïdes-monocytes (Laiosa et al., 2006).
D'autres réseaux plus complexes, impliquant plusieurs facteurs de transcription régulant de multiples cibles existent (Swiers et al., 2006). Dans cet article, les auteurs proposent une série de réseaux de gènes sous-jacents des CSH et de ses dérivés.
A partir de cette notion de réseaux de gènes, Tariq Enver et ses collègues ont défini un nouvel état appelé « multilineage » ou « multi-primed », état transitoire où les cellules comme les CSH expriment à faible niveau, les gènes de plusieurs lignées différentes avant de s'engager dans une voie unipotente (Cross and Enver, 1997; Enver and Greaves, 1998; Hu et al., 1997; Månsson et al., 2007). Comme dit précédemment, les réseaux de gènes sont modélisés par des calculs mathématiques ; les résultats obtenus sont une simple prédiction du devenir des cellules et sont donc à prendre avec précaution.

De tels événements de transition induits par le « bruit » peuvent être modulés par la forme du paysage, de sorte que les bassins attractifs ou les barrières élevées entre les bassins peuvent réduire la probabilité d'événements de transition spontanés et même imposer une direction sur les transitions (Wang et al., 2010).

III.B.1.b Transformations extrinsèques

Une autre alternative serait que les changements d'état seraient dus à une transformation du paysage, associés à des variations de certains paramètres du système et souvent médiés par des stimuli d'entrée dans les réseaux. En biologie, ces entrées peuvent représenter des signaux extracellulaires associés aux décisions du choix cellulaire comme les facteurs de croissance des fibroblastes, les signalisations nodales ou protéiques morphogènes. Le paysage peut aussi changer de manière qualitative : par exemple, la conversion d'un système monostable, avec un état stable ou un attracteur, à un système multistable. Ces changements d'état dynamique s'appellent des systèmes de bifurcations : ils peuvent prédire le comportement systémique à partir d'un réseau de gènes de régulation sous-jacent (Moris et al., 2016).

III.B.1.c Bifurcation et décision du destin cellulaire

Il existe différents types de bifurcation, chacune avec des caractéristiques spécifiques qui produisent des comportements différents. Seulement trois types de bifurcation pertinentes

pour l'étude de la prise de décision des cellules seront présentés ici (Huang et al., 2007; Verd et al., 2014).

La bifurcation la plus simple est celle qui correspond exactement au paysage classique de Waddington où la cellule passe d'un état monostable à un état bistable (Figure 11.A). Les décisions du destin apparaissent donc lisses et continues (au sens mathématiques), car une cellule reste dans un bassin stable tout au long de l'événement décisionnel. Un exemple de ce type de bifurcation est l'expression de gènes de GATA1 et PU.1 qui s'inhibent mutuellement (Huang et al., 2007). La cellule, une fois qu'elle a choisi un chemin, va exprimer soit l'un soit l'autre de façon stable. L'une des critiques que l'on peut faire sur ce modèle est que la cellule ne peut pas revenir à son état initial, le modèle est irréversible.

Une alternative à cette irréversibilité est de proposer une bifurcation en « nœud de selle » (Ferrell, 2012). Dans ce cas, au lieu de créer un nouvel état stable, tous les états stables préexistent déjà dans le système. Un état est alors supprimé lorsqu'un point instable, la « selle », et un état stable, le « nœud », se rencontrent, ce qui déstabilise efficacement l'état initial (Figure 11.B). Comme le précédent modèle, la bifurcation en « nœud de selle » est capable de reproduire la progression temporelle des états cellulaires et possède en plus les propriétés intrinsèques de l'irréversibilité. A noter que les états initiaux et finaux ne convergent jamais, ce qui signifie que le changement d'état temporel est lui-même discret.

Le troisième type de bifurcation contient à la fois des éléments de la première et de la deuxième bifurcation (Chang et al., 2008). Elle commence avec un état monostable localisé et passe ensuite par un état intermédiaire transitoire tristable avant de passer dans un état bistable (Figure 11.C). Un exemple de ce type de bifurcation est l'étude faite par Chang et ses collègues sur le marqueur de cellules souches de souris Sca-1 (Stem cell antigen 1). En triant la population de cellules exprimant Sca-1 à des niveaux différents, ils ont démontré que les cellules exprimant Sca-1 à des forts et des faibles niveaux étaient capables de reconstituer la distribution parentale de Sca-1 après plus d'une semaine de culture. Cette relaxation lente a été décrite dans un mixte de modèle gaussien qui incorpore des transitions stochastiques entre les sous populations discrètes, suggérant une stabilité multiple cachée dans un type cellulaire. Pour un rang de valeurs de paramètres, il existe donc un chevauchement entre les états initiaux et finaux. Ce modèle a des caractéristiques intéressantes incluant à la fois la réversibilité et la possibilité de changer d'état. Mais le point fort de ce type de bifurcation est que, dans un système stochastique, il y a une hétérogénéité et une interchangeabilité locale entre les états dans la plage de paramètres tristables car la force de chaque attracteur et les barrières correspondantes entre les états sont relativement faibles.

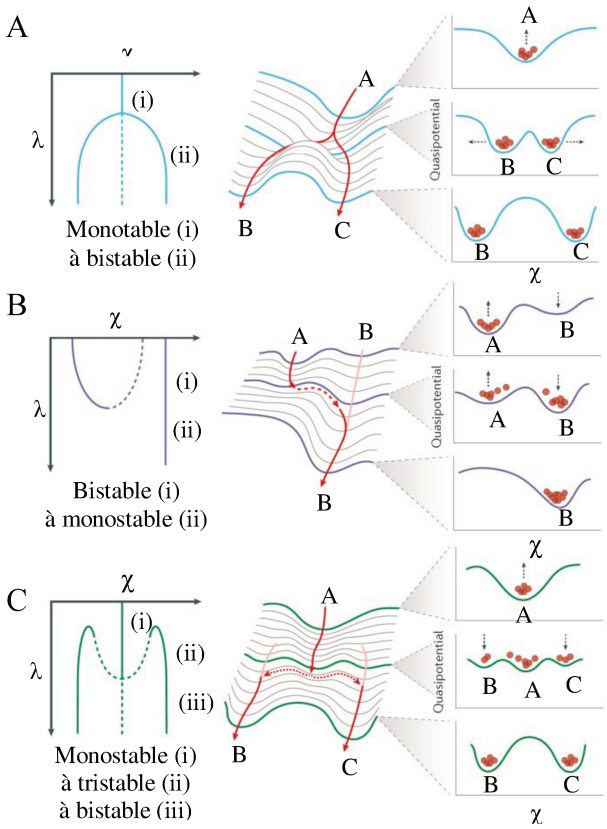


Figure 11 : Différents types de bifurcation.

A. Première bifurcation, la cellule passe d'un état monostable à un état bistable. B. deuxième bifurcation en « nœud de selle » où la réversibilité est possible. C. Troisième bifurcation qui est la combinaison des 2 autres avec un état transitoire tristable. Schéma adapté de Moris et al., 2016.

Alors que le modèle déterministe peut être défini grâce à des analyses populationnelles basées sur des mécanismes moléculaires, le modèle stochastique nécessite de regarder le comportement individuel de chaque cellule pour déterminer les probabilités de leur devenir dans les différentes voies de différenciation. Grâce aux nouvelles technologies de cellules uniques, il est possible de mieux caractériser et d'approfondir cette vision dynamique stochastique en dépit d'un modèle déterministe hiérarchique.

III.B.2 Technologie de cellules uniques : la fin du modèle déterministe

Récemment, grâce à la technologie de cellules uniques, les scientifiques remettent en cause ce modèle classique et pensent que les CMP n'existent pas (Notta et al., 2016; Paul et al., 2015; Yamamoto et al., 2013). En effet, Yamamoto et al. ont démontré que les progéniteurs érythroïdes-mégacaryocytes pouvaient être dérivés directement des CSH sans passer au préalable par les étapes conventionnelles de différenciation des CSH en MPP et CMP (Yamamoto et al., 2013). Il a aussi été montré in vivo que les CMP individuelles étaient hautement hétérogènes et produisaient soit des cellules myéloïdes, soit des cellules

érythroïdes après transplantation dans des souris (Perié et al., 2015; Perié and Duffy, 2016). Après marquage et isolement des MPP, ils ont démontré que l'engagement des cellules dans les lignées myéloïdes ou érythroïdes se faisait au niveau des MPP. Ces données prouvent que les MPP produisent en parallèle différents engagements de lignages distincts et que finalement il n'existe pas de CMP (Figure 12.A). Paul et ses collègues, eux, ont examiné l'expression des gènes de cellules uniques définies selon Akashi et al. comme étant des progéniteurs myéloïdes provenant de la MO de souris (Akashi et al., 2000; Paul et al., 2015). Ils ont montré qu'il existait différents sous-groupes de progéniteurs mais qu'aucun d'entre eux ne présentait un profil mixte. Les progéniteurs myéloïdes sembleraient donc être unipotents et non multipotents au niveau de l'expression des gènes. Des résultats similaires ont été trouvés in vitro chez l'homme. En effet, l'équipe de Notta a identifié par cytométrie dans le sang de cordon, dans la MO et dans le foie fœtal, 11 types cellulaires différents dans la population cellulaire des CMP humains (Notta et al., 2016). Les CMP isolés seraient finalement un mélange très hétérogène de différents types de cellules unipotentes qui seraient tous issus de la CSH multipotente (Figure 12.B). Dans toutes ces nouvelles données, aucun scientifique ne parle vraiment de modèle stochastique mais comme la hiérarchie disparait, je les considère comme tel ; c'est pourquoi j'ai intégré leur résultat dans ce paragraphe.

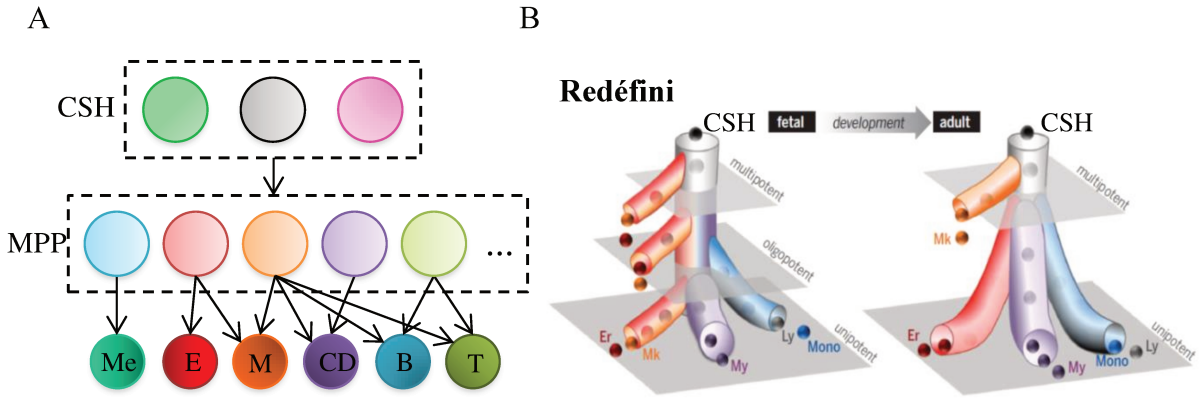


Figure 12 : Modèles redéfinis de Perié et al. (2016) et de Notta et al. (2016).

A. Modèle alternatif dérivé des résultats de cellule unique. Dans ce réseau, les CSH et les MPP sont un groupe hétérogène de cellules qui ont des biais intrinsèques vers certaines décisions de différenciation, même si les décisions sont réversibles (Perié and Duffy, 2016). B. Ce modèle propose un changement de développement dans l'architecture des cellules progénitrices du fœtus, où de nombreux types de cellules souches et progénitrices sont multipotents, jusqu'à chez l'adulte, où le compartiment de la cellule souche est multipotent mais les progéniteurs sont unipotents. Les plans grisés représentent des niveaux théoriques de différenciation (Notta et al., 2016).

Abréviation : CSH : cellule souche hématopoïétique. MPP : progéniteur multipotent. E/Er : érythroïde. Me/Mk : mégacaryocyte. M/My : myéloïde. Mono : monocyte. Ly : lymphocyte. CD : cellule dendritique. B : LB. T : LT.

Selon moi, le modèle stochastique est un bon modèle pour expliquer la différenciation des CSH car c'est un modèle qui ne nécessite pas de « connaître » à l'avance toutes les étapes intermédiaires du devenir des cellules. Il laisse une certaine « liberté » aux cellules dans leur choix du devenir cellulaire et ne suit pas une voie particulière. Quand on évoque le mot stochasticité ou aléatoire, on pense à des choses désordonnées, complètement imprévisibles or, le hasard peut être bien ordonné, organisé. En effet, le hasard se calcule avec des probabilités qui sont prédictibles et qui émergent du comportement hétérogène des cellules individuelles. Dans certains cas, on peut avoir des trajectoires qui auront une forte probabilité et d'autres voies qui seront quasiment nulles, voire presque improbables. C'est cette forte probabilité qui donne illusion d'une trajectoire déterminée ; et pourtant, celles qui ont une faible probabilité, même si ce sont des évènements très rares peuvent exister. Ce modèle est donc reproductible bien qu'il soit basé sur une part aléatoire. C'est le modèle que j'ai choisi d'utiliser et de défendre dans ma thèse. Bien que la population des CSH soit hétérogène, il a déjà été démontré qu'une population cellulaire isogénique mise dans un environnement homogène exprimait des protéines à des niveaux différents (Corre et al., 2014). Une explication à cette hétérogénéité observée est qu'il existe une part aléatoire dans le processus d'expression génique. Il est donc possible d'appliquer cette même vision probabiliste aux devenirs des CSH. De plus, les études précédentes étaient surtout réalisées au niveau de la population. Grâce à la technologie de cellule unique, il est possible aujourd'hui d'analyser chaque cellule dans une population hétérogène et de mieux comprendre les premières étapes de l'engagement des cellules. Dans la Partie 2 de ma thèse, l'intérêt et l'importance d'étudier les caractéristiques et le devenir des CSH à l'échelle de la cellule unique seront plus détaillés.

IV. Applications thérapeutiques des cellules souches hématopoïétiques

Les CSH ont été largement étudiées, en particulier, à des fins thérapeutiques pour soigner des maladies du sang et des maladies auto-immunes. D'ailleurs, c'est aussi grâce à la pratique comme la greffe de MO que l'on a pu faire des théories sur le développement et le comportement des CSH. Il existe deux grands types d'application thérapeutique :

- la thérapie cellulaire qui consiste à greffer des cellules du patient ou d'un donneur pour réparer/régénérer un organe ou un tissu endommagé ;
- la thérapie génique qui consiste à greffer des cellules du patient qui ont été au préalable modifiées génétiquement. Un transgène ou un gène thérapeutique est inséré dans les cellules du malade pour corriger le gène défectueux et produire ainsi à nouveau la protéine dont la déficience était la cause de la maladie. La thérapie génique est envisagée pour le traitement de certaines maladies génétiques et de certains cancers.

Une greffe à partir d'un donneur est possible dans le cas de la thérapie cellulaire : c'est ce qu'on appelle la greffe allogénique. Le problème dans ce type de greffe est qu'il n'existe pas de donneur compatible ou que le patient rejette la greffe du donneur. Pour pallier à cela, une autogreffe c'est à dire une greffe de cellules provenant directement du patient est possible. Dans ce cas, le patient peut subir soit une thérapie génique en modifiant génétiquement ses cellules et produire ainsi un grand nombre de cellules corrigées saines, soit une thérapie cellulaire en ne gardant que les CSH saines. Le choix de l'une ou l'autre des thérapies dans les maladies du sang se fait selon le type de maladie : la thérapie génique soigne plutôt des maladies acquises ou génétiques récessives héréditaires comme la maladie de WAS alors que la thérapie cellulaire soigne plutôt des maladies d'hémopathies malignes comme les leucémies (Delhommeau et al., 2009).

Seule la thérapie génique des maladies du sang sera développée dans ma thèse car c'est une technologie assez récente, encourageante que j'ai vu évoluer dans mon laboratoire de recherche, le Généthon mais qui pose certaines questions. Il existe différentes sources de CSH pour traiter les maladies du sang par les deux modes de thérapie.

IV.A Différentes sources de cellules souches hématopoïéiques

Il existe actuellement deux grandes sources de CSH utilisées pour la thérapie cellulaire et génique : la MO, le sang périphérique. Une troisième source, le sang de cordon ombilical, est utilisée uniquement pour des greffes allogéniques.

IV.A.1 CSH de la moelle osseuse

La plupart des recherches faites sur les CSH ont été motivées dans le but de transplanter les cellules hématopoïétiques pour traiter des leucémies et des aplasies médullaires (raréfaction de la MO). Pour rappel, les recherches sur les CSH ont été exponentielles pendant la Seconde Guerre Mondiale quand les effets biologiques des radiations ont été reconnus lors des bombardements nucléaires sur Hiroshima et Nagasaki en 1945. La MO a été la première source de cellules identifiée comme étant capable de greffer et de régénérer des cellules hématopoïétiques après irradiation.

En 1957, Thomas et ses collègues ont été les premiers à traiter des patients humains par irradiation ou chimiothérapie puis par injection de MO (Thomas et al., 1957). Leur résultats ont permis de montrer qu'il était possible d'administrer des cellules de MO par voie intraveineuse et qu'elles étaient correctement tolérées, bien que la greffe soit difficile chez l'Homme. Plus tard en 1959, cette même équipe a réussi à greffer 2 patients atteints de leucémie avancée après irradiation totale suivie de l'injection de MO et ont démontré la régénération des cellules hématopoïétiques même si les patients ont succombés plus tard à leur leucémie (Thomas et al., 1959). Beaucoup d'essais ont été faits à la suite de ces découvertes mais sans succès. La plupart du temps, la prise de greffe était bonne mais le patient ne survivait jamais à cause d'une rechute ou d'une maladie aiguë du greffon contre l'hôte (Bortin, 1970). Il a été vu plus tard que cela était dû à une mauvaise connaissance des systèmes d'histocompatibilité tissulaire et à un confinement pré-greffe insuffisant (Martin and Aulagner, 2009). Il aura donc fallu attendre les années 70 pour voir le premier succès d'une thérapie cellulaire d'une greffe allogénique, chez un enfant américain de 5 ans atteint d'un déficit immunitaire avec pour donneur un danois (Martin and Aulagner, 2009). Après cela, des registres de donneurs volontaires dont les antigènes HLA sont connus ont été créés comme le registre américain National Bone Marrow Donor Registry (NMDR) en 1986.

Dans les années 80, la greffe de MO de donneurs HLA identiques s'est étendue à d'autres maladies du sang comme la thalassémie (Lucarelli et al., 1984; Thomas et al., 1982), la drépanocytose (Johnson et al., 1984), les anémies de Fanconi (Deeg et al., 1983; Gluckman et al., 1980), les déficits immunitaires congénitaux (Fischer et al., 1994) et certaines maladies métaboliques comme les mucopolysaccharidoses (Hobbs, 1981).

Sur le plan technique, l'aspiration de la MO se fait dans la partie supérieure de l'os du bassin sous anesthésie générale. Les risques de cette méthode sont multiples pour le donneur ou le patient comme des saignements, des infections et des douleurs localisées (Panch et al., 2017).

A cause de ces risques et du manque de donneurs familiaux géno-identiques, d'autres sources de CSH ont dû être trouvées.

IV.A.2 CSH du sang périphérique

La présence de CSH dans le sang périphérique a d'abord été découverte en 1962 dans les modèles de souris par Goodman et Hodgson (Goodman and Hodgson, 1962). Plus tard, Epstein et ses collègues ont effectué, après irradiation, une circulation croisée entre des chiens malades atteints d'aplasie médullaire et des chiens sains et ont montré que les chiens malades avaient repeuplé la MO prouvant ainsi la présence de CSH dans le sang périphérique (Epstein et al., 1966). Grâce au dispositif d'aphérèse (prélèvement de certains composants sanguins) à flux continu développé en 1968 par Judson et al., McCredie et son équipe ont démontré que les leucocytes provenant du sang périphérique contenaient des CFU (Judson et al., 1968; McCredie et al., 1971).

En 1988, Kessinger et ses collègues ont été les premiers à montrer que le sang périphérique humain contenait suffisamment de cellules souches pour avoir une bonne prise de greffe après une ablation de la MO (Kessinger et al., 1988). En 1984, la découverte, l'isolement et la purification du facteur de stimulation des colonies de granulocytes (G-CSF) par Metcalf et al. ont montré que son utilisation augmentait la concentration des CSH circulantes (Metcalf, 1990). Ce sang mobilisé par G-CSF a permis d'avoir une prise de greffe meilleure et plus rapide que sans cette mobilisation (Juttner et al., 1990). Les CSH de sang périphérique sont utilisées dans des greffes allogéniques mais aussi dans des greffes autologues. En effet, la découverte de l'administration de G-CSF à des donneurs de cellules souches allogéniques sains pour augmenter le rendement des CSH a conduit à l'utilisation systématique de CSH de sang périphérique mobilisé dans des études cliniques de greffes autologues (Girbl et al., 2014; Karponi et al., 2015)et allogéniques (Schmitz et al., 1995).

IV.A.3 CSH du sang de cordon ombilical

La découverte de la troisième source de CSH dans la thérapie cellulaire, le sang de cordon ombilical, s'est produite lors d'un autre accident nucléaire en 1989 : la catastrophe nucléaire de Tchernobyl. Broxmeyer et al. ont été les premiers à suggérer d'utiliser le potentiel hématopoïétique des CSH du sang de cordon pour la greffe. Les CSH pourraient ainsi être maintenues en culture à long terme pendant plusieurs semaines avec un potentiel in vitro d'auto-renouvellement et de prolifération des cellules progénitrices (Broxmeyer et al., 1989). La collaboration de plusieurs cliniques a conduit en 1988 au premier succès de greffe de sang de cordon chez un enfant atteint de l'anémie de Fanconi (Gluckman et al., 1990; Gluckman

and Rocha, 2005). Aujourd'hui, plus de 30 000 greffes de sang de cordon ont été faites dans le monde entier à partir de donneurs apparentés ou non, chez l'enfant comme chez l'adulte (Laughlin et al., 2004; Locatelli et al., 2003; Rocha, 2016; Rocha et al., 2004; Wagner et al., 2002).

La collecte de sang de cordon est simple comparé aux autres sources de CSH impliquant une ponction veineuse du cordon ombilical et une récolte du sang dans un sac rempli d'anticoagulant. Les poches de sang sont ensuite cryopréservées dans des banques dans le monde entier pour une future greffe allogénique. La procédure de collecte est sans effet secondaire pour le donneur.

Pour conclure, le devenir du patient transplanté peut varier selon le type de sources utilisées (voir le Tableau 1) (Martin and Aulagner, 2009; Panch et al., 2017). En effet, les cellules CD34+ sont moins abondantes dans le sang de cordon et plus concentrées dans le sang périphérique mobilisé. Les 3 sources contiennent des cellules T qui peuvent induire des maladies du greffon contre l'hôte quand elles sont utilisées dans des greffes allogéniques. Le nombre de cellules T dans le sang périphérique est très élevé, entrainant des maladies aigues du greffon contre l'hôte. Comparé à la MO, le sang périphérique mobilisé avec du G-CSF contient plus de lymphocytes, de monocytes, de cellules T et de cellules dendritiques et la prise de greffe est plus rapide. Dans le sang de cordon, la prise de greffe est plus lente voire absente.

	Moelle osseuse	**Sang périphérique**	**Sang de cordon**
Type de greffe	Allogénique/ autologue	Allogénique/ autologue	Allogénique
Quantité cellules souches	Très nombreuse	Assez nombreuse	Peu nombreuse
Accessibilité	Peu accessible	Accessible	Très accessible
Quantité cellules T	faible	Très élevée	Très faible
Prise de greffe	moyenne	Très rapide	Lente voire nulle

Tableau 1 : Comparaison des différentes sources de CSH.

Malgré que le sang de cordon contienne moins de cellules CD34+, j'ai utilisé dans ma thèse cette source de CSH pour étudier les premières étapes de différenciation des cellules CD34+. La raison principale est en grande partie due à son accessibilité beaucoup plus simple que les autres sources surtout pour la recherche fondamentale et appliquée.

IV.B Thérapie génique

Les gènes sont des séquences spécifiques de l'ADN qui codent les instructions pour fabriquer des protéines. Si cette séquence est modifiée alors la protéine codée est incapable d'exercer sa fonction normale, entraînant des troubles génétiques. La thérapie génique consiste essentiellement à corriger les gènes défectueux responsables de la malformation de la protéine. Il existe plusieurs approches pour corriger les gènes :

- un transgène peut être inséré de façon non spécifique dans le génome pour remplacer le gène défectueux. C'est la méthode la plus courante ;
- un gène anormal peut être échangé avec un gène normal par recombinaison homologue ;
- un gène défectueux peut être réparé par mutation inverse sélective ;
- la régulation d'un gène particulier (activation ou inhibition) peut être modifiée.

Une anomalie génétique peut toucher le système hématopoïétique à différents endroits de l'hématopoïèse entrainant ainsi une immunodéficience primaire (Ghosh et al., 2015). Lorsque qu'elle a lieu dans les étapes très précoces de l'hématopoïèse, cela peut entraîner une immunodéficience combinée sévère (SCID).

Le premier essai clinique réalisé sur l'Homme était en 1990 chez un enfant atteint d'immunodéficiences combinés sévères par déficit en adénosine désaminase (ADA-SCID) traité par injection de cellules souches et de lymphocytes T génétiquement modifiées (Blaese, 1993). Dans les années 2000, un autre succès thérapeutique a vu le jour en France à l'hôpital Necker, encadré par le Dr Alain Fisher, chez des patients atteints d'une autre forme de déficit immunitaire, SCID-X1 lié au chromosome X, appelé aussi les « bébés bulles » (Hacein-Bey-Abina et al., 2003).

En 2017, plus de 2000 essais cliniques utilisant la thérapie génique sont en cours dont le domaine le plus important est la cancérologie avec 65%. Dans ce domaine, la thérapie génique serait essentiellement utilisée pour stimuler le système immunitaire du patient contre sa propre tumeur pour faciliter la reconnaissance des cellules cancéreuses et leur élimination. Seulement 10% des essais cliniques utilisant la thérapie génique sont des maladies monogéniques (comme les immunodéficiences, les maladies hématologiques, la mucoviscidose, etc.) et moins de 10% sont des maladies infectieuses, cardiovasculaires, ophtalmologiques, neurologiques ou inflammatoire (Figure 13).

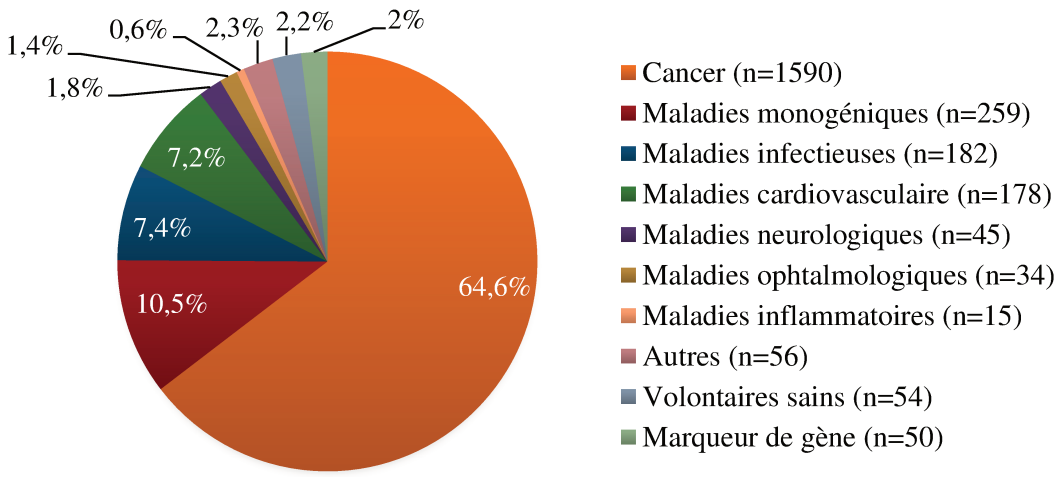


Figure 13 : Différents essais cliniques de thérapie génique.

Mise à jour 2017. Initialement, les études de thérapie génique étaient faites sur des maladies monogéniques mais peu après le cancer est devenu un intérêt majeur. (Source : The Journal of Gene Medicine Website (www.wiley.com//legacy/wileychi/genmed/clinical)).

IV.B.1 Deux stratégies de thérapie génique

La plupart du temps, un transgène est introduit dans le génome pour remplacer le gène défectueux. Son plus grand défi et aussi sa plus grande difficulté est d'arriver à livrer ce transgène dans les cellules cibles du patient (Misra, 2013). Pour cela, des vecteurs le plus souvent viraux sont utilisés comme moyen de transport pour introduire le gène. Ce vecteur devra être très spécifique, capable de livrer efficacement un ou plusieurs gènes de taille nécessaire à une application clinique. Il devra aussi ne pas être reconnu par le système immunitaire et pouvoir être purifié en grandes quantités à forte concentration. Une fois que le vecteur est inséré chez le patient, il ne doit pas induire une réaction allergique ou inflammatoire. Enfin, il devra être sans danger à la fois pour le patient et pour l'environnement. Deux méthodes existent pour délivrer les vecteurs chez le patient : la thérapie génique ex vivo et in vivo (Figure 14) (Kaufmann et al., 2013; Misra, 2013).

IV.B.1.a Thérapie génique ex vivo

Cette méthode de thérapie génique est la méthode la plus courante qui consiste à extraire les cellules du patient et de les modifier ex vivo. Dans le cas de maladies monogéniques, ce sont des CSH qui sont prélevées chez le patient. Pour les modifier, un vecteur délivre le transgène thérapeutique dans les CSH du patient, c'est ce que l'on appelle transduction, puis les cellules sont placées en culture pendant quelques jours. Une fois que les cellules modifiées génétiquement expriment le transgène, elles sont réinjectées chez le patient par intraveineuse où elles proliféreront dans l'organisme et le soigneront. L'avantage de cette méthode est de

pouvoir modifier un type cellulaire sans prendre le risque de transduire d'autres types cellulaires.

Dans le cas des maladies du sang, la thérapie génique ex vivo est la méthode la plus utilisée car il est assez facile de prélever les cellules cibles mais ce n'est pas toujours le cas notamment lorsqu'il faut modifier les cellules cardiaques ou des neurones. Pour cela, l'injection du vecteur se fait in vivo.

IV.B.1.b Thérapie génique in vivo

La thérapie génique in vivo n'extrait pas les cellules du patient. Le vecteur contenant le gène thérapeutique est directement injecté chez le patient par voie systémique comme par intraveineuse ou directement dans l'organe cible comme le foie, les reins, les poumons (aérosols), l'œil, le cœur, le muscle ou le système nerveux central. Néanmoins, avec cette méthode, le risque de dissémination du vecteur et du gène thérapeutique est moins maîtrisé.

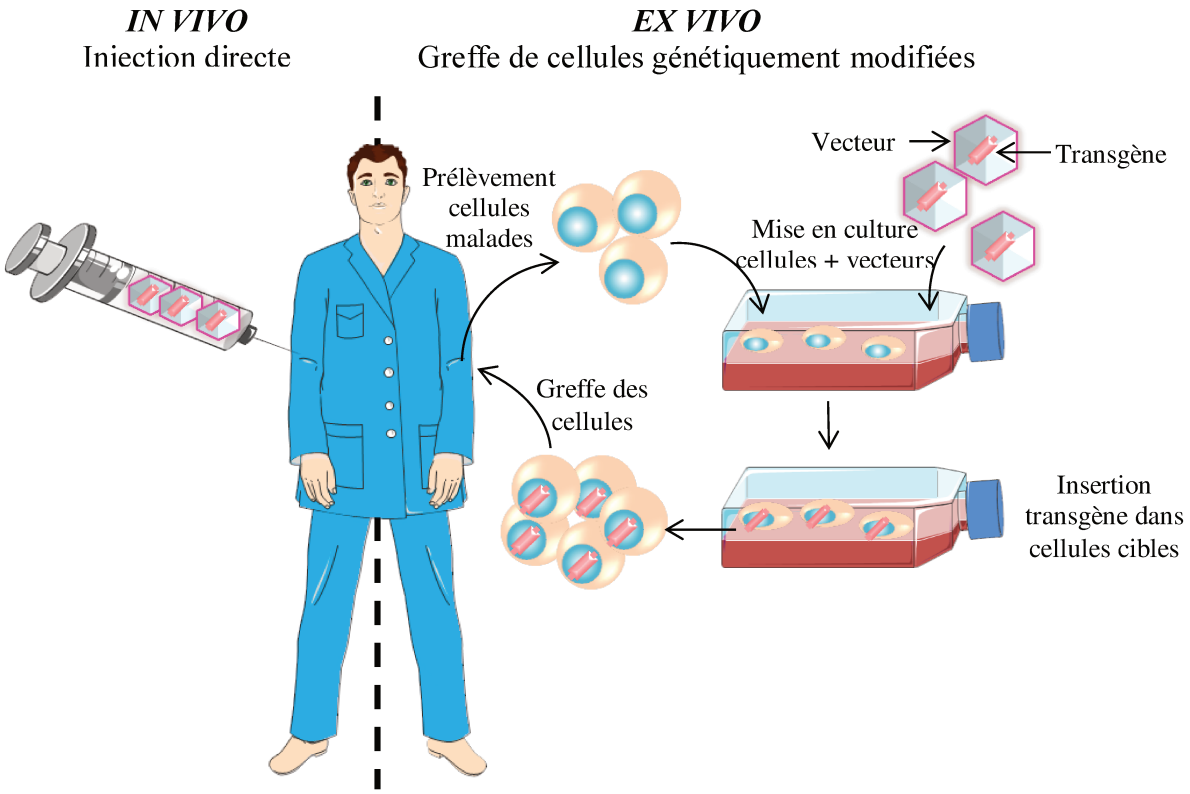


Figure 14 : Stratégies in vivo et ex vivo de la thérapie génique.

Dans la thérapie génique in vivo, le vecteur est introduit directement chez le patient. Dans la thérapie génique ex vivo, les cellules du patient sont extraites, mises en culture avec le vecteur contenant le gène thérapeutique. Le transgène va alors s'insérer dans les cellules du malade et les cellules modifiées génétiquement seront réinjectées chez le patient.

IV.B.2 Apport du transgène

Comme dit précédemment, pour apporter le transgène, il faut un vecteur. Il existe deux grandes classes de vecteurs : des vecteurs viraux et des vecteurs non viraux.

IV.B.2.a Vecteurs viraux

Les vecteurs viraux sont les vecteurs les plus utilisés. Ce sont des virus modifiés génétiquement. Les virus ont la capacité d'encapsuler et de délivrer leurs gènes aux cellules humaines de manière pathogène (Misra, 2013). Les scientifiques se sont inspirés de leur mode d'action pour pouvoir livrer le transgène à la cellule cible en remplaçant le génome viral sous forme d'ADN ou d'ARN, par le gène et en le rendant inoffensif (Walther and Stein, 2000). Une fois le gène intégré, la cellule peut alors produire la protéine manquante ou défectueuse. Il existe trois grandes classes de vecteurs viraux utilisés en thérapie génique : les rétrovirus, les adénovirus et les virus adéno-associés (AAV).

- Rétrovirus

Ce sont des virus intégratifs c'est-à-dire qu'ils insèrent leur ADN dans le génome de la cellule cible et peuvent ainsi transmettre le transgène aux cellules filles lors de la division cellulaire. Bien qu'ils aient été largement utilisés dans les premiers essais de thérapie génique, ils présentent certains désavantages comme une insertion aléatoire du transgène dans le génome en s'intégrant même dans les gènes pro-oncogènes (Baum et al., 2003). C'est ce qui a engendré des leucémies chez les « bébés bulles » ayant eu recours à la thérapie génique dans les années 2000 (Gaspar and Thrasher, 2005).

Pour éviter le problème d'insertion aléatoire, les chercheurs utilisent, de plus en plus, des lentivirus, virus modifiés dérivés du VIH (virus de l'immunodéficience humaine). Des essais utilisant ce vecteur ont été réalisés durant ces dernières années pour traiter des déficits immunitaires comme SCID-X1 (Cavazzana-Calvo et al., 2000), WAS (Aiuti et al., 2013; Charrier et al., 2007; Hacein-Bey-Abina et al., 2015), X-CGD (Ott et al., 2006), des maladies hématologiques comme la bêta-thalassémie (Cavazzana-Calvo et al., 2010) ou plus récemment la drépanocytose (Ribeil et al., 2017) mais aussi des maladies neurodégénératives comme l'adrénoleucodystrophie (Cartier et al., 2012, 2009).

- Adénovirus

Pour éviter les problèmes d'insertion aléatoire dans le génome, des virus non intégratifs comme l'adénovirus peuvent être utilisés. Le transgène reste alors libre dans la cellule malade

et ne s'intègre pas dans son génome. Ce type de virus cible les cellules qui ne se divisent pas comme les neurones ou les cellules de la rétine et disparait lors de la mort de la cellule. Ce virus est surtout utilisé dans les traitements contre le cancer comme le cancer du foie et des ovaires (Vorburger and Hunt, 2002). Il est assez peu utilisé pour les maladies monogéniques.

- Virus adéno-associé (AAV)

L'AAV est aussi un virus non intégratif mais à la différence de l'adénovirus, il peut infecter des cellules qui se divisent et qui ne se divisent pas. Ils permettent de transférer de petites séquences génétiques et sont peu inflammatoires. Ils sont maintenant utilisés in vivo pour traiter les maladies hématologiques héréditaires comme l'hémophilie B mais aussi les maladies musculaires et oculaires (Kay et al., 2000; Manno et al., 2006; Nathwani et al., 2011).

IV.B.2.b Vecteurs non viraux

Les vecteurs non viraux sont non intégratifs et n'utilisent pas de virus pour apporter le transgène mais a pour but d'injecter directement l'ADN. Soit l'ADN se trouve sous forme de complexe soit il est nu couplé à des particules physiques, naturelles ou organiques pour se protéger des enzymes cellulaires. Avec ce type de vecteur, il est possible de transporter une grande quantité d'ADN et d'avoir une meilleure sécurité des vecteurs. Néanmoins, ils sont moins performants que les vecteurs viraux dus à une expression temporaire du gène épisomal et sont donc moins utilisés dans les essais cliniques (Figure 15).

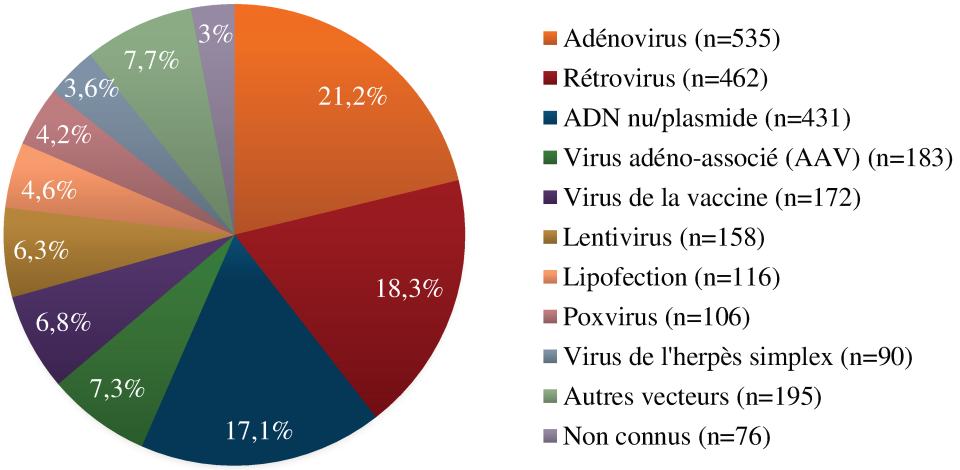


Figure 15 : Différents vecteurs viraux et non viraux utilisés dans les essais cliniques de thérapie génique. Mise à jour 2017. Les vecteurs viraux sont beaucoup plus utilisés que les vecteurs non viraux dans les essais cliniques de thérapie génique. (Source : The Journal of Gene Medicine Website (www.wiley.com//legacy/wileychi/genmed/clinical))

Bien que la thérapie génique soit une technique innovante et largement étudiée depuis deux décennies, elle n'est pas totalement sans danger et pose des questions. Nous allons voir dans le prochain paragraphe les limites de la thérapie génique uniquement dans le cas de la différenciation des CSH.

IV.B.3 Limites de la thérapie génique

Que ce soit des vecteurs viraux ou non viraux, la thérapie génique fait entrer un corps étranger chez le patient ce qui peut induire une réponse immunitaire. Pour rappel, le système immunitaire est constitué d'un ensemble de cellules qui coopèrent pour assurer les défenses de l'organisme contre les pathogènes et pour maintenir la tolérance au soi. Suite à l'administration de produits de thérapie génique, le système immunitaire peut provoquer une immunisation qui va conduire à la perte de l'effet thérapeutique. Les équipes de recherches cherchent à contourner ce problème en s'intéressant aux LTreg qui ont une action immunosuppressive sur les LT en les inhibant ou en les tuant (Bilate and Lafaille, 2012; Hippen et al., 2011; Neildez-Nguyen et al., 2015). C'est l'une des raisons pour laquelle, j'ai étudié au cours de ma thèse la prolifération et la différenciation des LT en LTreg.

De plus, il a été démontré que les manipulations faites in vitro sur les CSH en thérapie génique entrainaient des modifications épigénétiques (Aranyi et al., 2016; Paldi, 2013; Yamagata et al., 2012). Des études au préalable ont déjà démontré que les conditions de culture in vitro pouvaient provoquer des changements épigénétiques induits par l'environnement dans les cellules manipulées, telles que des modifications dans l'empreinte génétique ou dans la méthylation de l'ADN (Horii et al., 2010; Market-Velker et al., 2010). Des études récentes ont suggéré que de tels changements pouvaient avoir des effets à long terme sur les organismes résultants (Dumoulin et al., 2010; Fernández-Gonzalez et al., 2007; Nelissen et al., 2012). Yamagata et al. ont montré que la mise en culture in vitro de cellules CD34+ ainsi que l'utilisation de lentivirus dans ces cellules entrainaient des modifications épigénétiques (Yamagata et al., 2012). Plus exactement, des changements épigénétiques dans les cellules qui n'ont pas intégré le transgène mais qui ont été en contact avec le lentivirus ont été observés de façon systématique (Aranyi et al., 2016). Ces modifications semblent sans conséquence, mais les chercheurs tentent d'en connaître l'origine exacte afin de modifier les pratiques et de réduire ce phénomène.

IV.C Utilisation des cellules CD4+ et des cellules CD34+ dans ma thèse

Au cours de ma thèse, j'ai étudié deux types de différenciation cellulaire : la différenciation des LT en LTreg et la différenciation des CSH en cellules sanguines. Ces deux systèmes cellulaires sont de bons modèles pour étudier la différenciation cellulaire. En effet, dans le premier cas, nous avons un modèle simple avec une seule voie de différenciation, les LT en LTreg et de l'autre, un modèle beaucoup plus complexe avec plusieurs choix dans la différenciation, les CSH. Bien que ces processus soient une question fondamentale largement étudiée en biologie, notre connaissance sur les mécanismes de décision du développement reste insuffisante. La plupart des travaux ont étudié la différenciation au niveau de la population cellulaire. Cette approche a rendu possible la caractérisation des étapes majeures de la dynamique populationnelle mais elle n'a pas permis de comprendre pourquoi une cellule, à un moment donné, s'engage dans une voie de différenciation quand sa sœur, pourtant exposée aux mêmes stimuli, dans le même environnement, choisit une autre voie. Les explications typiques sur les mécanismes génétiques se basent la plupart du temps sur des études réalisées dans la population cellulaire et, par conséquent, ne reflètent qu'une moyenne qui ne peut pas être transposée directement sur les cellules uniques.
Des études sur cellule unique appartenant à la même catégorie phénotypique ont démontré sans ambiguïté que les cellules, prises individuellement, sont différentes les unes des autres car elles n'expriment pas exactement les mêmes gènes (Corre et al., 2014). Ainsi, les mécanismes définis à l'échelle de la population ne peuvent pas rendre compte de la grande variété des comportements cellulaires et ne peuvent pas expliquer le choix des cellules individuelles à s'orienter davantage dans une voie de différenciation que dans une autre. Durant ces dernières années, il est devenu évident que la cellule est soumise à d'importantes fluctuations stochastiques externes et internes. Ce sont, selon moi, ces fluctuations qui provoquent une hétérogénéité du phénotype. L'enjeu est alors d'expliquer comment malgré l'hétérogénéité des comportements individuels, la différenciation à l'échelle de la population cellulaire se déroule de façon ordonnée et prédictible.

Au cours de ma thèse, je vais aborder toutes ces questions en étudiant les premières étapes de différenciation des cellules CD4+ et des cellules CD34+ issues de sang de cordon à l'échelle de la cellule unique. Le choix de ces deux modèles n'est pas anodin ; comme nous avons vu précédemment, ces deux types cellulaires sont utilisés au Généthon pour des applications cliniques de thérapie génique. J'espère ainsi que les connaissances cumulées grâce à cette approche originale permettront d'améliorer les conditions expérimentales de thérapie génique.

PARTIE 2

Importance des études réalisées à l'échelle de la cellule unique

Bien que les CSH soient largement étudiées en biologie, les explications sur sa différenciation et son engagement dans une voie ne sont pas encore claires. Une exigence fondamentale pour toutes ces études faites sur les cellules souches est de pouvoir définir et identifier clairement ce qu'est une cellule souche (Etzrodt et al., 2014). Le potentiel d'auto-renouvellement et de différenciation d'une cellule doit donc être évalué à l'échelle d'une seule cellule.

En effet, les analyses faites au niveau populationnel ne reflètent qu'une moyenne de toutes les cellules et peuvent être surreprésentées à cause de sous populations cellulaires plus importantes (Figure 16). Par exemple, si dans une population cellulaire, il existe des sous populations qui expriment des gènes à des niveaux totalement opposés, en moyennant, on risque de tirer des conclusions fausses et de se tromper sur la véritable nature biologique de ces deux groupes (Dominguez et al., 2013). De plus, le profil des cellules rares dans une population hautement hétérogène est impossible à démasquer en utilisant des analyses de population. Ce type d'analyse est donc insuffisant pour résoudre l'hétérogénéité des populations de cellules souches, pour capturer la dynamique des molécules ou des phénotypes dans une cellule donnée et pour corréler les états des cellules actuelles avec les futurs destins.

Les récentes améliorations technologiques nous fournissent aujourd'hui des outils pour manipuler une cellule et effectuer des mesures multiparamétriques sur cette dernière. D'ailleurs, la montée exponentielle des études réalisées à l'échelle de la cellule unique est fortement liée à l'acceptation de la technique par les biologistes (Hu et al., 2016). Grâce à ces nouvelles technologies, il est possible d'acquérir des données génomiques, transcriptomiques, protéomiques et morphologiques à un instant donné ou en continu à l'échelle de la cellule unique. Les différentes technologies pour isoler une seule cellule et pour l'analyser vont être présentées dans la dernière partie.

C'est pour ces raisons que toutes les études que j'ai faites sur les LT et sur les cellules CD34+ ont été réalisées à l'échelle de la cellule unique. Pour analyser l'hétérogénéité de ces populations et pour comprendre la décision du destin de ces cellules dans les toutes premières étapes de différenciation, des études au niveau transcriptomique, protéomique et morphologique ont été faites. Dans un premier temps, j'ai étudié un modèle simple de différenciation, les LT en LTreg, puis un modèle plus complexe avec de multiple choix en étudiant les cellules CD34+ issues de sang de cordon.

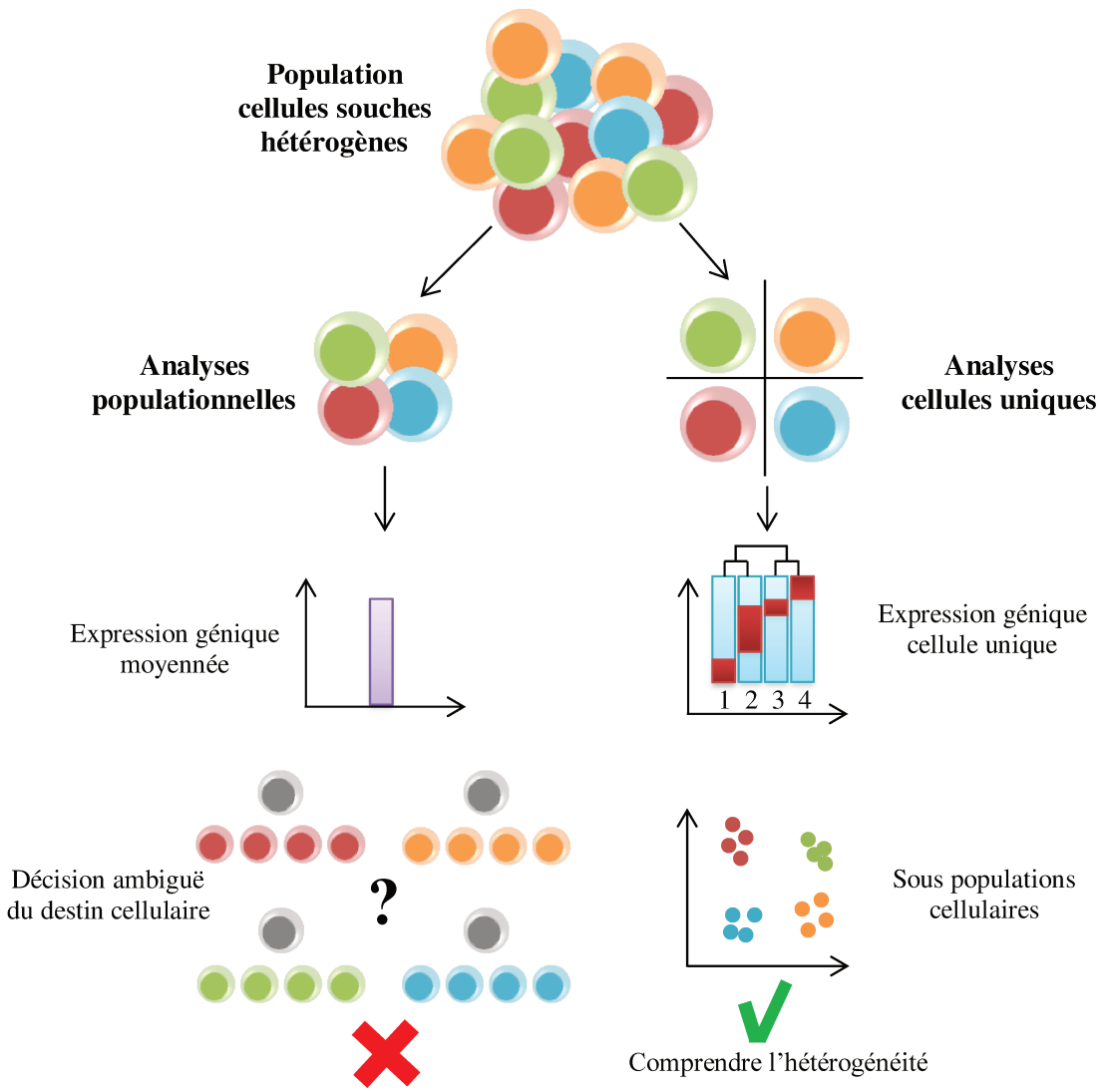


Figure 16 : Etudes faites à l'échelle de la cellule unique révèlent l'hétérogénéité populationnelle.
Les expériences traditionnelles en population masquent l'hétérogénéité entre les cellules individuelles et les sous populations. Schéma adapté de Ye et al., 2017.

Les analyses de cellules uniques ont donc un fort pouvoir pour révéler l'hétérogénéité dans les populations, les différentes voies de différenciation, les choix dans le lignage et dans le destin cellulaire et les réseaux de régulation ; c'est donc une méthode voire une démarche à suivre incontournable si on veut étudier la biologie des cellules souches (Etzrodt et al., 2014; Lunger et al., 2017; Ye et al., 2017).

I. Hétérogénéité des cellules souches hématopoïétiques

La production de cellules du sang dépend de l'auto-renouvellement des CSH et de leur capacité à se différencier. Cependant, ces cellules sont hétérogènes au niveau transcriptomique, protéomique et morphologique (Figure 17.1).

Perié et ses collègues, comme dit précédemment, ont démontré grâce au séquençage d'ARN de cellules uniques que les CMP individuelles étaient hautement hétérogènes et produisaient soit des cellules myéloïdes, soit des cellules érythroïdes après transplantation (Perié et al., 2015).

En étudiant les CSH individuelles provenant de jeunes et de vieilles souris, Kowalczyk et al. ont aussi démontré que les changements de transcription pendant le vieillissement de ces cellules sont inversement liés à ceux de leur différenciation, c'est-à-dire que les CSH anciennes à court terme ressemblent à des CSH jeunes à long terme.

Des données de comparaison d'expression génique à l'échelle de la cellule unique entre les cellules souches pluripotentes induites (CSPi) et des cellules souches embryonnaires (CSE) humaines ont montré une plus grande hétérogénéité dans l'expression des gènes dans les CSPi que dans les CSE humaines (Narsinh et al., 2011). Cela laisserait suggérer que les CSPi se trouveraient dans un état pluripotent alternatif, moins stable.

Pour mieux comprendre la régulation de l'expression de gènes des CSH hétérogènes de souris, Wilson et al. ont combiné le marquage membranaire et l'expression de gènes de plusieurs milliers de cellules uniques pour faire des analyses fonctionnelles (Wilson et al., 2015). Ils ont ainsi pu identifier des molécules clés associées à l'auto-renouvellement durable à long terme.

La microscopie en time lapse (TL) permet de suivre la dynamique de l'expression des protéines dans des cellules uniques vivantes. Généralement, les cellules sont transfectées avec un gène rapporteur fluorescent, comme la protéine GFP (green fluorescent protein), ou elles sont issues de souris transgéniques (Miller and Cornish, 2005; Sekulovic et al., 2011). Corre et ses collègues ont, par exemple, étudié la dynamique temporelle de la variation de l'expression des gènes en utilisant un modèle de gène à double rapporteur (Corre et al., 2014). Ils ont comparé deux types de clones cellulaires : des clones exprimant des ARNm hautement stables codant des protéines fluorescentes et des clones exprimant des ARNm et des protéines instables. Les deux types de clones ont montré une forte hétérogénéité des niveaux d'expression des gènes rapporteurs.

Duffy et al., eux, ont utilisé une souris transgénique Blimp1-GFP où Blimp1 est un facteur de transcription nécessaire pour la différenciation des LB en plasmablastes (Duffy et al., 2012).

Après 3 jours de mise en culture, ils ont suivi les LB de la 1ère, 3ème, 5ème et 7ème génération, qui n'expriment pas encore Blimp1 et ont analysé les temps de division, de différenciation et de mort cellulaire de ces cellules. Ces mesures ont indiqué que chaque destin cellulaire est poursuivi de manière autonome et stochastique.

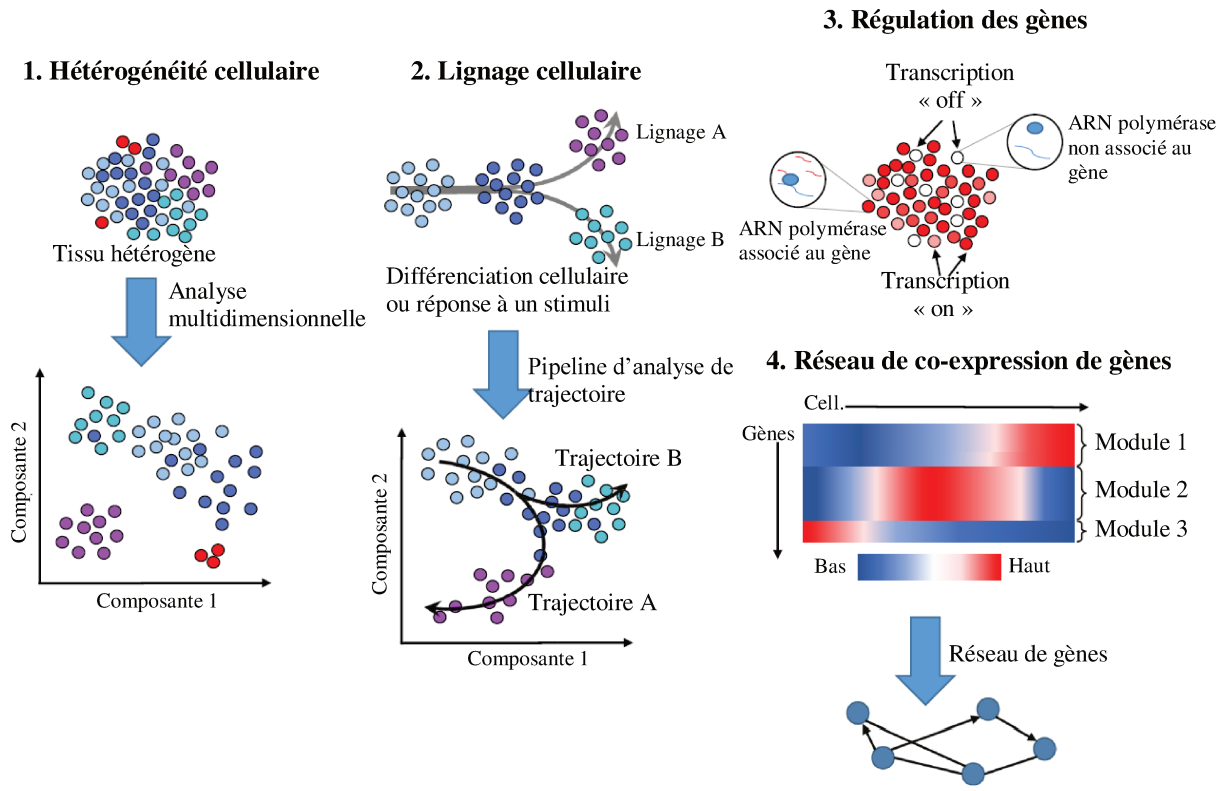


Figure 17 : Diverses applications réalisées à l'échelle de la cellule unique.

1. Mise en évidence de l'hétérogénéité dans la population cellulaire. 2. Analyse des trajectoires lors de la différenciation cellulaire. 3. La régulation des gènes par des transitions entre les états « on » et « off ». 4. Les réseaux de gènes. Schéma adapté de Liu and Trapnell, 2016.

L'hétérogénéité des CSH n'est pas seulement au niveau génomique ou protéique mais elle est aussi observable au niveau morphologique. L'imagerie en TL permet de suivre la dynamique des différentes formes des CSH et de ses descendantes. Sekhavati et al. ont étudié la différenciation des GMP en macrophage juste sur le critère de l'adhésion cellulaire et de leur mouvement brownien, sans marquage cellulaire (Sekhavati et al., 2015). En effet, la transition de l'état non adhérent à l'état adhérent des cellules a servi de marqueur de point de temps pour la différenciation des GMP. Pour les détecter, ils ont calculé le mouvement brownien des cellules entre les images et ont ainsi déterminé le temps de transition entre les 2 états sans avoir recours à un perturbateur extérieur comme le marquage.

II. Etude des voies de différenciation

Comme nous avons vu dans la partie précédente, plusieurs modèles de différenciation des CSH existent. Le modèle classique est un modèle déterministe hiérarchisé qui commence avec une CSH pouvant ensuite donner soit des CMP, soit des CLP. Cependant, les résultats de cellules uniques ont mis en cause ce modèle et ont permis de proposer d'autres modèles moins déterministes avec des progéniteurs plus hétérogènes (Figure 17.2) (Notta et al., 2016; Paul et al., 2015; Perié et al., 2015).

Grâce à l'analyse de l'expression d'une centaine de gènes dans plus de 1500 cellules hématopoïétiques uniques de souris, Guo et ses collègues ont montré une forte hétérogénéité dans tous les types de progéniteurs classiquement définis comme les CSH, les MPP, les CMP et les CLP (Guo et al., 2013). Ils ont revisité le modèle classique de l'hématopoïèse : la différenciation des cellules érythroïdes et mégacaryocytes est devenue la première branche de différenciation des CSH.

Néanmoins, en analysant un panel de gènes fixes choisi au préalable, notre analyse et interprétation peuvent être faussées. C'est pourquoi le séquençage des ARN totaux permet de mieux étudier les transitions cellulaires entre les différents états et de mieux caractériser les trajectoires cellulaires à travers des processus comme la différenciation.

Nestorowa et al. ont séquencé plus de 1000 CSH individuelles de 12 phénotypes différents et ont pu détecter plus de 6000 gènes codant pour une protéine (Nestorowa et al., 2016). La reconstruction des voies de différenciation révèle des changements d'expression dynamique associés à une différenciation lymphoïde, érythroïde et granulocyte-macrophage précoces.

Macaulay et al. ont étudié l'engagement du lignage des thrombocytes dans les poissons zèbres (Macaulay et al., 2016). Ils ont placé tous les points de données dans un continuum pour former une seule voie de lignage.

La cytométrie de masse est également une méthode largement utilisée pour étudier les voies de différenciation des CSH. Cette technologie puissante a permis d'analyser simultanément 34 paramètres dans les cellules humaines de MO uniques incluant 31 anticorps, un marqueur de viabilité, un marqueur de l'ADN et un marqueur de taille de cellule (Bendall et al., 2011). Cette étude a révélé une hétérogénéité parmi les populations cellulaires considérées comme étant homogènes et a également mis en évidence des réponses de signalisation redondantes dans différentes populations du système hématopoïétique. En utilisant toutes ces données, ils ont construit un nouvel arbre pour expliquer la différenciation des CSH. Dans cette même équipe, une autre étude a permis de reconstruire la différenciation des LB naïfs étape par étape dans la MO (Bendall et al., 2014).

La cytométrie de masse a également été utilisée pour cartographier les phases du cycle cellulaire des cellules hématopoïétiques humaines (Behbehani et al., 2012). Cela a permis de faire le profil des phases principales du cycle cellulaire simultanément dans des cellules individuelles.

Tous ces résultats ont remis en cause la hiérarchie actuelle de bifurcation du système hématopoïétique. L'hétérogénéité observée au niveau de la transcription dans les cellules progénitrices suggère que la différenciation cellulaire se déroulerait de manière plus sophistiquée, plus complexe et plus aléatoire.

III. Choix du destin cellulaire

Les CSH ont plusieurs voies de différenciation et de manière générale, on peut se demander comment une cellule « choisit » de s'engager dans une voie plutôt qu'une autre (Figure 18). Les analyses à l'échelle de la cellule unique peuvent nous aider à répondre à cette question avec une grande précision.

Comme nous avons vu précédemment, Chang et ses collègues ont montré que les cellules souches de souris exprimant Sca-1 à des forts et des faibles niveaux étaient capables de reconstituer la distribution parentale de Sca-1 après plus d'une semaine de culture, prouvant ainsi leur potentiel réversible (Chang et al., 2008).

En faisant des analyses fonctionnelles et moléculaires de cellules multipotentes individuelles qui s'auto-renouvellent, Pina et al. ont suggéré une transition moléculaire non coordonnée entre l'auto-renouvellement et les états engagés (Pina et al., 2012). Ils ont montré une dissociation entre l'auto-renouvellement et l'activation du transcriptome dans le lignage cellulaire. Ils ont suggéré que les cellules multipotentes avaient une activation indépendante des régulateurs individuels, ce qui entraîne une faible probabilité de transition vers l'état engagé.

Une étude récente a montré qu'il existait des intermédiaires de lignées mixtes présentant une expression simultanée de cellules souches/progéniteurs hématopoïétiques et des CMP (Olsson et al., 2016).

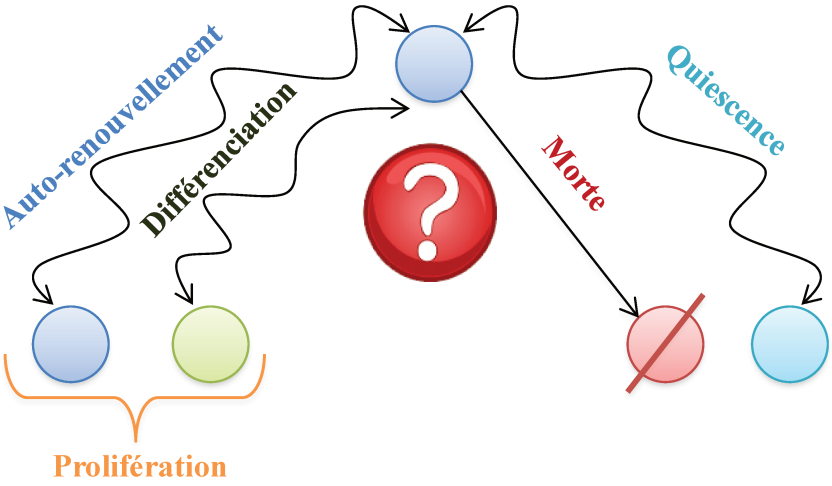


Figure 18 : Différents choix du destin cellulaire.

Une cellule souche a plusieurs choix : elle peut s'auto-renouveler, se différencier, rester quiescente ou mourir. Comprendre comment ce choix est orchestré est fondamental pour mieux connaitre le déroulement de l'hématopoïèse.

L'imagerie de cellules uniques fixes ou en TL sont également des outils puissants pour déterminer des lignées et comprendre leur engagement. Grâce à cela, il est possible de reconstruire l'arbre généalogique des cellules et de connaître les temps de division de chaque clone et de ces descendantes.

L'équipe de Bogdan et al. a ainsi pu identifier la coexistence de sous-populations se divisant rapidement et lentement avec des cellules quiescentes dans les cellules souches (Bogdan et al., 2014). Ils ont pu démontrer, par des analyses mathématiques, que les cellules souches ne se développaient pas de manière indépendante mais présentaient un comportement stochastique dépendant du temps lorsqu'elles interagissaient les unes avec les autres par des signaux moléculaires et tactiles.

Malgré les analyses réalisées à l'échelle de la cellule unique, la question sur le choix du devenir cellulaire reste encore non résolue. Est-ce un destin déterministe ou stochastique ? Bien que pour certains scientifiques, la fluctuation des gènes est la preuve d'un comportement aléatoire, pour d'autres la dynamique observée est contraire à un comportement stochastique comme Hoppe et ses collègues (Hoppe et al., 2016). Ils ont exploré grâce au TL la co-régulation des facteurs de transcriptions GATA1 et PU.1 dans la différenciation dynamique des CSH uniques. Selon eux, la dynamique d'expression observée entre PU.1 et GATA1 serait incompatible avec l'hypothèse que la commutation stochastique précède et initie la prise de décision entre le lignage mégacaryocyte-érythroïde et le lignage granulocyte-monocyte. Leurs résultats suggèrent que ces facteurs de transcription s'exécuteraient uniquement quand le choix du lignage est fait.

L'analyse de l'expression des gènes à l'échelle de la cellule unique a pu révéler que la stimulation des CSH hautement purifiées avec le facteur de stimulation des colonies de macrophages (M-CSF) entrainaît une activation du promoteur PU.1 et un grand nombre de cellules PU.1+ avec une signature génique myéloïde et un potentiel de différenciation (Mossadegh-Keller et al., 2013). Donc, M-CSF, une cytokine myéloïde libérée lors des infections et des inflammations, pourrait directement induire la régulation du gène PU.1 et changer le devenir des cellules myéloïdes dans les CSH de souris, indépendamment de la survie sélective ou de la prolifération.

Enfin, l'identification de chaque cellule individuelle tout au long de la différenciation est également cruciale pour révéler la fonction instructive des cytokines pendant le choix de la lignée hématopoïétique. Par exemple, l'équipe de Rieger et al. a démontré que M-CSF et G-CSF avaient une fonction instructive dans le choix de la lignée des GMP plutôt qu'une fonction sélective/permissive (Rieger et al., 2009). Pour cela, des cellules GMP uniques ont

été traitées soit avec des M-CSF, soit avec des G-CSF et ont été filmées. Le suivi des cellules uniques de toutes les descendantes de GMP en présence d'une seule cytokine a révélé l'absence de mort cellulaire dans la majorité des cas. Cela laisserait suggérer que les cytokines ne sélectionneraient pas une sous population par la survie d'un type de cellule mais qu'elles seraient plutôt une fonction instructive.

IV. Réseaux de régulation

L'explosion des données génomiques a permis de générer des réseaux de régulation dans divers systèmes biologiques (Ye et al., 2017). Grâce aux nouvelles technologies de cellules uniques, les réseaux de régulation peuvent mieux rendre compte du comportement des cellules individuelles et donc être « plus proches » de la réalité (Figure 17.3 et .4)

La variabilité intrinsèque de l'expression des gènes entre les cellules peut être aussi utilisée pour déduire des réseaux de gènes (Figure 17.4). A partir de ces données, la matrice de covariance gène par gène peut être calculée. Les couples de gènes avec un niveau d'expression hautement corrélé dans un même échantillon sont alors considérés comme étant co-exprimés (Stegle et al., 2015).

Moignard et al. ont analysé 18 facteurs de transcriptions hématopoïétiques dans une centaine de cellules souches/progéniteurs hématopoïétiques uniques provenant de la MO de souris (Moignard et al., 2013). Ils ont pu découvrir de nouveaux réseaux de gènes grâce aux corrélations positives et négatives entre 3 facteurs de transcription : GFI1, GFI1b et GATA2. Parmi ces corrélations, ils ont retrouvé des liens connus entre 2 facteurs de transcription comme PU.1-GATA1 et GFI1-GFI1b qui s'inhibent. Des corrélations négatives entre GATA2-GFI1 et positives entre GATA2-GFI1b ont permis de trouver de nouveaux liens entre ces gènes, suggérant une implication possible de GATA2 dans la modulation de la relation d'inhibition croisée entre GFI1 et GFI1b.

L'analyse par transcription des CSE uniques a montré un gradient et une hiérarchie d'expression de gènes pluripotents (Hough et al., 2009). Même les cellules qui ont des caractéristiques les moins différenciées n'expriment qu'un sous-ensemble de gènes caractéristiques des cellules souches. De nombreuses cellules co-expriment des gènes pluripotents et spécifiques de lignées. Finalement, seule une petite fraction de la population résiderait au sommet de la hiérarchie.

Pour explorer la relation possible entre l'autophagie et les voies de signalisation de Notch dans les pré-CSH de souris, l'équipe de Hu et al. a étudié l'expression de 4 gènes essentiels dans l'autophagie et 4 gènes impliqués dans la voie de signalisation de Notch (Hu et al., 2017). Pour rappel, Notch joue un rôle essentiel dans la différenciation et la prolifération des cellules. Ils ont trouvé que l'expression des gènes de l'autophagie était corrélée significativement négativement à l'expression des gènes de Notch. Finalement, l'activité d'autophagie dans les pré-CSH pourrait être corrélée avec une sous régulation des gènes de Notch durant l'émergence des CSH chez les souris.

Il est important de noter que ces réseaux de gènes ne sont que des suppositions calculées par des corrélations avec des méthodes différentes selon les articles (Pearson, Spearman, etc.) et dont le coefficient de corrélation est choisi de façon arbitraire. Ces résultats sont donc à prendre avec précaution.

V. Hématopoïèse développementale

L'origine des CSH chez les mammifères, comme nous l'avons vu dans la première partie, a été longtemps un mystère. L'ontogenèse du système hématopoïétique est caractérisée par deux grandes étapes : la première, l'hématopoïèse primitive située plus spécifiquement dans le sac vitellin et la deuxième étape, l'hématopoïèse définitive qui a lieu principalement dans la région AGM (Medvinsky and Dzierzak, 1996; Zovein et al., 2008). Cette dernière étape produit des CSH multipotentes avec une capacité de reconstitution à long terme.

Zhou et ses collègues ont analysé récemment des cellules endothéliales et des pré-CSH dans la région AGM et dans le foie fœtal. Ils ont montré que l'activation des cibles mécanistiques de la rapamycine (mTOR) était indispensable à l'émergence de CSH mais pas pour les progéniteurs hématopoïétiques (Zhou et al., 2016).

L'imagerie en TL a permis de confirmer que les cellules hématopoïétiques et les cellules endothéliales provenaient d'un même précurseur commun appelé endothélium hémogénique (Eilken et al., 2009; Lancrin et al., 2009). Les cellules hématopoïétiques et endothéliales vivantes ont été identifiées par la détection simultanée de la morphologie et plusieurs marqueurs moléculaires et fonctionnels. Cette étude observée ex vivo a été confirmée in vivo dans des poissons zèbres (Bertrand et al., 2010; Kissa and Herbomel, 2010). En effet, ils ont mis en place un microscope confocal en TL avec une résolution cellule unique et ont montré que les cellules hématopoïétiques des poissons zèbres provenaient directement de cellules endothéliales hémogéniques in vivo. Ces études ont aussi été confirmées in vivo chez la souris (Boisset et al., 2010).

Yvernogeau et al. ont suivi par imagerie la transition endothéliale vers l'hématopoïèse des cellules sanguines nouvellement formées à partir de la couche de cellules endothéliales hémogéniques (Yvernogeau et al., 2016). En modifiant les conditions de culture, ils ont montré qu'il était possible d'empêcher l'engagement des cellules endothéliales, d'accélérer la transition endothéliale vers l'hématopoïèse et de produire des cellules musculaires lisses au détriment des cellules endothéliales, ce qui démontre la polyvalence du système de culture cellulaire.

Ainsi, avec tous ces exemples, nous voyons bien l'intérêt d'analyser à l'échelle de la cellule unique les CSH pour mieux comprendre leur comportement et leur devenir. Les recherches futures sur la maintenance des CSH et sur leur engagement dans les lignées bénéficieront grandement de l'intégration des études réalisées à l'échelle d'une seule cellule, en tenant compte des réseaux transcriptomiques et métaboliques, des protéines de signalisation de

surface et intracellulaire, de la morphologie, de la structure chromatinienne et du paysage épigénétique de la cellule (Bendall et al., 2014; Guo et al., 2013; Levine et al., 2015; Moignard et al., 2013).

De plus, l'imagerie en TL à l'échelle d'une seule cellule est actuellement la méthode ultime pour quantifier les comportements hétérogènes sur le plan cellulaire, moléculaire et dynamique sur de longues périodes de temps. Plusieurs études ont déjà utilisé cette approche avec succès pour répondre à de vielles questions dans les systèmes de CSH ou autres (Costa et al., 2011; Ortega et al., 2011; Song et al., 2010).

Tout ceci est possible grâce aux nouvelles technologies. En effet, depuis quelques années, les technologies de cellules uniques ont augmenté de façon exponentielle afin de mieux étudier la population cellulaire. Comme j'ai étudié les LT et les cellules CD34+ à l'échelle de la cellule unique, nous verrons dans la prochaine partie les différentes méthodes qui existent aujourd'hui pour isoler une cellule (micromanipulateur, FACS (Flow Activated Cell Sorting), microfluidique, microgrille). Ensuite, nous verrons les différentes techniques que j'ai utilisées au cours de ma thèse pour étudier le transcriptome, le protéome et la morphologie des cellules T et des cellules CD34+ à l'échelle de la cellule unique.

PARTIE 3

Différentes techniques de captures et d'analyses de cellules hématopoïétiques uniques

I. Méthodes de capture de cellules uniques dans une population cellulaire

Les méthodes d'isolement de cellules uniques, bien que nécessaire pour étudier une population hétérogène, sont des technologies récentes et complexes. Des méthodes toujours plus novatrices ont dû être développées pour étudier le génome, le transcriptome et le protéome des CSH de façon individuelle.

Il existe plusieurs méthodes pour isoler des cellules uniques dans une population (Figure 19), chacune avec leurs avantages et leurs limites qui seront détaillés par la suite (Hu et al., 2016; Wu and Tzanakakis, 2013; Ye et al., 2017).

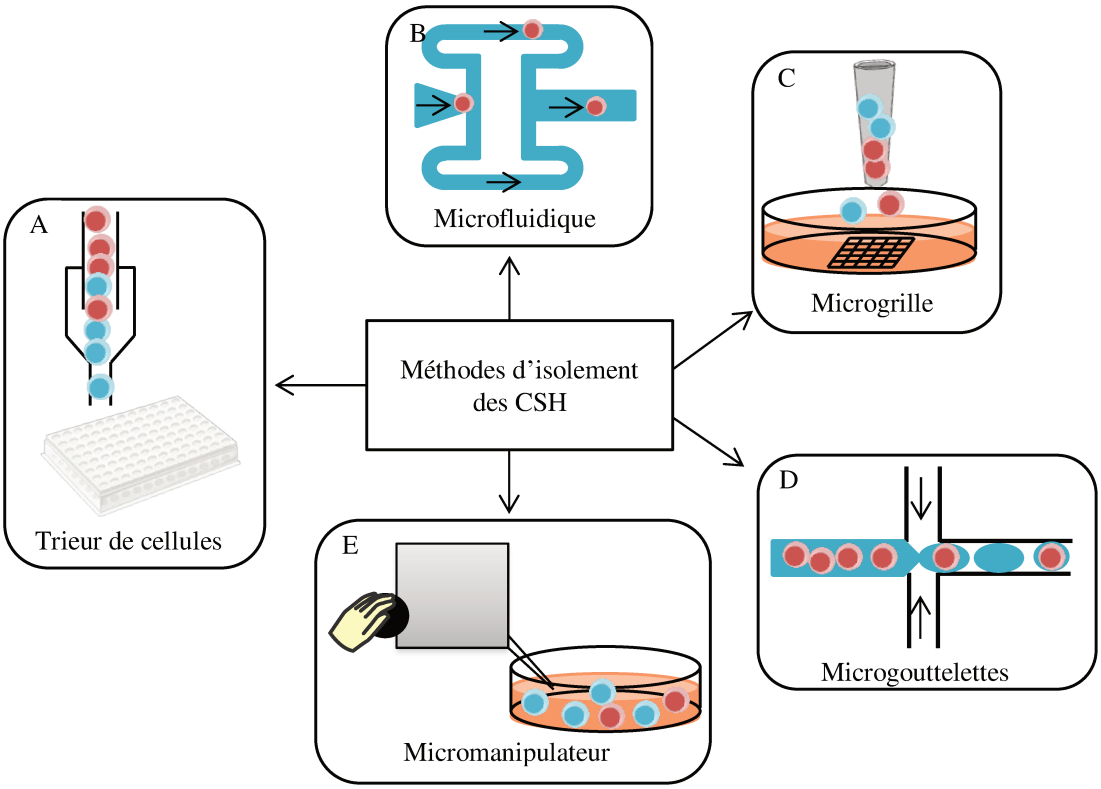


Figure 19 : Méthodes de capture des CSH uniques.

A. Trieur de cellules. B. Méthode par microfluidique, représentant ici la technologie Fluidigm. C. Microgrille faite en polydiméthylsiloxane (PDMS). D. Méthodes par microgouttelettes, les cellules sont encapsinées dans des gouttes. E. Micromanipulateur. Schéma adapté de Ye et al., 2017.

I.A Trieur de cellules par cytométrie en flux (FACS)

Grâce à des marqueurs fluorescents spécifiques ou des intercalants de l'ADN, la cytométrie en flux permet de donner des informations protéiques, génomiques et transcriptomiques d'une population cellulaire à l'échelle de la cellule unique (Fulwyler, 1980).

Le trieur de cellules par cytométrie en flux ou FACS permet, en plus, de séparer une population hétérogène selon un phénotype défini. Chaque cellule est exposée à un ou

plusieurs lasers qui permettent d'identifier les cellules selon les caractéristiques sélectionnées. L'instrument donne une charge, positive ou négative, à la goutte contenant une cellule, puis un système de déflecteur électrostatique distribue les gouttes chargées dans les tubes correspondants. Depuis ces dernières années, la capacité d'un trieur de cellules a été améliorée passant de un ou deux marqueurs de fluorescences à 10 ou 15 marqueurs simultanément (Ye et al., 2017). Cela a permis de redéfinir, ou de mieux définir l'hématopoïèse et de mieux caractériser les cellules hématopoïétiques (Görgens et al., 2013b; Notta et al., 2016).

Le FACS n'est pas seulement utilisé pour isoler une population cellulaire hautement purifiée mais aussi pour isoler une cellule et les mettre en un puits d'une plaque de culture cellulaire (Parks et al., 1979). C'est la méthode que j'ai utilisé au cours de ma thèse pour extraire une cellule parmi la population de cellules CD34+.

Bien que le trieur de cellules par cytométrie en flux soit largement utilisé en recherche et en clinique, il présente quelques inconvénients. Il nécessite une grande quantité de cellules pour faire de simples réglages. De plus, trier une population cellulaire qui a été marquée par un anticorps au préalable peut changer le devenir et le phénotype de la cellule en activant des voies de signalisation (Thaker et al., 2015).

I.B Microfluidiques

Bien que je n'aie pas utilisé la microfluidique au cours de ma thèse, il me parait essentiel de décrire cette technologie très en vogue depuis quelques années pour isoler une cellule.

La microfluidique est une technologie qui utilise très peu de fluide (des microlitres à des femtolitres) grâce à des câbles de diamètres compris entre dix à cent micromètres. Elle permet d'utiliser très peu de matériels biologiques et de réactifs en contrôlant la concentration des fluides (Whitesides, 2006) ce qui rend la microfluidique une technologie très peu couteuse. Les puces microfluidiques ont connu différents types de matériaux au cours des dernières décennies allant de la puce fabriquée en silicone/verre ou en polymère à la puce en thermodurcisseur ou en hydrogel (Ren et al., 2013). Il existe différentes méthodes pour isoler des cellules uniques : isolement par microgouttes, piège hydrodynamique, magnétique, optique, acoustique et diélectrophorétique (Lo and Yao, 2015).

Seules les 2 premières méthodes seront développées ici car ce sont les méthodes les plus utilisées aujourd'hui pour isoler une cellule. Brièvement, le piège magnétique est basé sur l'action de marquage immunomagnétique. Grâce à des anticorps couplés à des billes magnétiques, les cellules d'intérêt peuvent être isolées d'une population hétérogène.

Le piège optique, lui, utilise un faisceau laser pour générer une force atomique dans une chambre microfluidique (Ramser and Hanstorp, 2010). Les cellules sont soit immobilisées et exposées à des milieux aux propriétés chimiques différentes soit mobiles entre les différents compartiments d'un système microfluidique (Enger et al., 2004).

Le piège acoustique, quant à lui, n'utilise pas de contact physique pour isoler des cellules, permettant ainsi un contrôle parfait de l'environnement cellulaire par microfluidique durant l'expérience. Son principe est d'utiliser une onde stationnaire à ultrasons qui génère un gradient de pression dans le milieu environnant aqueux donnant ainsi une force sur les cellules (Lindström and Andersson-Svahn, 2010).

Enfin, les cellules étant non conductrices, elles peuvent être déplacées et isolées par des forces générées dans un champ électrique non uniforme : le piège diélectrophorétique. Un dipôle est alors induit dans la cellule par un champ électrique. Selon les propriétés telles que la conductivité du milieu environnant, une force diélectrophorétique attire la cellule sur le champ maximum ou au minimum (Jubery et al., 2014).

I.B.1 Piège hydrodynamique

Contrairement aux autres méthodes microfluidiques, il est possible de séparer des cellules en fonction de leur taille, de leur morphologie et de leurs propriétés mécaniques et chimiques sans utiliser des champs externes. Ces techniques sont dites passives et utilisent le système hydrodynamique. Cette approche est attractive car elle ne demande pas un système expérimental aussi sophistiqué que la diélectrophorétique ou des systèmes de valves. Il existe plusieurs sous catégories de capture par hydrodynamique.

Le premier principe utilise des obstructions physiques sur une puce telles que des coupelles en forme de U pour isoler physiquement des cellules individuelles (Carlo et al., 2006). L'équipe de Di Carlo a été la première équipe à utiliser ce type de puce pour isoler 100 cellules uniques de carcinome cervical humain et à les suivre pendant 24 heures (Figure 20).

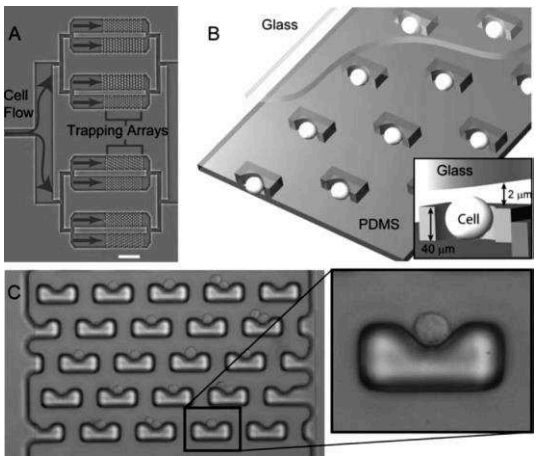


Figure 20 : Puce avec des coupelles en forme de U.

A. Photo du dispositif. B. Un schéma de l'appareil et du mécanisme de piégeage. C. Une micrographie à haute résolution du dispositif de piégeage. (Source : Carlo et al., 2006).

Ce dispositif représente un moyen pratique de regrouper des cellules individuelles pour étudier leurs réponses à divers stimuli. Elle permet aussi d'étudier la morphologie, la division et la mort cellulaire des cellules individuelles. Cette technique n'est pas faite pour suivre les cellules à long terme (pas plus de 24 heures) et le risque de perte des cellules dû au flux doit être pris en compte.

Depuis, cette approche a été largement modifiée et implémentée pour diverses applications comme la capture et la culture de cellules K562 et de fibroblastes dermiques humains ou l'exposition de différents courants sur des hépatocytes individuels (Benavente-Babace et al., 2014; Chen et al., 2015).

Le second principe est d'utiliser des canaux latéraux suffisamment petits pour piéger les cellules par aspiration insérés dans un canal principal (Tan and Takeuchi, 2007). Les doublons de cellules sont minimisés en ajustant la taille des canaux à la taille moyenne d'une cellule dans une population. Ainsi, le système hydrodynamique peut isoler un grand nombre de cellules issues d'un même échantillon en même temps grâce à l'implantation d'un grand nombre de canaux latéraux (Figure 21).

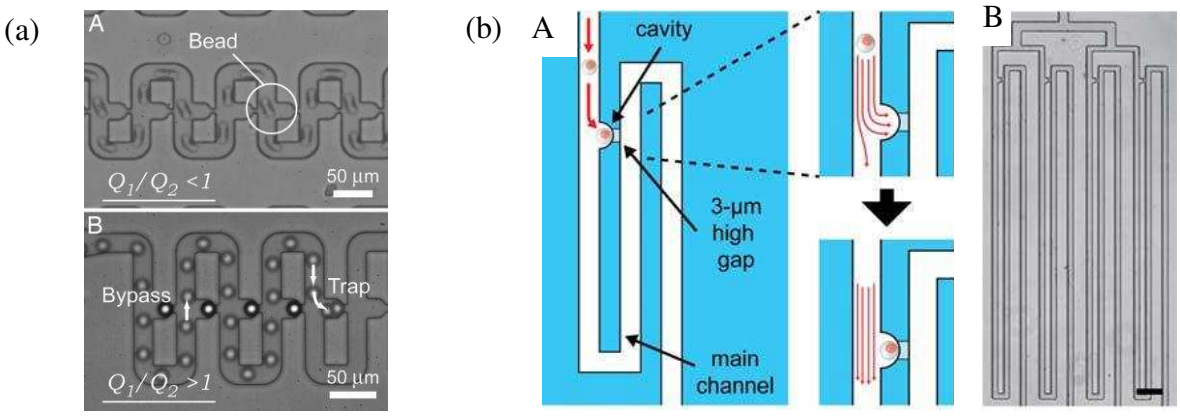


Figure 21 : Système hydrodynamique avec des canaux pour isoler des cellules uniques.

(a) Tan and Takeuchi, 2007. Ils sont les créateurs de ce type de système. A. et B. Time lapse montrant le chemin des billes dans le dispositif. (b) (Kobel et al., 2010). Contrairement à l'équipe de Tan et Takeuchi qui ont utilisé des billes, Kobel et al ont optimisé le système sur des cellules de mammifère. A. Illustration schématique d'un piège à cellule. B. Photographie du système de piégeage à cellules.

Par exemple, grâce au système C1 de Fluidigm, 96 cellules uniques peuvent être isolées et analysées en parallèle. Tsang et al. ont ainsi démontré que la délétion de Bcl11a dans les CSH élimine les lymphocytes et augmente l'expansion de deux sous population distinctes myélo-érythroïde (Tsang et al., 2015). Grover et al. eux ont montré qu'une surexposition de la cytokine érythropoïétine suppriment les niveaux phénotypiques et fonctionnels des progéniteurs non érythroïdes dans la MO et que la greffe de CSH exposées à l'érythropoïétine résulte à une lignée uniquement érythroïde (Grover et al., 2014).

Récemment, le système hydrodynamique peut être intégré dans des pipettes manuelles pour permettre de pipeter des cellules uniques de façon manuelle sans l'aide d'un micromanipulateur (Zhang et al., 2014).

I.B.2 Isolement par gouttes

La microfluidique à base de microgouttes est une technologie qui permet de séparer individuellement les cellules en les enfermant dans des microgouttes de volume compris entre le nanolitre et le femtolitre (Whitesides, 2006). Cette méthode utilise des canaux remplis d'un fluide non miscible comme l'huile pour permettre la séparation aqueuse des gouttes créées (telle que les émulsions) (Figure 22).

Le fait que les cellules soient chacune dans une goutte bien séparée évite les risques de contaminations biologiques et les cellules peuvent être analysées dans un large domaine d'application principalement par détection de fluorescence. Le petit volume des gouttes permet de réduire les réactifs tout en permettant une diffusion uniforme des réactifs dans toute la goutte.

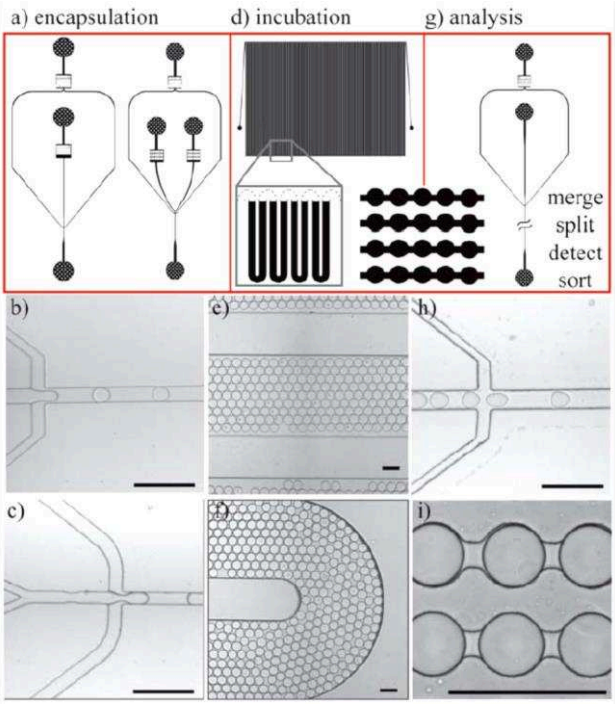


Figure 22 : Conception modulaire des composants microfluidiques.

Schémas et micrographies montrant la conception de puces d'encapsulation (a-c) avec une ou deux entrées aqueuses, des puces d'incubation (d-f) et des puces d'interface (g-h) pour analyser (fusionner, séparer, détecter, trier). (Source : Köster et al., 2008).

La plus grande limite de cette technique est le nombre de gouttes contenant une seule cellule ; environ 30% des gouttes chargées contiennent une cellule, c'est la loi de Poisson (Edd et al., 2008). Pour contourner ce problème, la solution a été d'espacer les cellules entre elles de façon uniforme dans le micro canal pour être sûr que les cellules entrent dans le générateur de gouttes avec la même fréquence que la formation de gouttes (Edd et al., 2008). La charge a pu ainsi atteindre 80% des gouttes contenant une cellule permettant un débit de plusieurs milliers de cellules par seconde.

L'avantage d'analyser des cellules uniques avec la microfluidique en gouttes vient de sa capacité de détecter rapidement des molécules sécrétées par des cellules grâce au petit volume autour d'elles (Lindström and Andersson-Svahn, 2010). Griffiths, Weitz et leurs collègues ont été les premiers à détecter la sécrétion d'anticorps de cellules hybridomes de souris isolées en gouttes après 6 heures de mise en culture (Köster et al., 2008). Récemment, la même équipe a démontré la détection de sécrétion d'anticorps après seulement 15 minutes d'encapsulation de ces mêmes cellules uniques avec des billes coatées à des anticorps spécifiques et des sondes fluorescentes (Mazutis et al., 2013). Ils ont aussi démontré la viabilité des cellules après plusieurs heures voire plusieurs jours d'encapsulation.

Une autre étape cruciale dans cette technologie est le maintien de la viabilité des cellules dans les gouttes. Les risques d'évaporation, d'épuisement de nutriments et d'accumulation de métabolites toxiques sont de réels obstacles qu'il faut prendre en considération avant de lancer des expériences sur du long terme. Des cellules Jurkat et HEK293T ont été cultivées dans des émulsions hors puce jusqu'à deux semaines et des analyses de viabilité ont été réalisées à différents temps, prouvant ainsi que les cellules restaient vivantes pendant quelques jours (Clausell-Tormos et al., 2008).

Ces mêmes équipes ont utilisé la microfluidique en gouttes pour étudier le transcriptome de cellules uniques à grande échelle : le Drop-seq et le inDrop (Ye et al., 2017).

Le Drop-seq met dans des nanogouttes une seule cellule avec une bille indexée. Pour chaque cellule, une bille différente va capturer les ARNm de la cellule (Figure 23). Ensuite, les billes sont mélangées et les ARNm sont rétro-transcrits. Les ADN complémentaires (ADNc) indexés provenant de milliers de cellules sont amplifiés ensemble et séquencés (Macosko et al., 2015).

Le inDrop, à la différence du Drop-seq, réalise la réverse transcription dans la goutte (Klein et al., 2015).

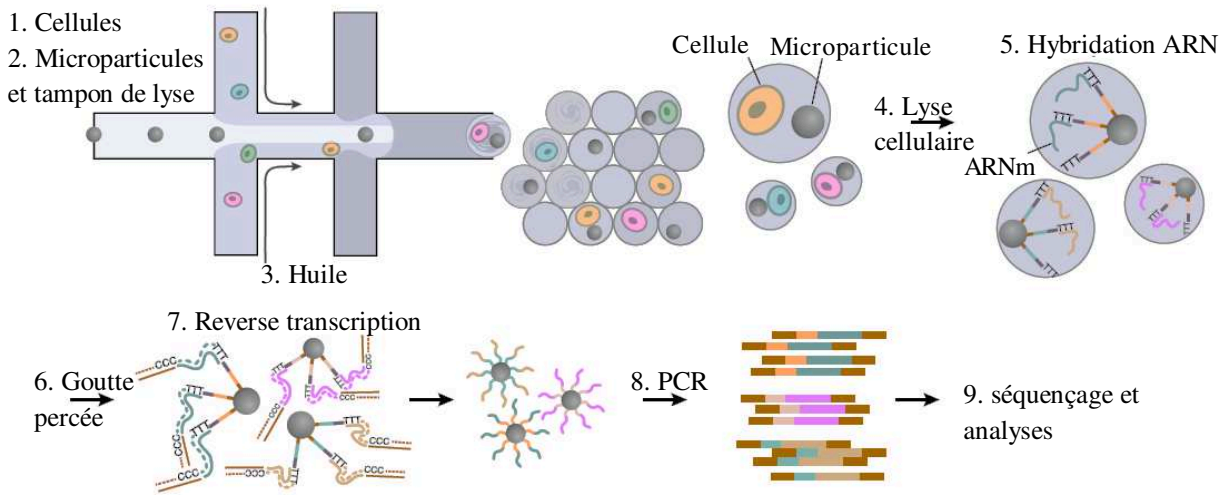


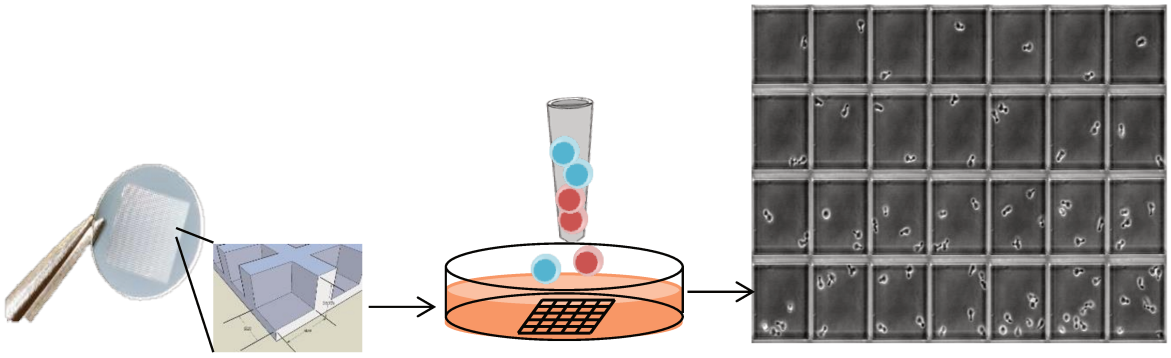
Figure 23 : Illustration de la préparation des ARNm-seq de cellule unique par la technique de Drop-seq. Le dispositif microfluidique possède 2 flux aqueux. L'un contient les cellules et l'autre contient les billes de primers barcodés dans un tampon de lyse. Immédiatement après la formation de gouttes, la cellule est lysée et libère ses ARNm qui s'hybrident aux primers situés à la surface des microparticules. Les gouttelettes sont cassées par l'ajout d'un réactif pour déstabiliser l'interface huile-eau et les microparticules sont collectées et lavées. Les ARNm sont ensuite rétro-transcrits ensemble et amplifiés par PCR. Schéma adapté de Macosko et al., 2015.

Les deux approches appliquent des indexes différents pour quantifier des ARNm individuels. Cependant, l'augmentation spectaculaire du débit est accompagnée d'un coût et le nombre de gènes détectés par cellule unique diminue en raison de la limitation de la profondeur du séquençage.

I.C Microgrilles

Créées dans les années 2000 en Australie par le Dr Daniel Day, les microgrilles, peu connues, sont une nouvelle méthode d'isolement de cellules vivantes non adhérentes (Day et al., 2008). Elles sont composées de plusieurs centaines voire milliers de puits allant de 50µm à 250µm de côté et 60µm de profondeur. Elles sont faites en polydiméthylsiloxane (PDMS), polymère de la famille des silicones. L'intérêt d'utiliser ce type de matériau est qu'il est non toxique et biocompatible avec les cellules, qu'il colle parfaitement sur une surface en verre ou en silicium et qu'il est hydrophobe.

Les microgrilles sont déposées dans une boite en plastique contenant une lame en verre et du milieu de culture (Figure 24). Les cellules sont ensuite déposées dans la microgrille et réparties selon la loi de Poisson. Grâce à un microscope inversé motorisé, une chambre d'incubation et une caméra, les puits contenant uniquement une cellule peuvent être sélectionnés ; les cellules peuvent donc être suivies et filmées individuellement sur plusieurs jours (Cosette et al., 2015; Duffy et al., 2012; Zaretsky et al., 2012). Ainsi, on peut suivre en temps réelle le temps de division de chaque cellule, leur dynamique morphologique et leur mouvement dans le puits. On peut aussi suivre l'expression génique fluorescente de cellules transgéniques (Duffy et al., 2012).



Microgrille en PDMS avec des micro-puits

Boite de culture

Time lapse durant plusieurs jours

Figure 24 : Schéma explicatif du mode de déroulement pour isoler des cellules grâce à la microgrille.

(Source image à gauche : Microsurfaces.com.au)

L'intérêt du TL, contrairement à un cytomètre en flux, est que l'on peut suivre l'hétérogénéité de prolifération, de mort cellulaire ou morphologique de chaque cellule dans une population, ce qui est complètement masquée lorsque l'on utilise un autre type de technique (Cosette et al., 2015). C'est d'ailleurs cette technique que j'ai utilisé au cours de ma thèse pour étudier la prolifération et la différenciation des LT en LTreg et pour étudier la division cellulaire et la morphologie des cellules CD34+.

L'inconvénient de cette technologie aujourd'hui est qu'on ne peut pas extraire les cellules à la fin ou au cours de l'expérience pour étudier leur transcriptome ou leur protéome. En effet, les cellules se trouvent dans des puits qui ne font pas plus de 250µm de côté et 60µm de profondeur, il est donc impossible de les extraire sans les perdre ou les perturber. De plus, les fichiers générés par le TL sont très volumineux, il faut donc avoir un logiciel assez puissant qui puisse supporter le traitement d'image. Enfin, l'analyse étant longue et fastidieuse, le nombre de cellules exploitées et exploitables est limité à des dizaines voire centaines de cellules, ce qui est un réel problème pour voir des populations rares.

II. Etude transcriptomique à l'échelle de la cellule unique

L'analyse du transcriptome à l'échelle de la cellule unique est un outil très puissant pour étudier l'hétérogénéité et le lignage cellulaire. Traditionnellement, la quantification de l'expression de gènes est réalisée en population, dans un échantillon contenant plusieurs milliers de cellules (Kanter and Kalisky, 2015).

Au cours de la dernière décennie, les avancées rapides dans la technologie génomique ont permis le développement de technologies à haut débit qui peuvent mesurer des centaines, voire des milliers de gènes, simultanément dans des centaines de cellules individuelles. À ce jour, les technologies transcriptomiques à l'échelle de la cellule unique ont permis de caractériser des populations de cellules rares comme des cellules tumorales circulantes ou des CSE dans les blastocystes de mammifères (Guo et al., 2010; Park et al., 2016). Elles ont aussi permis de revisiter les lignages cellulaires des CSE et des CSH (Nestorowa et al., 2016; Trott and Martinez Arias, 2013; Velten et al., 2017).

Grâce à ces nouvelles technologies, la variabilité de l'expression des gènes due à la stochasticité de l'expression des gènes ou dû à d'autres sources extrinsèques comme le cycle cellulaire a pu être mesurée (Elowitz et al., 2002; Hough et al., 2009; Richard et al., 2016; Zopf et al., 2013).

Après avoir réussi à isoler les cellules par l'une des méthodes décrites précédemment, le second défi ici est d'amplifier la petite quantité d'ARN qui est d'environ de 10 pg par cellule (Ye et al., 2017). Il existe plusieurs méthodes pour analyser des petites quantités d'ARN à l'échelle de la cellule unique comme la RT-qPCR (reverse transcription polymerase chain reaction), méthode par quantification de plusieurs ARNm simultanément et le séquençage des ARN totaux (ARN-seq).

II.A Mesure quantitative de l'expression de multiples gènes : la RT-qPCR

L'hétérogénéité de la population des cellules souches est largement reflétée dans les différences transcriptomiques parmi des cellules isogéniques individuelles. L'analyse du profil de transcription de cellules individuelles vise à établir des liens entre leur génotype et leur phénotype. En plus de la nature stochastique de la transcription d'un seul gène, l'hétérogénéité de l'expression génique chez les cellules ayant un génotype identique est due à des différences dans le cycle cellulaire, dans la morphologie, dans le microenvironnement et aux états épigénétiques. Donc le profil transcriptionnel à l'échelle de la cellule unique est crucial pour avoir une vision plus large sur l'hétérogénéité et son rôle dans l'auto-

renouvellement des cellules souches et de leur différenciation (Wu and Tzanakakis, 2013). C'est pourquoi, des méthodes pour mesurer simultanément plusieurs transcrits de gènes d'une seule cellule ont été développées. La RT-qPCR est une méthode quantitative qui mesure l'expression de plusieurs gènes simultanément et qui ne nécessite pas de séquencer l'échantillon.

II.A.1 Principales étapes de la RT-qPCR à l'échelle de la cellule unique

C'est cette méthode que j'ai utilisée au cours de ma thèse pour quantifier au cours du temps les ARNm de cellules CD34+ uniques. Je vais donc décrire brièvement les différentes étapes de cette technique.

Elle peut être appliquée à l'échelle de la cellule unique grâce aux méthodes de captures décrites précédemment telle que le FACS, la microfluidique ou l'isolement par gouttes. Après être isolées, les cellules sont lysées pour libérer le contenu cellulaire, puis les ARN sont purifiés et rétro transcrits (RT) de façon aspécifique avec des hexamères (Figure 25). Pour pouvoir quantifier l'expression des gènes d'une cellule unique, une étape de pré-amplification des gènes est nécessaire. Pour cela, une PCR de quelques cycles est faite sur les gènes choisis grâce à des primers. Les primers sont des petites séquences d'oligonucléotide qui se lient spécifiquement à l'ADNc du coté 5'et 3' du brin permettant à l'ADN polymérase d'amplifier les gènes. Une fois la pré-amplification réalisée, une PCR quantitative (qPCR) d'une trentaine de cycles peut être réalisée. À chaque cycle d'amplification, la quantité d'ADNc totale ou d'amplicon est mesurée grâce à un marqueur fluorescent. L'augmentation du signal d'émission fluorescente est directement proportionnelle à la quantité d'ADNc produits durant la réaction.

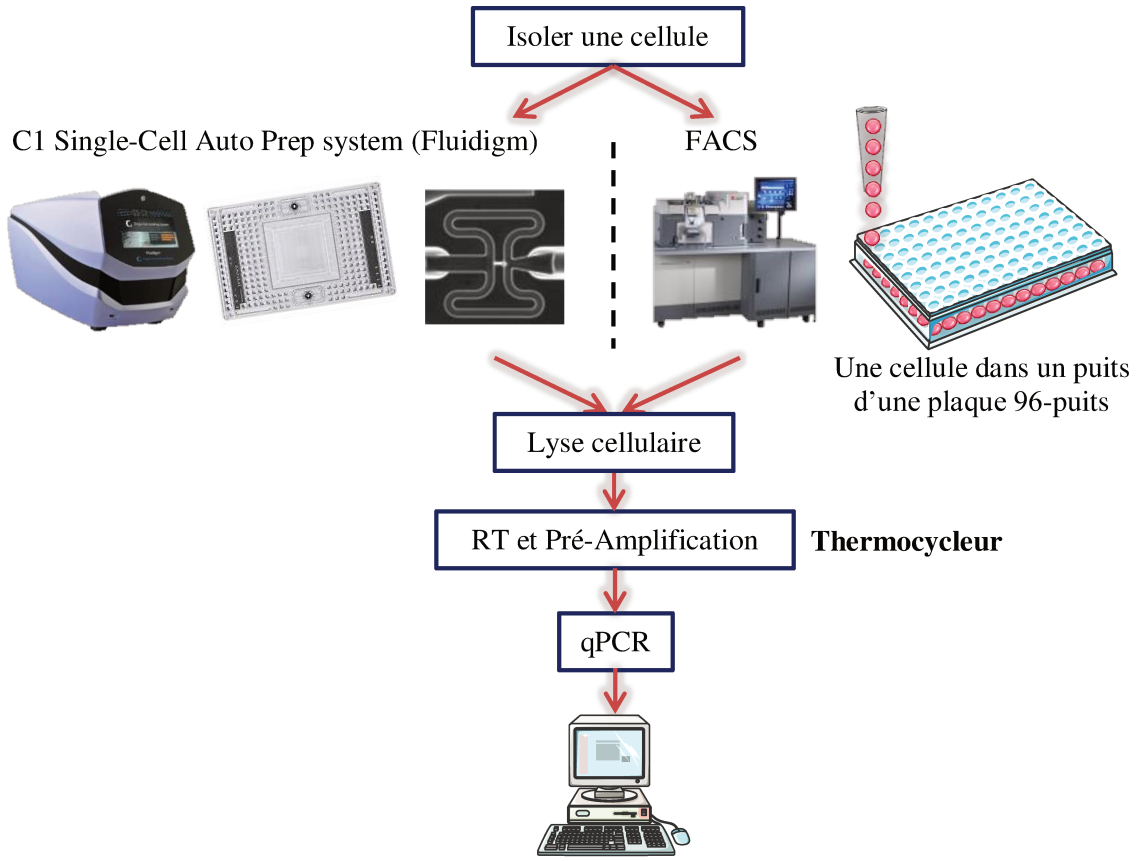


Figure 25 : Déroulement de la RT-qPCR à l'échelle de la cellule unique.

La cellule est d'abord isolée soit par microfluidique comme la technologie de Fluidigm avec son C1 soit par FACS et mise dans un puits d'une plaque 96 puits. Une fois isolée, la cellule est lysée, les ARN purifiés, rétro transcrits (RT) et les ADNc nouvellement formés sont pré-amplifiés par des primers préalablement choisis pour sélectionner les gènes d'intérêt. Enfin, une étape de qPCR de plusieurs cycles est réalisée.

Comme la qPCR en population, il existe deux principes pour détecter quantitativement les amplicons : des agents se liant à l'ADN double brin (ex. SYBR Green I) et des sondes fluorescentes (ex : sonde TaqMan) (Figure 26).

II.A.2 SYBR Green I

C'est la méthode que j'ai choisi d'utiliser pour quantifier les ARNm des cellules CD34+. Le SYBR Green I est une molécule qui détecte les nouveaux produits de PCR en se liant à l'ADNc double brin formé durant la PCR (Figure 26.A). Une fois lié à l'ADNc double brin, il devient un très bon fluorophore. Durant la PCR, l'ADN polymérase amplifie la séquence cible ce qui crée des nouveaux amplicons. Ainsi, le complexe ADN double brin/SYBR Green I émet une lumière verte (520 nm). Comme il y a plusieurs cycles de PCR, le nombre d'amplicons est multiplié par deux à chaque cycle ce qui entraine une augmentation de l'intensité de fluorescence.

L'avantage du SYBR Green I qui peut aussi être un inconvénient est qu'il peut être utilisé pour amplifier n'importe quelle séquence d'ADNc. De ce fait, il n'est pas nécessaire de créer des sondes mais il peut générer de faux positifs en se liant à des séquences d'ADN double brin non spécifiques.

II.A.3 Sonde TaqMan

La sonde TaqMan est une sonde oligonucléotide fluorescente qui permet la détection spécifique des nouveaux produits de PCR formés (Figure 26.B). Chaque sonde TaqMan s'hybride à une région d'ADNc spécifique amplifiée par des primers spécifiques. Elle contient un fluorophore à l'extrémité 5' du brin et un désactivateur (quencher) du coté 3' du brin.

Quand la sonde est intacte, la proximité du désactivateur réduit la fluorescence émise par le fluorophore. Quand la sonde est dégradée par l'enzyme Taq polymérase qui synthétise le brin néoformé, le fluorophore se trouvant au préalable à côté du désactivateur est alors libéré et peut ainsi émettre de la fluorescence. La fluorescence détectée est directement proportionnelle au fluorophore libéré et donc à la quantité d'ADNc présents dans la PCR.

L'avantage d'utiliser cette sonde plutôt que le SYBR Green est que seule une hybridation spécifique entre la sonde et l'ADNc d'intérêt génère une fluorescence. Il est aussi possible d'utiliser des fluorophores différents qui vont émettre des fluorescences dans d'autres couleurs ce qui permet d'amplifier diverses séquences dans le même tube de réaction. La limite de cette technologie est qu'il faut synthétiser des sondes différentes pour chaque séquence.

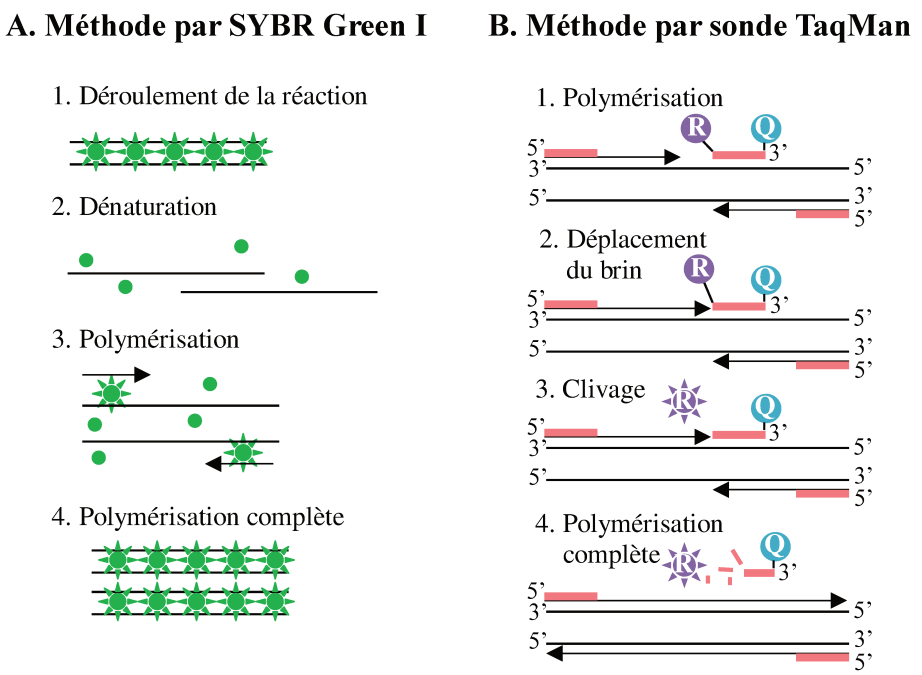


Figure 26 : Deux méthodes d'analyse quantitative de la PCR.

A. Méthode par SYBR Green I. 1. Déroulement de la réaction : la molécule SYBR Green fluoresce quand elle est liée à l'ADN double brin. 2. Dénaturation : quand l'ADN est dénaturé, la molécule SYBR Green est libérée et la fluorescence est réduite. 3. Polymérisation : durant l'étape d'extension, les produits de PCR sont générés. 4. Polymérisation complète : la molécule de SYBR Green est liée à tous les produits de PCR double brin ce qui conduit à une augmentation de la fluorescence.

B. Méthode par l'utilisation de sonde TaqMan. 1. Polymérisation : le fluorophore et le désactivateur sont liés respectivement à l'extrémité 5' et 3' de la sonde. 2. Déplacement du brin : quand la sonde n'est pas détruite, l'émission du fluorophore est désactivée. 3. Clivage : durant chaque cycle de PCR, la Taq polymérase clive le fluorophore de la sonde. 4. Polymérisation complète : une fois le fluorophore séparé du désactivateur, il émet de sa propre fluorescence.

L'analyse du transcriptome de cellules individuelles a donné de nouvelles informations concernant l'hétérogénéité dans la différenciation et l'auto-renouvellement des cellules souches. Grâce aux puces microfluidiques couplées à la technique PCR développées par la société Fluidigm, il est possible d'analyser en une seule fois 96 gènes de 96 cellules uniques, ce qui correspond à plus de 9000 réactions de qPCR (Livak et al., 2013; Porter et al., 2017; Richard et al., 2016).

Néanmoins, cette technologie présente quelques désavantages : il est impossible de conserver l'état de la cellule, une étape de lyse est forcément nécessaire pour récupérer le contenu de la cellule. Ainsi, le lien entre la morphologie, l'expression des gènes et la localisation des ARNm dans la cellule ne sont pas réalisables. De plus, cette technologie est chère ce qui limite le nombre de cellules analysables, la répétabilité et la reproductibilité des échantillons. Enfin, il faut choisir à l'avance les gènes que l'on veut analyser ce qui limite notre capacité à

découvrir de nouveaux gènes ou biomarqueurs. Il existe une méthode actuellement pour contourner ce problème : le séquençage des ARN totaux à l'échelle de la cellule unique.

II.B Etude des ARN totaux d'une cellule unique : le séquençage

Bien que je n'aie pas utilisé l'ARN-seq à l'échelle de la cellule unique au cours de ma thèse, il me parait essentiel de décrire cette technologie car elle est de plus en plus utilisée depuis quelques années pour séquencer tous les ARN d'une seule cellule.

L'hétérogénéité et la variation cellule-cellule sont des caractéristiques fondamentales et intrinsèques aux populations de cellules souches mais ces différences sont masquées quand l'étude est réalisée au niveau populationnel. Les technologies de séquençage à l'échelle de la cellule unique sont des outils puissants pour comprendre l'hétérogénéité des cellules et pour identifier des types cellulaires phénotypiques distincts même au sein d'une population de cellules souches « homogènes » (Wen and Tang, 2016).
L'étape critique dans le séquençage du transcriptome à l'échelle de la cellule unique est l'amplification du transcriptome entier qui comprend la RT des ARNm puis l'amplification des ADNc. Pour vraiment caractériser le transcriptome des cellules uniques, les ARNm doivent être complètement rétro-transcrits en ADNc et amplifiés de façon linéaire indépendamment de la quantité des ARN. Cependant, l'amplification par PCR introduit des variabilités surtout pour les ARN moyennement ou peu abondants (Wen and Tang, 2016). Par conséquent, plusieurs méthodes ont été développées pour rétro-transcrire toute la longueur de l'ARN en ADNc et éliminer les biais d'amplifications (Liang et al., 2014).

II.B.1 Méthode de Tang

Avant l'arrivée du séquençage à haut débit, des méthodes d'amplification du transcriptome entier à l'échelle de la cellule unique ont été développées pour analyser des profils d'expression de gènes dans des micro-puces à cellule unique (Kurimoto et al., 2007). L'équipe de Tang et al. a été la première à combiner la technologie de l'amplification du transcriptome entier dans des micro-puces à ARN à cellule unique avec le séquençage à haut débit, rendant ainsi possible l'analyse du transcriptome à l'échelle de la cellule unique (Tang et al., 2009).
Après avoir isolé la cellule, elle est directement lysée et les ARNm sont rétro-transcrits en ADNc en utilisant un primer oligo-dT avec une séquence d'ancrage, UP1 (Figure 27). Grâce à l'enzyme exonucléase I, les primers libres sont détruits et une autre enzyme, la terminale désoxynucléotidyl transférase, va venir ajouter une queue Poly(A) à l'extrémité 3' du premier

brin de l'ADNc. Le second brin de l'ADNc est synthétisé en utilisant un primer oligo-dT avec une autre séquence d'ancrage (UP2) et les ADNc sont amplifiées par PCR en utilisant les primers UP1 et UP2. Les ADNc amplifiés sont utilisables en l'état pour construire une librairie pour le séquençage à haut débit.

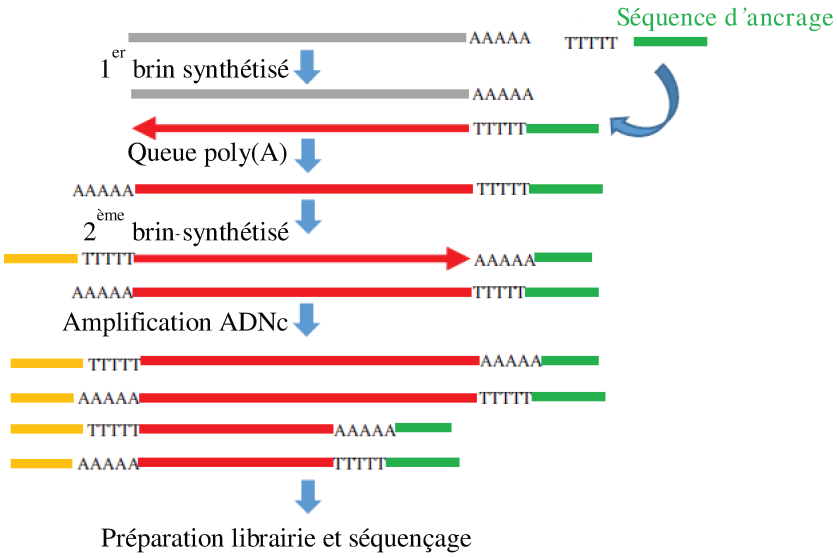


Figure 27 : Séquençage des ARN totaux par la méthode de Tang.

ARNm sont rétro-transcrits avec un primer oligo dT et une queue poly(A) est ajoutée à l'extrémité 3' du premier brin d'ADNc pour synthétiser le second brin d'ADNc. Schéma adapté de Liang et al., 2014.

Grâce à cette haute technologie, l'équipe de Tang a pu détecter l'expression de plus de 75% des gènes chez les blastomères de souris comparé à la technologie de micro-puces et découvrir plus de 1000 jonctions d'épissage jusqu'alors inconnues.

Cependant avec cette méthode, des biais sur les ARNm rétro-transcrits existent et sont principalement dus à la longueur limitée des produits de réverse transcription. Ainsi, elle peut être utilisée uniquement pour identifier des événements d'épissage distal et non proximal. Par conséquent, une méthode plus précise de séquençage du transcriptome à l'échelle de la cellule unique est nécessaire pouvant ainsi analyser toute la longueur des transcrits sans le biais de l'extrémité 3' du brin.

II.B.2 Smart-seq et Smart-seq 2

Smart-seq (switching mechanism at the 5' end of RNA template) génère et amplifie des ADNc entiers de cellules uniques grâce à la réverse transcriptase provenant du virus de la leucémie murine (Figure 28). L'activité transférase terminale de l'enzyme est essentielle au succès du Smart-seq. Quand l'enzyme arrive à l'extrémité 5' de l'ARNm, elle ajoute quelques nucléotides cytosines à l'ADNc (Zhu et al., 2001). Comme le mélange réactionnel contient un oligo marqué avec des résidus de guanines, la réverse transcriptase change de modèle, appelé

« template switching » et rétro-transcrit jusqu'à la fin de l'oligonucléotide. Cette méthode enrichit fortement les transcrits avec une extrémité 5' intacte et élimine le besoin de synthétiser un second brin (Ramsköld et al., 2012). Comparé à la méthode de Tang, Smart-seq augmente le nombre de transcrits entiers améliorant ainsi significativement l'identification d'isoformes de transcrits alternatives et de variantes génétiques.

Cependant, la limite principale du Smart-seq est une sensibilité plus faible que la méthode de Tang. Cela est dû à l'amplification sélective de l'ADNc de Smart-seq avec une extrémité 5' intacte.

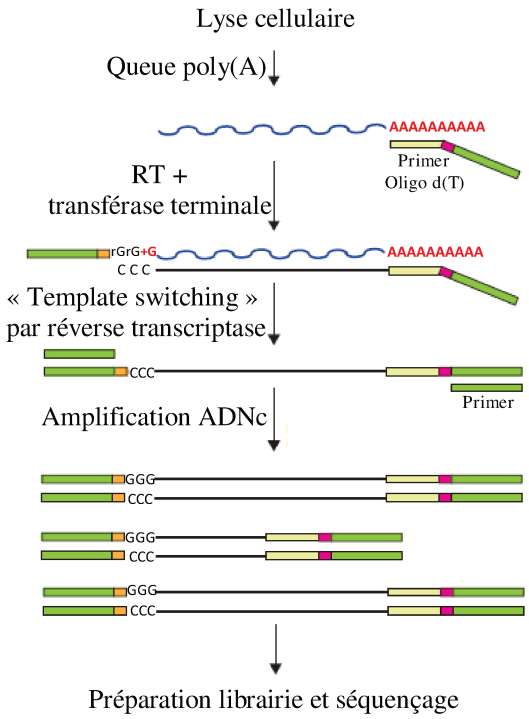


Figure 28 : Séquençage des ARN totaux par la méthode de Smart-seq 2.

Smart-seq et Smart-seq 2 utilisent la réverse transcriptase provenant du virus de la leucémie de murin avec une activité de « template switching » produisant ainsi des ADNc entiers. Schéma adapté de Picelli et al., 2014.

Smart-seq 2, développé par l'équipe de Picelli et al. a amélioré les étapes de RT, de « template switching » et de pré-amplification (Picelli et al., 2013). Smart-seq 2 a amélioré le rendement de la librairie d'ADNc, la sensibilité et a diminué les variabilités ainsi que les biais techniques (Picelli et al., 2014). Dans Smart-seq 2, les oligos du « template switching » portent deux riboguanosines et une guanosine modifiée pour produire un « acide nucléique bloqué » à l'extrémité 3' du brin (comparé à trois riboguanosines dans Smart-seq) (Figure 28). La modification introduit un verrouillage conformationnel dans la molécule mais conserve encore les propriétés physiques des acides nucléiques natifs. Les « acides nucléiques bloqués » permettent une amélioration de la stabilité thermique et de résister fortement à l'extension 3' non étirée de l'ADNc. De plus, dans Smart-seq 2, les dNTP sont ajoutés avant

la dénaturation de l'ARN ce qui augmente la longueur moyenne de l'ADNc préamplifié, probablement grâce à des mécanismes qui stabilisent les hybridations d'amorces d'ARN.

Smart-seq et Smart-seq 2 permettent de rétro-transcrire les ARNm en entier en ADNc, ce qui permet d'analyser les niveaux d'expression des gènes et de détecter les différentes variantes d'épissages même à l'extrémité 5' du brin.

II.B.3 Quartz-seq

La méthode de Quartz-seq est très similaire à la méthode de Tang (Sasagawa et al., 2013). La première étape est une RT avec un primer d'oligo-dT pour générer le premier brin de l'ADNc (Figure 29). La deuxième étape est une étape d'exonucléase pour éliminer un maximum de sous-produits. La troisième étape est l'ajout d'une queue poly(A) à l'extrémité 3' du premier brin d'ADNc. La quatrième étape est une étape différente de la méthode de Tang : le second brin de l'ADNc est synthétisé avec la même séquence d'ancrage que dans la synthèse du premier brin grâce au primer d'oligo-dT. La cinquième étape consiste en un enrichissement des ADNc par amplification avec une amorce PCR de suppression. Ainsi, les fragments de petites tailles comme les sous-produits forment une structure en épingle à cheveux due aux séquences d'ancrage identiques et ne sont donc pas amplifiés.

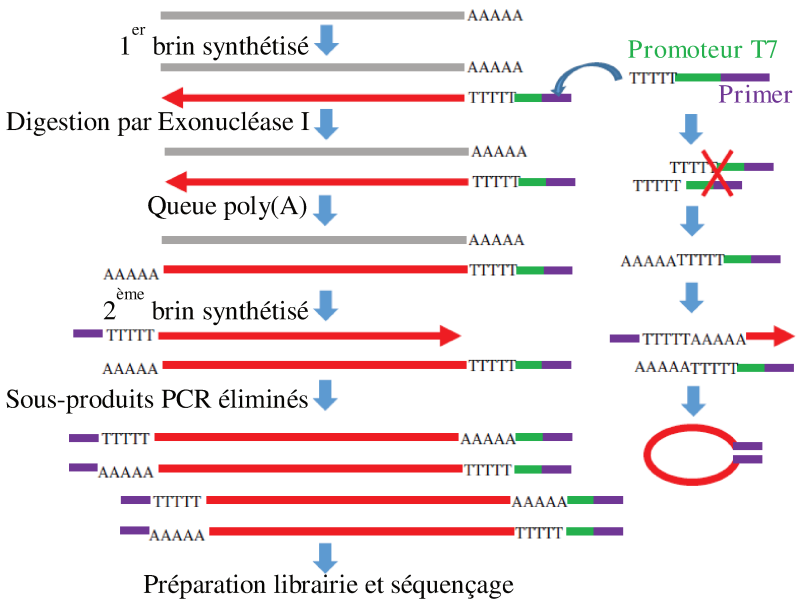


Figure 29 : Séquençage des ARN totaux par la méthode de Quartz-seq.

Après la synthèse du 1er brin d'ADNc, une queue poly(A) est ajoutée à l'extrémité 3' du 1er brin. La synthèse du 2ème brin est réalisée avec le même primer oligo-dT que pour la synthèse du 1er brin. Pour éliminer les résidus de primer restant, une PCR dite « suppresseur » va permettre de former des structures en épingle à cheveux pour les sous-produits dus aux séquences de primer identiques. Schéma adapté de Liang et al., 2014.

II.B.4 Expression cellulaire par amplification linéaire et séquençage (Cel-seq)

Cel-seq incorpore la transcription in vitro (TIV) pour amplifier le transcriptome à l'échelle de la cellule unique (Hashimshony et al., 2012). Les ARN de cellules individuelles sont rétro transcrits en ADNc avec un primer contenant des séquences d'oligo-dT, un barre-code unique pour chaque cellule, une séquence adaptatrice pour le séquenceur Illumina et un promoteur T7 (Figure 30). Après la synthèse du second brin, les ADNc de plusieurs échantillons sont rassemblées pour la TIV. L'ARN polymérase T7 va transcrire l'ADNc en ARN et les amplifier. Les ARN T7 transcrits sont ensuite fragmentés et un adaptateur à l'extrémité 3' est ajouté. Les ARN T7 sont ensuite rétro-transcrits en ADNc de sorte que seuls les fragments en 3' qui contiennent des adaptateurs Illumina et le barre-code sont sélectionnés. Puisque la TIV est une méthode d'amplification linéaire, Cel-seq peut théoriquement réduire les biais d'amplification par rapport à d'autres méthodes basées sur la PCR.

Cel-seq est une méthode hautement brin-spécifique parce que l'ADNc est généré par une amorce contenant un barre-code. Contrairement aux protocoles basés sur la PCR capables d'amplifier de l'ADNc en entier, Cel-seq montre un énorme biais dans l'extrémité 3' et ne parvient pas à détecter le spectre complet des variantes de transcription (Liang et al., 2014).

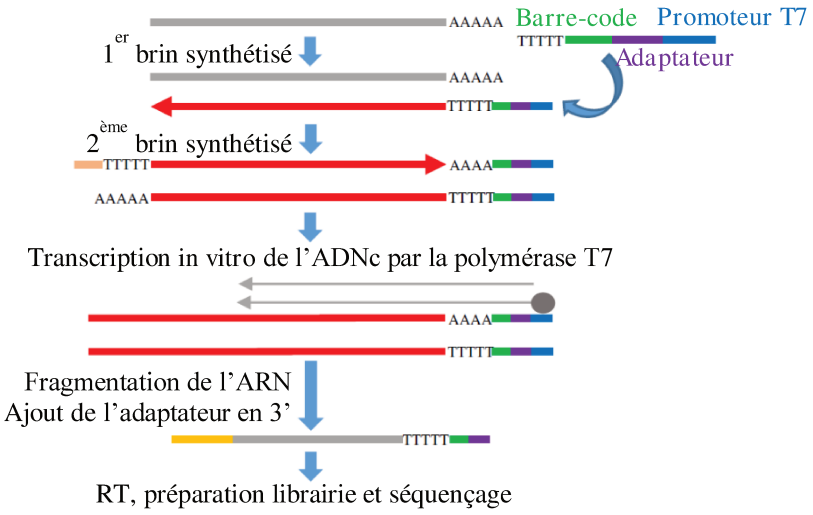


Figure 30 : Séquençage des ARN totaux par la méthode de Cel-seq.

Cel-seq ajoute un barre-code différent aux ADNc de chaque cellule unique et les rassemble pour réaliser une transcription in vitro (TIV) par la polymérase T7 pour amplifier de façon linéaire les ADNc. Schéma adapté de Liang et al., 2014.

II.B.5 Cellule unique marquée lors de la reverse transcription (STRT-seq)

La STRT-seq (single-cell tagged reverse-transcription) est une méthode proche de celle du Smart-seq (Islam et al., 2011).

Les ARNm sont rétro-transcrits en ADNc en utilisant des primers d'oligo-dT. Comme le mélange réactionnel contient un oligo marqué avec des résidus de guanines et un barre-code différent pour chaque cellule, les ADNc sont simultanément synthétisés et marqués à l'extrémité 5' grâce à la réverse transcriptase provenant du virus de la leucémie murine. Les ADNc provenant de différents échantillons sont rassemblés et amplifiés par PCR avec un unique primer relié à une molécule de biotine. Les ADNc amplifiés sont ensuite purifiés grâce à des billes de streptavidine et fragmentés (Figure 31).

Cette stratégie permet une comparaison juste et une analyse de l'expression des gènes entre plusieurs échantillons. C'est une méthode spécifique de séquençage du transcriptome à l'échelle de la cellule unique due aux barre-codes aux extrémités 5' des brins. Cependant, la méthode de STRT-seq détecte seulement les extrémités 5' des ARNm et non les ARNm en entier.

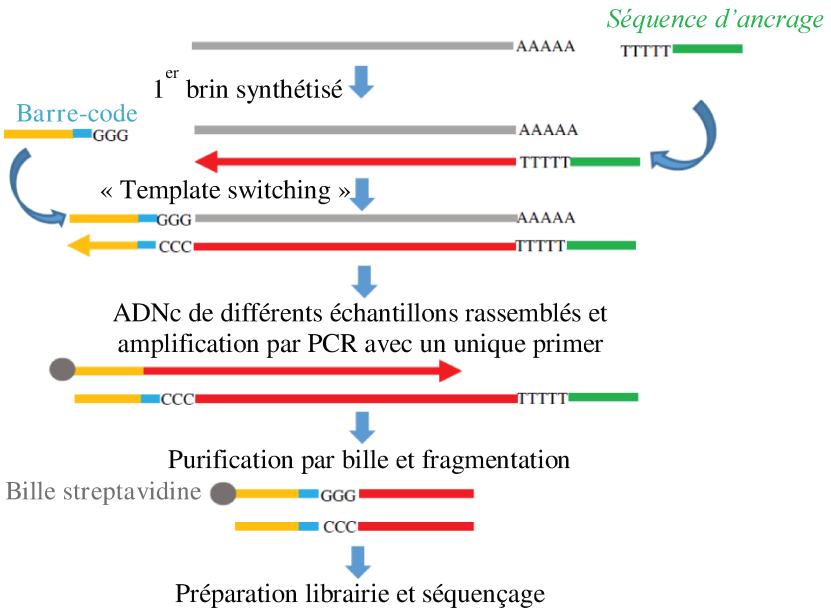


Figure 31 : Séquençage des ARN totaux par la méthode de STRT-seq.

STRT-seq marque les ADNc de chaque cellule unique avec un système de barre-code unique et les rassemble pour l'amplification dans le but de réduire les biais de comparaison des expressions de gènes entre les multiples cellules uniques. Schéma adapté de Liang et al., 2014.

II.B.6 ARNm-seq quantitative à l'échelle de la cellule unique

Pour pallier aux erreurs quantitatives provenant de la perte des ARN lors de la synthèse des ADNc et les biais dans l'amplification des ADNc, l'ARNm-seq quantitative à l'échelle de la cellule unique a apporté une nouvelle technique de marquage unique appelée identificateurs moléculaires uniques (IMU) (Islam et al., 2014). La synthèse des ADNc est réalisée en utilisant des primers oligo-dT en présence d'oligos de « template switching » portant un ensemble d'IMU (Figure 32). En ajoutant des séquences IMU aléatoires pour marquer les ADNc individuelles, le bruit d'amplification peut être éliminé. Les ADNc originaux sont identifiés par leurs IMU après amplification et un séquençage en profondeur.

En adoptant cette technique de marquage de l'IMU et le « template switching », cette méthode permet d'avoir un faible biais lors de la quantification des molécules d'ARNm, de générer et d'amplifier un ADNc complet similaire à la méthode de Smart-seq.

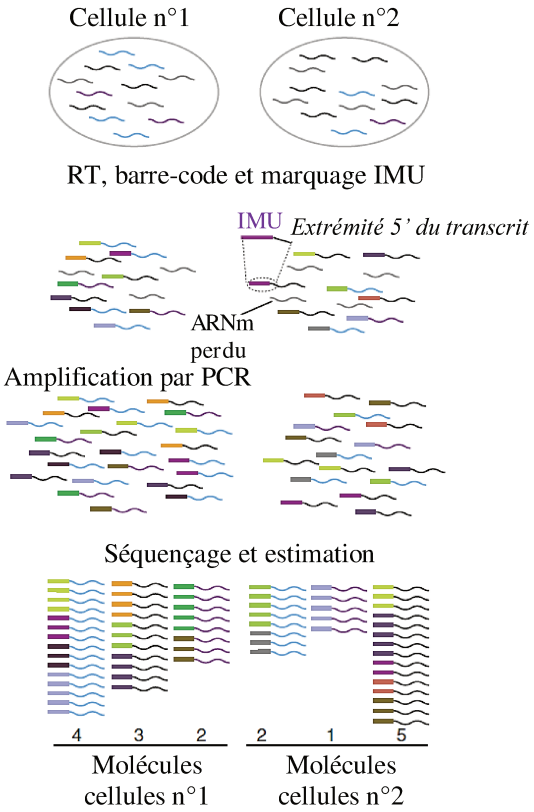


Figure 32 : Séquençage des ARN totaux par la méthode des ARNm-seq quantitative.
L'ARNm-seq quantitative à l'échelle de la cellule unique ajoute des séquences d'IMU (identificateurs moléculaires uniques) aléatoires pour marquer les molécules ADNc individuelles pour éliminer les biais d'amplification. Schéma adapté de Islam et al., 2014.

Grâce à ces nouvelles technologies de cellules uniques, l'expression des gènes des CSH individuelles peut être mieux caractérisée ; cela peut permettre de mieux comprendre et définir ce qu'est une CSH.

III. Etude protéomique à l'échelle de la cellule unique

L'analyse des ARN et de l'ADN à l'échelle de la cellule unique peuvent donner des informations qualitatives sur l'expression de la protéine mais ne peuvent pas donner des informations sur sa concentration, sa localisation ou ses interactions avec d'autres protéines ou molécules (Figure 33). La technologie protéomique à l'échelle de la cellule unique donne des informations cruciales sur l'hétérogénéité des phénotypes d'une population cellulaire et sur la différenciation des CSH.

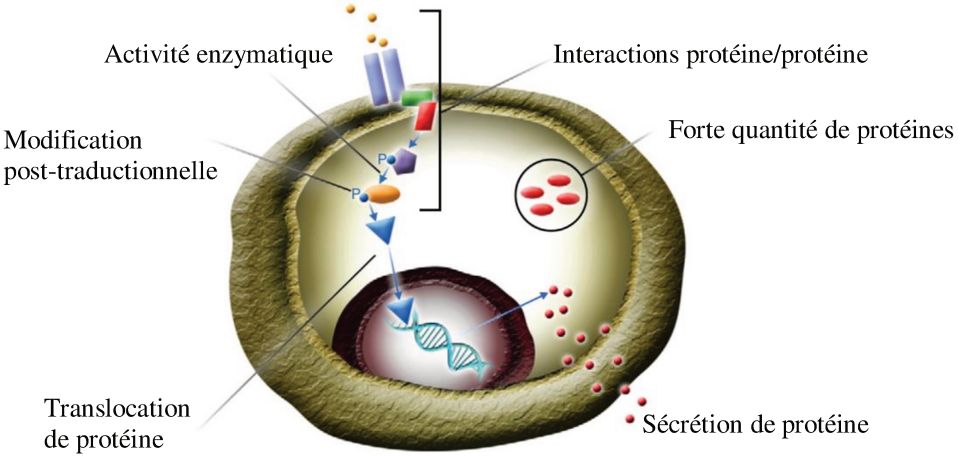


Figure 33 : Analyse protéique dans une cellule individuelle.

Le protéome d'une cellule est complexe et nécessite de mesurer la concentration de protéines cytoplasmiques et transmembranaires, d'étudier les interactions protéiques avec d'autres protéines ou molécules, la translocation des protéines, les modifications post-traductionnelles et les activités enzymatiques. Les formes colorées représentent différentes protéines. P : phosphorylation. Schéma adapté de Wu and Singh, 2012.

Les principales difficultés dans l'analyse des protéines à l'échelle de la cellule unique sont la détection d'un très faible nombre de protéines individuelles exprimées et le manque de méthodes d'amplification. Cependant, des avancées récentes dans les technologies multiparamétriques comme la cytométrie en flux, la microfluidique, la cytométrie de masse, la spectrométrie de masse et bien d'autres encore ont permis de réaliser des études protéomiques avec une grande spécificité et sensibilité à l'échelle de la cellule unique.

La cytométrie en flux est un outil que j'ai utilisé au cours de ma thèse pour analyser le phénotype des cellules des CD34+ uniques au cours du temps. Depuis sa création il y a un demi-siècle, la cytométrie en flux est devenue un outil majeur pour les analyses protéiques de cellule unique alimentant ainsi les progrès dans la compréhension d'une variété de systèmes cellulaires complexes (Mair et al., 2016). C'est une technologie puissante à haut débit (10^4 -

10^5 cellules/seconde) qui permet d'analyser plusieurs paramètres à la fois de façon qualitative et quantitative à l'échelle de la cellule unique (Piyasena and Graves, 2014). Elle permet de donner des informations sur l'hétérogénéité de la population cellulaire basée sur la taille, la granulosité et l'auto-fluorescence (Fulwyler, 1965; Gucker et al., 1947; Moldavan, 1934).

Le principe est le suivant : un échantillon de suspension cellulaire est injecté dans une cavité en forme de cône rempli d'un liquide de gaine sous pression (Figure 34). Ce liquide contraint l'échantillon injecté à un seul flux d'écoulement laminaire. Ce système fluidique assure le passage d'une seule cellule devant un laser. Comme le laser analyse chaque cellule, la dispersion de la lumière est enregistrée dans une direction opposée et perpendiculaire au laser. Ces deux paramètres sont appelés foward scatter (FSC) et side scatter (SSC) indiquant ainsi la taille et la granulosité de la cellule. Les cellules marquées avec un fluorophore génèrent un signal d'intensité de fluorescence. Ainsi, la cytométrie en flux permet la caractérisation phénotypique de cellules uniques d'une suspension cellulaire (Hosic et al., 2016).

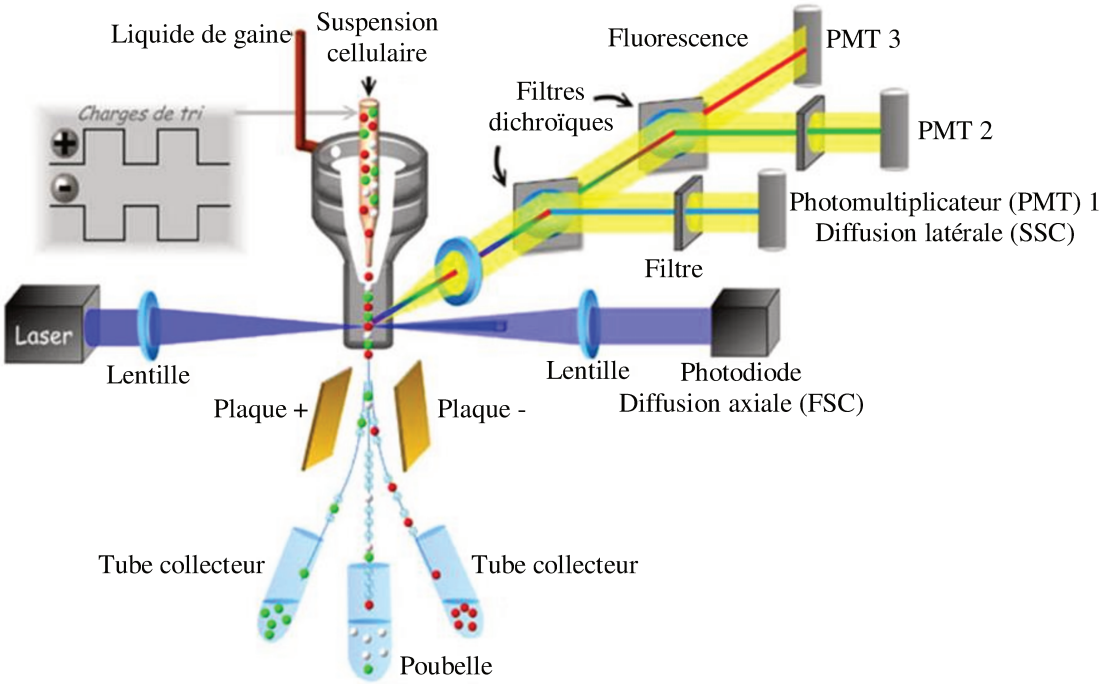


Figure 34 : Représentation schématique d'un cytomètre en flux.

Le cytomètre en flux comprend trois parties : 1. un réseau fluidique constitué d'un liquide de gaine qui entraîne et focalise un deuxième flux liquide contenant la suspension cellulaire. 2. un banc optique avec une ou plusieurs sources lumineuses, des détecteurs du type photodiode (pour la diffusion de la lumière), des photomultiplicateurs (PMT) et des filtres optiques qui permettent de détecter et quantifier les diverses fluorescences émises par chaque objet. 3. un microprocesseur qui convertit les signaux électriques en signaux numériques.(Source : Brown et al., 2001).

Il est aussi possible d'utiliser la cytométrie en flux intracellulaire en fixant et en perméabilisant les cellules pour analyser les protéines cytoplasmiques en se servant d'anticorps appropriés (Hoppe et al., 2014). En combinant la cytométrie en flux intracellulaire avec des marqueurs de surface d'antigènes, la population rare d'intérêt n'a pas besoin d'être purifiée avant l'analyse. Les inconvénients de cette technologie sont l'incapacité de détecter différents isoformes de protéines, le chevauchement des spectres d'émission des fluorophores, ce qui limite le nombre de fluorophores différents dans une expérience et la perte d'information sur la morphologie. Néanmoins, ce dernier problème peut être surmonté avec des technologies combinant la cytométrie en flux et l'imagerie (voir paragraphe V). Pour pallier au problème de chevauchement des spectres, récemment une nouvelle génération de cytomètre en flux avec une résolution spectrale améliorée, appelé « cytomètre spectral » a été mise sur le marché. Ce nouvel instrument contient un monochromateur multiprisme qui remplace les diviseurs de faisceau et les filtres des cytomètres en flux classiques. De cette façon, la perte de signal est minimisée et la détection jusqu'à une quinzaine de couleurs avec seulement deux lasers devient possible (Grégori et al., 2014).

La cytométrie en flux est la méthode la plus simple et la plus sensible pour analyser des protéines de cellules individuelles. Elle permet de maintenir intact l'état de la cellule. Depuis sa création, cette technique a été amélioré à la fois au niveau de l'appareil mais aussi au niveau de la disponibilité d'un grand nombre d'anticorps hautement spécifiques (Wu and Singh, 2012). Il est possible aujourd'hui d'analyser des dizaines de protéines à la fois dans une même cellule (Majeti et al., 2007; Notta et al., 2016, 2011).

Il existe depuis quelques années une autre approche permettant d'augmenter sensiblement le nombre de protéines analysables simultanément : la cytométrie de masse appelé CyTOF (Cytometry by Time-Of-Flight) (Mullard, 2011). Elle combine le débit de la cytométrie en flux avec la sensibilité du spectromètre de masse. Dans ce cas, les cellules sont marquées avec des anticorps couplés à des isotopes de métaux lourds. Les cellules marquées sont injectées, nébulisées (liquide transformé en nuage de gouttelettes) et les tags métalliques sont quantifiées à l'aide d'un spectromètre de masse (Wu and Singh, 2012). L'intensité du signal est proportionnelle au nombre d'anticorps liés à la cellule (Bendall et al., 2014, 2011). L'avantage d'utiliser la spectrométrie de masse comme moyen de quantification est dans sa capacité à pouvoir distinguer différents rapporteurs et à pouvoir discriminer les isotopes de différentes masses atomiques avec une grande précision (Spitzer and Nolan, 2016). Cette différence fondamentale permet significativement d'analyser les caractéristiques cellulaires

simultanément en utilisant une plate-forme de masse. Bien que cette technologie soit révolutionnaire, elle présente néanmoins des inconvénients comme l'impossibilité de trier et récupérer les cellules vivantes à la fin de l'expérience dus à l'atomisation et à l'ionisation des cellules et un débit plus lent en raison de la dynamique du vol ionique dans le spectromètre de masse (Spitzer and Nolan, 2016).

Dans toutes ces nouvelles technologies transcriptomiques et protéomiques de cellules uniques, le défi le plus important est de savoir analyser et interpréter les résultats multiparamétriques. En ce qui concerne la cytomètre en flux, la plupart des études ne prennent pas en compte cet état unicellulaire et ne font que des analyses bidimensionnelles successives des populations cellulaires à l'aide de « gating » manuel, régions d'intérêt dessinées. De nouveaux algorithmes existent et leur nombre augmente de jour en jour pour pouvoir analyser et interpréter ces résultats multidimensionnels sans omettre des informations essentielles. Dans le prochain paragraphe, nous allons justement voir les différents outils qui existent pour analyser ces données multiparamétriques.

IV. Analyse des données multiparamétriques

Avec l'arrivée des technologies de cellule unique, les données multiparamétriques générées en transcriptome et en protéome ont été exponentielles mais leur analyse n'est pas toujours simple. Les données sont parfois très complexes ; c'est pourquoi des algorithmes multiparamétriques ont été créés pour analyser et comprendre les résultats obtenus comme l'ACP (analyse en composantes principales), le t-SNE (t-distributed stochastic neighbor embbeding), SPADE (spanning-tree progression analysis of density-normalized event), etc. Une description brève de ces différents outils multiparamétriques va être faite dans ce paragraphe (voir Tableau 2 pour leur comparaison).

En cytométrie, malheureusement, la plupart des analyses sont faites à l'aide de fenêtres bidimensionnelles manuellement, appelées « gating » manuel et les résultats obtenus sont ensuite représentés sous forme d'histogrammes ou de nuages de points. Dans le prochain paragraphe, nous allons voir pourquoi cette méthode d'analyses « traditionnelle » n'est pas la plus efficace ni la plus appropriée pour ce genre de données.

IV.A Inconvénients de la méthode « gating » manuel en cytométrie

Historiquement, les données de cytométrie en flux étaient/sont encore analysées en utilisant des représentations en 2D (2 dimensions) et des « gating » manuels. Cette manière d'analyser est simple et intuitive et permet de sélectionner des populations spécifiques d'intérêt qui peuvent être isolées en créant des sous-groupes sur des représentations bivariants différents et en reportant les mesures qualitatives ou quantitatives associées. Cependant, elles n'offrent pas toutes les possibilités de visualisation des données produites par la cytométrie en flux. En effet, elle n'est pas facilement évolutive car le nombre de diagrammes en 2D augmente de façon exponentielle avec le nombre de paramètres mesurés et ne peut plus être comprise (par exemple, une expérience de 9 paramètres nécessiterait un total de 36 diagrammes en 2D pour afficher chaque combinaison possible). De plus, toutes les approches de « gating » manuel demandent d'avoir des connaissances antérieures, introduisant ainsi un biais vers les résultats « attendus ». Par conséquent, une grande partie des informations nouvelles et potentiellement pertinentes pourrait ne pas être identifiée et être complètement ignorée passant ainsi devant la découverte de nouvelles populations hétérogènes ou de nouveaux destins cellulaires (Lugli et al., 2010; Mair et al., 2016).

Afin de palier à ces difficultés et permettre l'extraction de toutes les informations des données, des algorithmes ont été développés depuis quelques années pour analyser les données multidimensionnelles de cytométrie et d'expression de gènes.

Méthodes d'analyse	Applications principales	Avantages	Limites	Exemples (non exhaustifs)	
				transcriptomique	protéomique
ACP	Visualisation des relations dans les données multidimensionnelles	Familier et établie	Technique de réduction des dimensions linéaire	Buettner et al., 2015; Moignard et al., 2013; Trott and Martinez Arias, 2013	Kalina et al., 2009; Kosugi et al., 1988; Mann et al., 1984; Newell et al., 2012
viSNE	Visualisation des relations dans les données multidimensionnelles	Technique de réduction des dimensions non linéaire	Difficulté de comparer des échantillons	Grün et al., 2015; Macosko et al., 2015; Saadatpour et al., 2014; Wilson et al., 2015	Acuff and Linden, 2017; Amir et al., 2013; Shekhar et al., 2014
t-SNE					Becher et al., 2014; Saadatpour et al., 2014
SPADE	Trouver des différences relatives dans les niveaux d'expression entre les échantillons	Visualisation des données de changement relatif	Perte résolution cellule unique et extraction manuelle des caractéristiques de groupes	-	Bendall et al., 2011; Bodenmiller et al., 2012; Jobin et al., 2015; Lee et al., 2015; Qiu et al., 2011
Citrus	Trouver groupes de cellules qui diffèrent entre échantillons et qui corrèlent ou prédisent les résultats expérimentaux	Analyse de régression intégrée pour comparer les échantillons et identifier les groupes différents entre les points finaux	Besoin d'env. 10 échantillons par groupe pour comparaisons valides	-	Bruggner et al., 2014; Gaudilliere et al., 2014; Hansmann et al., 2015
Diffusion map	Visualisation des trajectoires continues de développement cellulaire	Analyse non linéaire	Données uniquement d'expression génique	Haghverdi et al., 2015; Moignard et al., 2015	-
Wanderlust	Visualisation des trajectoires continues de développement cellulaire	Visualisation croissance, différenciation et morphogénèse des cellules	Processus de développement linéaires et non ramifiés	-	Bendall et al., 2014

Tableau 2 : Différentes méthodes d'analyses multiparamétriques.

Tableau adapté de Chester and Maecker, 2015 et Mair et al., 2016

IV.B Différentes méthodes d'analyses multiparamétriques

Les méthodes d'analyses multiparamétriques sont utilisées pour n'importe quel type de données de cellules uniques. En général, elles sont fondées sur des méthodes mathématiques classiques utilisées pour analyser la relation entre plusieurs variables.

IV.B.1 Réduction multidimensionnelle en 2D : ACP et t-SNE

La réduction multidimensionnelle permet de visualiser et d'analyser des données multiparamétriques dans un espace à 2D ou 3D en utilisant des dimensions de substitution. Cette représentation en dimension réduite donne une visualisation accessible des données où chaque cellule est représentée par un point. Il existe deux outils qui permettent de réduire les données multidimensionnelles en 2D/3D : l'ACP et le t-SNE.

IV.B.1.a Analyse en composantes principales (ACP)

L'ACP est une méthode de réduction linéaire de dimensions non supervisées qui génère un nouvel ensemble de variables décorrélées, appelées composantes principales, comme des combinaisons linéaires des variables d'origines (Figure 35.A). Elle réduit les données multiparamétriques en 2D ou 3D qui représentent les deux vecteurs propres les plus longs de la matrice de covariance ou de la matrice de corrélation. La perte des informations encourue avec cette transformation est minimale et permet de classifier les différents échantillons expérimentaux.

L'analyse des données de cytométrie en flux en ACP n'est pas récente (et pourtant peu employée aujourd'hui) ; elle a été utilisée pour la première fois en 1984 et en 1988 sur des LT (Kosugi et al., 1988; Mann et al., 1984).

Kalina et ses collègues, par exemple, ont étudié le phénotype des LB (marquage à 6 couleurs) chez des donneurs sains et chez des patients atteints du syndrome d'immunodéficience commune variable (Kalina et al., 2009). Avec la représentation des données en ACP, ils ont trouvé qu'une population de cellules spécifiques était présente chez les patients malades et absente chez les donneurs sains. Après « gating », ils ont montré que cette population représente des cellules B transitoires (CD19+CD27-CD38++CD24++) qui sont présentes spécifiquement chez les patients atteints d'immunodéficiences.

Pour analyser plus en profondeur le phénotype des LT CD8+, l'équipe de Newell et al. a utilisé la cytométrie de masse pour étudier 25 paramètres simultanément (Newell et al., 2012). Avec l'ACP en 3D, ils ont observé que les données de LT CD8+ obtenues formaient un « Y » avec à sa base des LT naïfs et des cellules effectrices à courte durée de vie et au niveau de ses

bras, des cellules T de mémoires centrales (Figure 35.A). Cette projection a été utilisée pour illustrer la différenciation continue des LT (Chester and Maecker, 2015).

IV.B.1.b t-distributed Stochastic Neighbor Embedding (t-SNE)

Le t-SNE permet de visualiser les données multiparamétriques à travers une réduction non linéaire (Van der Maaten and Hinton, 2008). Il permet de visualiser des similarités de cellules à haute dimension en les réduisant à 2D ou 3D (Figure 35.B). Les cellules qui ont des similarités vont être localisées à côté sur la représentation permettant ainsi de voir l'hétérogénéité des cellules.

Brièvement, les étapes de calcul pour représenter t-SNE sont les suivantes : premièrement, la matrice de similarité par paire des données est calculée en fonction de leur distance à haute dimension. Deuxièmement, une matrice de similarité à faible dimension est calculée basée sur la localisation initialement aléatoire pour chaque cellule en 2D ou 3D. Dans un processus itératif, l'algorithme essaie de minimiser les différences entre les matrices de similarités à haute et faible dimension ce qui permet d'ajuster la position de chaque cellule dans un espace à 2D (Mair et al., 2016).

Pour illustrer cette méthode, je vais vous présenter des analyses t-SNE issues des données de cytométrie, méthode peu couramment utilisée dans ce domaine. Au contraire, cette méthode d'analyse est largement employée dans les analyses de données transcriptomiques. Je l'ai moi-même utilisée au cours de ma thèse.

Pour comparer des échantillons de cellules de MO saines et cancéreuses, l'équipe d'Amir et al. a caractérisé ces cellules grâce à la représentation des données de cytométrie de masse en viSNE (Visualization of t-distributed Stochastic Neighbor Embedding) (Amir et al., 2013). Ils ont constaté que les cellules de MO saines se regroupaient automatiquement en une forme cohérente alors que les cellules des échantillons leucémiques se regroupaient en différentes formes qui sont distinctes des cellules de MO saines. Ils se sont aussi servis de cette représentation pour comparer les échantillons de leucémie avec les échantillons de rechute et ils ont ainsi pu identifier une population de leucémie rare rappelant une maladie résiduelle minimale.

Pour caractériser la diversité du phénotype des cellules myéloïdes, l'équipe de Becher et al. a développé un panel de 38 marqueurs pour la cytométrie de masse (Becher et al., 2014). En utilisant des souris transgéniques (Csf2rb-/-) dépourvues de récepteur au facteur de stimulation de la colonie macrophage-granulocyte (GM-CSF), ils ont confirmé grâce au t-SNE des différences de fraction de cellules dendritiques, macrophages pulmonaires et

éosinophiles dans les tissus. Cette méthodologie a aussi permis d'identifier des variations dans les compartiments des monocytes et des cellules lymphoïdes innées.

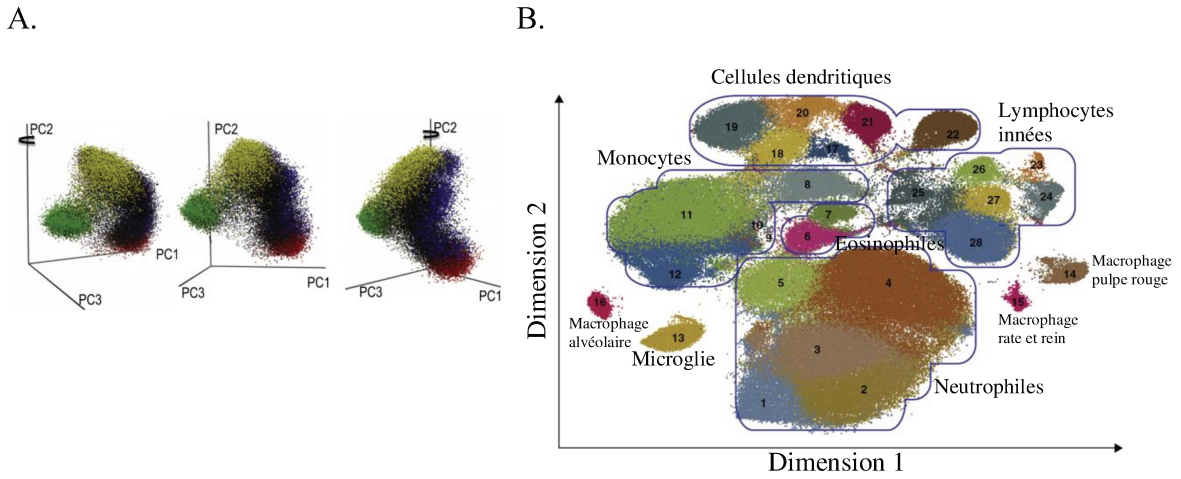


Figure 35 : Exemple de réduction multidimensionnelle en 2D ou 3D : ACP et t-SNE.
Chaque point représente une cellule avec ses différents paramètres. A. Représentation des données de Newell et al. (2012) avec l'analyse en composantes principales (ACP) en 3D. Les populations cellulaires T CD8+ naïves (vert), mémoires centrales (jaune), mémoires effectrices (bleu) et effectrices à court durée de vie (rouge) sont superposées pour identifier les principaux groupes phénotypiques. B. Représentation des données de Becher et al. (2014) en t-distributed stochastic neighbor embedding (t-SNE). Les couleurs représentent les différents groupes non supervisés trouvés par l'algorithme DensVM.

IV.B.2 Approche par regroupement : SPADE et Citrus

Les algorithmes de regroupement permettent de visualiser des données multidimensionnelles en définissant et en cartographiant des groupes discrets de cellules. Cette méthode fonctionne en rassemblant des observations similaires dans un seul groupe et en séparant celles qui sont différentes en les mettant dans des groupes différents. Dans une analyse cellulaire, les cellules avec des phénotypes similaires sont regroupées. Les inconvénients de cette approche par regroupement sont la perte de la résolution cellule unique et de devoir définir à l'avance le nombre de groupes souhaités ce qui introduit un biais. Il existe deux méthodes d'approche par regroupement : la méthode SPADE et la méthode Citrus.

IV.B.2.a Spanning-tree Progression Analysis of Density-normalized Event (SPADE)

L'un des premiers outils pour analyser et visualiser les données de cytométrie de masse est la méthode SPADE. Elle représente des populations cellulaires sous forme d'arbre généalogique, appelé dendrogramme, visualisant ainsi des données multiparamétriques en 2D. L'arbre SPADE consiste à relier des nœuds qui représentent des groupes de cellules, donnant ainsi des informations sur les relations des cellules (Figure 36.A). La distance entre 2 nœuds ne représente aucune information et peut être réarrangée par l'utilisateur. Généralement, les

branches de l'arbre sont annotées manuellement pour donner un aperçu immédiat des lignées cellulaires principales. L'un des avantages du SPADE est sa visualisation codée par couleur des intensités du signal (niveau d'expression) et des changements relatifs en se référant à un échantillon contrôle défini par l'utilisateur. C'est un outil parfaitement approprié pour visualiser des données de signalisation. Néanmoins, cette représentation en arbre avec des nœuds fait perdre la résolution cellule unique.

SPADE a été utilisé dans plusieurs publications, la plupart pour donner un aperçu des différentes populations immunes et de l'expression de leurs protéines de surface ou des molécules de signalisation intracellulaire (Bendall et al., 2011; Bodenmiller et al., 2012; Qiu et al., 2011). L'équipe de Jobin et al. a montré que des cellules CD34+ mises en culture dans deux conditions atmosphériques et physiologiques différentes, l'une à 21% d'O2 et l'autre à 8% d'O2, pendant 7 jours donnaient des phénotypes et des devenirs cellulaires différents (Jobin et al., 2015). Grâce à cette représentation, ils ont pu démontrer que lorsque ces cellules étaient soumises à des conditions atmosphériques à 8% d'O2, les cellules progénitrices érythroïdes et mégacaryocytes étaient moins abondantes.

IV.B.2.b Méthode Citrus

L'algorithme « Citrus » donne un processus non supervisé et automatisé qui combine l'identification de sous populations cellulaires avec différentes associations de modèles afin de révéler les groupes les mieux prédictives ou les mieux corrélées avec le critère expérimental (donneur sain/cellule cancéreuse, avant/après les soins thérapeutiques, etc...). Cette méthode reçoit en entrée les données de cytométrie avec les groupes expérimentaux prédéfinis. Ensuite, elle définit automatiquement des populations cellulaires et calcule des caractéristiques de regroupement. Après avoir appliqué une association de modèles choisis, elle donne à l'utilisateur une liste de groupes différents selon les d'échantillons expérimentaux, ainsi que ses caractéristiques représentées sous forme d'histogrammes ou de courbes (Figure 36.B).

Citrus a été utilisé pour analyser la signalisation des leucocytes sanguins et a identifié une corrélation avec la récupération des patients après remplacement de la hanche (Gaudilliere et al., 2014). Les patients opérés ont complété un questionnaire mesurant des paramètres cliniques incluant la douleur et la déficience fonctionnelle et ont donné des échantillons de sang périphérique. Cette méthode a identifié une corrélation entre les paramètres cliniques de la récupération chirurgicale et les réponses de signalisation dans les sous-ensembles des monocytes CD14+ (Figure 36.B).

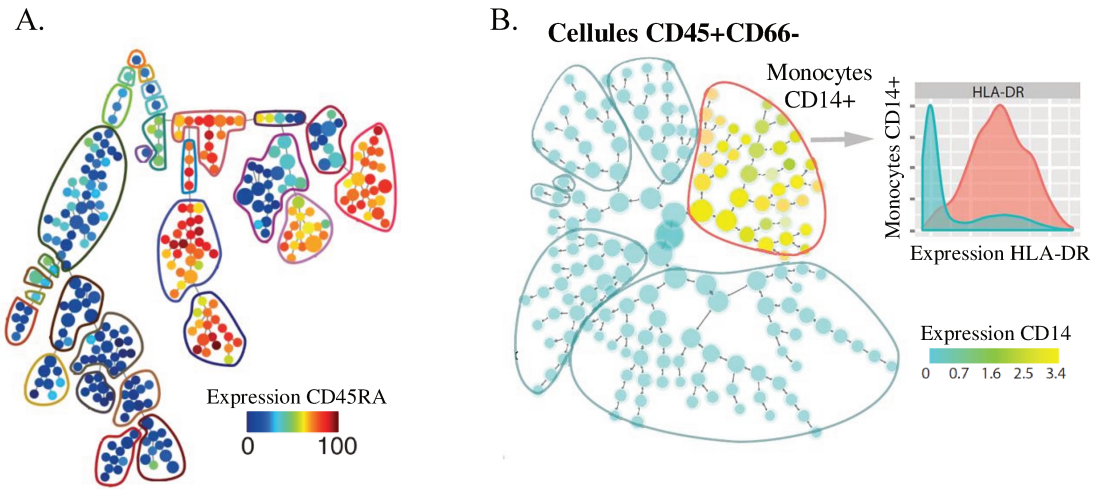


Figure 36 : Exemple de représentation des données multiparamétriques avec la méthode SPADE et citrus. A. Représentation des données de Bendall et al. (2011) avec l'algorithme SPADE. Le gradient de couleur représente l'expression de CD45RA dans les cellules de moelle osseuse. B. Représentation des données de Gaudilliere et al. (2014) avec l'algorithme Citrus. Le gradient de couleur représente l'expression de CD14 dans les cellules CD45+CD66-. L'histogramme compare l'expression de HLA-DR dans le groupe CD14+ (rouge) à l'expression de HLA-DR dans toutes les cellules CD45+CD66- (bleu).

Grâce à la représentation en Citrus des données de cytométrie de masse de cellules cancéreuses de plasma, Hansmann et ses collègues ont pu montrer que la fréquence des LT, des LB et des cellules NK étaient significativement différentes entre le groupe des cellules cancéreuses et le groupe des cellules saines (Hansmann et al., 2015). Ils ont ainsi pu conclure que les cellules cancéreuses du plasma avaient un effet sur les phénotypes des cellules mononucléaires du sang périphérique des patients.

IV.B.3 Détection de la trajectoire développementale

D'autres outils ont été développés pour analyser les données multiparamétriques à l'échelle de la cellule unique dans le contexte de la différenciation et du développement cellulaire : la carte de diffusion et Wanderlust.

IV.B.3.a Carte de diffusion

Haghverdi et ses collègues ont été les premiers à proposer la méthode de cartes de diffusion (ou en anglais diffusion map) pour résoudre le problème de la définition des trajectoires de différenciation (Haghverdi et al., 2015). Ils ont adapté cette méthode aux données de cellules uniques en incluant des incertitudes ou des valeurs de mesure manquantes, ce qui a permis d'établir un ordre pseudo-temporel de cellules isolées dans un espace d'expression génique à haute dimension (Coifman et al., 2005).

La carte de diffusion conserve la structure non linéaire des données sous forme de continuum et est robuste au bruit. Avec la normalisation de la densité, les cartes de diffusion sont résistantes aux hétérogénéités de densité d'échantillonnage et peuvent capturer des populations rares et abondantes. En tant qu'outil de réduction dimensionnelle non linéaire, les cartes de diffusion peuvent être appliquées sur les données omiques d'une seule cellule pour effectuer une réduction de dimension et ordonner les cellules le long du chemin de différenciation en une seule étape, fournissant ainsi un aperçu de la dynamique de différenciation (ou d'autres concepts avec une dynamique continue).
Grâce à l'analyse transcriptomique en carte de diffusion de plusieurs milliers de cellules uniques intervenant dans le développement précoce du sang, Moignard et ses collègues ont pu reconstruire les voies moléculaires contrôlant le développement des organes jusqu'alors mal comprises (Moignard et al., 2015).

IV.B.3.b Méthode de Wanderlust

L'algorithme Wanderlust est un outil créé par l'équipe de Bendall et permet d'analyser les données dans le contexte de différenciation et développement cellulaire (Bendall et al., 2014; Mair et al., 2016). A partir d'un processus de développement à un unique point de temps, l'algorithme de Wanderlust explore les relations de développement entre les cellules et construit une trajectoire qui ressemble aux voies de développement d'une lignée cellulaire. La trajectoire de Wanderlust ordonne les cellules uniques selon leur état les plus immatures à leur état les plus matures utilisant une échelle provisoire allant de 0 à 1 souvent appelée échelle de progression ou de développement. Cette trajectoire prédictible in situ peut ensuite être utilisée pour suivre la régulation de tous les autres marqueurs à travers les voies de développement.

La méthode de Wanderlust commence par convertir les données en un graphe de k voisins les plus proches (Figure 37). Chaque cellule correspond à un nœud et est connectée à ces k voisins par un lien. La distance la plus courte entre deux cellules est définie comme la longueur du chemin entre deux nœuds. Pour créer la trajectoire, l'utilisateur choisit une cellule initiatrice qui commencera la trajectoire et les autres cellules seront ordonnées en fonction du chemin le plus court par rapport à la cellule initiatrice. Pour rendre le calcul de la trajectoire plus robuste, Wanderlust inclut des points aléatoires et itère les calculs des chemins les plus courts sur plusieurs sous graphes. L'algorithme de Wanderlust produit enfin une trajectoire qui peut être tracée sur un axe d'un graphe en 2D.
Grâce à cet algorithme, Bendall et al. ont pu définir des nouvelles voies de signalisations régulatrices dans le développement des cellules B (Bendall et al., 2014).

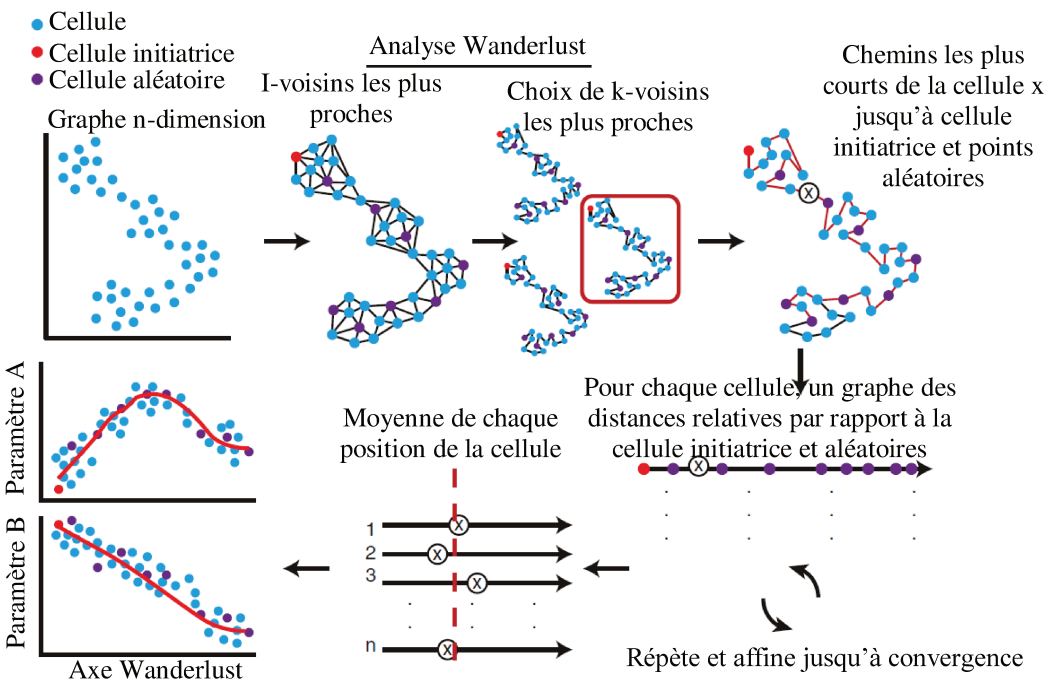


Figure 37 : Description de l'algorithme de Wanderlust.

Wanderlust transforme les données en un ensemble de graphes et sélectionne des points aléatoires. Chaque graphe est analysé indépendamment où une cellule initiatrice (point rouge) définie par l'utilisateur est utilisée pour calculer une trajectoire orientée. La trajectoire est répétée et affinée en utilisant les cellules aléatoires. La trajectoire finale est une moyenne de tous les graphes. Schéma adapté de Bendall et al., 2014.

Toutes ces méthodes d'analyse et de représentation des données de cellules uniques multidimensionnelles ont la capacité d'approfondir les connaissances sur la diversité des populations cellulaires et sur le devenir des cellules dans la différenciation.

Néanmoins, ces technologies transcriptomiques et protéomiques à l'échelle de la cellule unique ne permettent pas d'identifier les morphologies des cellules ni de suivre leur dynamique dans le temps. Une méthode complémentaire est donc nécessaire pour accroitre nos connaissances sur les cellules hématopoïétiques. La microscopie est l'outil idéal pour avoir des informations à la fois sur la morphologie et sur le phénotype des cellules. Grâce aux images prises successivement (TL), on peut avoir des informations importantes sur la dynamique des cellules durant la différenciation hématopoïétique.

V. Lien entre phénotype et morphologie des cellules individuelles

Toutes les technologies décrites précédemment ne permettent pas de prendre en compte la morphologie des cellules (Tableau 3). Les cellules étant toutes différentes, elles possèdent leur propre morphologie qui peut se transmettre de génération en génération. Grâce à la microscopie, il est possible de faire le lien entre la forme de la cellule et son génome, son transcriptome ou son protéome. Les cellules sont soit fixées, soit vivantes et isolées. La fixation des cellules permet ainsi de caractériser le phénotype des cellules à un instant donné et le suivi temporel en continu des cellules individuelles vivantes permet de caractériser sa dynamique morphologique. Ces deux procédures sont importantes et se complètent parfaitement dans l'étude de l'hétérogénéité des CSH et de leur devenir. Au cours de ma thèse, j'ai eu recours à ces deux modes pour étudier la morphologie des cellules CD34+ et leur dynamique.

Techniques \ Résolution	Niveau protéique	Cellules vivantes	Imagerie	Haut débit	Analyse en continu
ARNm FISH			X		
RT-qPCR					
Séquençage ARN totaux				X	
Cytométrie en flux ou de masse	X	X		X	
Imagerie en time lapse	X	X	X	X	X

Tableau 3 : Récapitulatif des différentes technologies de cellule unique et de leur résolution.

Seule l'imagerie en time lapse permet de quantifier à long terme le comportement des cellules individuelles. Tableau adapté de Kokkaliaris et al., 2012.

V.A Images instantanées de cellules uniques avec marqueurs phénotypiques

Pour faire le lien entre le protéome ou le transcriptome d'une cellule et sa morphologie à un instant donné, il existe plusieurs méthodes comme l'hybridation fluorescente in situ (ARNm FISH) pour la quantification des ARNm ou la cytométrie en flux en images (CFI) pour la quantification des protéines. Seule la cytométrie en flux en images sera détaillée ici car c'est la technologie que j'ai utilisé au cours de ma thèse pour analyser la morphologie d'un très grand nombre de cellules CD34+ et pour faire le lien entre leur morphologie et la localisation spécifique d'une protéine membranaire : le CD133.

Brièvement, l'ARNm FISH est une technologie qui permet de compter et de marquer les molécules d'ARNm de cellules fixées ou de tissus grâce à des sondes fluorescentes spécifiques (Cui et al., 2016; Kanter and Kalisky, 2015). Elle est souvent utilisée pour analyser quelques gènes exprimés et n'est donc pas appropriée pour analyser un grand nombre de gènes (Akel et al., 2000; Levsky et al., 2002; Visconti et al., 2006; Zeng et al., 2005). C'est pourquoi, elle est souvent complémentaire à la technique de RT-qPCR qui elle permet d'analyser et de quantifier des centaines de gènes simultanément dans une cellule.

La CFI, contrairement à l'ARNm FISH, est une technologie qui permet d'analyser à la fois la morphologie et les protéines des cellules uniques. C'est une technologie récente, à haut débit qui permet d'analyser le protéome et sa localisation dans des milliers de cellules. L'imagerie à l'échelle de la cellule unique à haut débit est une technologie essentielle pour l'étude de la biologie moléculaire et cellulaire. La CFI combine les capacités d'imagerie d'un microscope à l'échelle de la cellule unique avec les capacités à haut débit d'un cytomètre en flux classique (Han et al., 2016).

Le principe est le suivant : les cellules individuelles en suspension marquées par des sondes fluorescentes ou non sont introduites dans un système fluidique où elles sont focalisées dans un flux cellulaire. Les cellules passent ensuite devant une source de lumière à diode électroluminescente lumineuse et devant au moins un laser qui va créer des signaux de lumière transmise et dispersée ainsi que des signaux fluorescents (Wilkins et al., 2017). Avec un CFI, les signaux lumineux, transmis, dispersés et fluorescents sont collectés par une lentille d'ouverture numérique élevée (objectif 20x, 40x et 60x) et décomposés en champs spécifiques en fonction de leurs longueurs d'onde. Ces intervalles de longueurs d'onde sont ensuite focalisés dans différents canaux d'une caméra de dispositif à transfert de charge (caméra CCD) dans un champ spectral compris entre 430 et 800 nm. Ces canaux capturent des sous images qui peuvent être combinées pour analyser la co-localisation des signaux (Figure 38). L'avantage d'utiliser un CFI comparé à un cytomètre en flux classique est la possibilité d'utiliser des algorithmes d'analyse d'images sur un grand nombre d'échantillons simultanément. Une fois établie, les modèles d'analyses peuvent être facilement réutilisables et partagés entre les systèmes similaires. Cette approche a permis d'identifier un sous ensemble de cellules myéloïdes dans la MO qui interagit avec les CSH et induit la sur-régulation de COX-2 sous le stress (Ludin et al., 2012).

Elle a aussi permis de découvrir que les cellules CD34+ issues de sang de cordon étaient plus résistantes à l'apoptose endogène et induite par irradiation que les lymphocytes CD34- (Durdik et al., 2017). Enfin, la CFI a aussi permis d'étudier en profondeur les

mégacaryocytes, cellules très rares dans la MO (McGrath, 2015). La forme, la taille et la teneur élevée en ADN de ces cellules sont très similaires à des amas de cellules. En utilisant la forme, la texture et le chevauchement des signaux fluorescents, cette équipe a pu limiter l'analyse aux mégacaryocytes en éliminant les contaminants.

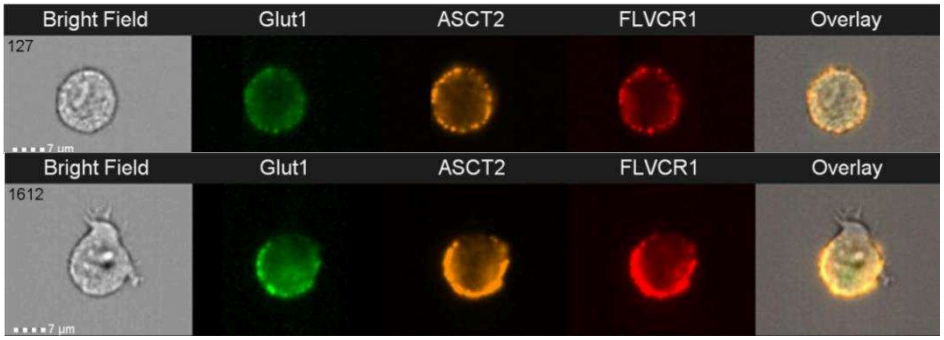


Figure 38 : Images obtenues avec un cytomètre en flux en images (ImageStream®).

Cellules CD34+ issues de sang de cordon, marquées avec des marqueurs métaboliques. De gauche à droite : image en lumière blanche, marquage avec Glut1 (Glucose Transporter 1), marquage avec ASCT2 (Solute carrier family 1 member 5), marquage avec FLVCR1 (Feline Leukemia Virus Subgroup C Cellular Receptor 1) et superposition des images pour voir la co-localisation des marqueurs.

Il faut toutefois noter que la CFI à l'heure actuelle ne peut pas remplacer la microscopie confocale due à sa plus faible résolution mais qu'elle est complémentaire à cette technologie de par son nombre important de cellules analysables simultanément.

Depuis quelques années, la cytométrie de masse peut aussi être couplée à de l'imagerie. Pour gagner en informations 3D, Giesen et al. ont couplé les méthodes d'immunohistochimie et d'immunocytochimie avec la cytométrie de masse CyTOF (Giesen et al., 2014). Cette méthode a l'avantage d'utiliser les mêmes équipements que la cytométrie de masse et permet de prendre des images de cellules marquées simultanément avec une trentaine de protéines. Grâce à cette nouvelle approche, ils ont pu ainsi imager des échantillons de cancer du sein humain ce qui leur a permis de délimiter les sous-populations cellulaires, les interactions cellule-cellule et de mettre en évidence l'hétérogénéité tumorale.

Une autre approche, appelée imagerie à faisceau ionique multiplexée, utilise un faisceau d'ions pour libérer les rapporteurs d'ions métalliques qui sont quantifiés par le spectromètre de masse (Angelo et al., 2014). Cette technologie est capable d'analyser jusqu'à 100 cibles simultanément sur une échelle de 5 log. Bien que cette méthode nécessite un équipement plus spécifique, elle offre une haute sensibilité et une grande résolution.

Les images prises instantanément donnent des informations sur la localisation de l'expression des gènes et des protéines dans la cellule et sur l'hétérogénéité de la morphologie d'une population cellulaire. De par son caractère instantané, elles ne donnent qu'une partie des informations et ne révèlent pas la dynamique morphologique des cellules. L'idéal (mais malheureusement impossible) serait donc de pouvoir observer in vivo les fluctuations de l'expression des gènes et des protéines au cours de la vie de la cellule afin de comprendre l'engagement des cellules dans les différentes voies de différenciation. Il est néanmoins possible de suivre le déplacement de la cellule in vitro et d'observer sa dynamique morphologique au cours du temps grâce à la microscopie en TL.

V.B Suivi temporel de la dynamique phénotypique des cellules uniques in vitro : l'imagerie en time lapse

L'imagerie en TL, bien que sous exploitée, n'est pas une technologie récente ; la chronophotographie a été inventé en 1891 par Etienne-Jules Marey, passionné par les mouvements des animaux (Landecker, 2006). Les images de cellules uniques prises à un instant donné donnent des informations précieuses sur le phénotype des cellules mais ne suffisent pas pour étudier la dynamique cellulaire et moléculaire (Figure 39). La compréhension du processus dynamique dans les populations hétérogènes passe forcément par une observation continue des CSH individuelles sur une période de temps définie (Rieger and Schroeder, 2008). L'imagerie en TL à long terme est donc indispensable pour une compréhension du destin cellulaire (Coutu and Schroeder, 2013). Elle permet d'étudier le comportement des cellules individuelles et de ces progénitures comme le temps de division, la différenciation, la mort cellulaire, la dynamique morphologique, la migration, etc., sur un lapse de temps continu (Kokkaliaris et al., 2012).

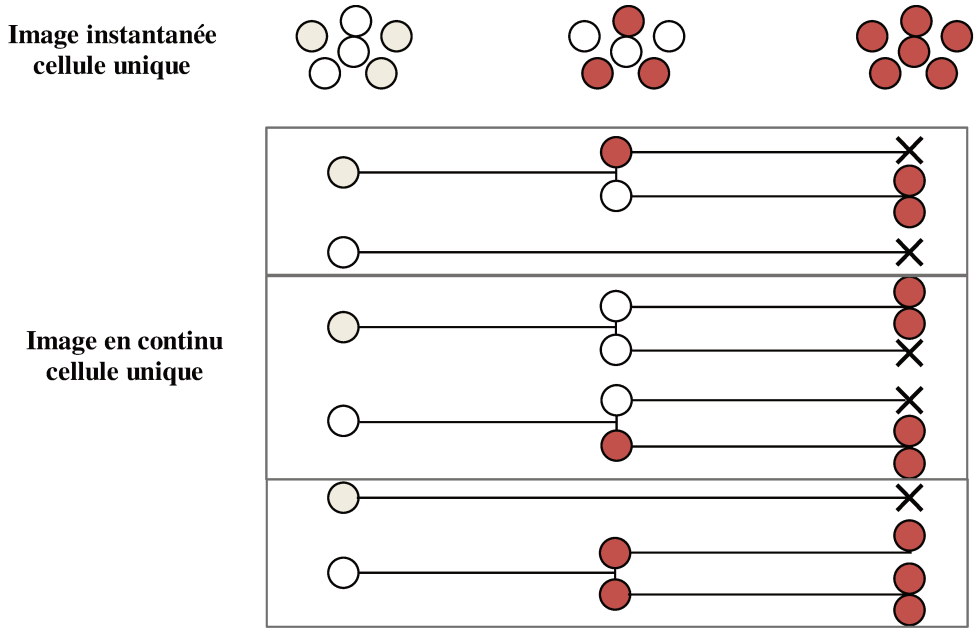


Figure 39 : Comparaison entre des images prises à des instants fixes et des images prises en continu. Les analyses de cellules uniques prises à un instant donné révèlent l'hétérogénéité mais ne donnent pas d'informations précises sur le comportement cellulaire. Seules les images prises en continu permettent une analyse complète du comportement cellulaire et moléculaire. Schéma adapté de Kokkaliaris et al., 2012.

Après avoir isolé les cellules avec différentes méthodes comme le FACS, la microgrille, la microfluidique, etc. et mises dans une boîte de culture adaptée, elles sont ensuite filmées grâce à un microscope inversé motorisé comprenant une caméra digitale à haute résolution. Grâce à la microgrille, j'ai pu isoler et suivre la prolifération et la différenciation des cellules CD4+. J'ai aussi pu recréer l'arbre généalogique des cellules CD34+ et suivre leur dynamique morphologique au cours du temps. Pour les maintenir en culture durant toute la durée de l'expérience, le système comprend aussi une chambre d'incubation avec un contrôle environnemental comme la température, l'humidité et les concentrations de gaz (CO_2). Grâce à la platine motorisée, des images en contraste de phase ou en fluorescence de différents champs peuvent être acquises ; plusieurs conditions expérimentales peuvent ainsi être suivies en parallèle pour un même échantillon. Ce genre d'expérience génère des fichiers de grande taille (environ une centaine de gigabits) qu'il faut ensuite analyser la plupart du temps de façon manuelle grâce à des logiciels de traitement d'images comme ImageJ, FijiJ, CellProfiler, etc. La plupart des équipes scientifiques crée leur propre logiciel pour analyser leurs données (Buggenthin et al., 2013; Chakravorty et al., 2014). Avec ces données, les informations sur le lignage cellulaire, la dynamique morphologique et le mouvement des cellules peuvent être extraites.

La nécessité d'une analyse de cellule unique pour résoudre les systèmes biologiques complexes devient de plus en plus évidente et de nombreuses études sur la biologie

moléculaire et cellulaire se sont tournées vers ces approches expérimentales à l'échelle de la cellule unique. L'imagerie en TL est une technologie puissante pour observer l'hématopoïèse à haute résolution spatiale et temporelle sur de longues périodes. En combinant le suivi de clones uniques et de leurs descendants, cette technologie permettra de nouvelles connaissances sur le comportement des cellules souches. Au fur et à mesure que de nouvelles souris transgéniques avec des rapporteurs fluorescentes spécifiques des lignées sont créées, la visualisation précise de petites étapes de développement dans l'hématopoïèse deviendra possible. Le suivi des cellules et le comportement des cellules individuelles ainsi que ceux de leurs descendantes pourront être analysés in vivo en continu.

Pour conclure, les analyses de cellules uniques ont eu une incidence importante sur notre compréhension actuelle de l'hématopoïèse et de la biologie des CSH. Malgré les énormes progrès réalisés dans le développement de technologies pour obtenir des données à l'échelle de la cellule unique à n'importe quel niveau, du génome au protéome en passant par l'état chromatinien, l'acquisition de deux procédés ou plus simultanément sur les mêmes cellules permettrait d'alimenter notre compréhension sur les processus consécutifs des CSH (Buenrostro et al., 2015). Les progrès récents dans les multi-omiques à une seule cellule offrent de nouvelles possibilités. En déterminant les états génomiques et transcriptomiques en parallèle d'une seule cellule, il est possible de révéler la signature transcriptionnelle, et potentiellement des cibles moléculaires, qui régulent la diversité dans la réactivité à la thérapie du cancer, par exemple. Une nouvelle méthode de séquençage à une seule cellule permettrait d'avoir en parallèle le méthylome et le transcriptome d'une même cellule afin d'évaluer ses processus de régulation (Angermueller et al., 2016).

Les données obtenues à un unique point de temps ne représentent qu'une image instantanée et sont souvent moins informatives que les données obtenues en continu comme le TL, car ni les fluctuations, ni les niveaux relatifs d'expression ne peuvent être déduits à partir d'un seul point de temps. Malheureusement, la plupart des analyses moléculaires nécessitent une lyse cellulaire ce qui rend impossible son suivi temporel. Des analyses en continu d'un nombre limité de gènes ou de protéines dans les cellules vivantes et le lien entre le niveau d'expression et le destin cellulaire ont été nécessaires pour répondre aux questions de longue date qui n'ont pas été répondues avec des analyses faites à un instant donné (Hoppe et al., 2016).

Bien que les premières études faites à partir d'une seule cellule datent depuis quelques décennies dans la recherche sur l'hématopoïèse, les nouvelles technologies de cellules uniques vont nous aider à mieux comprendre le choix du destin cellulaire.

OBJECTIFS ET STRATEGIES

L'objectif de ma thèse est d'étudier et de comprendre le choix du destin cellulaire des premières étapes de différenciation et de prolifération des cellules hématopoïétiques à l'échelle de la cellule unique.
On peut alors se poser plusieurs questions sur la différenciation cellulaire : est-ce un processus déterminé ou stochastique ? La part du hasard est-elle importante dans le devenir de ces cellules ou tout est connu, prédéterminé à l'avance ? Est-ce que tout est déjà figé avec une seule voie possible ou les cellules peuvent avoir un comportement dynamique, fluctuant ? On peut aussi se demander quand et à quelle vitesse ces évènements se déroulent. Est-ce que le choix du destin cellulaire se passe dans les premières heures de prise de décision ou dans les premiers jours ? Quelle est la part de la division cellulaire dans la différenciation ? La division asymétrique est-elle la seule, l'unique raison qui explique l'auto-renouvellement des CSH et de leur engagement ?
Pour répondre à toutes ces questions, il faut donc analyser les cellules à l'échelle individuelle car ce sont les cellules uniques qui se différencient et non pas la population.

Dans un premier temps, j'ai étudié un modèle simple de différenciation c'est-à-dire un modèle où la différenciation correspond à une transition entre deux phénotypes : la différenciation des LT CD4+ de souris en LTreg grâce à un gène rapporteur du facteur de transcription Foxp3. Dans notre équipe, nous avons déjà observé en population que la majorité des LT CD4+ se différenciait en LTreg après une semaine de culture, soit après une à deux divisions, soit après plus de 10 divisions (Neildez-Nguyen et al., 2015). Cette hétérogénéité de prolifération est plutôt surprenante dans une population cellulaire où les cellules se trouvent toutes dans les mêmes conditions. Mon but a été d'étudier cette hétérogénéité de prolifération mais aussi de différenciation et de mort cellulaire en analysant les comportements de ces cellules à l'échelle d'une seule cellule, en continu sur plusieurs jours grâce à la microscopie en TL. Afin de connaître et de comprendre l'origine du choix du destin cellulaire, stochastique ou déterministe, un modèle mathématique a été l'outil nécessaire pour simuler le comportement de ces cellules.

Dans un deuxième temps, j'ai étudié un modèle plus complexe de différenciation avec de multiple choix : les cellules CD34+ humaines issues de sang de cordon. Comme pour les LT CD4+, mon but a été d'étudier, à l'échelle individuelle, la différenciation et la prolifération de ces cellules mises en culture avec des cytokines pendant plusieurs jours afin de comprendre l'engagement d'une cellule dans une voie de différenciation plutôt qu'une autre. En reliant toutes les données obtenues, nous proposons un nouveau modèle de différenciation

stochastique avec des fluctuations d'expression génique possibles dans toutes les lignées avant de s'engager réellement dans une voie.

RESULTATS

PARTIE 1

Modèle simple de différenciation : la différenciation des lymphocytes T en lymphocytes T régulateurs

On peut se demander comment le choix du devenir cellulaire s'opère dans les cellules, en général. Est-ce un choix purement déterministe ou une part aléatoire existe dans cette prise de décision ?

Des études préliminaires, réalisées au sein de notre équipe, ont démontré que la prolifération des LTreg était hétérogène à l'échelle de la population (Figure 40) (Neildez-Nguyen et al., 2015). Les LTreg sont une sous population des LT caractérisés par l'expression du CD3, CD4 et CD25 chez la souris et de Foxp3. Les LTreg ont une fonction régulatrice du système immunitaire grâce, entre autres, à la sécrétion de cytokines immunosuppressives.

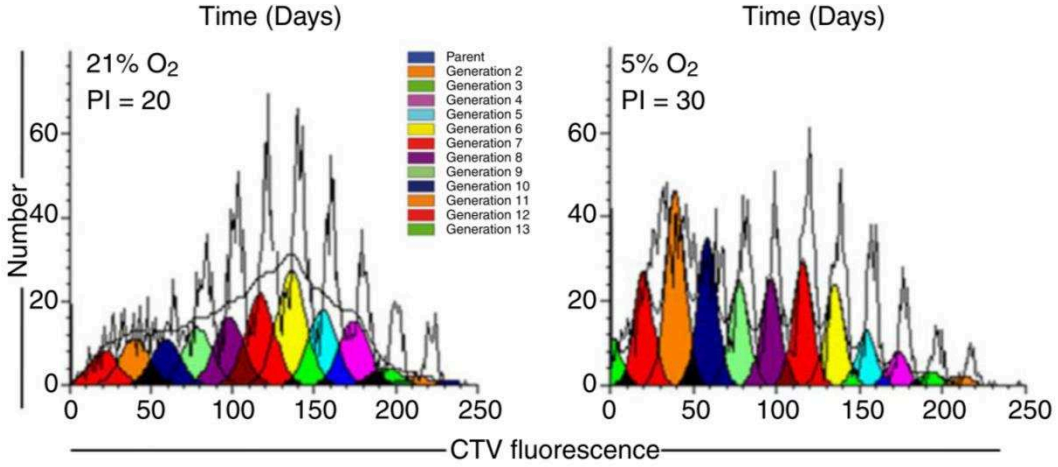


Figure 40 : Comparaison de la prolifération des LTreg entre 21% et 5% d'O_2.

Evaluation par cytométrie en flux de la capacité proliférative des LTreg après 11 jours de mise en culture sous les conditions d'O_2 à 21% et à 5%. Les différentes couleurs représentent les différentes générations. CTV = Cell Trace Violet. Figure de Neildez-Nguyen et al., 2015.

Nous nous sommes alors demandé d'où pouvait provenir cette hétérogénéité. Est-ce qu'elle préexistait déjà dans ces cellules ? Est-ce qu'elle était due exclusivement à l'hétérogénéité de prolifération, de différenciation, peut-être des deux ou elle était due uniquement à la mort cellulaire de certains clones ?

Pour répondre à toutes ces questions, nous avons étudié ce phénomène à l'échelle de la cellule unique. Pour analyser la prolifération et la différenciation de ces cellules, j'ai observé le comportement dynamique des LT CD4+ uniques, provenant de souris transgéniques, en utilisant la microscopie en TL. Les cellules ont été mises en culture dans un milieu contenant des cytokines et des billes d'activation recouvertes d'anticorps anti-CD3 et anti-CD28 et ont ensuite été déposées dans une microgrille de PDMS contenant 4096 puits, faisant chacun 50 µm de côté et 60 µm de profondeur. Nous avons suivi et analysé ces cellules toutes les heures pendant sept jours, nous permettant de mesurer le taux et le temps de prolifération, de

différenciation et de mort cellulaire. L'engagement des LT CD4+ en LTreg a été étudié grâce à l'expression d'un gène rapporteur GFP sous le contrôle du promoteur endogène du facteur de transcription Foxp3, exprimé uniquement chez les LTreg. L'analyse des images a permis de quantifier la prolifération, la différenciation et la mort cellulaire de plus d'une centaine de clones. Grâce à un modèle mathématique, nous avons pu mieux comprendre le comportement hétérogène des LT CD4+.

ARTICLE 1

“Single Cell Dynamics Causes Pareto-Like Effect in Stimulated T Cell Populations”

Alice Moussy*, Jérémie Cosette*, Fanny Onodi, Adrien Auffret-Cariou, Thi My Anh Neildez-Nguyen, András Páldi and Daniel Stockholm

* Co-premiers auteurs

Scientific Reports 5:17756 (Décembre 2015)

DOI: 10.1038/srep17756

SCIENTIFIC REPORTS

OPEN

Single Cell Dynamics Causes Pareto-Like Effect in Stimulated T Cell Populations

Received: 16 July 2015
Accepted: 27 October 2015
Published: 09 December 2015

Jérémie Cosette[1,*], Alice Moussy[2,3,*], Fanny Onodi[1,2], Adrien Auffret-Cariou[1,2], Thi My Anh Neildez-Nguyen[2,3], Andras Paldi[2,3] & Daniel Stockholm[1,2,3]

Cell fate choice during the process of differentiation may obey to deterministic or stochastic rules. In order to discriminate between these two strategies we used time-lapse microscopy of individual murine CD4+T cells that allows investigating the dynamics of proliferation and fate commitment. We observed highly heterogeneous division and death rates between individual clones resulting in a Pareto-like dominance of a few clones at the end of the experiment. Commitment to the Treg fate was monitored using the expression of a GFP reporter gene under the control of the endogenous Foxp3 promoter. All possible combinations of proliferation and differentiation were observed and resulted in exclusively GFP–, GFP+ or mixed phenotype clones of very different population sizes. We simulated the process of proliferation and differentiation using a simple mathematical model of stochastic decision-making based on the experimentally observed parameters. The simulations show that a stochastic scenario is fully compatible with the observed Pareto-like imbalance in the final population.

Naïve CD4+T lymphocytes are able to take multiple fate-decisions; they can give rise to various specialized cell types such as T helper effector or regulatory T (Treg) lymphocytes[1]. They do this in response to stimulations of their T-cell receptors (TCR) and various cytokines. Although studied for decades, the mechanisms of cell fate choice between different options remain elusive. The hypothesis of stochastic fate choice of hematopoietic cells was proposed 50 (fifty!) years ago[2]. Yet, the debate between the stochastic and deterministic mechanisms is still not settled. Some consider that the acquisition of the differentiated phenotype occurs via a predetermined pathway[3], where each signal induces a defined cell fate. Others argue in favour of a stochastic mechanism[4]. According to this view, a cell responds to a signal by randomly choosing between two or more options. It is the collective action of the individual stochastic choices that creates non-random regularities at the level of the whole cell population. We have previously observed that phenotypic heterogeneity may appear spontaneously and contribute to the fate decisions in a clonal population without the action of external signals[5–8]. In the present study we aimed to evaluate the stochastic contribution to T cell differentiation on the basis of single-cell observations obtained in an *in vitro* system. When the naïve T cells are stimulated *in vitro* with anti-CD3 and anti-CD28 antibody-coated beads, IL-2 and TGF-β, they proliferate and preferentially acquire the Treg phenotype making this artificial system convenient for the study of cell fate decision-making mechanisms[9]. These conditions are highly selective, because essentially all cells acquire Treg phenotype after a week or so of culture. One can therefore consider that there is little room left for chance during this process. Nevertheless, the cells take at least two different decisions: they divide or they differentiate. It is not known whether these two decisions are independent or whether they are taken in a fixed pre-determined order. Recently, we observed that the majority of the cells are displaying a Treg phenotype after a week of culture, some cells reach this stage after only one or two divisions while others divide up to ten times[7].

[1]Genethon, 1bis rue de l'Internationale, 91000 Evry, France. [2]Unit INSERM INTEGRARE UMRS 951-UEVE-EPHE, 91000 Evry, France. [3]Ecole Pratique des Hautes Etudes, Paris, France. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to D.S. (email: stockho@genethon.fr)

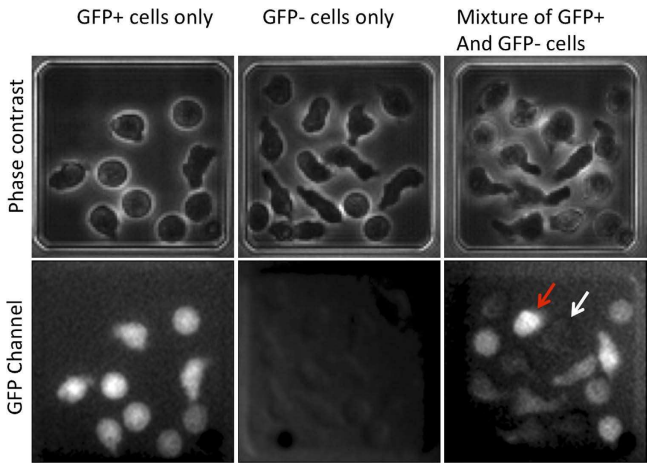


Figure 1. Cell division and phenotype detection using phase-contrast and fluorescent imaging. Phase contrast and green fluorescence channel representative images of the three main scenarios are observed: homogenous GFP+ Treg phenotype, homogenous GFP– phenotype and heterogeneous mixture of GFP+ (red arrow) and GFP– (white arrow) phenotypes.

This substantial proliferation heterogeneity is surprising in a cell population where each cell encounters identical conditions.

In order to get insight in the origin of this heterogeneous behavior we used a single-cell time-lapse approach coupled to mathematical modeling. Single-cell observations were successfully used to demonstrate the stochastic nature of fate decisions in B-cell differentiation[10]. Here, we used primary CD4+ cells from Foxp3-GFP knock-in mice so the acquisition of the Treg phenotype could be monitored in living cells using the expression of the GFP protein[11]. We observed substantial heterogeneity in the proliferation, differentiation and death rates leading to an unequal contribution of clonal cell lineages to the final population. Data-driven modeling of stochastic cell decision allowed us to show that the observed Pareto-like effect essentially results from the cumulative effect of stochastic cell decisions and events. Variations of cell cycle length and cell death rate are the key factors contributing to the phenotypic heterogeneity of the final cell population. The initial differences between the cells in the starting population may reinforce this effect but alone is insufficient to fully account for it.

Our observations show that due to the heterogeneity of proliferation and death rate, the final cell population is composed essentially from cells derived from a small number of initial founder cells.

Results

Time-lapse microscopy enables the exploration of single cell fate and shows that different scenarios of clonal differentiation/proliferation co-exist. Individual micro-wells containing a single cell were imaged every hour over a period of a hundred hours using time-lapse microscopy. We used the phase contrast channel to identify the cells, detect and count divisions, and detect the presence of the activation bead (Fig. 1). The fluorescence channel (GFP) was used to monitor the GFP fluorescence level that indicated the expression of the Foxp3 gene. The cells expressing GFP were considered as a sign of fate decision to acquire the Treg phenotype. No functional analyses were performed on these cells in this study. However, identical conditions were used in our previous work where the functional characteristics of these cells were confirmed[7]. Cell death was detected with the help of different parameters: the shape on the phase contrast image, and strong auto-fluorescence in the Rhodamine channel (not shown).

High rate of mortality was observed during the first 24 hours. In the 630 micro-wells containing a single living cell and at least one activation bead at $t = 0$h, 260 (41%) cells died during this period (Fig. 2). In the 310 micro-wells containing a single cell but no beads 131 cells (42%) died during the first 24 hours of culture. There was no significant difference in the death rate between the two groups (Kolmogorov-Smirnof test $p = 0.952$). Overall, the distribution and rate of the cell mortality was very similar in wells with or without beads during the whole period of the experiment. Therefore, the high death rate at the beginning of the cell culture is not due to the lack of activation. It could be rather due to the stress related to the transition from *in vivo* to *in vitro* culture.

Surviving cells in micro-wells without activation beads did not divide and did not differentiate demonstrating that stimulation is essential for both processes. Therefore, we only considered individual cells that survived in the presence of at least a single bead. After a lag period of varying length, the majority of these cells started to proliferate and differentiate. Cell death also occurred sporadically at different stages of the experiment. In order to determine if differentiation occurs in a predetermined way, we recorded the cell number and phenotype in each individual micro-well every hour. Each clone displayed a different proliferation and differentiation dynamics (Fig. 3). More than 50% of the cell clones

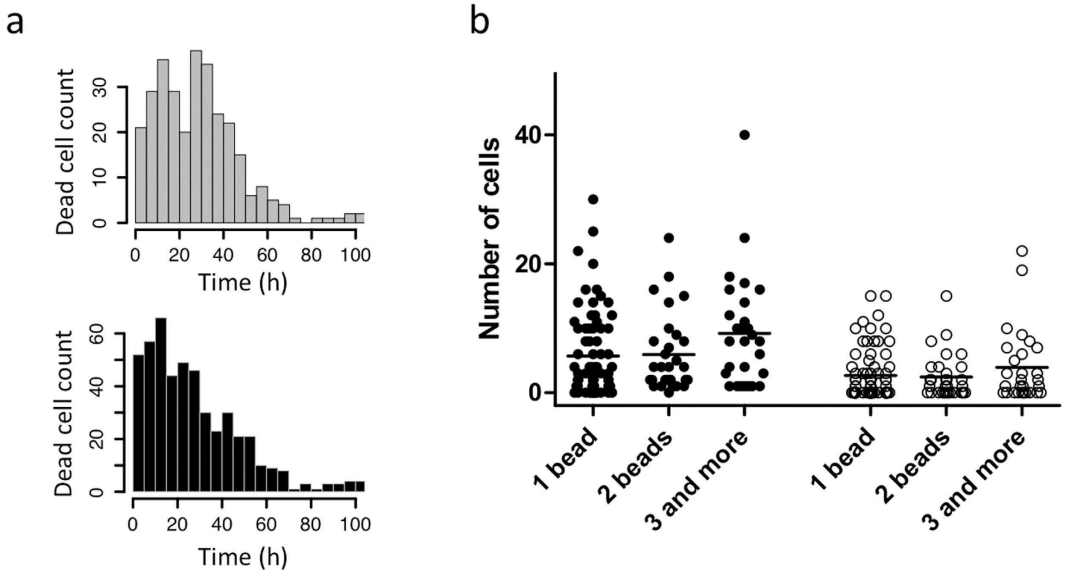


Figure 2. Influence of cell stimulation on mortality and proliferation. (**a**) Distribution of cell mortality during the experiment in absence (upper panel) and in presence (lower panel) of an activating bead in the micro-well shows no statistical difference (Kolmogorov-Smirnoff test $p = 0.9952$). (**b**) Clone size distribution (left – black dots) and number of GFP+ cells per clone (right – white dots) at the end of the experiment (100 h) in the presence of 1, 2 or 3 + beads. Each dot represents a single clone. The non-parametric ANOVA Kruskal-Wallis test was used to compare the different groups. Neither the total cell number ($p = 0.0847$) nor the number of GFP+ cells ($p = 0.5529$) were found significantly different. Same conclusion was reached using the Dunn's multiple comparison test.

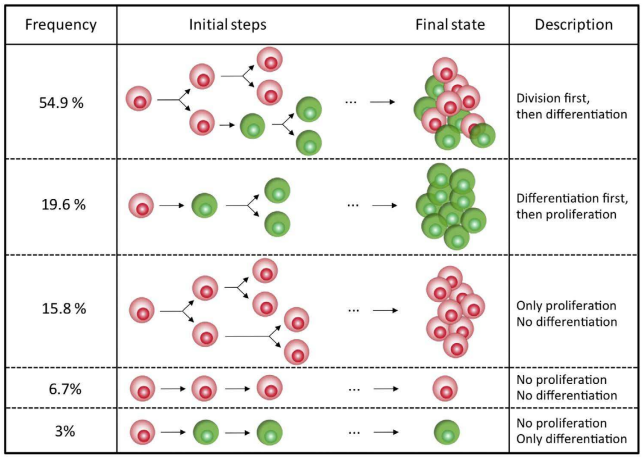


Figure 3. Different coexisting scenarios were observed during the first hundred hours suggesting that there is no pre-established sequence of events.

were composed of varying ratios of GFP+ and GFP– cells. About 20% of the clones were composed of only differentiated cells and 15% were pure GFP–. Interestingly, about 10% of the initial cells remained mitotically inactive in spite of the contact with the activation bead. Among them, some differentiated and others did not (Fig. 3).

In addition to the clones described above, we could observe micro-wells that randomly received a single starting cell and two or more activation beads. These cells had more chance to interact with the beads and received a stronger stimulus than those with a single bead. Nevertheless, no significant difference was observed between the dynamic proliferation/differentiation behaviors of the two categories of cells (Fig. 2).

In summary, all combinations of clonal differentiation/proliferation scenarios were simultaneously observed. Importantly, more than 50% of the clones were mixed with varying proportions of differentiated and undifferentiated cells, indicating that daughter cells within the same clones showed divergent

proliferation and differentiation patterns. This observation suggests that the decision to differentiate is taken by individual cells according to a stochastic rather than deterministic mechanism and independently of the decision to divide or not.

The unequal contribution of clones to the population increases with time as quantified by the Gini coefficient. We observed substantial heterogeneity in the proliferation capacity of the clones. The time to the first division after stimulation varied from 15 to more than 100 hours. The subsequent cell cycles also varied in length, but much less than the lag period. In addition, cell death occurred sporadically during the whole period of the experiments leading to the extinction of more than 20 clones. By the end of the experiment, some micro-wells contained more than 30 cells while the majority less than 12. This means that only a small fraction of the founder cells contributed the majority of the cells to the final population defined as the total number of living cells at the end of the experiment. Such unequal contribution, that we chose to call Pareto-like effect, is frequently observed in nature or in social and economic life. A suitable parameter to measure this inequality is the Gini coefficient. The Gini coefficient varies between 0 and 1. It is equal to 0, if the population is composed by cells derived from each initial cell at equal proportion. On the other hand, the Gini coefficient is equal to 1 if all cells of the population descend from a single founder. In order to evaluate how the unequal contribution of the founder cells is established we calculated the Gini coefficient at every hour. As shown on the Fig. 4, the Gini coefficient started to increase concomitantly with the cell divisions. However, about 70 hours after the beginning of the experiment the increase of the Gini coefficient slowed down and appeared to converge to a value between 0.55 and 0.6 with a final value of 0.581 at the end of the experiment at 100 h. However, as indicated on the Fig. 4, during this final period the population continues to increase steadily. This tendency indicates that the unequal clonal contribution to the final population is essentially established during the first few cell divisions. We have separately calculated the Gini coefficient for the differentiated GFP+ cells only. We found that it increased following a similar dynamics and reached the value of 0.544 by the end of the experiment (Fig. 4). This value is almost identical than that found for the whole population and indicates that the undifferentiated GFP– and differentiated GFP+ cells follow a similar proliferation dynamics.

A stochastic mathematical model can explain the contribution of proliferation and cell death to the Pareto-like effect. Although it is clear that both differential proliferation and death rates may lead to a Pareto-like effect in the cell population, the simple description of the experimental system is insufficient to determine the relative contribution of these two potential sources. In order to address this issue, we designed a computer model, where each parameter can be independently manipulated. The model allows the simulation of the proliferation, death and differentiation of the cells based on a stochastic decision-making mechanism. Each cell can either, divide, die, differentiate or remain unchanged. The probabilities of these cell-fate events are the input parameters; they can be varied. In the basic version, we used probability values determined on the basis of the experimentally measured frequencies. Once the cell fate is chosen, the timing of the event is determined by drawing randomly from the experimentally determined distribution of the corresponding event. Since the first cell cycle was significantly longer than the subsequent cell cycles and the probability of the cell death and differentiation during this period was also significantly different, the probabilities were calculated separately for the first cycle (first event) and all subsequent cycles (subsequent events). We simulated simultaneously 133 cells in every run and we performed 100 independent runs. The Gini coefficient was calculated and plotted against the time in the same way as we did for the experimental values. The evolution of the median curve + /– with the 2 x standard deviation interval shown on the Fig. 5a is very close to the experimentally determined one. The median of the Gini Coefficient is equal to 0.6 at the end of the simulation and is almost equal to the experimentally measured value. In addition, our model was able to reproduce a similar profile of proliferation assay using dye labeling and flow cytometry on bulk culture (Fig. 5b,c). Therefore, the model based on a stochastic fate-determination mechanism can faithfully reproduce the experimentally observed Pareto-like effect when the input probabilities are calculated on the basis of the observed frequencies.

In order to determine the relative contribution of the cell division and death probabilities to the Pareto-like effect, we performed simulations using fixed input parameters. First, we suppressed the cell death. Second, we used a constant first cell cycle length leading to synchronized first division of all cells. Then we tested the effect of constant cell-cycle length in the subsequent cell divisions followed by the synchronization of all cell divisions. Finally, we combined the first and the fourth cases, i.e. all divisions were synchronized and the cells were immortal (see Supplementary Table S1). We performed 100 runs for each case and calculated the average Gini coefficients with their standard deviations at the end of the simulation. Figure 6 show that each of the individual parameters contributes significantly to the Pareto-like effect ($p < 0.001$). When all three parameters were assigned a fixed value, the Gini coefficient dropped to a very low level. This observations show that the stochastic nature of the cell division and death together are the major causes of the unequal contribution of the individual clones to the final population. The remaining weak Pareto-like effect is presumably due to the variable time of cell differentiation that resulted in a modest but measurable difference in the doubling time of individual clones.

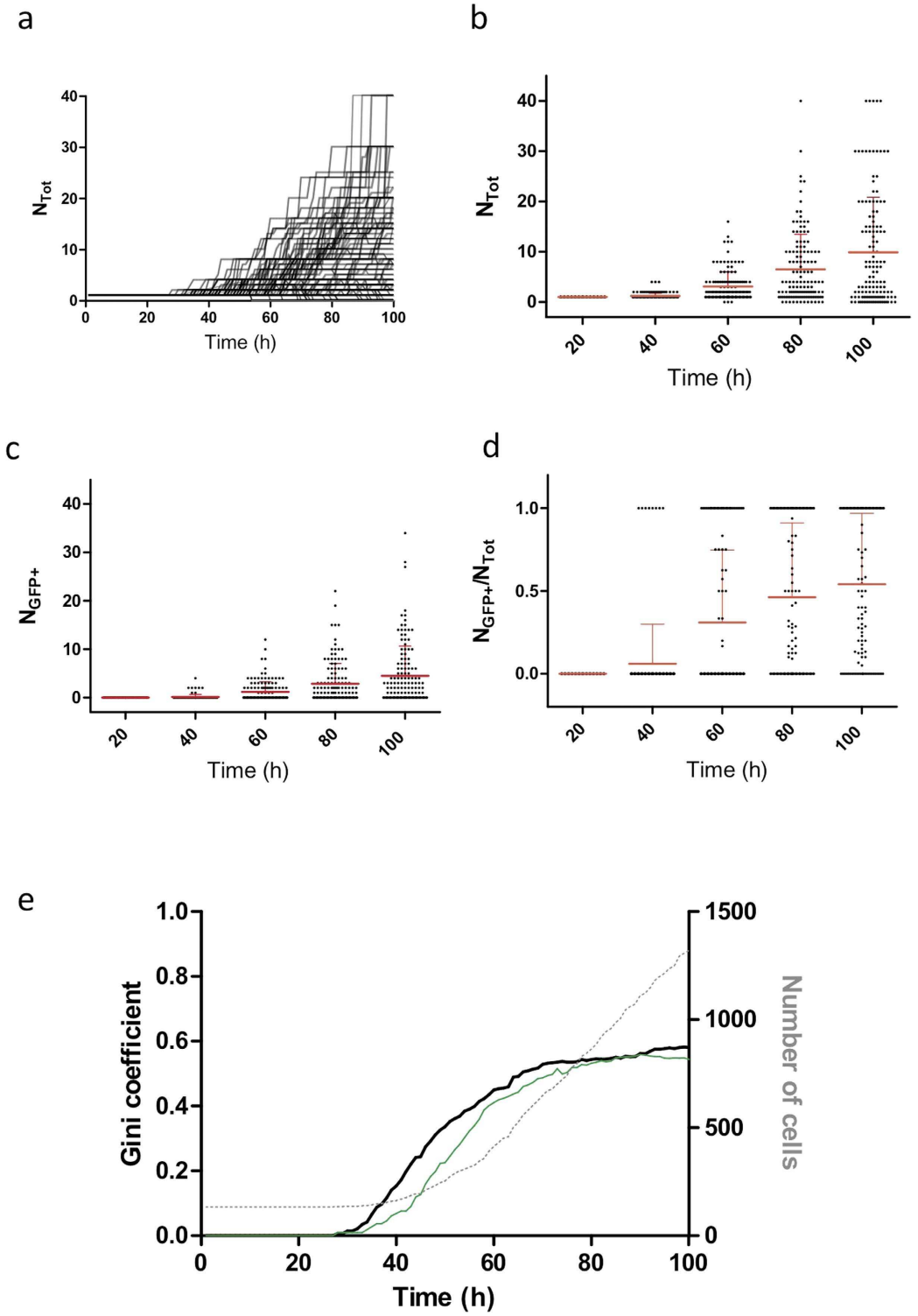


Figure 4. Evolution of the Gini coefficient and the population size in the *in vitro* experiment. (**a**) Individual proliferation curves show highly variable cell division rates. (**b**) Total number of cells per clone at five different time points during the experiment. (**c**) Number of GFP+ cells per clone at the same five different time points during the experiment. (**d**) Ratio of GFP+ to total number of cells per clone at the five time points of the experiment. Each point represents a single clone. The means and standard deviations are indicated with a thick and thin red bands respectively. (**e**) Evolution of the Gini coefficient calculated on the basis of the total cell number (black bold line) or GFP+ cells only (thin green line) in the cell population. Both curves increase rapidly between 30 h and 70 h and remain almost stable during the lasts stage with a final value of 0.581 and 0.544 at the end of the experiment. Interestingly, the total number of cells keeps increasing at the same rate (dashed line).

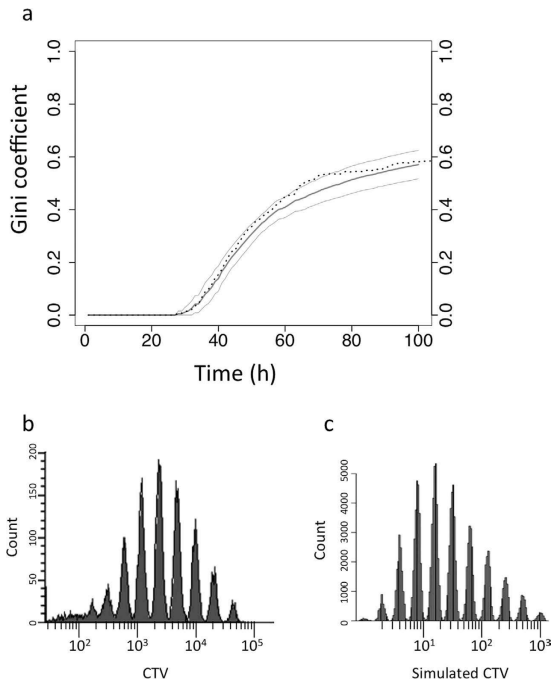


Figure 5. Evolution of the Gini coefficient and cell proliferation in the model. (**a**) The Gini coefficient in the model shows similar evolution (bold line) as in the experiment (dotted line). The +/− twice the standard deviation interval calculated on the basis of 100 simulations is shown by thin lines. (**b**) Cell proliferation profile obtained experimentally using CellTrace Violet (CTV)-fluorescent dye shows heterogeneous proliferation profile. The diagram represents the number of cells (Count) as a function of CTV fluorescence intensity (in arbitrary units). (**c**) The simulated proliferation profile is similar to that obtained experimentally.

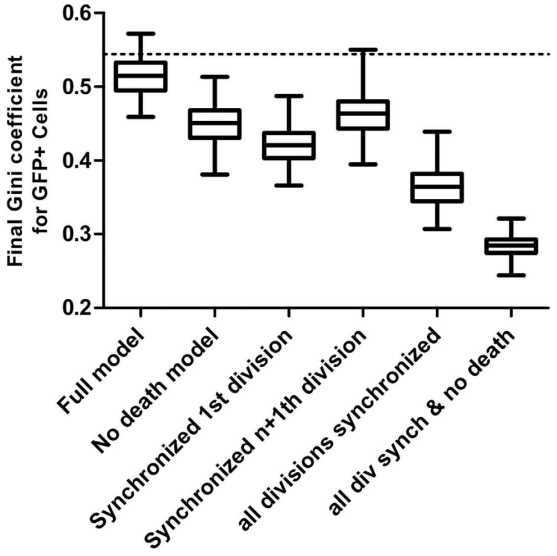


Figure 6. Variation of the Gini coefficient depending on the stochastic or deterministic nature of the input parameters. The average values of the Gini coefficients and their variations calculated under different conditions of simulation at 100 h are shown by boxplots. The horizontal dotted line shows the experimental value. The Kruskal-Wallis test (non-parametric ANOVA) on the 6 groups revealed that the medians varied significantly with a p-value < 0.0001. The Gini coefficient calculated in the full model is significantly higher than in the other models ($p < 0.001$) using the Dunn's multiple comparison test.

Discussion

In this study, we have quantified the proliferation, differentiation and death rates of individual CD4+ T-cell clones after *in vitro* stimulation. We have observed that the high death rate during the first 24 hours following the isolation of the cells is independent of the TCR stimulation by the anti-CD3 and anti-CD28 antibody-coated beads. The initial high rate of cell death in *in vitro* experiments is traditionally attributed either to the overstimulation or the insufficient stimulation of the cells[12]. Our observations show that the cell death is independent from the TCR stimulation. We assume that the major cause of the high mortality may be associated to the cell stress induced by the manipulation and culture. The slow recovery from the stress may also explain why the lag before the first division of the surviving cells varied so widely. As expected, physical contact between the cell and an activating bead was mandatory for proliferation and differentiation. However, we could clearly observe in our time-lapse records that a permanent contact was not necessary. After a lag period of varying length, the cells started to proliferate and differentiate. We observed all possible combinations of proliferation and differentiation events (Fig. 3) that resulted in individual clones with only undifferentiated (GFP–), differentiated (GFP+) and mixed cells. The latter type represented more than the half of the clones. We also observed cells that acquired the GFP+ phenotype without cell division. These finding show that differentiation may occur independently of cell division, and more specifically, without asymmetric division. It is clear from these observations that there is no pre-determined order for division and differentiation; all options are possible and the timing of differentiation is not determined neither by the stimulation nor the proliferation rate.

Although there is no correlation between the cell divisions frequencies and the lag period before the first division, the clones grow at different rates. This is essentially due to the variation in the cell-cycle length of individual cells and sporadic cell death. As a result, 100 hours after the beginning of the experiment some clones are composed of a large number of cells (>>40) while others become extinct or are represented by a few cells only. The majority of the final population is composed of cells from a small number of clones. We used the Gini coefficient to describe the unequal clonal contribution to the whole population. The value of the Gini coefficient is equal to zero if all individual cells contribute equally to the clonal composition of the final population. The Gini coefficient is equal to 1 if the final composition is made up of a single clone. In our experiment, we calculated a Gini coefficient equal to 0.6, which means that a small number of clones outgrew the others and dominate the final population. Indeed, the evolution of the Gini coefficient over time describes well the dynamics of the population. The Gini coefficient starts to increase sharply after the lag period when the first cells start to divide and reaches a plateau around the final value of 0.6. The clonal structure of the final population is established during this period and in spite of the continuous growth of the population size it remains essentially unchanged at later stages. There are several factors that could contribute to the establishment of the Pareto-like clone size distribution. First, the time between the stimulation to the first division varies widely between cells. The rate of cell divisions is also variable. This is a well-known phenomenon[13] and one would expect that positive and negative deviations would equilibrate each other conferring a stable average proliferation rate. Presumably, this happens at later stages when the number of cells in the clones is large enough to stabilize the average proliferation rate. However, at earlier stages the number of cells per clone is small and the clone size is very sensitive to fluctuations of division rates and fortuitous cell death. After the first wave during the first 24 hours, cell death occurs sporadically at all stages and in all clone types. Some clones even become extinct after the first few divisions. This is the third factor that can impact the clone composition of the population at the early stages of proliferation. Overall, these findings suggest that the cell-fate decisions to divide and differentiate are taken separately by individual cells using stochastic mechanisms. The individual random decisions lead in a few days to a biased clonal composition of the whole cell population characterized by the dominance of a few clones.

In order to estimate the relative contribution of proliferation/differentiation/death to the Pareto-like distribution we designed a simple mathematical model based on stochasticity. The cells were assigned a separate probability to accomplish the first and subsequent divisions or to die. The actual values of the probabilities were calculated on the basis of the frequencies observed during the experiment. Differentiation also occurred as a chance event with a probability calculated on the basis of the observed frequencies and distributions. As an output, we calculated the evolution of the Gini coefficient during the population growth. Comparing the results of the simulations and the observations, we can see that a stochastic model of cell division and death can account for the Pareto-like distribution of the clonal composition. The model allowed us to determine the relative contribution of the first division timing, division rate and cell death by setting them respectively to fixed values or zero. The simulations show that all three factors contribute to the Pareto-like effect and none of them is determinant.

In conclusion, our experimental observations and the results of the computer modeling indicate that no complicated deterministic mechanisms are needed to explain the proliferation and differentiation dynamics of naive CD4+ cells *in vitro*. A simple stochastic mechanism is fully sufficient. Further studies are required to determine if the Pareto-like effect found in the clonal composition plays a role during cell proliferation and differentiation *in vivo*. In *in vitro* studies however, this phenomenon can introduce a significant bias in our understanding of biological mechanisms. Particular characteristics of a limited number of clones selected by the culture conditions and transmitted by epigenetic mechanisms through cell divisions can hide and be confounded with cellular mechanisms induced by the experimental treatment.

Methods

Mice. All procedures on animals were performed in accordance with the European guidelines for the protection of vertebrate animals used for experimental purposes (Directive 2010/63/EU of 22 Septembre 2010). All experimental protocols were approved by the ethics committee of Genethon. Transgenic mice with fluorescent Treg cells (Tg(TcraH-Y,TcrbH-Y)1Pas, Ptprc,Foxp3, Rag2) were generated and housed in our facility by crossing Foxp3-GFP-KI mice (B6.Cg-Foxp3tm1Mal/J) with Marilyn mice (B6.129-Ptprca Rag2tm1Fwa Tg(TcraH-Y,TcrbH-Y)1Pas/Pas)[7]. Animals were used in experiments at between 6 and 9 weeks of age.

Cells and activation. CD4 + T-cells were harvested from spleen of Foxp3-GFP$^{+/+}$ transgenic mice and sorted using magnetic beads as described[7]. CD4 + T-cells were activated with activation beads (Gibco – Life Technologies, Dynabeads Mouse T-Activator CD3/CD28) coated with anti-CD3 and anti-CD28 antibodies.

Micro-grid cell culture. A polydimethylsiloxane (PDMS) micro-grid array (Microsurfaces, Australia) of 4096 micro-wells (50 μm-sided) was placed in a MatTek dish (MatTek Corporation, Slovak Republic). It was covered by complete Roswell Park Memorial Institute (RPMI) 1640 Medium Glutamax plus (RPMI; Gibco-Invitrogen, France) supplemented with 10% fetal calf serum (FCS; PAA Laboratories, Germany), 100 μg/mL penicillin, and 100 μg/mL streptomycin (both from Eurobio, France), 50 μM 2-mercaptoethanol (Gibco-Invitrogen), and 30 U/mL IL-2 (Prometeus Laboratories) and 20 ng/mL TGF-β (Miltenyi Biotec SAS, France). The suspension of activating beads (see above) was added at a concentration that allowed the highest possible number of micro-wells with a single bead. Then, the suspension of 1.5×10^4 cells was added, again at a concentration to get high number of wells with a single cell. As a result of this procedure, a large number of micro-wells with a single bead and a single cell were obtained together with micro-wells with cells without beads or two or more beads. We define a clonal population as being the cells in a micro-well that had one cell at the beginning of the experiment. The cells were maintained at 37 °C, 5% CO_2.

Proliferation assay. The CD4 + T cells were labeled with 1 μM of CellTrace Violet (CTV) (Life Technologies) and cells were analyzed by flow cytometry (LSR II – BD Biosciences, Le Pont de Claix, France) as described by Neildez *et al.*[7].

Time-lapse microscopy. Time-lapse acquisitions were performed with the Biostation IM time-lapse microscope (Nikon Instruments, Europe). 99 field positions were recorded covering 20 micro-wells each. Images were acquired every hour for 100 hours in phase contrast, GFP fluorescence channel (505–550 nm bandpass filter) and Rhodamine fluorescence channel (560–605 nm bandpass filter) at 20 X magnitude.

Image analysis. Images were analyzed using ImageJ 1.47 g 64-bits (Rasband, W.S., ImageJ, U.S. National Institutes of Health, Bethesda, Maryland, USA, http://imagej.nih.gov/ij/, 1997–2014.). The cells were counted manually at each step of the time-lapse to determine the total number of cells, the number of differentiated cells (GFP channel) and the number of dead cells (identified on the basis of their auto-fluorescence in the Rhodamine channel). Micro-wells containing a single cell and at least one activation bead were considered in the analysis; micro-wells containing no beads and one single cell were used as controls.

Experimental distribution and frequencies. Three types of events were considered: cell death, cell division and cell differentiation. For each micro-well a first event was determined and its time associated. By compiling all the micro-wells, 3 frequencies and 3 time distributions were calculated for the first event. The same approach was used to determined and calculated frequencies and time distributions of the second event according to the type of the cell, GFP– or GFP+ (See Supplementary Fig. S1).

Gini coefficient calculations. The Gini coefficient enables the quantification of the Pareto-effect. It is calculated on the basis of the cell numbers in each clone ($x_1, x_2, \ldots x_N$) as a sum of all combinations of absolute differences between the cell numbers in two clones, $|x_i - x_j|$, normalized by N and μ.

μ is the arithmetic mean :

$$\sum_{i=1}^{N} \frac{x_i}{N} \tag{1}$$

The Gini coefficient is given using the following formula [11]:

$$\frac{\sum_{i=1}^{N}\sum_{j=1}^{N}\left|x_i - x_j\right|}{2N^2\mu} \tag{2}$$

The Gini coefficient ranges from a minimum value of 0, when all clones contribute equally to the final cell population, to a theoretical maximum of 1 when the entire population is derived from a single founder cell and all other clones die out.

The Gini coefficient was calculated separately for GFP positive and GFP negative clones. We considered a clone as GFP positive if there was at least one GFP positive cell at the end of the experiment. For the sake of calculability, we considered that a GFP positive cell was initially in the micro-well. This consideration is only a mathematical artifice to avoid having a zero vector in the calculation of the Gini coefficient.

***In sillico* model and simulations.** Our model simulates individual cells that are able to divide, differentiate, die or remain unchanged. The cell fate is determined at each step using a decision tree. The model is divided into two parts: the first cell cycle of the founder cell and all subsequent cell cycles. In the first part, the model simulates the fate of the founder cell. We considered three possible fate decisions: division, differentiation or remaining unchanged till the end of the simulation. We calculated the probability of the first event on the basis of the observed experimental frequencies. Once the event is chosen for a given founder cell, the time to this particular event is assigned on the basis of its experimentally observed cumulative time distribution. We used the same process for every cell derived from the division of a founder cell except that cell death was added as a fourth possible event (See Supplementary Fig. S2). The probabilities and the time to the event were calculated on the basis of the cumulative distributions obtained experimentally on the cells after the first division of the founders.

In the model, each cell is considered as a vector containing parameters concerning its fate:

$$Cell_k = (Event,\ t_{Event},\ Gen,\ Age,\ Diff,\ Death,\ PD) \tag{3}$$

The parameters are:

Event	is the event the cell is attributed
t_{Event}	is the relative time when the Event is going to occur (in hours)
Gen	is the generation of a given cell (e.g. generation for the single mother cell is 1, and is 2 for the daughters cells)
Age	is the age of the cell in hour. This parameter is incremented at each step and set to 0 at each division.
Diff	is a binary parameter: 1 if the cell has differentiated, 0 if not.
Death	is a binary parameter: 1 if the cell is dead, 0 if not.
PD	is a parameter that corresponds to a virtual proliferation dye (PD) labelling. It is set to 1000 + v where v is a value drawn in a normal distribution N(0,50). This value is divided by two at each step.

Each time point is associated to a matrix M that contains row vectors of the cells in the clone. Each step in the simulations represents an increment of 1 h. At each step, the matrix M is updated according to the previous state:

$$M_{n(t),7}(t) = \begin{bmatrix} Event_1 & \dots & PD_1 \\ \dots & \dots & \dots \\ Event_n & \dots & PD_n \end{bmatrix} \tag{4}$$

t is the time in hours.

The number of columns is 7 and the number of rows *n(t)* depends on the number of steps and on divisions.

The output is the number of total number living cells N_{tot} and the number of differentiated and living cells N_{diff} .

We simulated 133 clonal populations for 100 steps (100 hours) in each run and we performed 100 independent runs. We calculated the corresponding Gini coefficient either for all clones or only for clones containing differentiated cells.

The model was implemented using custom-made R (R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/). Scripts are available on request (stockho@genethon.fr).

References

1. Pollizzi, K. N. & Powell, J. D. Integrating canonical and metabolic signalling programmes in the regulation of T cell responses. *Nat Rev Immunol* **14,** 435–446 (2014).
2. Till, J. E., Mcculloch, E. A. & Siminovitch, L. A stochastic model of stem cell proliferation, based on the growth of spleen colony-forming cells. *Proc. Natl. Acad. Sci. USA* **51,** 29–36 (1964).
3. Reiner, S. L. & Adams, W. C. Lymphocyte fate specification as a deterministic but highly plastic process. *Nat. Rev. Immunol.* **14,** 699–704 (2014).
4. Hodgkin, P. D., Dowling, M. R. & Duffy, K. R. Why the immune system takes its chances with randomness. *Nat. Rev. Immunol.* **14,** 711 (2014).
5. Stockholm, D. *et al.* The origin of phenotypic heterogeneity in a clonal cell population *in vitro*. *PLoS ONE* **2,** e394 (2007).
6. Stockholm, D. *et al.* Bistable cell fate specification as a result of stochastic fluctuations and collective spatial cell behaviour. *PLoS ONE* **5,** e14441 (2010).
7. Neildez-Nguyen, T. M. A. *et al.* Hypoxic culture conditions enhance the generation of regulatory T cells. *Immunology*, doi: 10.1111/imm.12388 (2014).
8. Corre, G. *et al.* Stochastic fluctuations and distributed control of gene expression impact cellular memory. *PLoS ONE* **9,** e115574 (2014).
9. Fantini, M. C., Dominitzki, S., Rizzo, A., Neurath, M. F. & Becker, C. *In vitro* generation of CD4+ CD25+ regulatory cells from murine naive T cells. *Nat Protoc* **2,** 1789–1794 (2007).
10. Duffy, K. R. *et al.* Activation-induced B cell fates are selected by intracellular stochastic competition. *Science* **335,** 338–341 (2012).
11. Zaretsky, I. *et al.* Monitoring the dynamics of primary T cell activation and differentiation using long term live cell imaging in microwell arrays. *Lab Chip* **12,** 5007–5015 (2012).
12. Marchingo, J. M. *et al.* T cell signaling. Antigen affinity, costimulation, and cytokine inputs sum linearly to amplify T cell expansion. *Science* **346,** 1123–1127 (2014).
13. Golubev, A. Exponentially modified Gaussian (EMG) relevance to distributions related to cell proliferation and differentiation. *J. Theor. Biol.* **262,** 257–266 (2010).

Acknowledgements

This work was supported by funds from the French Muscular Dystrophy Association (AFM), ANR Stochagene n° BSV6 014 02, the Fond d'Innovation pour la Recherche de l'Université d'Evry. Equipment funds were from Genopole Research-Evry, University Evry Val d'Essonne, Conseil General de l'Essonne and from Region Ile de France. A.M. is a recipient of a PhD scholarship from E.P.H.E.

Author Contributions

J.C. and A.M. performed experiments and analyzed data. F.O. performed the spleen extraction and CD4+cells purification. A.A.-C. performed preliminary experiment and was involved in the concept. T.M.A.N.-N. performed preliminary results and proliferation assays. A.P., J.C. and D.S. wrote the main manuscript. D.S. supervised the project and edited the manuscript. All authors discussed the results and reviewed the manuscript.

Additional Information

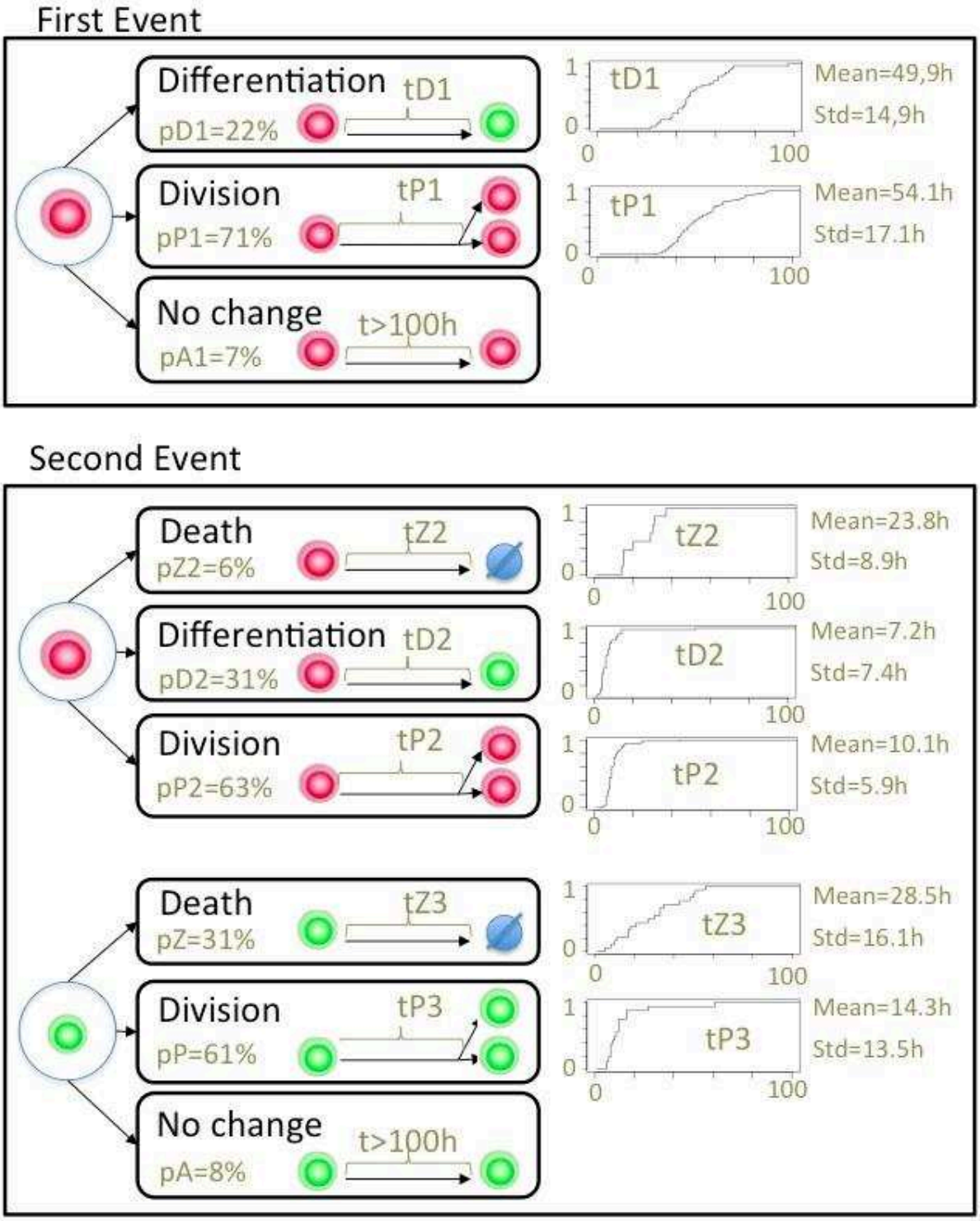
Supplementary information accompanies this paper at http://www.nature.com/srep

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Cosette, J. *et al.* Single Cell Dynamics Causes Pareto-Like Effect in Stimulated T Cell Populations. *Sci. Rep.* **5**, 17756; doi: 10.1038/srep17756 (2015).

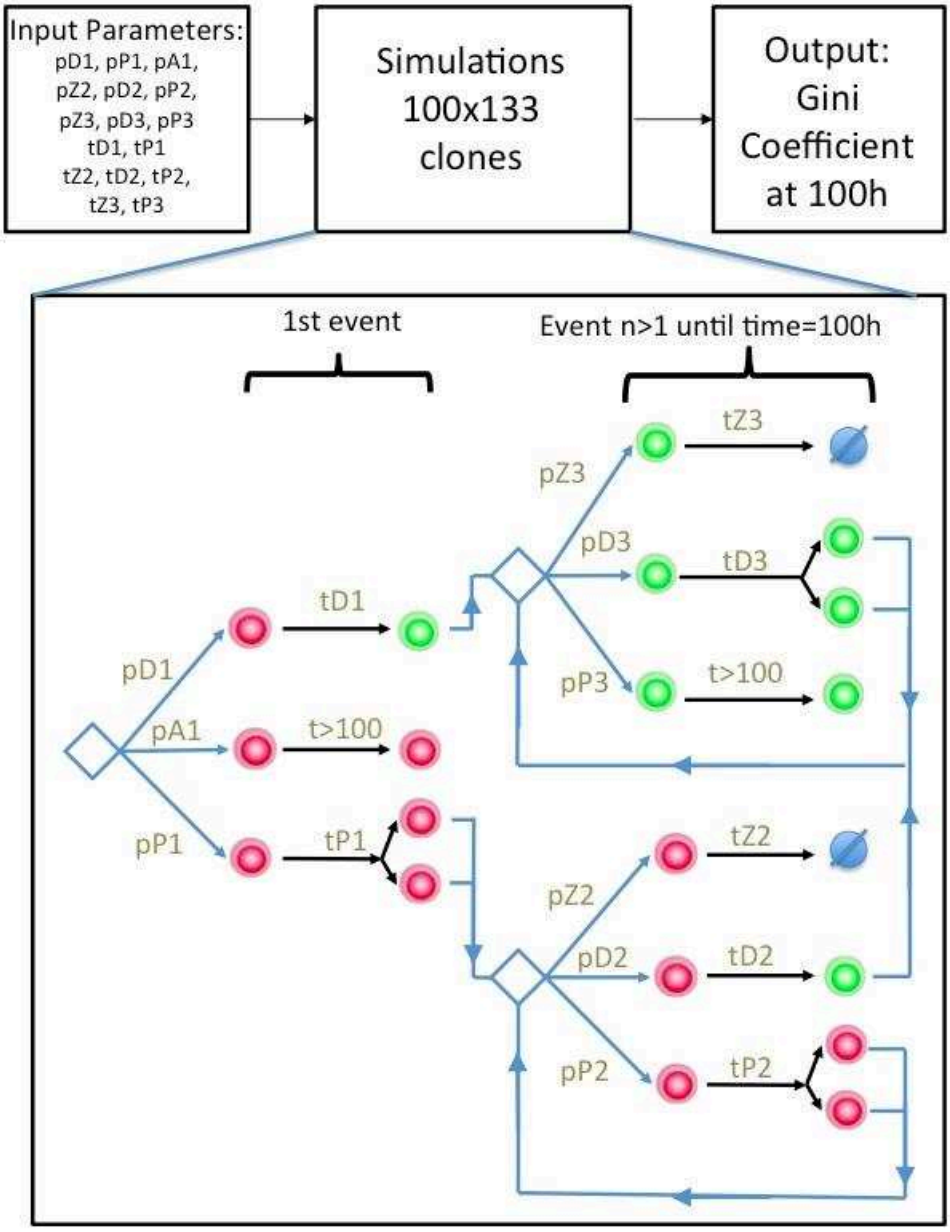
This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/

SUPPLEMENTAL DATA



Supplementary Figure S1. **Experimental frequencies and time distributions**. Three types of first events were considered: differentiation (D), cell division (P) and "no change" (A). If a cell died before division or differentiation, it was excluded from the calculation of the frequencies. In the second and subsequent events "death" (Z) was included. The probabilities

(pD1, pP1, pA1 for the first events, pZ2, pD2, pP2 for the second event of non-differentiated and pZ3, pD3, pP3 for the second event of differentiated cells) were obtained from the frequencies calculated on the 133 clonal populations. A cumulative time distribution (tD1, tP1, tZ2, tD2, tP2, tZ3, tP3) is presented with the mean and standard deviation on the right.



Supplementary Figure S2. **The structure of the model** and the flow chart of the simulation of a clonal population is represented. Simulations rely on 9 probabilities (pD1, pP1, pA1, pZ2, pD2, pP2, pZ3, pD3, pP3) and 7 time distributions (tD1, tP1, tZ2, tD2, tP2, tZ3, tP3). Simulation starts with a non-differentiated cell. The first event (differentiation, division or "no

change" until the end of the simulation at 100h) is chosen randomly according to the probabilities pD1, pP1 and pA1 (pD1+pP1+pA1=1). A new cell is created if a division occurs. The fate of each cell is then calculated according to the probabilities of the "second event" as shown on the right part of the diagram. A simulation step corresponds to 1h. In each run, 133 cell clones are simulated and the Gini coefficient is calculated for time=100h.

	Inputs conditions					
Name in figure	**Prolif. Proba. pP1**	**Prolif. times tP1**	**Proliferation Proba. pP2 and pP3**	**Proliferation times tP2 and pP3**	**Cell Death Proba. pZ2 and pZ3**	**Cell death times tZ2 and tZ3**
Full model	exp.	exp.	exp.	exp.	exp.	exp.
No death model	exp.	exp.	exp.	exp.	**0**	**NA**
Synchronized 1st division	exp.	**Const.**	exp.	exp.	exp.	exp.
Synchronized n+1th division	exp.	exp.	exp.	**Const.**	exp.	exp.
All divisions synchronized	exp.	**Const.**	exp.	**Const.**	exp.	exp.
All div synch & no death	exp.	**Const.**	exp.	**Const.**	**0**	**NA**

Supplementary Table S1. Combinations of input parameters used for the simulations. Initially, the simulations were performed using probabilities calculated on the basis of experimental observations (exp). In other the other variants of the model, one or several parameters were fixed to a constant value. (Const.). When the probability of cell death is fixed to "0", the time parameter tZ becomes irrelevant, thus not applicable (NA).

Lors des analyses des clones, nous avons constaté une division cellulaire et un taux de mortalité hautement hétérogènes dans les cellules CD4+ : c'est ce que l'on a appelé l'effet Pareto ; peu de clones du départ ont contribué à la population finale. Cet effet est souvent oublié par les biologistes, pensant que la prolifération des cellules est plutôt homogène. Il suffit pourtant d'avoir un clone dominant pour obtenir et définir une population finale déséquilibrée en regard de la population initiale. L'effet Pareto n'est pas spécifique aux LT et peut aussi exister dans d'autres systèmes de différenciation, plus complexes, comme les CSH.

Pour quantifier cette hétérogénéité, nous avons utilisé le coefficient de Gini qui varie entre 0 et 1. Si le coefficient de Gini est égal à 0 alors les clones ont contribué de façons égales à la population cellulaire finale avec des temps de division, de différenciation et de mort cellulaire égaux pour toutes les cellules. A contrario, si le coefficient de Gini est égal à 1 alors les clones ont contribué de façons inégales à la population cellulaire finale et toutes les cellules de la population finale descendent d'un seul fondateur. En le calculant toutes les heures, nous avons vu que le coefficient de Gini augmentait de façon exponentielle durant les 70 premières heures de l'expérience (une fois que les cellules commençaient à se diviser) puis se stagnait et convergeait vers une valeur comprise entre 0,55 et 0,6 alors que la population cellulaire, elle, continuait d'augmenter régulièrement. Nous en avons conclu que la contribution inégale des clones à la population finale était essentiellement établie au cours des premières divisions cellulaires.

Nous avons aussi observé toutes les combinaisons possibles de prolifération et de différenciation à la fin de l'expérience dans des proportions différentes :

- des clones uniquement composés de cellules différenciées, GFP+ (19,6% des cas) ;
- des clones uniquement composés de cellules non différenciées, GFP- (15,8% des cas) ;
- des clones mixtes avec les deux types de phénotypes (54,9% des cas) ;
- des clones qui n'ont pas proliféré et/ou différencié (9,7% des cas).

Nous avons aussi observé qu'il n'y avait pas d'ordre prédéterminé pour la division et la différenciation de ces cellules : toutes les options étaient possibles (d'abord la prolifération puis la différenciation et inversement). De plus, le temps de différenciation n'est déterminé ni par la stimulation, ni par le taux de prolifération. Ces résultats ont bien montré que la différenciation cellulaire pouvait se produire indépendamment de la division cellulaire, et plus particulièrement, sans division asymétrique.

Enfin, pour pouvoir au mieux comprendre la prise de décision de ces cellules, une simulation du processus de prolifération et de différenciation a été faite en utilisant un modèle mathématique simple de prise de décision stochastique, basé sur les paramètres observés expérimentalement. Chaque cellule « pioche » un temps de prolifération, de différenciation et de mort dans la distribution expérimentale. On calcule ensuite le coefficient de Gini de la population générée. L'intérêt de cette simulation est que nous pouvons contrôler les différents phénomènes et imposer une synchronisation des temps de prolifération et de différenciation ou d'enlever la mort cellulaire du système. Grâce à cela, l'influence de chaque paramètre sur le coefficient de Gini, et donc sur l'hétérogénéité de la population finale, peut être quantifiée. En synchronisant les paramètres, on arrive toujours à un coefficient de Gini plus petit que la valeur expérimentale alors que l'approche stochastique donne des valeurs similaires. Ces observations prouvent que la nature stochastique de la division et de la mort cellulaire expliquent facilement la contribution inégale des clones individuels à la population finale et qu'un scénario stochastique est totalement compatible avec les observations expérimentales.
Dans l'ensemble, ces résultats suggèrent que les décisions du destin cellulaire pour la division et la différenciation cellulaire sont prises séparément par des cellules individuelles à travers des mécanismes stochastiques.

Grâce à cette première étude sur les LT CD4+ réalisée au cours de mon stage de fin d'études, j'ai pu acquérir des connaissances mais aussi des techniques d'études sur cellules uniques qui m'ont permises de mieux me préparer pour l'étude du comportement des cellules CD34+. En effet, lors de cette première étude, j'ai pu voir et comprendre l'importance d'étudier les cellules à l'échelle individuelle tant au niveau de leur prolifération qu'au niveau de leur différenciation et me rendre compte qu'il n'y avait pas de règles prédéfinies et que les cellules étaient toutes différentes.
L'étude des LT a été, pour moi, un premier modèle « simple » de différenciation avec une seule voie de différenciation : LT en LTreg. Cela m'a permis de mieux appréhender un modèle de différenciation plus complexe que sont les cellules CD34+ avec de multiple choix dans les voies de lignage. Pour caractériser la différenciation des cellules CD34+ humaines issues de sang de cordon, nous avons analysé la dynamique morphologique de ces cellules au cours du temps reliée à des analyses transcriptomiques et protéomiques à l'échelle de la cellule unique.

PARTIE 2

Modèle plus complexe de différenciation : les cellules CD34+ humaines issues de sang de cordon

I. Etude de la dynamique morphologique des cellules CD34+ issues de sang de cordon à l'échelle de la cellule unique

La différenciation des CSH est un processus largement étudié en biologie dont la compréhension du choix du devenir de la cellule n'est pas encore résolue. Afin de mieux comprendre cette différenciation, nous avons étudié et analysé la dynamique morphologique des cellules CD34+ issues de sang de cordon pendant quatre jours après stimulation. En effet, la différenciation cellulaire est largement étudiée au niveau moléculaire mais beaucoup moins au niveau morphologique alors qu'elle peut avoir aussi toute son importance dans le devenir du choix de la cellule. On peut même se demander si la différenciation cellulaire est un processus dynamique. Pour répondre à cette question, au cours de ma thèse, j'ai étudié le comportement morphologique de ces cellules et de leurs descendantes au cours du temps à l'échelle de la cellule unique sans marquage spécifique. En effet, les marqueurs comme des anticorps peuvent jouer un rôle voire changer le comportement des cellules (Gilner et al., 2007; Schneider et al., 1997). Par conséquent, nous n'avons pas marqué les cellules car nous ne pas voulions pas ajouter un autre biais dans l'analyse du comportement des cellules CD34+. Cette analyse nécessite donc des outils combinant une résolution précise dans le temps et une forte puissance statistique.

Pour étudier cette dynamique morphologique, nous avons combiné les données obtenues en microscopie en TL et en CFI. Avec la microscopie en TL, nous avons pu suivre pendant quatre jours la dynamique morphologique de clones de cellules et de leurs descendantes grâce à des images prises toutes les minutes. Le TL étant une technologie statistiquement peu puissante, nous avons combiné ces résultats avec les données de CFI réalisées tous les jours pendant quatre jours. La CFI nous a permis d'obtenir de nombreuses images (en nombre suffisamment grand pour être statistiquement significatif) de cellules CD34+ individuelles dans une population à un instant donné.

ARTICLE 2

"Combination of imaging flow cytometry and time-lapse microscopy for the study of label-free morphology dynamics of hematopoietic cells"

Jérémie Cosette, Alice Moussy, András Páldi and Daniel Stockholm

Cytometry Part A. 91A: 254-260 (Mars 2017)
DOI: 10.1002/cyto.a.23064

BRIEF REPORT

Cytometry PART A
Journal of the International Society for Advancement of Cytometry

MIFlowCyt

Combination of Imaging Flow Cytometry and Time-Lapse Microscopy for the Study of Label-Free Morphology Dynamics of Hematopoietic Cells

Jérémie Cosette,[1] Alice Moussy,[2] Andras Paldi,[2] Daniel Stockholm[2*]

[1]Genethon, Evry, 91002, France

[2]Ecole Pratique des Hautes Etudes, PSL Research University, UMRS_951, Genethon, Evry, 91002, France

Received 12 September 2016; Revised 6 December 2016; Accepted 15 January 2017

Additional Supporting Information may be found in the online version of this article.

*Correspondence to: Daniel Stockholm; Ecole Pratique des Hautes Etudes, PSL Research University, UMRS_951 INTEGRARE Research Unit, Genethon, 1bis rue de l'Internationale, Evry 91002, France. E-mail: stockho@genethon.fr

Published online 1 March 2017 in Wiley Online Library (wileyonlinelibrary.com)

DOI: 10.1002/cyto.a.23064

© 2017 International Society for Advancement of Cytometry

• **Abstract**

Cell differentiation is a longitudinal and dynamic process. Studying and quantifying such a process require tools combining precise time resolution and statistical power. Imaging flow cytometry (IFC) provides statistically significant number of microscopy images of individual cells in a sample at a given time point. Time-lapse microscopy (TLM) is the method of choice for studying the dynamics of cell processes at a high temporal, but low statistical resolution. In this work, we show that the dynamic changes of cord-blood derived CD34+ cells in response to cytokine stimulation can be successfully studied, in a label-free way, by the combination of the IFCs statistical power and the TLM's high time resolution. Cell morphology phenotypes were quantified through roundness and surface area, measured both in IFC and with a home-made segmentation algorithm in TLM. Two distinct morphologies—polarized and round—were observed in cord-blood derived CD34+. We show that some cells have the ability to fluctuate between these morphologies, suggesting that the apparent stable composition of round and polarized cells may actually represent a dynamic equilibrium. This example demonstrates that the different resolutions and modalities of IFC and TLM are complementary and allow the study of complex dynamic biological processes. © 2017 International Society for Advancement of Cytometry

• **Key terms**

cytometry; imaging; morphology; time-lapse; imaging flow cytometry

Introduction

Cell differentiation and development is by essence a longitudinal process involving dynamic features. Quantitative tools are required to monitor these features. Most phenotype analyses rely on fluorescent labelling of surface markers (1). However engagement of surface markers has been shown to influence cell fate (2). Morphology, as observed by light microscopy is a label-free and easily accessible phenotype (3). Moreover it is suitable for long-term follow-up.

Different technologies can be used for quantitative high-throughput morphology analysis. Although flow cytometry is the gold standard for high-throughput phenotypic analyses (1), without the use of fluorescent labeling, it is limited to indirect shape parameters (Side Scatter-SSC and Forward Scatter-FSC). Imaging flow cytometers (IFC) overcome this limitation by providing light microscopy images of individual cells and direct measurement of morphology parameters while keeping the statistical power of flow cytometry (4–6). Still, these devices produce snapshots of individual cell states within a population and cannot measure the longitudinal phenotypic dynamics of individual cells. Time-lapse microscopy (TLM) is recognized as the method of choice for quantifying single cell phenotypic dynamics (7–9). It is

ISAC International Society for Advancement of Cytometry

obviously statistically less powerful than IFC but way more time resolved. We chose the combination of IFC and TLM for a complete study of dynamic morphology integrating population and single cell analyses.

We applied the combination of IFC and TLM to follow the morphological phenotype of human cord blood-derived CD34+ cells. These cells are widely used in clinics in the context of bone marrow transplantation. Although differentiation of these cells has been extensively studied (10), the process of initial cell fate decision is still debated and remains unclear. Morphology is a good indicator of cell state and lineage commitment as it has been described that CD34+ cells in culture exhibit differential phenotypes in terms of morphology (11,12). Stability or fluctuations of these morphology phenotypes are poorly described. High-resolution dynamics studies could fill the gap and give insights on the process.

In this article, we describe quantitative methods to explore dynamics of the morphological phenotypic space at different degrees of resolution, i.e., population or single cell. Using IFC and TLM, we monitored cell shape parameters of label-free human cord blood CD34+ cells revealing highly dynamic phenotypes with two stable states.

Materials and Methods

Cell Isolation and Culture

Human umbilical cord blood (UCB) was collected from placentas and/or umbilical cords obtained from Etablissement Français du Sang (EFS), Saint Louis Hospital, France or from Centre Hospitalier Sud Francilien, Evry, France in accordance with international ethical principles and French national law (bioethics law n°2011–814) under declaration N°DC-201-1655 to the French Ministry of Research and Higher Studies. Human CD34+ cells were isolated using autoMACS pro (Miltenyi Biotec, France) by immunomagnetic selection within the mononuclear cell fraction of UCB samples. Cells were cryopreserved in Cryostor (StemCell, France) and stored in liquid nitrogen.

Immediately after thawing, cells were put in an activation medium, whose base is X-VIVO medium (Lonza) supplemented with 100 U/ml penicillin, 100 μg/ml streptomycin (Gibco, Thermo Scientific), and that contains activation cytokines: 50 ng/ml h-FLT3, 25 ng/ml h-SCF, 25 ng/ml h-TPO, and 10 ng/ml h-IL3 (Miltenyi Biotec, Paris, France). Cells were cultured in this medium for the whole study in a humidified 5% CO_2 atmosphere at a temperature of 37°C.

Imaging Flow Cytometry

Acquisitions were performed everyday from Day 0 to Day 3. Day 0 corresponds to the time point immediately after thawing. 4.10^5 cells were thawed: 1.10^5 cells were preserved for Day 0 analysis; the remaining 3.10^5 cells were equally distributed in 3 wells—respectively corresponding to Day 1, Day 2 and Day 3—containing 200 μL of culture medium. We used glutaraldehyde as a fixative known for preserving cell shape (13). For each time point, 100 μL of 3% glutaraldehyde solution (1% final) was added directly in the cell medium to avoid any perturbation in cell shape. After 10 min, cells were PBS-washed twice and resuspended in 25 μL PBS. Images of cells were then acquired with the Image Stream MKII® with the 40× magnitude objective using only the bright field channel and the SSC channel (fluorescence lasers off) and using INSPIRE® software.

Analysis was done using IDEAS® software. Cells were gated using the Area versus Aspect Ratio scatterplot. The Aspect Ratio is the ratio between the minor axis and the major axis of the object image fitting ellipse; it is commonly used to identify cells, beads, doublets, etc. Cells related features (after cells gating) were calculated using a custom cell mask which is the 2-pixel eroded MC mask (mask combined—preexisting in the software). MC mask is the result of the logical OR operation among all channel masks. In our case for instance, MC mask = bright field mask OR SSC mask. Our custom cell mask is closer to cell membrane and allows better comparison of the measured data between the different modalities (Supporting Information Fig. S1). *k*-means algorithm (14) was applied using the *k*-means function of R software (R Core Team, 2016). R: A language and environment for statistical - computing. R Foundation for Statistical Computing, Vienna, Austria——URL; https://www.R-project.org/) on days 0, 1, 2, and 3 data.

Time-Lapse Microscopy

Five thousands cells were loaded onto a polydimethylsiloxane (PDMS) microgrid array (microsurfaces—Australia) containing 1024 125 μm-sided square micro-wells in cell culture medium (15,16). The micro-grid was placed in a Biostation (Nikon) at 37°C 5% CO_2. Only micro-wells containing a single cell were imaged at a rate of one frame per minute during 72 h. Images were acquired with a 20× magnitude phase contrast objective.

Image Processing of Time-Lapse Sequences

All images were analyzed using ImageJ 1.47 g 64-bits (Rasband, W.S., ImageJ, U.S. National Institutes of Health, Bethesda, Maryland, USA; http://imagej.nih.gov/ij/, 1997–2016). Images of micro-grid were cropped to isolate single microwells containing initially a single cell. Images were segmented using custom made scripts based on the following method. First, to remove any immobile structure in the micro-well, a median projection is performed using the Image/Stack/Z-project/median function. This projection results in an image containing the immobile structures in the micro-well. All the images from the stack (i.e., the time-lapsed sequence), are divided by this projection image using the image calculator feature resulting in 32-bit float stack. To remove edge structures of the micro-well, a Fast Fourier Transform (FFT) filtering using a custom filter that removes high frequency signals in the image in the horizontal and vertical directions is applied. To enhance contrast, the rolling ball (radius of 10 pixels) algorithm described by Sternberg (17) is used on the 8 bit converted stack. To eliminate sporadic remaining background noise, a 2-pixels radius median filter was applied. Finally, in order to obtain the final segmentation of the cells (i.e., a binary image with only the cells in black and the background in white), stack was thresholded using

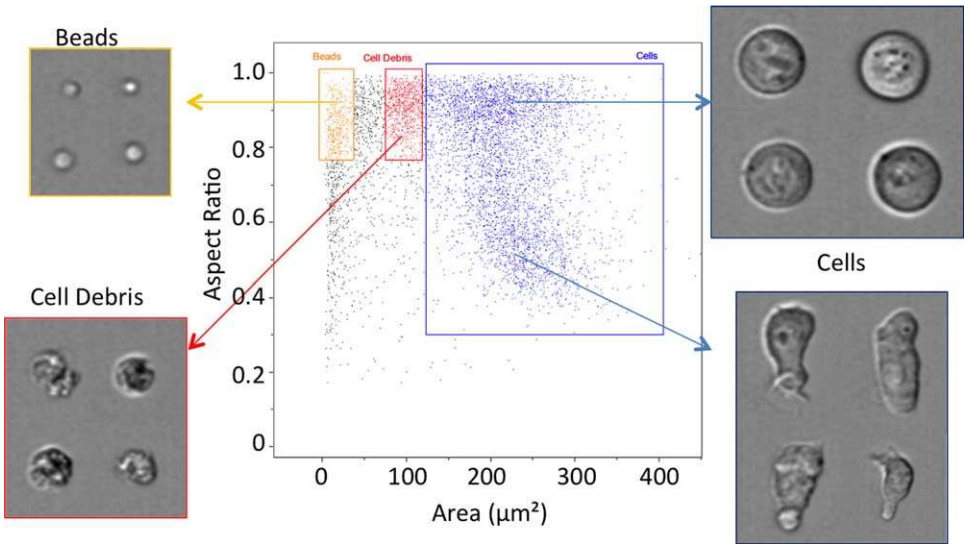


Figure 1. Gating strategy for including all types of morphology in imaging flow cytometry (IFC). The area versus aspect ratio (AR) scatterplot shows different types of population. Beads are very round (AR close to 1) and are very small (10–20 μm^2). Cell debris are around 50–100 μm^2. The cell gate is defined for AR from 0.3 to 1 and surface area from 150 to 500 μm^2. This choice of wide gate usually including doublet of cells is justified because polarized cells exhibit very close morphology parameters than doublet of cells. No doublet was observed in this plot thanks to a strategy based on cell concentration. [Color figure can be viewed at wileyonlinelibrary.com]

Otsu thresholding algorithm (18). These different steps are shown graphically in Supportive Information Figure S2.

Fiji Trackmate plugin (19) was used to track the cells and have their lineage. For each cell, this plugin allowed recording the positions of its center of mass. Those coordinates were used to measure the roundness and the area of each cell using the Wand feature and the Analyze Particles feature. *k*-means clustering was applied similarly to IFC on the combination of all cell data.

Results

Gating in IFC

The cells analyzed in this study were found to be mainly in two forms: round and elongated (polarized). In order to differentiate them, we used the common Area versus Aspect Ratio scatter plot. However, we had to modify our gating strategy because some of our elongated cells shared characteristics in terms of Area and Aspect Ratio with doublets of cells. Various strategies were tested to efficiently separate cell doublets from single cells. Gating on SSC intensity, expected to be higher for doublets than for singlets or morphology-based selection was not satisfactory after visual control of event images. Other possibilities were also ruled out. We therefore decided to adopt an upstream strategy by adapting the cell concentration of the sample. We used Poisson law to theoretically calculate the probability to have 0, 1, or 2 cells at the same time in the observed volume (Supporting Information Table S1). This probability depends on cell concentration and we chose to set a threshold probability at 5% doublet of cells. We confirmed our theoretical calculations using cells suspensions at different concentration (Supporting Information Figs. S4, S5, and Table S2). Figure 1 presents a scatter plot of area versus aspect ratio including gates for various event types. The single cells gate is wider than usual, but do not exclude any cell shape.

Different Phenotypes Distinguished on the Basis of Roundness and Area in IFC

In order to classify automatically and represent graphically the distribution of cell shapes, we chose the morphological roundness parameter, calculated as follows:

$$\text{Roundness} = 4 \times \frac{\text{Area}}{\pi \times (\text{Major Axis})^2}$$

Major axis is the longest dimension of the mask's best fitting ellipse. Roundness quantifies how close to a circle an object is: it is perfectly round when it is equal to 1.

A scatterplot of Roundness versus Area is shown in Supporting Information Figure S6. Two populations can be distinguished visually: round cells have roundness close to 0.9 and polarized cells have roundness close to 0.5. To avoid any thresholding subjectivity, we pooled all IFC roundness data from days 0 to 3 and used the *k*-mean algorithm (14) to distinguish the two groups. The aim of the *k*-means algorithm is to divide a set of points into k clusters (in our case $k = 2$) so that the within-cluster sum of squares is minimized. The calculation enabled us to set a roundness threshold of 0.67.

Dynamics of Morphology at Population Level

Four time points were analyzed using IFC on a stimulated CD34+ cell population. A dramatic change in the overall aspect of the scatter plots is observed between Day 0 and Day 1 (Fig. 2). The cells increase in size and the mean roundness of the population decreases while standard deviation of both parameters (size and roundness) increases. It results in a global increase of morphology heterogeneity. Scatterplots for Day 1, Day 2, and Day 3 look similar. The proportion of polarized

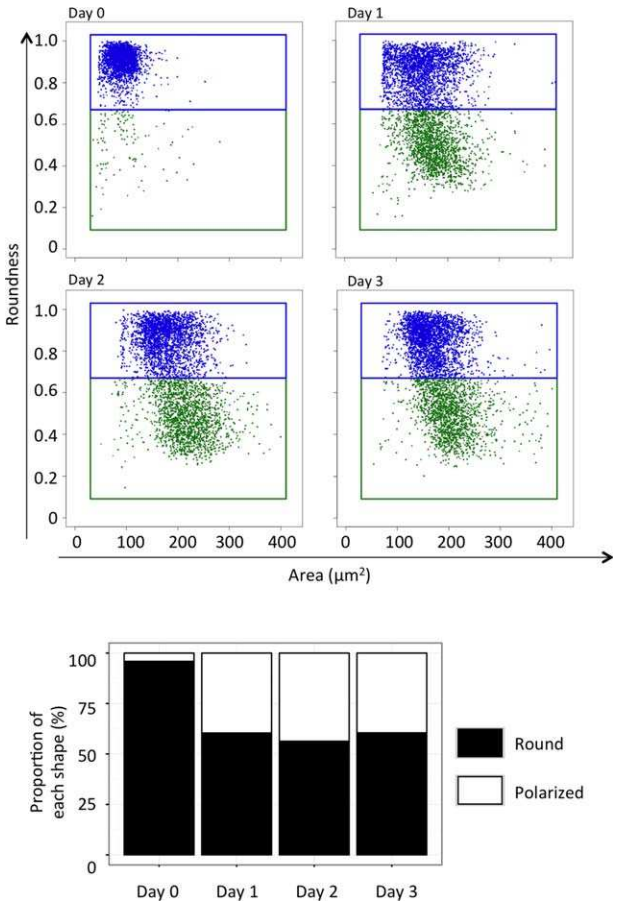


Figure 2. Evolution of human cord blood derived CD34+ cell population phenotype using imaging flow cytometry. Scatterplots of roundness against area (upper panel) for four time points. The threshold displayed for separating round and polarized phenotype was set to 0.67 (calculated using the *k*-means clustering algorithm). Proportion of each phenotype. At Day 0, cells exhibit a very homogenous phenotype and most of cells (96%) are round. A phenotypic change occurs between Day 0 and Day 1 leading to an increase in the percentage of polarized cells up to 40% at Day 1. The proportions barely change from Day 1 to Day 3. [Color figure can be viewed at wileyonlinelibrary.com]

cells increases from 3% at Day 0 to 36% at Day 1, whereas from Day 1 to Day 3, the proportions of each phenotypic state remain stable (Fig. 2). We consequently used TLM in order to determine whether this stability is the result of the stability of individual cell phenotypes or reflects a dynamic equilibrium between the two forms continuously transforming into each other.

Single Cell Dynamics by TLM

TLM allows monitoring of the morphology at a high sampling rate during the duration of the whole cell cycle of the same individual cell. Individual CD34+ clones isolated in micro-wells were followed by TLM over three generations. We recorded the cell shape changes at a rate of one frame per minute.

Based on a homemade automatic image processing method, roundness and area were calculated for each cell at each time point. We observed different cell profiles of roundness plotted against time (Fig. 3). In order to quantify the dynamics of cell shape, we pooled all TLM roundness data and used the previously described *k*-means algorithm to determine a threshold. It resulted in a roundness threshold of 0.65. Interestingly, *k*-means clustering with the two modalities gave similar threshold. Under this threshold, cells are considered as "polarized", over it, they are assigned to the category "round". Importantly, every cell changed morphology during its cycle. The dynamics of these changes differed between the cells. Cells 1 and 2 (Fig. 3), kept their phenotype for relative long periods (mean time for polarized shape = 30.3 min for cell 1 and 26.7 min for cell 2) whereas others switched between the two phenotypes with higher frequencies (mean time for polarized shape: 20.1 min for cell 3 and 9.5 min for cell 4). Overall, this analysis shows that morphology is a dynamic phenotypic feature. These cells dynamically explore the state space of possible morphologic phenotypes.

We therefore chose to represent TLM data with the same scatterplot as for IFC. For TLM, the scatterplot corresponds to a single cell, each point being the state of the cell at one time point. Unlike IFC, points can be connected with lines representing the time trajectories of individual cells. To visualize the dynamic behavior of a clone, individual trajectories were integrated in the genealogic tree of the clone (Fig. 4).

The first scatterplot of Figure 4 represents the phenotypic path of the founder cell (FC), which has a long cell cycle (68 h). At the beginning of the experiment, the founder cell is small and fluctuates between round and polarized morphologies. It gradually increases in size before division. Daughter cells (DC1 and DC2) have comparable trajectories with a longer time spent in the polarized phenotype. Granddaughter cells (GC1–4) present four highly dispersed but similar trajectories showing larger sizes with a more balanced phenotype. These four trajectories and the scatterplot obtained in IFC at Days 1,2, and 3 are very similar.

Discussion

We optimized a label-free methodology to study dynamics of cell morphology by combining two complementary approaches. IFC is powerful to make statistics of the morphological forms in a cell population and TLM can follow the real dynamics of a single cell.

IFC is the ideal solution for studying the highest number of cells in a population. Yet, this technique has drawbacks. First it is not compatible with the tracking of individual cells. Second, as for any flow method, IFC suffers from hydrodynamics effects that dramatically modify cell shape even when using paraformaldehyde fixation. To keep cell shape integrity, we optimized the use of glutaraldehyde, a fixative well known for preserving shape, commonly used in electron microscopy. The polarized cell shapes observed by this approach have similar properties to doublets of cells, hence requiring other methods to eliminate them. We chose to optimize cell suspension concentration using the probabilities of events given by a Poisson distribution. Having the correct concentration is indeed

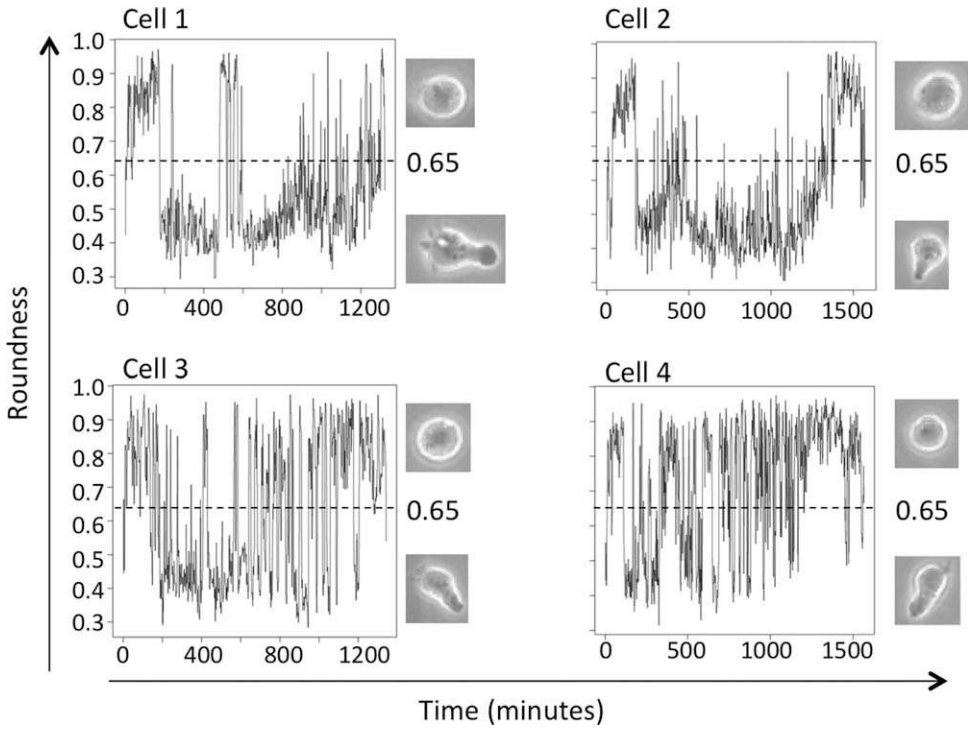


Figure 3. Evolution of four cells phenotype (roundness) during their cell cycle using time-lapse microscopy. Roundness was plot against time for four cells. Cells fluctuate between the two morphologies (round and polarized) separated by a 0.65 threshold (calculated using the *k*-means clustering algorithm). Cell 1 and cell 2 spend more time in the polarized phenotype: polarized mean time = 28.5 min. Cell 3 and cell 4 have smaller mean time: polarized mean time = 14.8 min.

crucial to avoid considering doublets of cells as polarized cells. With such optimized approach, IFC gave us access to morphological parameters of the cell population at the considered time points.

TLM is the method of choice for the study of cell dynamics. It can track the same cell during its cell cycle and enables the reconstruction of a complete cell lineage. The main limitation of TLM is the low number of cells that can be tracked and analyzed. It depends on a compromise between the number of fields acquired (physical positions of the objective or stage) and the frame rate. Cell tracking consists in identifying each cell frame after frame. Thus, the delay between two frames has to be adapted to cell mobility. In our case, CD34+ cells are so mobile that the delay should not exceed one minute. The other drawback of TLM is the time necessary to analyze the TLM images. Although we have used an efficient image-processing algorithm for the automatic segmentation, time consuming manual corrections and validation might be required.

TLM and IFC are experimental techniques that provide information at different scales and different resolutions. Both modalities rely on different optics, camera and image processing algorithm. However, when automatically calculating the optimal threshold using *k*-means algorithm for each modality, we obtained very close values: 0.67 for IFC and 0.65 for TLM. This underlines the robustness of the combination of IFC and TLM.

The phenotype of a cell population is frequently displayed as a 2D scatterplot showing a typical cloud of points. The pattern displayed by this cloud can look similar through time giving the impression of stability but all the points (individual cells) may simultaneously move within the cloud without changing the average shape of the cloud. A system (cell population), in such an apparent steady state with dynamics at the individual level, can be considered to be in a dynamic equilibrium. With TLM we showed that some cells fluctuate between the two phenotypes. It is therefore plausible that the observed cell behavior reflects the majority of cell's behavior inside the population. If this were the case, our cell population would be in a state of dynamic equilibrium. Such view is fundamentally different from a population with a stable cell phenotype, and might change population 2D scatterplot interpretations.

Several studies have highlighted dynamic effects comparable with dynamic equilibrium using cell sorting on muscle cells (20), hematopoietic stem cells (21), or pluripotent cells (22). Isolation of specific cell subpopulation belonging to a heterogeneous population resulted indeed, after several days of cell culture, in a population of which heterogeneity was similar to the initial population. These studies indirectly demonstrated phenotypic fluctuations whereas our approach enables direct quantification of phenotypic dynamics.

Yin et al. (23–25) developed a quantitative method to describe cell morphology based on fluorescent labeling of adherent cells. Hundreds of morphological features are recorded and analysed using specific clustering algorithms. The results of this analysis show that in some cases, cells exhibit typical discrete shapes. The cells populations explore

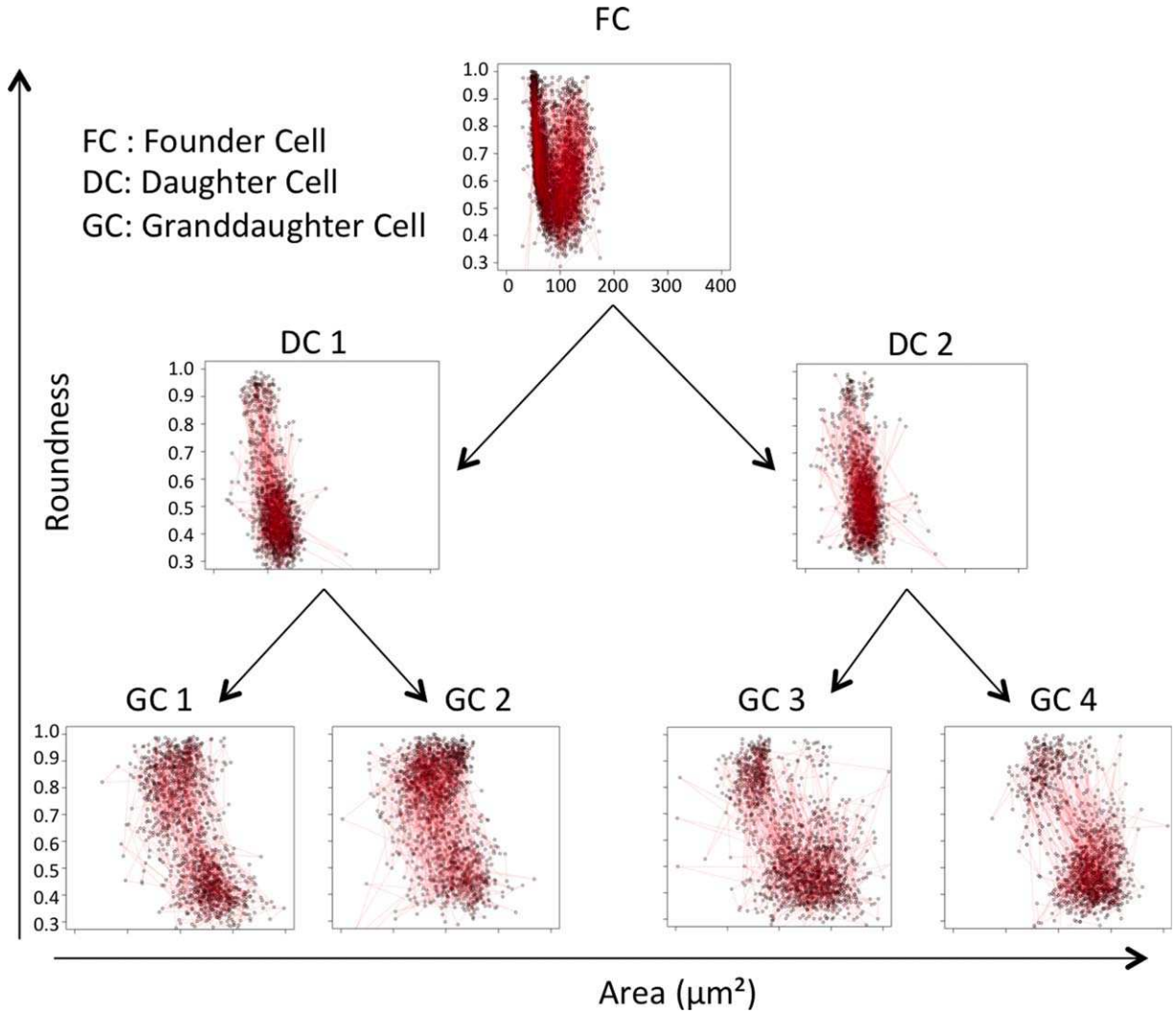


Figure 4. Lineage tree and phenotypic space exploration (roundness-area) of a CD34+ clone for three generations. The founder cell (FC) has a 4080 min (68 h) cell cycle; it stays small in size and round for several hours until it eventually grows and divides. Daughter cell 1 (DC1) has a similar trajectory to daughter cell 2 (DC2) and almost the same cell cycle of 1440 min (24 h); both are very different from the FC. They spend most of their lifetime in the polarized phenotype. The grand daughter cells 1 and 2 (GC1 and GC2) also have similar patterns: they explore quite equally round and polarized phenotype. Interestingly, grand daughter cell 3 and 4 (GC3 and GC4) have also a similar pattern with more time spent in the polarized phenotype. GC1, GC2, GC3, and GC4 have a wider pattern than their mother and the FC, and have cell cycle lengths between 1200 and 1500 min (20–25 h), showing scatterplots looking like the population scatterplot obtained with IFC. [Color figure can be viewed at wileyonlinelibrary.com]

the potential shape space in a discrete switch-like manner with few intermediate shapes (24). Our observations suggest that non-adherent human hematopoietic cells display a very similar morphological behavior. In addition, we show that a label-free approach with lower resolution is also suitable to detect the switch-like cell behavior and makes possible the continuous time-lapse observation of an individual cell.

In summary, the present study highlights the advantages of a label-free multimodal approach at different scales of sampling and time: population level, clone level and single cell level. The strategy based on high time-resolution TLM data, obtained from a small number of cells, anchored to IFC snapshots of a large number of cells overcomes the drawbacks of both modalities. This combination of TLM and IFC shows that cell morphology reveals high fluctuations compatible with a dynamic equilibrium scenario. Our approach paves the way for label-free morphological cell dynamics and offers insights for the understanding and interpretation of population data widely used in flow cytometry.

Acknowledgments

This study has been sponsored and promoted by Genethon, a nonprofit organization and the Fond d'Innovation pour la Recherche de l'Université d'Evry. Equipment funds were from Genopole Research-Evry, University Evry Val d'Essonne, Conseil General de l'Essonne and from Region Ile de France. A.M. is a recipient of a PhD scholarship from

E.P.H.E. We thank the Imaging-Cytometry Platform of Genethon.

Literature Cited

1. Brown M, Wittwer C. Flow cytometry: Principles and clinical applications in hematology. Clin Chem 2000;46:1221–1229.
2. Gilner JB, Walton WG, Gush K, Kirby SL. Antibodies to stem cell marker antigens reduce engraftment of hematopoietic stem cells. Stem Cells 2007;25:279–288.
3. Hou HW, Petchakup C, Tay HM, Tam ZY, Dalan R, Chew DEK, Li KHH, Boehm BO. Rapid and label-free microfluidic neutrophil purification and phenotyping in diabetes mellitus. Sci Rep 2016;6. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4933935/. Accessed August 2, 2016.
4. Basiji DA, Ortyn WE, Liang L, Venkatachalam V, Morrissey P. Cellular image analysis and imaging by flow cytometry. Clin Lab Med 2007;27:653–viii.
5. Blasi T, Hennig H, Summers HD, Theis FJ, Cerveira J, Patterson JO, Davies D, Filby A, Carpenter AE, Rees P. Label-free cell cycle analysis for high-throughput imaging flow cytometry. Nat Commun 2016;7:10256.
6. Vorobjev IA, Barteneva NS. Quantitative functional morphology by imaging flow cytometry. Methods Mol Biol 2016;1389:3–11.
7. Kokkaliaris KD, Loeffler D, Schroeder T. Advances in tracking hematopoiesis at the single-cell level. Curr Opin Hematol 2012;19:243–249.
8. Megyola CM, Gao Y, Teixeira AM, Cheng J, Heydari K, Cheng E, Nottoli T, Krause DS, Lu J, Guo S. Dynamic migration and cell–cell interactions of early reprogramming revealed by high resolution time-lapse imaging. Stem Cells 2013;31:895–905.
9. Kong Z, Zhu X, Zhang S, Wu J, Luo Y. Phase contrast microscopy of living cells within the whole lens: Spatial correlations and morphological dynamics. Mol Vis 2012;18:2165–2173.
10. Eaves CJ. Hematopoietic stem cells: Concepts, definitions, and the new reality. Blood 2015;125:2605–2613.
11. Giebel B, Corbeil D, Beckmann J, Höhn J, Freund D, Giesen K, Fischer J, Kögler G, Wernet P. Segregation of lipid raft markers including CD133 in polarized human hematopoietic stem and progenitor cells. Blood 2004;104:2332–2338.
12. Görgens A, Ludwig A-K, Möllmann M, Krawczyk A, Dürig J, Hanenberg H, Horn PA, Giebel B. Multipotent hematopoietic progenitors divide asymmetrically to create progenitors of the lymphomyeloid and erythromyeloid lineages. Stem Cell Rep 2014;3:1058–1072.
13. Auinger S, Small JV. Chapter 14 correlated light and electron microscopy of the cytoskeleton. In: Biology B-M in C, editor. Introduction to Electron Microscopy for Biologists, Vol. 88. New York: Academic Press; 2008. pp 257–272. http://www.sciencedirect.com/science/article/pii/S0091679X08004147. Accessed August 12, 2016.
14. Artigan JA, Wong MA. A *k*-means clustering algorithm. Appl Stat 1979;100–108.
15. Cosette J, Moussy A, Onodi F, Auffret-Cariou A, Neildez-Nguyen TMA, Paldi A, Stockholm D. Single cell dynamics causes Pareto-like effect in stimulated T cell populations. Sci Rep 2015;5:17756.
16. Day D, Pham K, Ludford-Menting MJ, Oliaro J, Izon D, Russell SM, Gu M. A method for prolonged imaging of motile lymphocytes. Immunol Cell Biol 2009;87:154–158.
17. Sternberg. Biomedical image processing. Computer 1983;16:22–34.
18. Otsu N. A threshold selection method from gray-level histograms. IEEE Trans Syst Man Cybern 1979;9:62–66.
19. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, et al. Fiji: An open-source platform for biological-image analysis. Nat Meth 2012;9:676–682.
20. Stockholm D, Edom-Vovard F, Coutant S, Sanatine P, Yamagata Y, Corre G, Le Guillou L, Neildez-Nguyen TMA, Pàldi A. Bistable cell fate specification as a result of stochastic fluctuations and collective spatial cell behaviour. PLoS ONE 2010;5:e14441.
21. Chang HH, Hemberg M, Barahona M, Ingber DE, Huang S. Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. Nature 2008;453:544–547.
22. Galvin-Burgess KE, Travis ED, Pierson KE, Vivian JLTGF. Beta superfamily signaling regulates embryonic stem cell heterogeneity: Self-renewal as a dynamic and regulated equilibrium. Stem Cells 2013;31:48–58.
23. Yin Z, Zhou X, Bakal C, Li F, Sun Y, Perrimon N, Wong STC. Using iterative cluster merging with improved gap statistics to perform online phenotype discovery in the context of high-throughput RNAi screens. BMC Bioinform 2008;9:264.
24. Yin Z, Sadok A, Sailem H, McCarthy A, Xia X, Li F, Garcia MA, Evans L, Barr AR, Perrimon N, et al. A screen for morphological complexity identifies regulators of switch-like transitions between discrete cell shapes. Natl Cell Biol 2013;15:860–871.
25. Yin Z, Sailem H, Sero J, Ardy R, Wong STC, Bakal C. How cells explore shape space: A quantitative statistical perspective of cellular morphogenesis. BioEssays 2014;36:1195–1203.

SUPPLEMENTAL INFORMATION

Custom Mask for Imaging flow cytometry

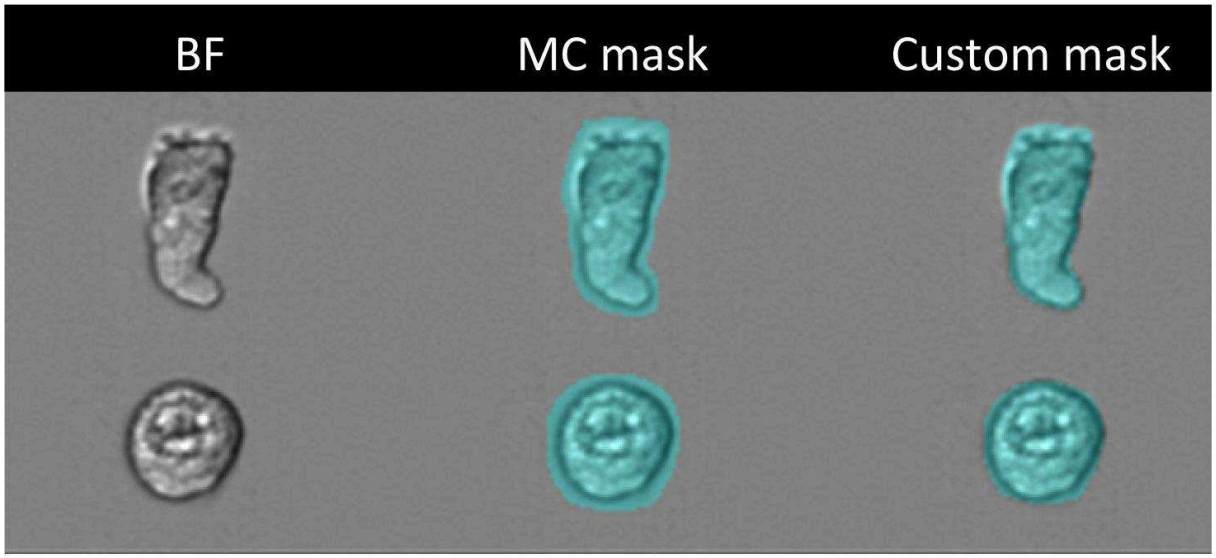


Figure S1: Illustrative images of the differences between the MC mask (Mask Combined) and our Custom mask. The Custom mask is closer to the cell contour than the MC mask, allowing accurate quantification of shape and morphology parameters.

Image processing algorithm.

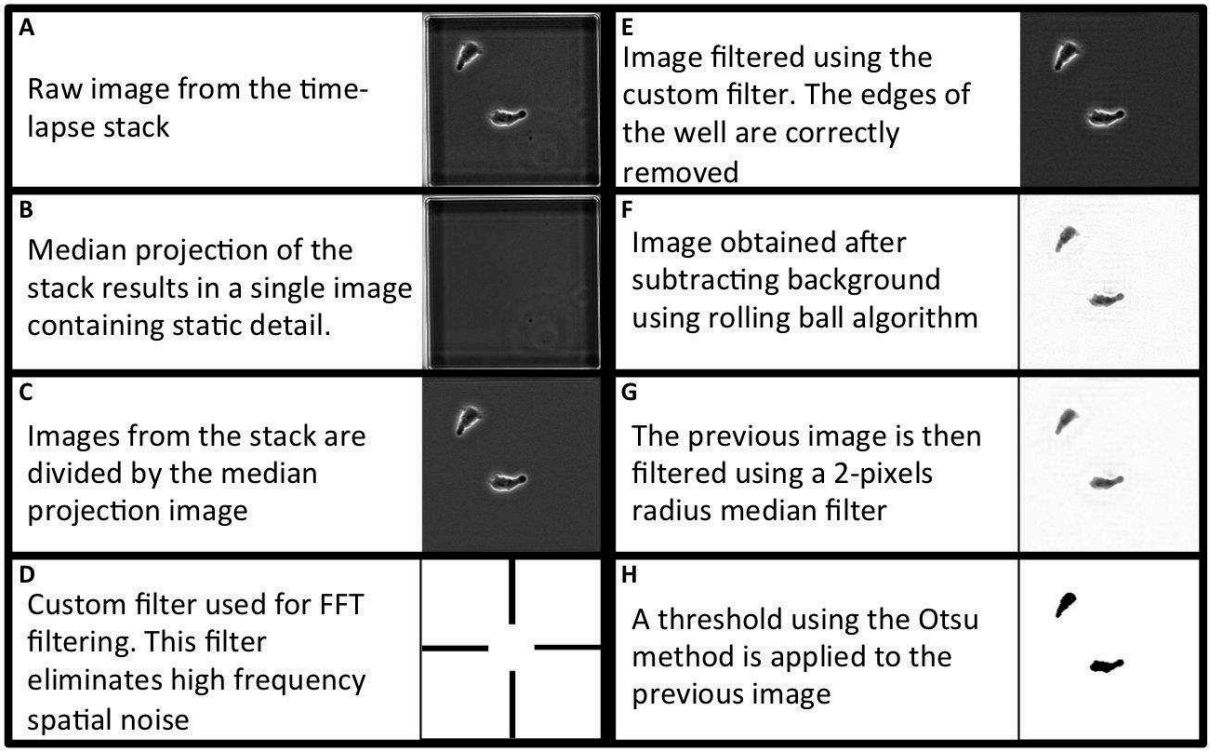


Figure S2: Image processing algorithm. A. Raw Image. B. Median projection of the entire time lapsed sequence. C. Result of images of the stack divided by the median projection. D. Custom filter used for FFT filtering resulting in E. F. Background substraction (rolling ball algorithm). G. 2-pixel median filtering result. H. Final segmented image using Otsu thresholding.

Calculation of cell suspension optimal concentration

The issue here is to set the optimal concentration that maximizes the probability to have only one cell inside the observed area and that minimizes the probability to have 2 cells (a doublet) in the same area of the flow cell.

According to the supplier, the field of view with the 40X objective is 40µm x 170µm, and the depth of field is 4µm.

This means that the observed volume is

$$V_{obs} = 40.10^{-6} \times 170.10^{-6} \times 4.10^{-6}$$
$$V_{obs} = 2.72\ 10^{-14}\ m^3 = 2.72\ 10^{-5} \mu L$$

The concentration corresponding to 1 cell in the observed volume (C_{1cell}) is 1 cell every 2.72E-5 µL.

$$C_{1cell} = \frac{1}{2.72\ 10^{-5}}\ Cells\ /\ \mu L$$

$$C_{1cell} \approx 36800\ Cells\ /\mu L$$

Calculation of the probability to have 0, 1 or 2 cells in the explored volume according to the C_{1cell} concentration is given by the Poisson law.

$$P(X = k) = \frac{\lambda^k\ e^{-\lambda}}{k!}$$

λ being the expected value

In Figure S3, we show the evolution of the probability to find 0, 1 or 2 cells in the explored volume calculated with the Poisson distribution.

We established that for an optimal concentration C_{opt} = 0.4 x C_{1cell} , the probability to have 2 cells in the explored volume is 0.05, whereas the probability to have 1 cell is 0.29. The probability to have 0 cells is increased dramatically, increasing then the time of acquisition but minimizing the doublets of cells.

$$C_{opt} = 0.4 \times C_{1cell}$$
$$C_{opt} = 0.4 \times 36800$$

$$C_{opt} \approx 15000\ Cell\ /\mu L$$

$$C_{opt} = 3.10^5\ Cells\ /\ 20\ \mu L\ sample$$

To quantify theoretically the acquisition time, the volume flow in the cytometer needs to be calculated.
With our setting (40X, "low" velocity), the flow velocity v is 66 mm.s^{-1}.

The volume flow is then: $Q = v \times S$

S being the surface area of the observed volume $S = 40 \times 4\ \mu m^2 = 0.04 \times 0.004\ mm^2$

$$Q = 66 \times 0.04 \times 0.004\ mm^3 s^{-1}$$
$$Q = 0.0105\ mm^3 s^{-1} = 1.05\ 10^{-2} \mu L\ s^{-1}$$

On the other hand, $V_{obs} = 2.72\ 10^{-5} \mu L$

The corresponding time τ for an observed volume is:

$$\tau = \frac{V_{obs}}{Q} = \frac{2.72\ 10^{-5}}{1.05\ 10^{-2}}$$
$$\boldsymbol{\tau = 2.60\ ms}$$

Using Poisson law's probabilities, we can calculate the number of empty images N_{empty}, or doublet containing images $N_{doublets}$, and the total number of "observed volumes" N_{total}, based on an acquisition of 5000 cells (singlets)

$$N_{total} \times P(X = 0) = N_{empty}$$

$$N_{total} \times P(X = 1) = 5000$$

$$N_{total} \times P(X = 2) = N_{doublet}$$

$$T_{acq} = N_{total} \times \tau$$

This mean that for each concentration we can calculate the corresponding $N_{doublets}, N_{total}, N_{empty}$ and acquisition time T_{acq}. All of these information are in the following table S1 for the concentrations of 0.1*C_{1cell}, 0.2*C_{1cell}, 0.4*C_{1cell}, 0.8*C_{1cell} and 1* C_{1cell}.

Concentration (xC_{1cell})	Empty event Probability	Singulet Probability	Doublet Probability	N_{total}	N_{empty}	$N_{singulets}$	$N_{doublets}$	T_{acq} (s)
0.1	0.9048	0.0905	0.0045	55258	50000	5000.0	250.0	143.7
0.2	0.8187	0.1637	0.0164	30535	25000	5000.0	500.0	79.4
0.4	**0.6703**	**0.2681**	**0.0536**	**18647**	**12500**	**5000.0**	**1000.0**	**48.5**
0.8	0.4493	0.3595	0.1438	13909	6250	5000.0	2000.0	36.2
1	0.3679	0.3679	0.1839	13591	5000	5000.0	2500.0	35.3

Table S1: Probability to have empty, one-cell or two-cells containing images, according to Poisson's law. The total number of explored volume and the acquisition time can be therefore calculated. The

concentration of 0.4*C1cell is a good compromise with an intermediate time of acquisition while maximizing the one-cell containing image probability and minimizing the doublet containing images.

We established that for an optimal concentration C_{opt} = 0.4 x C_{1cell}, the probability to have 2 cells in the explored volume is 0.05, whereas the probability to have 1 cell is 0.27. The probability to have an empty event is dramatically increased, increasing then the time of acquisition but minimizing the doublets of cells.

$$C_{opt} = 0.4 \times C_{1cell}$$

$$C_{opt} = 3.10^5\ Cells\ /\ 20\ \mu L\ sample$$

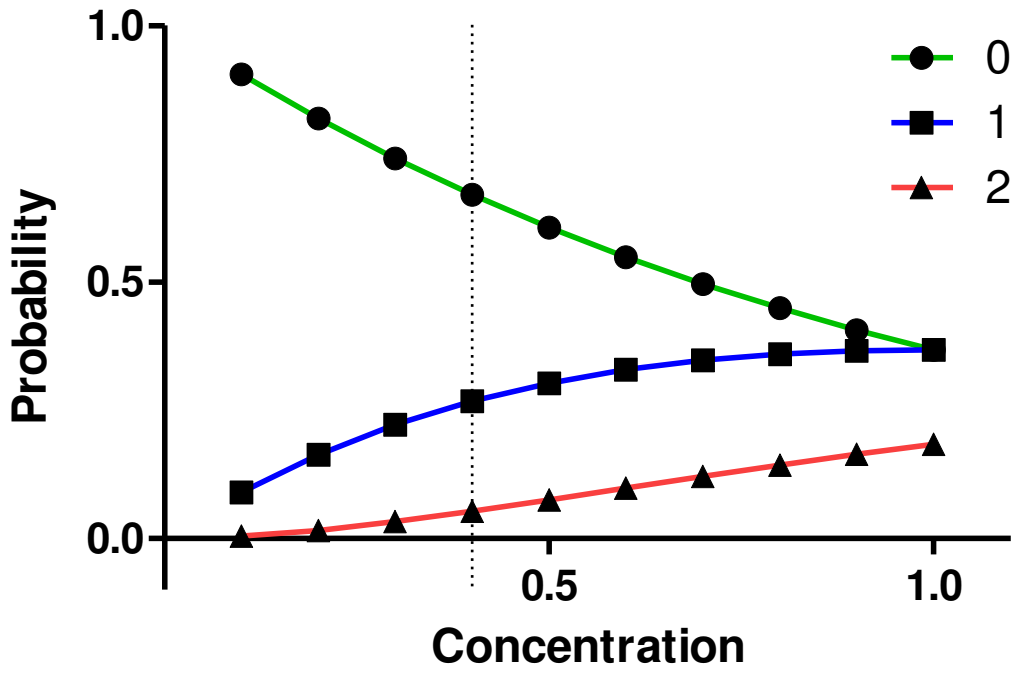


Figure S3: Probabilities to have 0, 1 or 2 cells in the observed volume, according to concentration (expressed in relation to C_{1cell}).

Experimental confirmation of optimal concentrations

Human leukemia cells derived from hematopoietic progenitor cells (K562 cells) were used to verify the theoretical calculations of doublet proportions, time of acquisition, and to exhibit doublets and triplets absolute number increase at very high cell concentration.

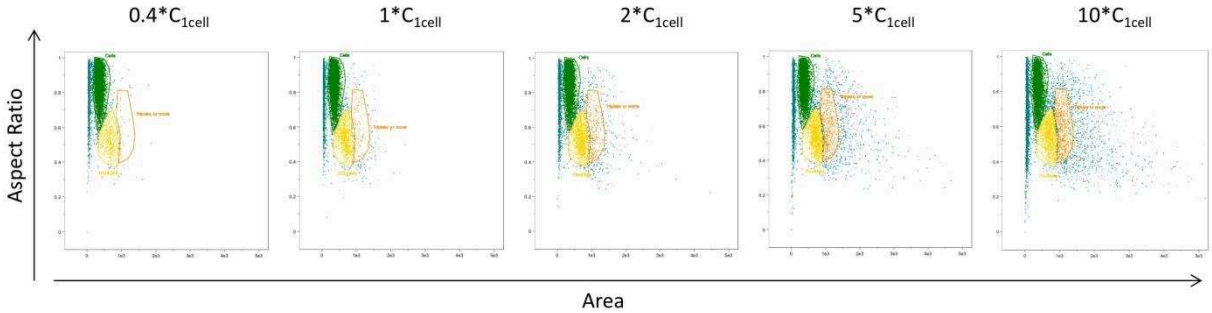


Figure S4: Area vs Aspect ratio dot plots of K562 cell suspensions at different concentrations. For high concentrations the cloud corresponding to cell population is modified due to an increase in doublets, triplet and even multiplets proportions.

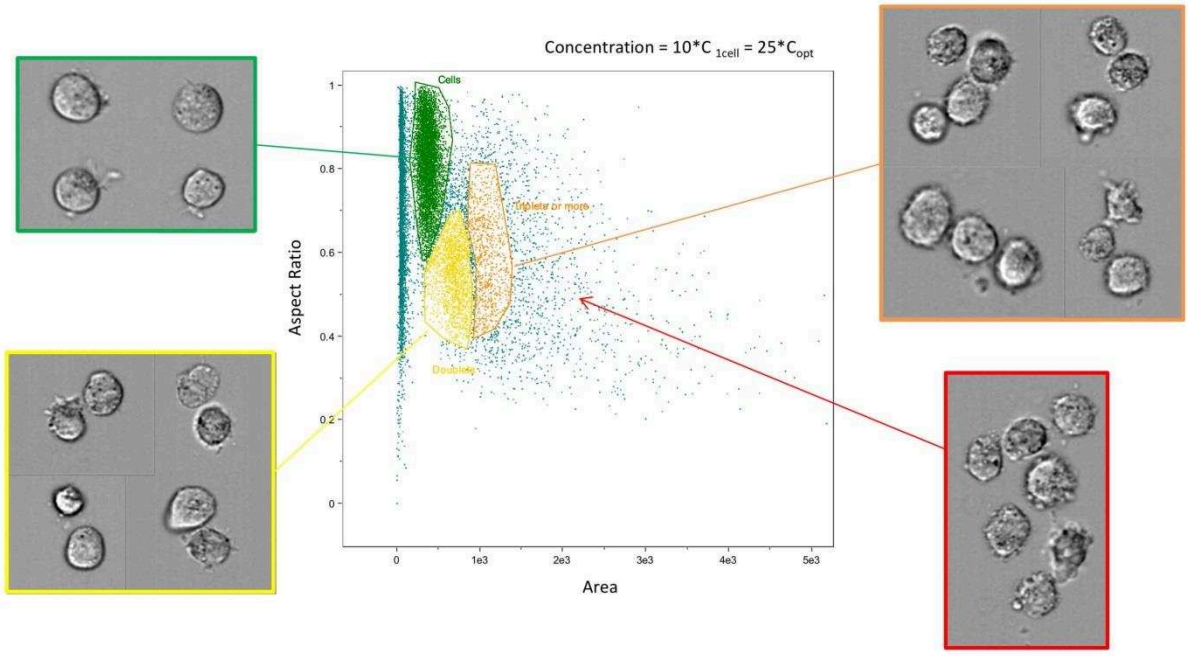


Figure S5: Area vs Aspect ratio dot plot for a high concentration cell (K562) suspension $10*C_{1cell} = 1.75x10^8$ cells / mL. The proportion of doublet is dramatically increased, as well as the triplet proportion. Moreover, a lot of multiplets (> 4 cells) are observed.

Concentration (xC1cell)	Acquisition time (s)	Total number of events	Time per event (ms)	% of singulets	% of doublets	% of triplets or more
0.4	46	12205	3.77	75.3	5.33	0.25
1	22	12186	1.80	70.7	9.19	1.08
2	15.5	13119	1.18	61	10.8	2.26
5	15	13522	1.11	55.5	11.9	4.13
10	19.5	16544	1.18	40.6	10.8	4.99

Table S2: Acquisition time, number of events, percentage of 'singlets', 'doublets' and 'triplets or more' for K562 cell suspension at $0.4*C_{1cell}$, $1*C_{1cell}$, $2*C_{1cell}$, $5*C_{1cell}$, and $10*C_{1cell}$. This table shows that singlets proportion decrease when concentration increases whereas 'doublet' and 'triplets or more' proportion increase when concentration increases.

Imaging flow cytometry scatterplot of Roundness vs Area.

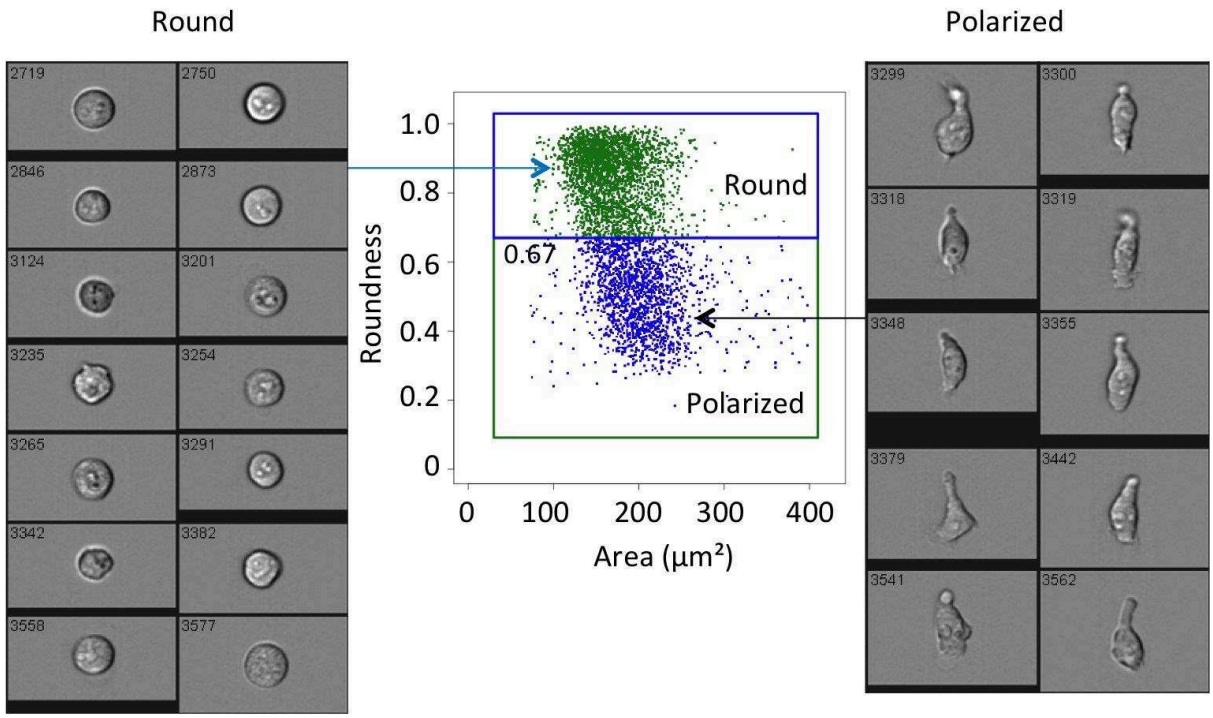


Figure S6: Imaging flow cytometry scatterplot of Roundness vs Area. Two populations are distinguishable on this graph and were found separated at a 0,67 threshold using the k-means algorithm. The upper part (roundness > 0.67) of the scatterplot corresponds to the round cells, whereas the lower part (roundness < 0.67) corresponds to polarized cells.

Grâce à la microscopie en TL, nous avons pu observer deux types de morphologies distinctes, les cellules rondes et les cellules polarisées.
Grâce à la CFI nous avons pu quantifier ces deux types de morphologie. A temps 0h, les cellules sont d'abord presque toutes rondes. A partir de 24h, les cellules sont de plus en plus hétérogènes, grossissent et changent de morphologie, elles sont davantage polarisées. Dans la population, à temps 0h, 3% des cellules sont polarisées et à 24h, 48h et 72h, elles sont à environ 35% polarisées. Afin de déterminer si cette stabilité est le résultat d'une stabilité phénotypique cellulaire individuel ou si elle reflète un équilibre dynamique continu entre les deux formes, nous avons utilisé la microscopie en TL.
Nous avons pu ainsi observer que certains clones de cellules gardaient le même phénotype entre deux cycles cellulaire sur une longue période alors que d'autres clones avaient la capacité de fluctuer entre ces deux types de morphologies. Cela laisserait suggérer que la composition stable des cellules rondes et polarisées pourrait réellement représenter un équilibre dynamique. De plus, nous avons montré que les cellules sœurs et les cellules cousines avaient un comportement dynamique similaire mais qu'elles étaient presque toutes différentes de la cellule mère.

Par la combinaison de ces deux types de technologies, nous avons pu montrer que la morphologie des cellules CD34+ issues de sang de cordon, en réponse à la stimulation d'un cocktail de cytokines, était dynamique et pouvait être étudiée avec succès sans marquage spécifique. La stratégie basée sur des données de microscopie en TL à haute fréquence d'un petit nombre de cellules, combinée avec des données d'un grand nombre de cellules obtenues avec la CFI à des instants donnés a permis de surmonter les inconvénients des deux technologies. La combinaison de ces deux méthodes montre que la morphologie des cellules CD34+ est dynamique et fluctue beaucoup entre les phénotypes ronds et polarisés. Cette dynamique morphologique, bien qu'elle soit peu exploitée dans les analyses de différenciation, pourrait avoir un rôle important dans le devenir des CSH mais aussi dans d'autres types cellulaires.

Cette approche novatrice offre des idées pour la compréhension et l'interprétation des données populationnelles des cellules CD34+ et plus généralement des cellules souches qui sont largement utilisées dans la cytométrie en flux. Pour comprendre au mieux le destin « choisi » par les cellules CD34+ dans la différenciation hématopoïétique, l'idéal serait, en plus d'étudier leur dynamique morphologique, de pouvoir suivre leur dynamique transcriptomique et protéomique. Cela est impossible à l'heure actuelle car il faut détruire la

cellule pour accéder à l'information. Pour faire face à ce problème, nous avons exécuté à différents temps des analyses transcriptomiques des cellules CD34+ uniques issues de sang de cordon. En reliant la morphologie dynamique et la transcription à un instant donné de ces cellules, nous avons ainsi étudié la complexité du destin non linéaire de ces cellules.

II. Etude cinétique des changements phénotypiques des cellules CD34+ humaines issues de sang de cordon à l'échelle de la cellule unique

Les CSH sont des cellules capables de donner toutes les cellules du sang. Les cellules prises individuellement s'engagent dans des voies de lignages spécifiques qui ne sont pas nécessairement uniformes ou prédéterminées. Au cours de ma thèse, j'ai abordé cette question en caractérisant les changements transcriptomiques des cellules CD34+ humaines issues de sang du cordon à l'échelle de la cellule unique au cours du temps. J'ai couplé ces données avec des analyses sur la généalogie des cellules et sur les changements dynamiques morphologiques réalisées grâce à la microscopie en TL et à la CFI.

Le profil transcriptomique des cellules CD34+ individuelles a été déterminé à temps 0h, 24h, 48h et 72h après stimulation avec des cytokines. Nous avons analysé 90 gènes :

- plus de la moitié a été choisi de façon aléatoire dans une liste de gènes (Cattoglio et al., 2010; Livak et al., 2013) ;
- un tiers des gènes a été sélectionné pour leur fonction connue dans la différenciation précoce des cellules hématopoïétiques ;
- 5 autres gènes ont été choisis pour leur rôle dans le maintien de l'état pluripotent dans les CSE.

Cette liste de gènes nous a permis d'avoir une estimation globale de l'activité transcriptionnelle du génome.

Parallèlement, nous avons analysé la dynamique morphologique de ces cellules en microscopie en TL comme décrit dans le papier précédemment couplé à des analyses effectuées manuellement. Nous avons suivi les cellules CD34+ uniques toutes les minutes pendant 7 jours, après stimulation, isolées grâce à une microgrille de PDMS contenant 1024 puits faisant chacun 125 µm de côté et 60 µm de profondeur.

Notre plan expérimental a combiné à la fois des observations en continu avec la microscopie en TL et des observations ponctuelles associées à des analyses transcriptomiques à l'échelle de la cellule unique.

ARTICLE 3

"Integrated time-lapse and single-cell transcription studies highlight the variable and dynamic nature of human hematopoietic cell fate commitment"

Alice Moussy, Jérémie Cosette, Romuald Parmentier, Cindy da Silva, Guillaume Corre, Angélique Richard, Olivier Gandrillon, Daniel Stockholm and András Páldi

PLoS Biology 15(7): e2001867. (Juillet 2017)
DOI: 10.1371/journal.pbio.2001867

PLOS | BIOLOGY

RESEARCH ARTICLE

Integrated time-lapse and single-cell transcription studies highlight the variable and dynamic nature of human hematopoietic cell fate commitment

Alice Moussy[1,2], Jérémie Cosette[2], Romuald Parmentier[2], Cindy da Silva[1], Guillaume Corre[2], Angélique Richard[3], Olivier Gandrillon[3], Daniel Stockholm[1], András Páldi[1]*

1 Ecole Pratique des Hautes Etudes, PSL Research University, UMRS 951, INSERM, Univ-Evry, Evry, France, **2** Genethon, Evry, France, **3** Laboratoire de Biologie et de Modélisation de la Cellule, Ecole Normale Supérieure de Lyon, CNRS, Université de Lyon, Lyon, France

* andras.paldi@ephe.sorbonne.fr

Check for updates

OPEN ACCESS

Citation: Moussy A, Cosette J, Parmentier R, da Silva C, Corre G, Richard A, et al. (2017) Integrated time-lapse and single-cell transcription studies highlight the variable and dynamic nature of human hematopoietic cell fate commitment. PLoS Biol 15 (7): e2001867. https://doi.org/10.1371/journal.pbio.2001867

Academic Editor: Sui Huang, Institute for Systems Biology, United States of America

Received: December 21, 2016

Accepted: June 23, 2017

Published: July 27, 2017

Copyright: © 2017 Moussy et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The quantitative RT-PCR, time-lapse and cytometry data presented in the paper and used to generate the graphical representations are available in the "Supporting information" as S1-S3 Data files respectively. Raw Q-RCR data and original time-lapse files are available on request.

Funding: Ecole Pratique des Hautes Etudes www.ephe.fr (grant number 11REC/BIMO). Received by AP. The funder had no role in study design, data

Abstract

Individual cells take lineage commitment decisions in a way that is not necessarily uniform. We address this issue by characterising transcriptional changes in cord blood-derived CD34 + cells at the single-cell level and integrating data with cell division history and morphological changes determined by time-lapse microscopy. We show that major transcriptional changes leading to a multilineage-primed gene expression state occur very rapidly during the first cell cycle. One of the 2 stable lineage-primed patterns emerges gradually in each cell with variable timing. Some cells reach a stable morphology and molecular phenotype by the end of the first cell cycle and transmit it clonally. Others fluctuate between the 2 phenotypes over several cell cycles. Our analysis highlights the dynamic nature and variable timing of cell fate commitment in hematopoietic cells, links the gene expression pattern to cell morphology, and identifies a new category of cells with fluctuating phenotypic characteristics, demonstrating the complexity of the fate decision process (which is different from a simple binary switch between 2 options, as it is usually envisioned).

Author summary

Hematopoietic stem cells are classically defined as a specific category of cells at the top of the hierarchy that can differentiate all blood cell types following step-by-step the instructions of a deterministic program. We have analysed this process, and our findings support a much more dynamic view than previously described. We apply time-lapse microscopy coupled to single-cell molecular analyses in human hematopoietic stem cells and find that fate decision is not a unique, programmed event but a process of spontaneous variation and selective stabilisation reminiscent of trial–error processes. We show that each cell explores (at its own pace and independently of cell division) many different possibilities before reaching a stable combination of genes to be expressed. Our results suggest,

collection and analysis, decision to publish, or preparation of the manuscript. Agence Nationale de la Recherche www.anr.fr (grant number BSV6 014 02 « Stochagene »). Received by AP and OG. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Genethon www.genethon.fr. Received by AP and DS. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Abbreviations: CLP, common lymphoid progenitor; CMP, common myeloid progenitor; CTV, CellTrace Violet; HDAC, histone deacetylase; HSPC, hematopoietic stem and progenitor cell; MPP, multipotent progenitor; PCA, principal component analysis; qRT-PCR, quantitative reverse transcription polymerase chain reaction; t-SNE, t-distributed stochastic neighbour embedding; VPA, valproic acid.

therefore, that multipotency seems to be more like a transitory state than a feature of a specific cell category.

Introduction

Hematopoietic stem and progenitor cells (HSPCs) give rise to all the cellular components of blood. The major stages of differentiation and the key genes participating in this process are now well characterised [1]. According to the classical view, haematopoiesis is a hierarchically organised process of successive fate commitments, in which differentiation potential is progressively restricted in an orderly way over cell divisions. There are several variants of the model [2–6]. In all cases, the first fate decision is a binary choice taken by multipotent progenitors (MPPs), which leads to 2 different committed progenitors (for the purpose of simplicity, these progenitors are designated here as common myeloid progenitors [CMP] and common lymphoid progenitors [CLP]). In molecular terms, the choice is believed to be the result of the strictly regulated activation of master regulator genes and their underlying transcriptional network [7]. However, the strict hierarchical logic of classical models has recently been challenged by a number of in vivo and in vitro studies [8–10]. Single-cell gene expression studies have revealed a much higher heterogeneity of cell subtypes than can be detected using a combination of surface markers [11]. It is not surprising that the number of identifiable cell types increases with the resolution of the detection method. Although correct cell type classification is a key step in understanding the cell fate decision issue, it cannot reveal the dynamic features of the fate commitment process and leaves a number of unanswered questions. Do different phenotypic forms represent different cell types or different stages of the same process? How does the transition between the forms occur? How long does it take?

Until recently, fully deterministic explanations were predominant, but recent studies have suggested other alternatives. Two different possibilities have been put forward. According to the first, the commitment process starts with the sporadic, independent activation of genes within the same cell. The simultaneous stochastic expression of regulatory genes specifying different lineages creates a multiprimed intermediate state that enables these cells to choose 1 of the lineages [12–16]. A coherent lineage-specific expression profile would then emerge from this metastable state. According to the second, commitment is preceded by transcriptome fluctuations between different lineage-biased states [17–19]. Surprisingly, the time scale of transformations related to the cellular fate decision process remains largely unexplored. The transcriptome of the same cell can be analysed only once, because the cell is destroyed by RNA extraction. Therefore, indirect approaches are required to identify trends and patterns in time series.

We addressed the issue of the dynamics and the time scale of the commitment process by integrating single-cell quantitative reverse transcription polymerase chain reaction (qRT-PCR), cell division history, and morphological changes determined by time-lapse analysis. Contrary to the common strategy consisting of isolating defined cell subpopulations with the help of specific surface markers and characterising their gene expression profiles at the single-cell level [20], we used an alternative approach. Individual cells were randomly isolated from the heterogeneous cord blood CD34+ cell fraction at different time points after cytokine stimulation, and their gene expression profiles were determined using single-cell qRT-PCR. The data provided a series of snapshots, showing the actual statistical distribution of single-cell gene expression patterns across the whole population. The structure of the population at the successive time points was revealed by unsupervised classification of the expression profiles

according to their similarity using multiparametric approaches. The progression of the fate commitment process was deduced from the evolution of the population structure. At the same time, using time-lapse microscopy, we tracked randomly isolated individual CD34+ cells and their progeny for several days after cytokine stimulation. We recorded the division history and the morphological changes of each cell in the clones. The population structure was deduced on the basis of the statistical analysis of these observations. The efficiency of the time-lapse approach in investigating cell fate decisions has recently been shown [21]. To reinforce this approach, the time-lapse and gene expression data were integrated into a coherent scenario. This was done by using CD133 protein expression levels to isolate cells with 1 transcription profile or the other and directly record their dynamic phenotype, thereby providing a direct link between dynamic phenotype and transcription profile.

Altogether, our results revealed that fate decision is a dynamic, fluctuating process that is more complex than a simple binary switch between 2 options, as it is usually envisioned.

Results

Single-cell gene expression

The transcriptional profile of individual cord blood CD34+ cells was determined at 0, 24, 48, and 72 h after the beginning of cytokine stimulation (Fig 1A). Single-cell qRT-PCR was used to quantify the mRNA levels of 90 different genes. A set of 32 genes was selected for their known function in the early differentiation of hematopoietic cells and was expected to inform on the functional state of the cells (see S1 Table). A second set of 54 genes was chosen randomly from a list of genes known to be expressed in the hematopoietic lineage [22,23]. These genes provided an assessment of the overall transcriptional activity of the genome. Five additional genes were added to the list for their role in maintaining the pluripotent state in embryonic stem cells. A heat map of all data and a violin plot of the expression profile of each gene at the 4 separate time points are shown in S1 Fig and S2 Fig. The normalized single-cell quantitative gene expression data obtained for the different time points were merged into a single database and screened for subpopulations by k-means clustering. The number of statistically distinguishable groups was inferred using gap statistics [24]. The groups were visualised on heat maps and on a 2D plot using t-distributed stochastic neighbour embedding (t-SNE) [25]. Although every cell had a unique gene expression pattern, this approach enabled us to clearly identify subgroups of cells in the population on the basis of the statistical similarity of their gene expression patterns (Fig 1B).

Nonstimulated CD34+ cells isolated from cord blood represented the t = 0 h time point. A heat map of the single-cell transcriptional profiles of genes contributing significantly to the identification of subgroups (S1 Fig) showed that this population of cells was heterogeneous. Several genes reported to play a role in self-renewal, quiescence, and other stem cell functions (CD71, CD133, CXCR4, GATA2, and FLT3) were expressed sporadically and at variable levels in a fraction of cells. Genuine pluripotent stem cell genes were also expressed at low levels in a fraction of cells (NANOG, OCT4, KLF4). Nevertheless, no correlation was found between these genes (Fig 1D), and the statistical analysis did not reveal distinguishable expression patterns that could define cell types. The only detectable differences were donor-associated and probably reflected differences related to the processing of individual blood samples. Donor-specific differences disappeared at later stages.

The gene expression profile 24 h after the onset of cytokine stimulation was found to be fundamentally different to t = 0 h cells. Almost every cell responded to cytokine stimulation by increasing transcript levels and generating a unique gene expression pattern (Fig 1C). When represented on the 2D t-SNE (Fig 1B) and principal component analysis (PCA) plots (S3 Fig),

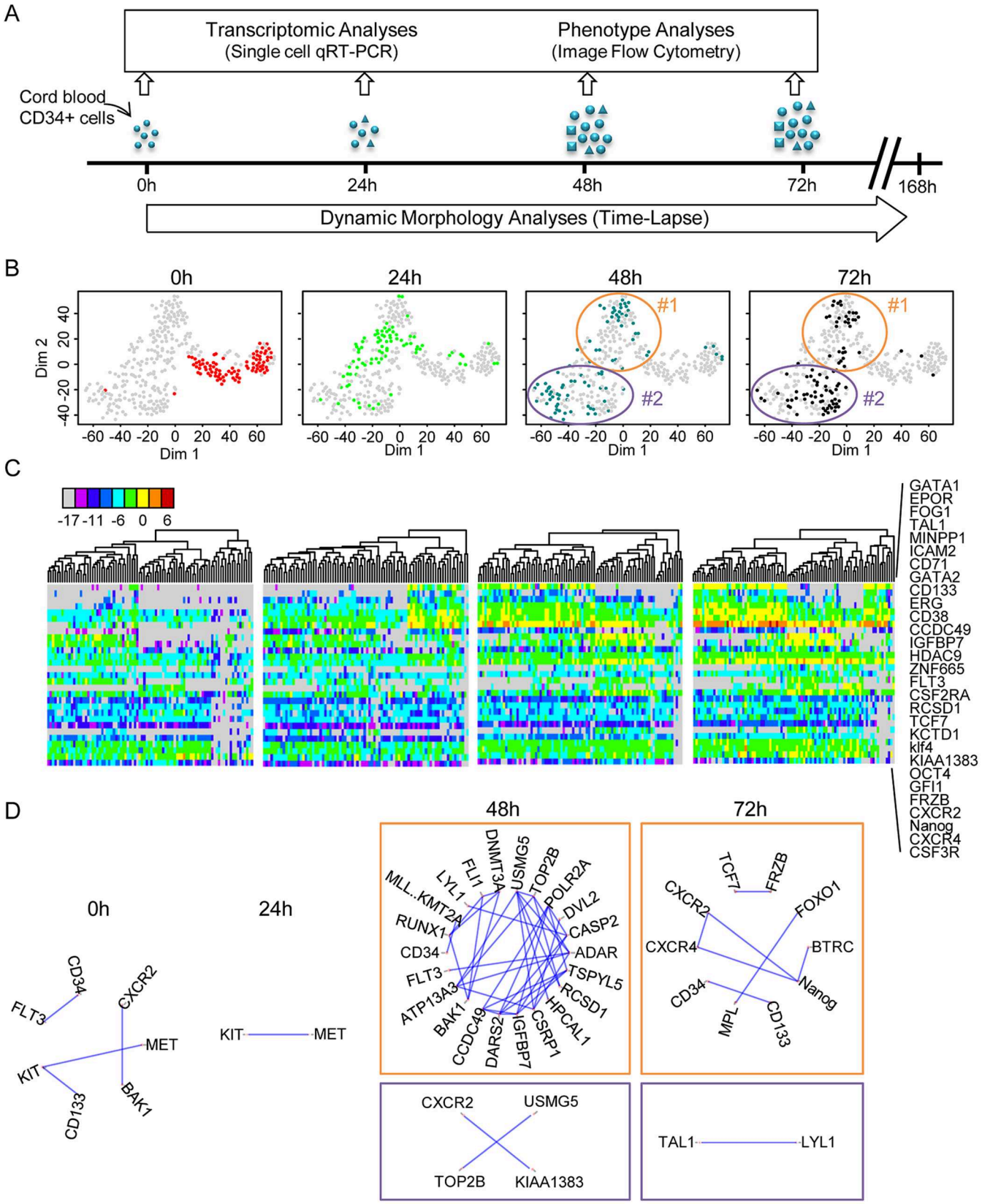

A
Transcriptomic Analyses
(Single cell qRT-PCR)
Phenotype Analyses
(Image Flow Cytometry)
Cord blood
CD34+ cells
0h
24h
48h
72h
168h
Dynamic Morphology Analyses (Time-Lapse)
B
0h
24h
48h
72h
#1
#2
Dim 1
Dim 2
C
-17 -11 -6 0 6
GATA1
EPOR
FOG1
TAL1
MINPP1
ICAM2
CD71
GATA2
CD133
ERG
CD38
CCDC49
IGFBP7
HDAC9
ZNF665
FLT3
CSF2RA
RCSD1
TCF7
KCTD1
klf4
KIAA1383
OCT4
GFI1
FRZB
CXCR2
Nanog
CXCR4
CSF3R
D
0h
24h
48h
72h
CD34
FLT3
CXCR2
MET
KIT
CD133
BAK1
KIT
MET
DNMT3A
USMG5
TOP2B
FLI1
LYL1
MLL..KMT2A
POLR2A
DVL2
RUNX1
CASP2
CD34
ADAR
FLT3
TSPYL5
ATP13A3
RCSD1
BAK1
HPCAL1
CCDC49
CSRP1
DARS2
IGFBP7
TCF7
FRZB
CXCR2
FOXO1
CXCR4
BTRC
CD34
Nanog
MPL
CD133
CXCR2
USMG5
TOP2B
KIAA1383
TAL1
LYL1

Fig 1. Transcriptional profile of cord blood-derived CD34+ cells at t = 0 h, t = 24 h, t = 48 h, and t = 72 h after the beginning of the experiment. (A) CD34+ cells were isolated from human cord blood and cultured in serum-free medium with early acting cytokines. Single-cell quantitative reverse transcription polymerase chain reaction (qRT-PCR) was used to analyse single-cell transcription at 0 h, 24 h, 48 h and 72 h. At the same time, individual clones were continuously monitored using time-lapse microscopy. (B) t-distributed stochastic neighbour embedding (t-SNE) map of single-cell transcription data. The 4 panels show analysis of the same data set, with each point representing a single cell highlighted in different colours depending on the given time point. The 2 clusters identified by gap statistics at t = 48 h and t = 72 h are surrounded by an ellipse and numbered #1 and #2 for multipotent and common myeloid progenitor (CMP)-like cells. Note the rapid evolution of the expression profiles. (Underlying data can be found in S1 Data.) (C) A heat map representation of the expression levels of a subset of genes that strongly contributed to the differentiation of the different groups (as detected by principal component analysis [PCA]; see S2 Fig) and cluster analysis of expression profiles at the different time points show the rapid evolution of gene expression. The list of the genes used (shown on the right) includes well-known genes acting during hematopoietic differentiation but also many randomly selected genes. The colour code for expression levels is indicated below. (Underlying data can be found in S1 Data.) (D) Pairwise correlations between the genes analysed using single-cell quantitative reverse transcription polymerase chain reaction (qRT-PCR). Only the gene pairs with a Pearson correlation coefficient higher than 0.8 are indicated for each time point. The 2 clusters identified at t = 48 h and t = 72 h are represented separately. Note the transient increase of the average correlation in cluster #2 at t = 48 h, indicating a state transition. (Underlying data can be found in S1 Data.)

https://doi.org/10.1371/journal.pbio.2001867.g001

the cells formed a single but dispersed cluster, well separated from the t = 0 h cells. In a fraction of cells, moderate to high transcription of previously nonexpressed hematopoietic regulator genes was observed in addition to that already seen at t = 0 h. For example, the expression of GATA1, GATA2, PU1, CD71, FOG1, CD133, or EPOR increased or was more frequent than at t = 0 h. In some cells, all these genes were expressed simultaneously. Nevertheless, no distinct subpopulations could be identified at the resolution of our approach. The pairwise correlation coefficients between genes remained low (Fig 1D). It is therefore likely that the patterns observed at 24 h resulted from essentially uncoordinated up-regulation of gene transcription and led to a highly heterogeneous cell population. This is a transition state reminiscent of the reported multilineage primed state with simultaneous expression of lineage-affiliated genes specifying alternative cell fates [12,15].

The first signs of coordinated differential gene expression appeared at t = 48 h after cytokine stimulation. At this stage, 2 distinct gene expression patterns emerged from the highly variable background of earlier stages. The 2 clusters are clearly distinguishable on the t-SNE plot (Fig 1B) and identified by gap statistics. They are also easily seen on the heat map representing gene expression levels (Fig 1C). Cluster #2 comprised cells with simultaneous expression of genes characteristic of CMPs such as GATA1 and EPOR [7]. The expression profile of the cells in cluster #1 was characterised by the strong expression of genes reported for multipotent cells (CD133, GFI1, KLF4, or FLT3) and the lack of expression of GATA1 and EPOR. Although this pattern is reminiscent of a self-renewing, multipotent profile, it is difficult to determine the exact identity of these cells at the level of resolution used in our study [26]. Typical genes for pluripotent stem cells like NANOG and OCT4 were expressed at moderate levels in many cells from both clusters (Fig 1C, S1 Fig and S2 Fig). Randomly selected genes were also good predictors for the 2 groups of cells. Only a small fraction of cells could not be classified in 1 of the 2 main clusters at t = 48 h (Fig 1B). The tendency observed at t = 48 h was further reinforced by t = 72 h. The cells in cluster #2 with CMP-like profiles represented more than half of all cells (Fig 1B and 1C). We observed a strong but transient increase in the number of highly correlated genes in this group (Fig 1D). Such an increase in the overall gene-to-gene correlation is a typical hallmark of imminent state transition in these cells [27–29]. Indirectly, this suggested that the cluster #1 profile was more in continuity with the previous profile observed at t = 24 h and that the cluster #2 profile at t = 48 h represented a transition to a new pattern.

Taken together, these single-cell gene expression observations revealed that the cell fate decision process in cytokine-stimulated CD34+ cord blood cells occurred during the first 2 d. Initially, each cell responds to cytokine stimulation with an uncoordinated change in gene expression, which is followed by the emergence of 2 distinct gene expression patterns reminiscent of the 2 known major types of hematopoietic progenitor cells. Although indications of

this second change may appear as early as 24 h after stimulation, the 2 distinct gene expression patterns are clearly distinguishable at 48 h and consolidated by 72 h. By this stage, almost every cell seems to have adopted 1 profile or the other.

Time-lapse tracking studies

In order to integrate the gene expression snapshots into a dynamic scenario, we made time-lapse records of individual CD34+ cells under in vitro conditions identical to those in the single-cell gene expression studies. We imaged individual cells in microwells at a rate of 60 frames per hour for 7 d (Fig 2A). Using a semiautomatic image analysis approach, we established individual clonal pedigrees and recorded cell cycle durations and major morphological changes. As shown in Fig 2B and 2C, the pedigrees of individual clones were highly variable but shared some general features. Some clones produced only a few cells during the observation period, while others proliferated faster and produced up to 30–40 siblings. We focused our attention on the first 3 generations. As reported for cells cultured in early acting cytokines [30], the first cell cycle was exceptionally long in all clones. The division of the founder cell occurred between 35 h and 80 h after the start of culturing, with the median cell cycle length being 58 h (Fig 2). We questioned whether the culturing of isolated cells in microwells, in which direct contact with the other cells was not possible, influenced cell cycle length. To measure the division rate in a population context, the cells cultured together were labelled using CellTrace Violet (CTV). The results (S4 Fig) showed that the cells had similar division profiles regardless of whether they were cultured individually or in population. The unusually long first cell cycle was particularly important when interpreting results. It implied that the transition from the initial to the multilineage primed transcription profile followed by 1 of the 2 types of progenitor-like profiles observed at 24 h and 48 h after CD34+ cell stimulation occurred during the life of the founder cell, before the first mitosis.

Previous studies have demonstrated that there is a connection between cell morphology and the differentiation potential of CD34+ cells. Two major morphological forms have been described in the CD34+ cord blood cell fraction. Polarised cells are capable of active motion with the help of lamellipodia and possess, on their opposite end, large protrusions called uropods. These cells have been found to retain primitive self-renewing and stem cell functions [31,32]. The second morphological type is round. These cells have been considered as already engaged in differentiation [31,32].

Time-lapse records revealed that the 2 cell morphologies were not permanent; most cells were able to switch between forms several times during the cell cycle. After recovering from the stress of isolation and manipulation, founder cells acquired polarised morphologies within a few hours, developing uropods and starting to move actively (see S1–S3 Movies). During the first cell cycle, cells mostly conserved the polarised form, and switches between the 2 morphologies were rare. As indicated above, the first cell division occurred (on average) at 58 h, and the average lengths of subsequent cell cycles were around 20 h to 22 h. The daughter (second generation) and granddaughter (third generation) cells were able to switch between the 2 morphologies at a much higher frequency compared to the founder cells. In order to quantify these events, we manually tracked each cell and recorded each switch. Representative profiles are presented in S5 Fig.

In order to compare quantitatively the dynamic phenotype of cells, we calculated 3 parameters based on their dynamic profiles. The first parameter was calculated as a ratio of the time a given cell spent in a round shape compared to the time spent in a polarised shape. This parameter was close to 0 for stable polarised cells and 1 for stable round cells. Intermediate values correspond to the fraction of time cells spent in round shape. The second parameter was the

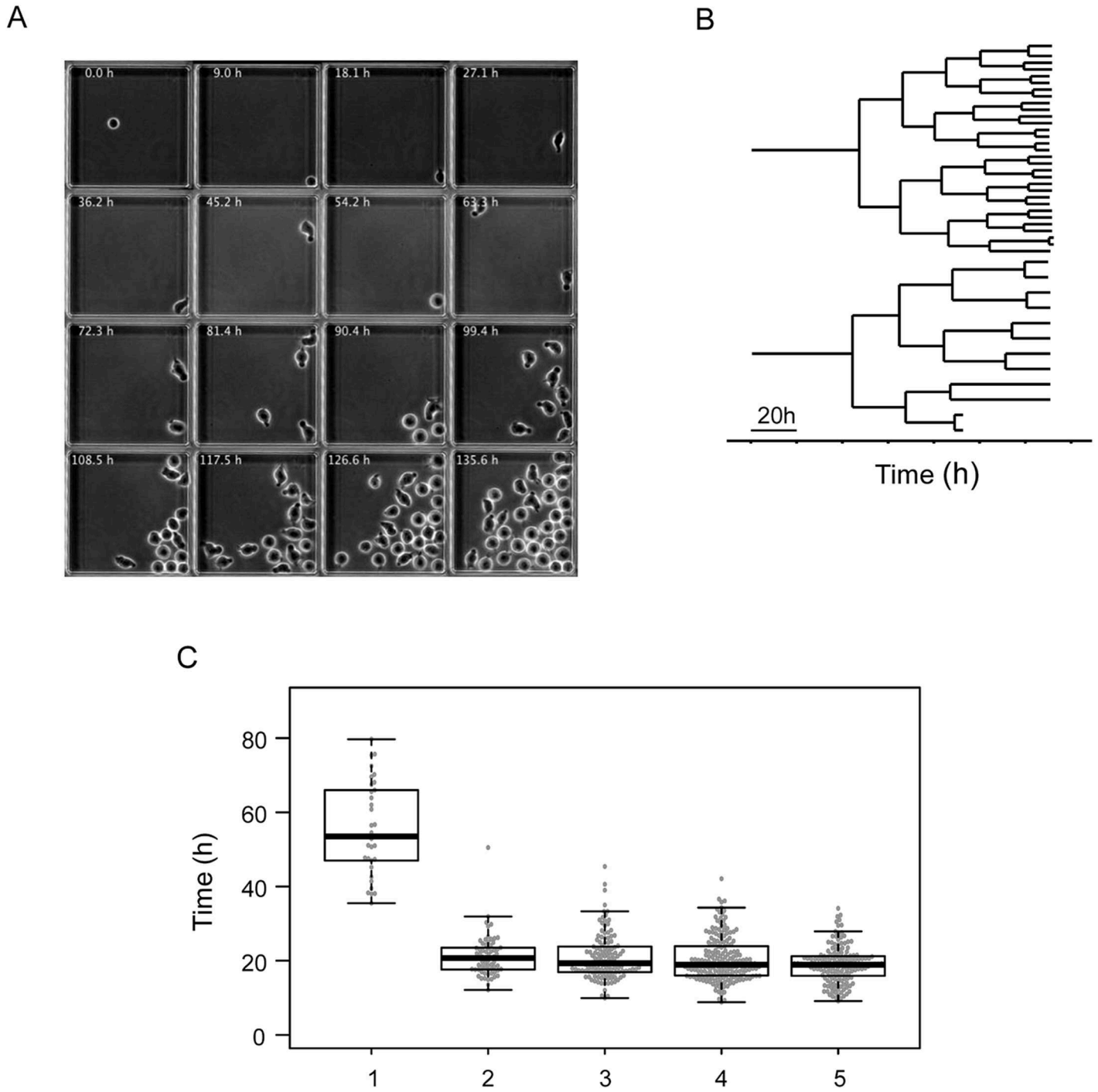


Fig 2. Time-lapse tracking of individual clones. (A) These frames, extracted from a representative time-lapse record, show a microwell containing a single founder cell, which divides to produce a clone. Each individual cell was tracked, and their morphological characteristics were recorded. (B) Two representative lineage pedigrees obtained from the time-lapse record. The strong difference in clone size observed at the end of the record is established gradually after the third cell division. (C) Box plot representation of cell cycle lengths obtained from the time-lapse records of every clone. Note the long first cell cycle. Subsequent cell cycles have comparable lengths, with a slight tendency to become shorter. (Underlying data can be found in S2 Data.)

https://doi.org/10.1371/journal.pbio.2001867.g002

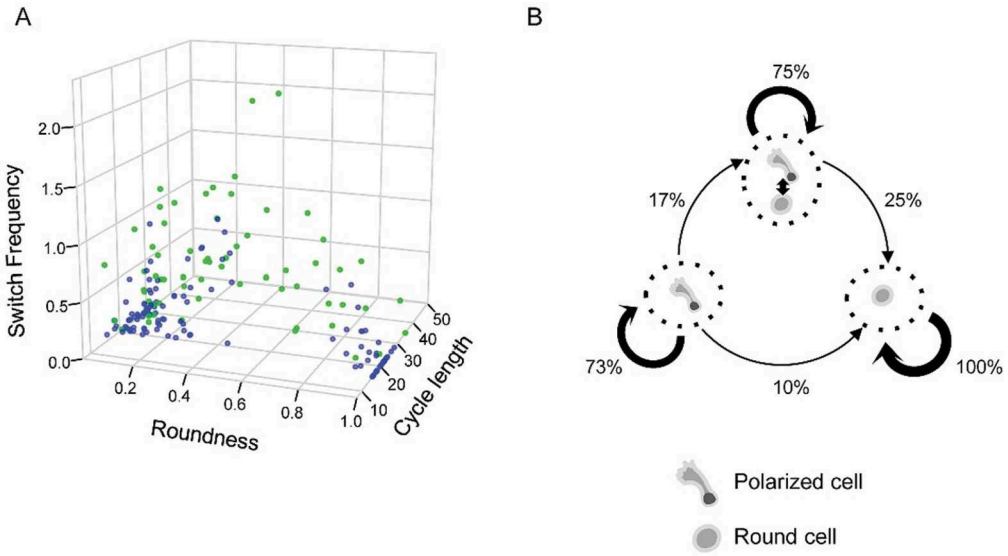


Fig 3. Quantitative analysis of dynamic phenotypes as determined by time-lapse data. (A) Association between the morphology, switch frequency, cell cycle length, and the type of cell divisions of second- and third-generation cells. Each point represents a single cell. Siblings with different dynamic behaviour and morphology (in green) are usually characterised by high switch frequencies. Siblings with similar dynamic behaviour and morphologies are shown in blue. The morphology is given as a ratio of time spent in round/polarised shape by a cell during the cell cycle. Switch frequency is given in number of morphological transformations per hour. Cell cycle length is in hours. (B) Dynamic phenotype change during the first 2 cell divisions as determined on the basis of time-lapse records. Three different dynamic phenotypes were identified: stable polarised, frequent switchers, and stable round. Cells tended to transmit dynamic phenotypes to daughter cells during cell division. Polarised and frequent switchers produced round cells, and frequent switchers were always produced by polarised mothers. Phenotypic change is not associated with asymmetric division; it can occur at any time in the cell cycle. Since round cells always produce round daughters, the whole process is biased and the proportion of this phenotype increases. (Underlying data can be found in S2 Data.)

https://doi.org/10.1371/journal.pbio.2001867.g003

frequency of morphological switches during the cell cycle. This parameter expressed the cell's ability to maintain a stable morphology. The third parameter was the cell cycle length. When cells were represented as individual points in the space determined by the 3 parameters, we identified 3 major categories (Fig 3A). The first category included cells with mainly polarised shapes; the second category was composed of cells with predominantly round shapes; the cells in the third category switched shape frequently, generally fluctuating between both morphologies (Fig 3A).

When sister cell pairs were examined, it became obvious that many displayed very similar dynamic phenotypes. In some cases, periods of stable morphology and switching events coincided almost perfectly (S5 Fig). In other cases, the 2 sister cells behaved differently. In the most extreme cases, 1 sister cell adopted a stable round form and the other a stable polarised form immediately after division.

We calculated the frequency with which a cell with a given dynamic phenotype was produced by a mother cell with similar or dissimilar phenotype (Fig 3B). Maternal cells clearly tended to transmit the dynamic phenotype to daughter cells. We also observed the regularity with which phenotype changes occurred in daughter cells. Polarised cells were systematically produced by polarised cells. At lower probabilities, both polarised and fluctuating cells could produce stable round phenotype cells. Round cells always gave rise to round siblings (Fig 3B). Following these simple rules, the cumulative outcome of the process was the gradual increase of round cells in the population. Cells with fluctuating morphologies appeared to be an intermediate form between polarised and round cells. Since 25% of daughter cells conserved this

phenotype, the fluctuating intermediate cells persisted in the population. On static snapshots, however, this category remained undetectable: only polarised and round cells were observed. A polarised form was considered to be a feature of multipotent cells and the round form a committed myeloid progenitor phenotype [31,32].

Coupling the molecular and cellular scales

The dynamically fluctuating behaviour we have described here for the first time represents a transition between the 2 states and reflects a 'hesitant', incomplete fate-determination process. Since we detected only 2 major transcription profiles but observed 3 different dynamic behaviours, it is possible that 'hesitant' cells are not characterised by a clearly distinct transcription pattern. Morphology fluctuations may be accompanied by fluctuations in the transcript or protein levels of at least some key genes.

To test this assumption, we took advantage of the observation that the gene coding for the CD133 cell surface protein was expressed preferentially in 1 of the 2 transcription patterns detected at 48 h (Fig 1B and 1C). Previous reports have established that CD133 protein is typically present in cells with polarised forms and accumulates in the uropod [31–33]. We confirmed this using image cytometry and immunohistochemistry on fixed cells (S6 Fig). Cells expressing high levels of CD133 were mostly polarised, while those with low levels of CD133 were round (S6 Fig). This observation explicitly established a direct link between the cell morphology and the transcription patterns detected by single-cell qRT-PCR.

We therefore used the CD133 protein as a proxy for the isolation of a cell fraction enriched in either polarised or round cells and recorded their dynamic phenotype. The 'high', 'medium', and 'low/negative' CD133-expressing cell fractions were isolated 48 h after cytokine stimulation, put in culture, and tracked using time-lapse microscopy for an additional 48 h (Fig 4A). The fraction of the time the cell spent in round or in polarised shape, the switch frequency, and the division asymmetry of the tracked cells were quantified. The 'high' CD133-expressing cells and their progeny reproduced the 3 types of cells observed previously but in different proportions (Fig 4B). Most of the cells displayed stable polarised morphologies or were frequent switchers; only a few cells displayed stable round morphologies (Fig 4B). By contrast, the 'low/negative' cells produced either stable round progeny or cells with fluctuating morphologies (Fig 4B). The 'medium' CD133-expressing cells had a higher switch frequency, and both shapes were represented in a more equilibrated manner (Fig 4B).

These observations confirmed the idea that cells with stable round shapes were derived from cells with polarised shapes and high CD133 levels following a period in which they had a fluctuating phenotype. The fluctuations occurred in a bi-stable manner; the cells switched from 1 morphology to another and back rapidly without stable intermediate states (see S1–S3 Movies). The process of transformation did not correlate with the cell cycle; some cells reached a stable morphology rapidly, while others fluctuated over several cycles. The process was accompanied by a gradual decrease in CD133 protein levels in cells. We found no evidence that asymmetric divisions played a direct role in this process.

Next, we isolated individual cells with high, medium, and low/negative CD133 protein levels using a cell sorter and performed single-cell qRT-PCR analysis using the same gene panel used previously on the unsorted population. Fig 5A shows the t-SNE representation of the single-cell gene expression profiles. The heat map representation of the full set of gene expression results is shown in S7 Fig together with the PCA analysis. The 'high' and 'low' CD133 cells displayed different transcription profiles similar to clusters #1 and #2, respectively, found in cells of the unsorted population at t = 48 h (Fig 1). The cell fraction isolated on the basis of intermediate CD133 levels contained a large number of cells with intermediate transcription profiles,

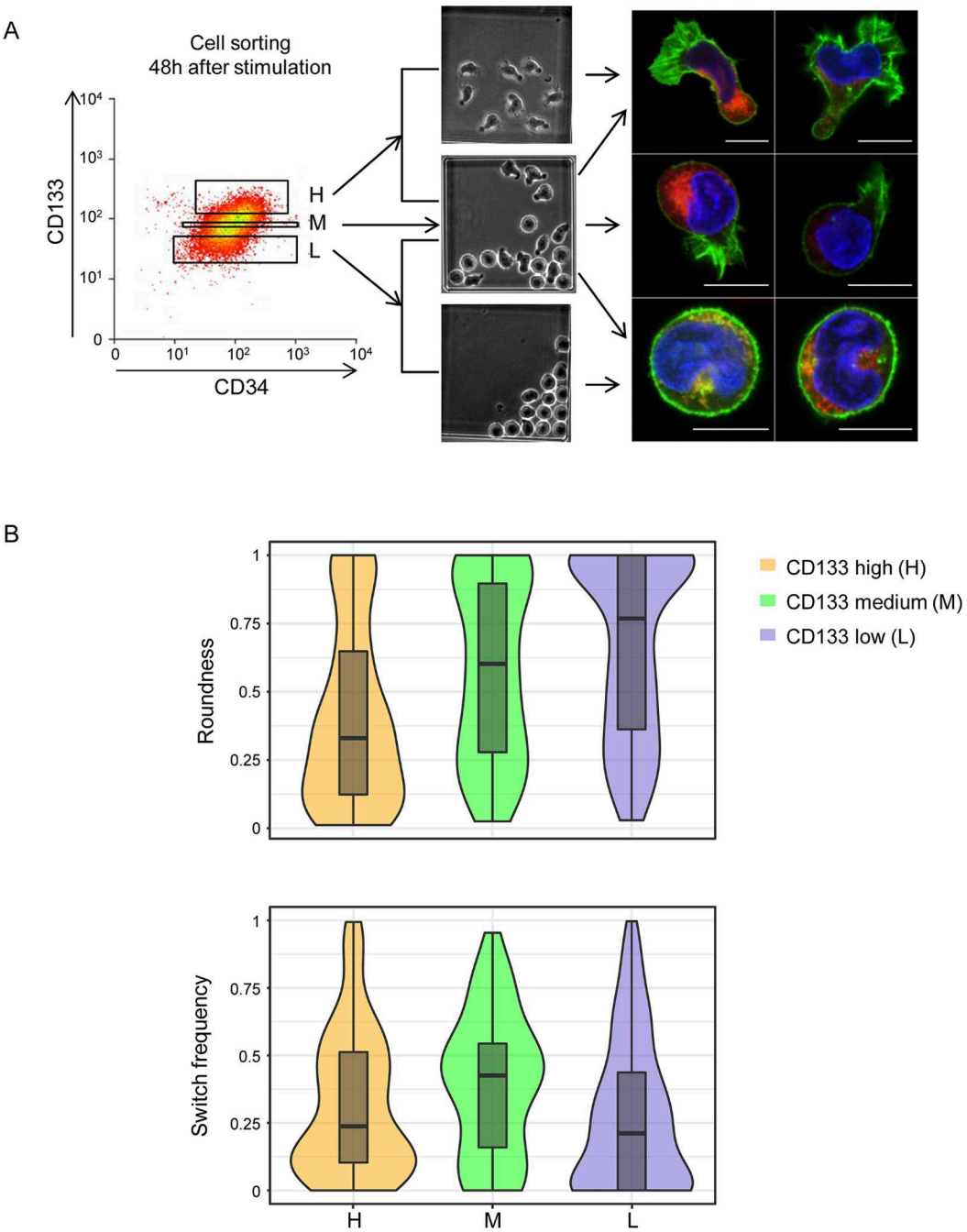


Fig 4. Isolation and time-lapse analysis of 'high', 'medium', and 'low' CD133-expressing expressing cells. (A) Cell-sorting strategy to isolate cells on the basis of the CD133 surface protein level. The sorted cells were cultured individually and tracked by time-lapse microscopy. They produced cells with a polarised, round, or fluctuating dynamic phenotype (illustrated by the middle panel). Examples of cells with different morphologies are shown on the right, as detected by confocal microscopy. Red: CD133 protein. Green: actin filaments detected by phalloidine. Blue: DNA. Note the preferential localisation of the CD133 protein in the uropods of polarised cells. Actin is concentrated in lammelipodia or evenly distributed in the periphery of round cells. (B) Quantitative evaluation of cell morphology and switch frequency. Distribution of the 'roundness' parameter (upper panel) indicates a gradual increase of the proportion of round cells between the 'high' and 'low' fraction. Distribution of the switch frequency as switch/h of sorted 'high', 'medium', and 'low' CD133 cells is shown in the lower panel. Note that the switch frequency is the highest in 'medium' CD133 cells. (Underlying data can be found in S2 Data.)

https://doi.org/10.1371/journal.pbio.2001867.g004

again supporting their dynamic transitory phenotype. When expression pattern of individual genes is examined (for example, CXCR4, CXCR2, DVL2, FOXO3, NANOG, ZNF665, or TSPYL5 [S8 Fig]), some of them displayed a very different distribution in the 'medium CD133' compared to the 'high' and 'low' CD133 cells. This further demonstrates that the 'medium' CD133 cells are more than simple intermediates between the 'high' and 'low' cells; they have their own dynamic transcription profile. One can conjecture that if the cell shape correlates with the CD133 level that correlates with the transcription profile, then those cells that change shape must also change transcription profile.

PU1 and GATA1 are well-known transcription factor-coding genes that play an important role in the specification of granulocytic–monocytic and erythroid–megakaryocitic cells [34]. It has been proposed that PU1 and GATA1 can cross-inhibit each other's activity and generate a bi-stable switch between the 2 lineages [35], but more recent observations challenged this model [21]. Our analysis showed that 'high' CD133 cells express only PU1; 'medium' CD133 cells express PU1 only or coexpress the 2 genes; and 'low' CD133 cells express PU1 only, coexpress the 2 genes, or express GATA1 only (Fig 5B). This observation places the 'medium' CD133 cells as a possible intermediate between the cells expressing only PU1 and cells committed to different pathways, without providing evidence either for or against a direct competition between them.

Single-cell transcription profile of the multipotent stage

In order to determine which of the observed phenotypes correspond to the multipotent stage, we took advantage of recent observations demonstrating that the inhibition of histone deacetylase (HDAC) activity with a pharmacological agent resulted in a substantial increase in their incidence in the CD34+ cord blood population [13,36,37]. We anticipated that this would increase the proportion of cells with transcription profiles typical of the multipotential phenotype. Since valproic acid (VPA) was shown to be the most efficient [36], we used this agent to treat CD34+ cord blood cells stimulated by cytokines as above, before sampling transcription profiles. The increase of the CD90 marker (as analysed by flow cytometry) confirmed that the VPA effect was already visible after 24 h and gradually grew stronger during subsequent steps (Fig 6A and S9 Fig). The expression of CD34 and CD38 markers remained unchanged (S9 Fig). Although we did not analyse the in vivo potential of these cells, based on previous reports, we considered them enriched for bona fide multipotent cells. We performed single-cell qRT-PCR at 0 h, 24 h, 48 h, and 72 h after the start of the experiment, as in control cells. At all 4 time points, cell populations were very heterogeneous. At each time point, the cells displayed a unique transcription profile (Fig 6B), and no identifiable transcription patterns appeared during the 72 h of the experiment, despite slight profile evolutions. Overall, transcription patterns in individual cells were reminiscent of the uncoordinated multilineage primed profile detected in control cells at 24 h, but the 2 groups clustered separately on t-SNE maps (Fig 6C). Since the cells did not divide during the first 48 h, the increase observed in the multipotent cell fraction could not result from the selective proliferation of an initially small subpopulation of cells. Instead, this occurred because cells already present in the population changed the expression of many genes in response to the VPA.

Discussion

In this study, we aimed to identify the initial stages of fate commitment in the CD34+ cell fraction of human cord blood and determine the typical time scale for these events. Without cytokine stimulation, CD34+ cells remain quiescent and die after a few days in culture. Early acting cytokines allow these cells to survive, become metabolically active, and enter the cell cycle [38] without showing overt signs of differentiation during the first few days. This creates ideal conditions for studying early events. Our experimental design combined continuous time-lapse observations

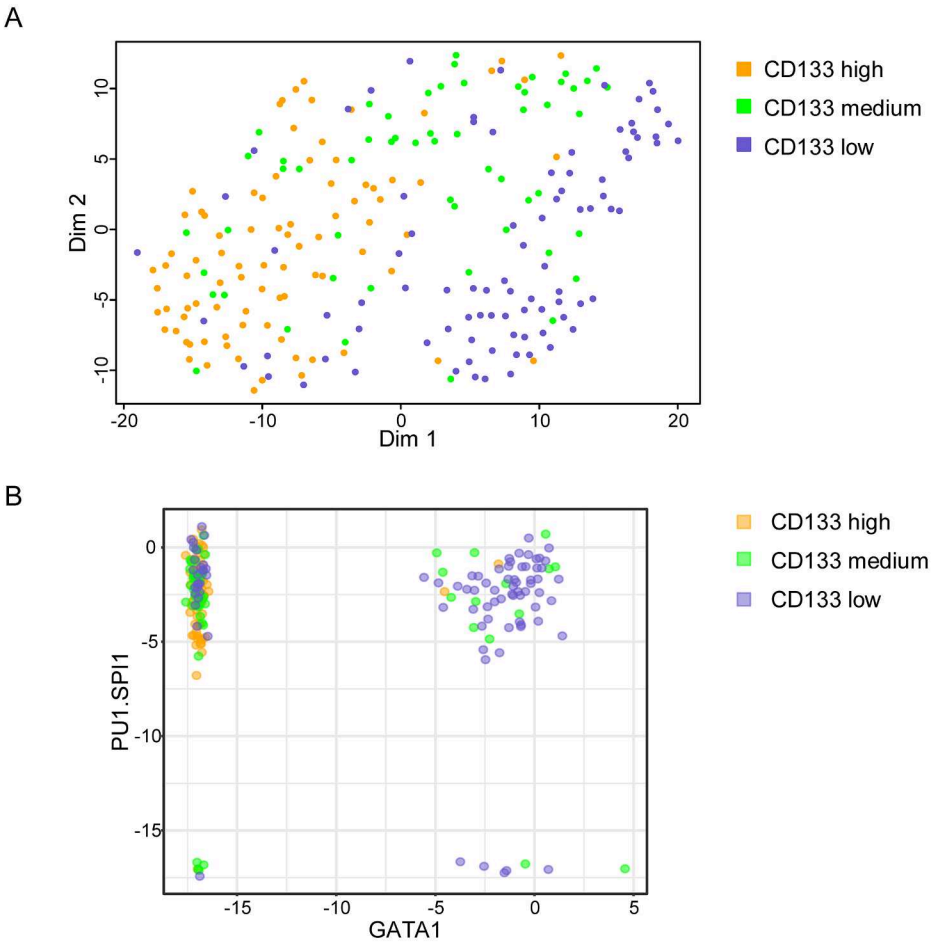


Fig 5. Single-cell gene expression in 'high', 'medium', and 'low' CD133 cells. (A) t-stochastic neighbour embedding (t-SNE) map of single-cell transcriptional data. Each point represents a single cell highlighted in a different colour for 'high', 'medium', and 'low' CD133 cells. 'High' and 'low' cells are in separated clusters corresponding to cluster #1 and #2 in Fig 1B. 'Medium' CD133 cells are distributed in and between these 2 clusters, indicating their intermediate character. (B) Scatter plot representation of PU1 and GATA1 expression in individual cells of the 'high', 'medium', and 'low' CD133 fraction. Note that GATA1 is not expressed in 'high' cells. Coexpression of the 2 genes is observed only in some 'medium' and 'low' cells. (Underlying data can be found in S1 Data.)

https://doi.org/10.1371/journal.pbio.2001867.g005

with snapshots of high-resolution single-cell transcriptome analysis. The data can be integrated in a dynamic fate-decision process scenario. Fate decision is necessarily accompanied by a change in the gene expression pattern. This is a multistep process. First, upon stimulation, cells rapidly reach the multiprimed state, which is characterised by a promiscuous gene expression pattern and predominantly polarised morphologies. This is an unstable phase, and 2 distinct transcription profiles start to emerge before the end of the first cell cycle. The process by which cells relax from a multiprimed to a more stable state is continuous and of variable length. Some cells reach stable morphology and a coherent, lineage-affiliated transcription profile by the end of the first cell cycle, which they then transmit to daughter cells. Other cells divide into unstable daughter cells with dynamic, 'hesitant' behaviour. This behaviour is characterised by fluctuations between polarised, actively moving amoeboid and round morphologies over several cell cycles, suggesting that instability can be transmitted mitotically. Although we have no formal evidence that the

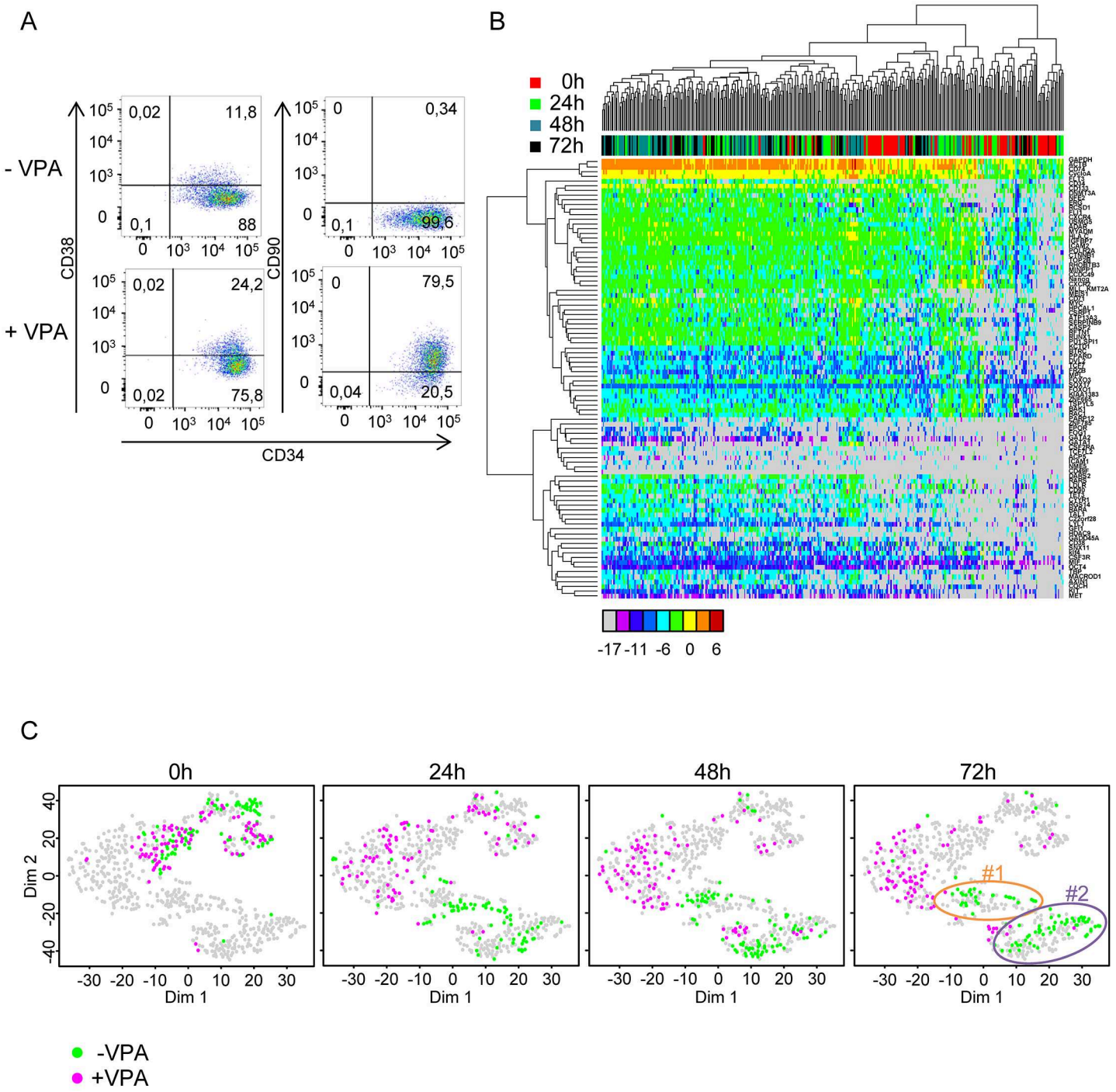


Fig 6. Transcriptional profile of cord blood-derived CD34+ cells treated with valproic acid (VPA) at t = 0 h, t = 24 h, t = 48 h, and t = 72 h after the beginning of the experiment as compared to untreated, normal control cells. (A) A cytometric analysis of the effect of VPA on cord blood CD34+ cells shows an increase in the CD90 protein in most cells, while the CD34 and CD38 markers remain essentially unchanged. (B) Heat map representation of the expression levels of 90 genes as determined by single-cell quantitative reverse transcription polymerase chain reaction (qRT-PCR) in VPA-treated cells at t = 0 h, t = 24 h, t = 48 h, and t = 72 h. The colour codes for the time points of cells are indicated on the right; the colour codes for expression levels are indicated below the heat map. Note the high heterogeneity and lack of clear clustering of the expression patterns. (C) t-distributed stochastic neighbour embedding (t-SNE) plot representation of transcription data obtained for VPA-treated cells compared to untreated normal cells (data for these cells are the same as in Fig 1). The gene expression data obtained in the 2 experiments were mapped together. Each point represents a single cell, and the cells at t = 0 h, t = 24 h, t = 48 h, and t = 72 h are highlighted separately in the 4 panels. The colour codes for VPA-treated (+VPA) and VPA-untreated (−VPA) are indicated below the panels. Clusters #1 and #2, identified at t = 48 h and t = 72 h in −VPA cells (see Fig 1), are indicated on the t = 72 h panel. Note the clear separation

of the +VPA and −VPA cells at every time point except t = 24 h. Note also that +VPA cells do not contribute to clusters #1 and #2, indicating that they do not acquire expression profiles typical of these cells. (Underlying data can be found in S1 Data.)

https://doi.org/10.1371/journal.pbio.2001867.g006

transcriptome of these cells also fluctuates, 2 observations suggest that this could be the case. First, we only found 2 established transcription profiles that correspond to polarised and round morphologies with 'high' and 'low' CD133 protein levels (Fig 1B). However, we observe 3 dynamic phenotypes, 1 of which is fluctuating. Second, cells isolated on the basis of having 'medium' CD133 protein levels represent a transition between the stable polarised and round morphology. We propose a dynamic scenario in which the initial stochastic multilineage primed state is followed by a period of relaxation and uncertain 'hesitant' phase of variable length with fluctuating transcriptomes [15,17] before a stable lineage committed state is reached. In addition, this scenario is also in accordance with the recent proposal that there exists a fraction of low-primed, undifferentiated cells called 'CLOUD'-HSPCs in which that can fluctuate without passing through fixed, discrete states [10].

Increased stochastic variation in gene expression may be responsible for the rapid shift away from the initial quiescent state and lead to the uncommitted multilineage primed state [10,12,15,16]. Cell division is not required for this process; it occurs during the first cell cycle following stimulation. Cells on the path toward the new phenotype represent the committed state. The critical moment in this process is the transition between the 2 phenotypes, when the old gene network has broken down but the new network is not yet assembled. We consider that cells with fluctuating morphologies represent this transition state. The rapidity of the transition may be dependent on the time required for the new gene expression network to settle into a stable state. Since phenotypic stability of a cell lineage largely depends on the frequency of transcription initiation and the stability of the resulting mRNAs and proteins [39,40], the observed 'hesitant' phenotype might be the consequence of stochastic fluctuations due to rapid mRNA and protein turnover. The consolidation of the chromatin structure appears to be an essential element in this process, because, as shown in single-cell transcription studies, the HDAC inhibitor VPA delays the transition and blocks cells in a promiscuous gene expression pattern typical of a multilineage primed state. Indeed, HDAC inhibitors have been shown to increase gene expression stochasticity by increasing chromatin acetylation [41].

In summary, in this study, we identified the earliest phases of fate commitment in human cord blood CD34+ cells and assigned a time scale to this process. We demonstrated that the rapid initiation of the process occurs within a single cell cycle and is followed by a dynamic transition state of variable length that may span several cell cycles. Since experimental conditions were constant, the changes observed are likely to reflect cell-intrinsic processes, whereas the convergence toward a similar endpoint may reflect the constraints imposed by these conditions. From this perspective, fate decision appears to be a process of spontaneous variation/selective stabilisation reminiscent of trial–error learning, in which each cell explores many different possibilities at its own pace by expressing a large variety of genes before finding a stable combination corresponding to the actual environment. This is in remarkable agreement with earlier theoretical predictions and experimental work [42–46]. At least 3 independent theoretical models predicted the existence of an initial fluctuating phase during differentiation. According to the first theory, cell differentiation is a variation/selection process analogous to evolution [42,43]. Variations are created by stochastic fluctuations of gene expression, and some patterns are selectively stabilised through interactions with the environment and neighbouring cells. Another approach envisions cell phenotype as an attractor state in the parameter space defined by the gene expression network [47]. Differentiation is seen as a transition from 1 attractor to another and governed by the stochastic dynamics and self-organisation of the

gene network. Finally, a dynamic system view of differentiation was independently proposed by Kaneko [48]. A common theme of these approaches is the prediction that differentiating cells must necessarily go through a dynamically fluctuating phase with oscillating gene expression. Several recent studies have reported on the existence of gene expression fluctuations during the critical state transitory phase of the differentiation process [15,29,46,49–51]. Our study goes a step further by demonstrating that the cellular phenotype also fluctuates during the critical transitory phase.

Materials and methods

Ethics statement

Human umbilical cord blood (UCB) was collected from placentas and/or umbilical cords obtained from Etablissement Français du Sang (EFS), Saint Louis Hospital, France or from Centre Hospitalier Sud Francilien, Evry, France in accordance with international ethical principles and French national law (bioethics law n° 2011–814) under declaration N° DC-201-1655 to the French Ministry of Research and Higher Studies.

Human sample and cell culture

Human CD34+ cells were isolated from the mononuclear fraction of UCB samples using the autoMACSpro (Miltenyi Biotec, Paris, France) immunomagnetic cell separation system. They were then cryopreserved in Cryostor (StemCell, Paris, France) and stored in liquid nitrogen or used directly without freezing.

Cells were cultured at 37°C in a humidified atmosphere containing 5% CO_2 in a 24-well plate in X-VIVO (Lonza) supplemented with 100 U/ml penicillin, 100 μg/ml streptomycin (Gibco, Thermo Scientific), 50 ng/ml h-FLT3, 25 ng/ml h-SCF, 25 ng/ml h-TPO, and 10 ng/ml h-IL3 (Miltenyi Biotec, Paris, France) final concentration. VPA (Sigma Aldrich) was used at a final concentration of 1.25 mM.

Single-cell qRT-PCR

Single-cell qRT-PCR was carried out using the BioMark HD System (Fluidigm). Deltagenes assays (Life Technologies) were used at a final concentration of 500 nM for each of the 96 assays. Individual cells were sorted directly into a reverse transcription RT mix solution and spikes (Life Technologies) in a 96-well plate. RNA was denatured and reverse-transcribed. Twenty cycles of preamplification of 96 specific cDNA were performed by denaturing the cDNA at 96°C for 5 seconds, followed by annealing and extension at 60°C for 4 min. Unincorporated primers were cleaned up by Exonuclease I, and the preamplified products were diluted 5-fold. Amplification was performed with Evagreen supermix with low ROX (Bio-Rad) and inventoried DeltaGenes assays in 96.96 Dynamic Arrays on a BioMark HD System (Fluidigm). Cycle threshold (Ct) values were calculated from the system's software (BioMark Real-Time PCR Analysis, Fluidigm).

Single-cell data normalisation

Ct values obtained from the BioMark HD System (Fluidigm) were normalised with the help of 2 externally added controls (spike 1 and spike 4, Life Technologies) according to a set of rules provided below. For each gene, inconsistent readings or 'Failed' quality control readings were removed. Cells with failed or inconsistent detection of spikes were removed. Expression values were calculated by subtracting the gene Ct value from the geometric average of Ct values from

spike 1 and spike 4 in the corresponding cell. An arbitrary differential cycle threshold (dCt) value of −17 was assigned for all the genes with a dCt value less than −17.

Single-cell qRT-PCR data analysis

Analyses of qRT-PCR single-cell data were done with R software (R Core Team [2016]. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, http://www.R-project.org/) using Heatmap3 [52], factomineR [53], k-means, and ggplot2 packages [54]. Correlation calculations were performed using custom R scripts. t-SNE and gap statistics calculations were performed as described by Grun et al. [55]

Confocal microscopy

Images were obtained with a spectral confocal Leica SP8 scanning microscope (Leica Microsystems, Germany). 5.10^4 cells were cultured in a 48-well plate in 200 μL prestimulation medium. After 72 h, 100 μL of 3% glutaraldehyde was added to the cell-containing well (1% final) for 15 min. Cells were washed twice with PBS 1X and incubated 2 h with 2 mg/mL $NaBH_4$ at room temperature. Fcɣ receptors were saturated with Gamma Immune (Sigma Aldrich) for 5 min at 4˚C (1:2 dilution). The cells were permeabilised with the fix/perm kit (BD-Biosciences), labeled for 20 min at 4˚C with a 1:10 dilution of the mouse anti-human CD133-APC antibody (clone Ac133, Miltenyi Biotec), a 1:1,000 dilution of phalloidin–Tetramethylrhodamine B isothiocyanate (Sigma Aldrich) and stained with DAPI.

The images were acquired using a 63X PL APO CS2 1.40 NA oil immersion objective (Leica Microsystems, Germany). DAPI was excited with a 405-nm laser, TRITC with a 552-nm laser, and APC with a 635-nm laser. Finally, images were processed with a contrast enhancement algorithm (histogram equalisation) and a home-designed background subtraction algorithm.

Microgrid cell culture

A polydimethylsiloxane (PDMS) microgrid array (Microsurfaces, Australia) of 1,024 microwells (125-μm width, 60-μm depth) was placed in a specialised culture dish divided into 4 parts (Hi-Q4, Ibidi, Germany). Each part of the dish was filled with cell culture medium. A suspension of 5×10^3 cells per case was added at a concentration likely to lead to a high number of wells with a single cell.

Time-lapse microscopy

The time-lapse microscopy protocol was previously described [56]. Time-lapse acquisitions were performed with the Biostation IM time-lapse microscope (Nikon Instruments, Europe). Twenty field positions were recorded covering 4 microwells each. Images were acquired every minute for 2 d to 7 d using a 20X magnitude phase contrast objective. Only microwells containing a single cell were considered in the analyses.

Image analyses

Images were analysed using ImageJ 1.47g 64-bits software (Rasband, W.S., ImageJ, U.S. National Institutes of Health, Bethesda, Maryland, USA, http://imagej.nih.gov/ij/, 1997–2014). Cell tracking was performed manually using the ImageJ TrackMate plugin. The morphologies of first, second, and third generation cells were analysed semiautomatically with Fiji (ImageJ 1.50e). A cell counter plugin was used to identify the moment when the cell switches from a round to a polarised morphology.

Analysis of the time-lapse records

Analyses of time-lapse data were performed using R software. Cell lineage representations, cycle length, roundness, and switch frequency were calculated with custom R-made scripts. Euclidean distances of the last 3 parameters (cycle length, roundness, and switch frequency) between the 2 sister cells were calculated. Cells were classified into 2 groups using the k-means algorithm: with similar or divergent dynamic phenotypes. Box plot representation combined with individual points was calculated with the beeswarm package (Aron Eklund [2016]. beeswarm: The Bee Swarm Plot, an Alternative to Stripchart. R package version 0.2.3. https://CRAN.R-project.org/package=beeswarm). The ggplot2 package was used to represent the roundness and switch frequency of cells sorted on CD133 protein.

Proliferation assay

CD34+ cells were labeled with 2.5 μM of CTV (Life technologies) at t = 0 h and analysed using flow cytometry (LSRII–BD biosciences, France) after 24 h, 48 h, and 72 h with ModFit LT software as described previously by Neildez et al. [57]

Image flow cytometry

Image flow cytometry analysis was performed using Image Stream MKII (Amnis, Proteigen, Merk Millipore). 5.10^4 cells were cultured in a 48-well plate in 200 μL prestimulation medium. After 72 h, 100 μL of 3% glutaraldehyde was added to the cell-containing well (1% final) for 15 min. Glutaraldehyde offers good preservation of cell shape. Cells were washed twice with PBS 1X and incubated 2 h with 2 mg/mL $NaBH_4$ at room temperature. Fcɣ receptors were saturated with Gamma Immune (Sigma Aldrich) for 5 min at 4˚C (1:2 dilution). Cells were labeled for 20 min at 4˚C with a 1:10 dilution of mouse anti-human CD133-APC antibody (clone AC133, Miltenyi Biotec). Cells were then suspended in PBS and analysed with the image flow cytometer. Bright Field and APC channels were recorded (Bright Field: 745-nm laser; APC: 642-nm laser) with the 40X magnitude objective. Analyses of image stream data were performed with the IDEAS software (Amnis, Proteigen, Merk Millipore).

Cell sorting

The $CD34^+CD133^{high}$, $CD34^+CD133^{medium}$ and $CD34^+CD133^{low/neg}$ cells were sorted at t = 48 h. Prior to labeling, Fc receptors were saturated with Gamma Immune (Sigma Aldrich). The CD34 + cells were labeled with CD34-PE (Miltenyi Biotec), CD45-APC-H7 (Beckman Coulter) and CD133-APC (clone AC133, Miltenyi Biotec) antibodies and 7—Aminoactinomycine D (Sigma Aldrich). Isotype controls were used for the gating strategy. Cells were purified using a MoFlo Astrios cell sorter (Beckman Coulter, France) and analysed with Kaluza software.

Flow cytometric analysis

The CD34+ cells were labeled using the following cell-surface markers: CD34-PE (Miltenyi Biotec), CD38-Pacific Blue (Beckman Coulter), and CD90-APC-Cy7 (Beckman Coulter) antibodies and 7-AAD marker (Sigma Aldrich). Isotype controls were used for gating strategies. Cells were analysed at 72 h after prestimulation by flow cytomety (LSRII–BD biosciences, France) and analysed with FlowJo (v10.1) software.

Supporting information

S1 Fig. Full set of gene expression data obtained using single-cell qRT-PCR in cord-blood CD34+ cells cultured in vivo with early-acting cytokines. Extended heat map of the

transcriptional profiles of cord blood-derived CD34+ cells at t = 0h, t = 24h, t = 48h and t = 72h after the beginning of the experiment. The color codes for the time points of cells are indicated on the right, the color code for expression level are indicated below the heat-map. Note the tendency of cells with the same time-points to cluster. (Underlying data can be found in S1 Data.).
(TIF)

S2 Fig. Violin plot representation of individual gene expression levels at the four time points. (Underlying data can be found in S1 Data.).
(TIF)

S3 Fig. Principal component analysis of single-cell expression profiles. A. 2D PCA plot. Each point represents a single cell and the different time-points are coloured differently. Color codes are in the box to the right of the plot. B. Contribution of individual genes to principal component 1 and 2. Only the 40 highest contributions are indicated. (Underlying data can be found in S1 Data.).
(TIF)

S4 Fig. Analysis of cell division rates. A. The number of cells at t = 24h, t = 48h and t = 72h as observed by time-lapse microscopy. The cells of different generations are color coded in the histogram. Note that none of the cells has divided after 24 hours and only 11 of the 32 cells underwent one division after 48 hours. At t = 72h, three of the founder cells have not undergone division. (Underlying data can be found in S2 Data) B. Cell division analysis using Cell Trace Violet labelling. Cells were labelled at t = 0h (not shown) and analyzed using flow cytometry at t = 24h, t = 48h and t = 72h. When divided, the average fluorescence intensity of the two daughter cells is reduced by half compared to the maternal cell. Therefore, the peak on the right represents the parental generation. The number of the peaks to the left indicates the number of cell generations in the culture and the size of the peaks is indicative of the number of cells in each generation. Note that after 24h no cell division is detected and after 72h a fraction of undivided cells can still be detected. Most of the cells underwent one or two divisions. Overall, the profile is very similar to that detected by time lapse. (Underlying data can be found in S3 Data.).
(TIF)

S5 Fig. Representations of morphological profiles of cells in three representative clones. Each horizontal box in the three panels represents the morphology of an individual cell. The cell morphology–polarized or round–is shown with a horizontal line, the length of which is proportional to the time spent in the corresponding form. Vertical lines show the transitions between forms. The length of the horizontal lines is proportional to duration of the cell cycle and the time scale in hours is the same for each cell. The founder cell is numbered Cell_1, the two daughter cells Cell_11 and Cell_12 and granddaughter cell pairs as Cell_111, Cell_112 and Cell_121, Cell_122 respectively. In clone number 1 the polarized founder cell gives rise to frequent switcher daughters and granddaughters. Note the striking similarity of the time profiles for the morphological switches that can be observed in sister cells. In clone number 2 the polarized founder cell gives rise to stable polarized siblings. In clone number 3 the founder cell and its progeny are round. The two daughter cells switch to polarized shape for short periods. Note again the striking similarity of the sister cells' switch profiles. (Underlying data can be found in S2 Data.).
(TIF)

S6 Fig. Cell morphology and CD133 localisation. Image-based cytometry analysis shows correlation between the CD133 protein expression level and cell morphology at t = 72h. The middle plot shows the CD133 protein density detected in glutaraldehyde-fixed cells.

Representative examples of the morphologies of "high" (upper frame) and "low" (lower frame) expressing cells are shown on the left and right respectively.
(TIF)

S7 Fig. The full set of the gene expression data obtained on "high," "medium," and "low" CD133 expressing individual cells. A. Heat-map representation of the expression levels of 90 genes as determined by single-cell qRT-PCR. Color codes for the "high", "medium" and "low" fractions are indicated on the right, and the color codes for expression levels are indicated below the heat-map. Note the intermediate expression pattern of the "medium" cells. B. Principal component analysis of the single-cell gene expression data shown on the panel A. "Medium" cells are intermediate. (Underlying data can be found in S1 Data.).
(TIF)

S8 Fig. Violin plot representation of individual gene expression levels in the "high," "medium," and "low" CD133 cells. The color code is identical to that on S7 Fig. (Underlying data can be found in S1 Data.).
(TIF)

S9 Fig. Cytometry analysis of the effects of valproic acid on CD34+ cells. The histogram in the left panel indicates the proportion of CD34+/CD38- cells in VPA+ and VPA- cell cultures at different time points. Note that there is no substantial difference between the two. The right panel indicates the proportion of CD34+/CD90+ cells in the same cultures. Note the increasing proportion of CD34+/CD90+ cells in VPA+ culture. This rapid increase cannot be explained by the selective proliferation of the CD90+ cells and is the result of the de novo synthesis of the CD90 protein, because as indicated in Fig 2, and S4 Fig, cells do not divide before 72h. (Underlying data can be found in S3 Data.).
(TIF)

S1 Table. List of genes analyzed and primer sequences used for single-cell qRT-PCR amplification.
(XLSX)

S1 Data. RTqPCR normalized.
(XLSX)

S2 Data. Timelapse.
(XLSX)

S3 Data. Cytometry.
(XLSX)

S1 Movie. Time-lapse video of a cell clone with cells conserving polarized morphologies. The video has been accelerated to 5 frames per second.
(MOV)

S2 Movie. Time-lapse video of a cell clone with cells conserving round morphologies. The video has been accelerated to 5 frames per second.
(MOV)

S3 Movie. Time-lapse video of a cell clone with cells changing morphology at high frequency (dynamic phenotype of frequent switchers). Only a period between 61 and 81 h is shown. Note that individual snapshots taken at different moments may show a population composed of only polarized, only round or cells with mixed morphology. The video is the original speed.
(MOV)

Acknowledgments

The authors are grateful to Jean-Jacques Kupiec and François Delhommeau for helpful discussions, to Peggy Sanatine and Simon Jimenez (Imaging-Cytometry Platform of Genethon) for the precious help with cell sorting, to Nadège Brunel-Meunier for her help in the use of Biomark HD System (FluExGen Platform of Saint Antoine Hospital [Paris, France]), and to Rhonda Campbell for the language revision of the manuscript. The authors are also grateful to the mothers and the staff of the Centre Hospitalier Sud Francilien (Evry, France) for UCB samples.

Author Contributions

Conceptualization: Olivier Gandrillon, Daniel Stockholm, András Páldi.

Data curation: Alice Moussy, Jérémie Cosette, Romuald Parmentier, Cindy da Silva, Daniel Stockholm, András Páldi.

Formal analysis: Alice Moussy, Olivier Gandrillon, Daniel Stockholm, András Páldi.

Funding acquisition: Olivier Gandrillon, András Páldi.

Investigation: Alice Moussy, Jérémie Cosette, Romuald Parmentier, Cindy da Silva, Angélique Richard.

Methodology: Alice Moussy, Jérémie Cosette, Romuald Parmentier, Guillaume Corre, Angélique Richard.

Software: Jérémie Cosette, Guillaume Corre.

Supervision: András Páldi.

Validation: Alice Moussy.

Visualization: Alice Moussy, Jérémie Cosette, Daniel Stockholm, András Páldi.

Writing – original draft: András Páldi.

Writing – review & editing: Alice Moussy, Jérémie Cosette, Angélique Richard, Olivier Gandrillon, Daniel Stockholm, András Páldi.

References

1. Kawamoto H, Katsura Y. A new paradigm for hematopoietic cell lineages: revision of the classical concept of the myeloid-lymphoid dichotomy. Trends in immunology. 2009; 30(5):193–200. https://doi.org/10.1016/j.it.2009.03.001 PMID: 19356980.
2. Akashi K, Traver D, Miyamoto T, Weissman IL. A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. Nature. 2000; 404(6774):193–7. https://doi.org/10.1038/35004599 PMID: 10724173.
3. Iwasaki H, Somoza C, Shigematsu H, Duprez EA, Iwasaki-Arai J, Mizuno S, et al. Distinctive and indispensable roles of PU.1 in maintenance of hematopoietic stem cells and their differentiation. Blood. 2005; 106(5):1590–600. https://doi.org/10.1182/blood-2005-03-0860 PMID: 15914556; PubMed Central PMCID: PMC1895212.
4. Kondo M, Weissman IL, Akashi K. Identification of clonogenic common lymphoid progenitors in mouse bone marrow. Cell. 1997; 91(5):661–72. PMID: 9393859.
5. Nutt SL, Metcalf D, D'Amico A, Polli M, Wu L. Dynamic regulation of PU.1 expression in multipotent hematopoietic progenitors. The Journal of experimental medicine. 2005; 201(2):221–31. https://doi.org/10.1084/jem.20041535 PMID: 15657291; PubMed Central PMCID: PMC2212785.
6. Terszowski G, Waskow C, Conradt P, Lenze D, Koenigsmann J, Carstanjen D, et al. Prospective isolation and global gene expression analysis of the erythrocyte colony-forming unit (CFU-E). Blood. 2005; 105(5):1937–45. https://doi.org/10.1182/blood-2004-09-3459 PMID: 15522951.

7. Wolff L, Humeniuk R. Concise review: erythroid versus myeloid lineage commitment: regulating the master regulators. Stem cells. 2013; 31(7):1237–44. https://doi.org/10.1002/stem.1379 PMID: 23559316.
8. Notta F, Zandi S, Takayama N, Dobson S, Gan OI, Wilson G, et al. Distinct routes of lineage development reshape the human blood hierarchy across ontogeny. Science. 2016; 351(6269):aab2116. https://doi.org/10.1126/science.aab2116 PMID: 26541609; PubMed Central PMCID: PMC4816201.
9. Sun J, Ramos A, Chapman B, Johnnidis JB, Le L, Ho YJ, et al. Clonal dynamics of native haematopoiesis. Nature. 2014; 514(7522):322–7. https://doi.org/10.1038/nature13824 PMID: 25296256; PubMed Central PMCID: PMC4408613.
10. Velten L, Haas SF, Raffel S, Blaszkiewicz S, Islam S, Hennig BP, et al. Human haematopoietic stem cell lineage commitment is a continuous process. Nature cell biology. 2017; 19(4):271–81. https://doi.org/10.1038/ncb3493 PMID: 28319093.
11. Paul F, Arkin Y, Giladi A, Jaitin DA, Kenigsberg E, Keren-Shaul H, et al. Transcriptional Heterogeneity and Lineage Commitment in Myeloid Progenitors. Cell. 2015; 163(7):1663–77. https://doi.org/10.1016/j.cell.2015.11.013 PMID: 26627738.
12. Hu M, Krause D, Greaves M, Sharkis S, Dexter M, Heyworth C, et al. Multilineage gene expression precedes commitment in the hemopoietic system. Genes & development. 1997; 11(6):774–85. PMID: 9087431.
13. Milhem M, Mahmud N, Lavelle D, Araki H, DeSimone J, Saunthararajah Y, et al. Modification of hematopoietic stem cell fate by 5aza 2'deoxycytidine and trichostatin A. Blood. 2004; 103(11):4102–10. https://doi.org/10.1182/blood-2003-07-2431 PMID: 14976039.
14. Nimmo RA, May GE, Enver T. Primed and ready: understanding lineage commitment through single cell analysis. Trends in cell biology. 2015; 25(8):459–67. https://doi.org/10.1016/j.tcb.2015.04.004 PMID: 26004869.
15. Pina C, Fugazza C, Tipping AJ, Brown J, Soneji S, Teles J, et al. Inferring rules of lineage commitment in haematopoiesis. Nature cell biology. 2012; 14(3):287–94. https://doi.org/10.1038/ncb2442 PMID: 22344032.
16. Olsson A, Venkatasubramanian M, Chaudhri VK, Aronow BJ, Salomonis N, Singh H, et al. Single-cell analysis of mixed-lineage states leading to a binary cell fate choice. Nature. 2016. https://doi.org/10.1038/nature19348 PMID: 27580035.
17. Chang HH, Hemberg M, Barahona M, Ingber DE, Huang S. Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. Nature. 2008; 453(7194):544–7. https://doi.org/10.1038/nature06965 PMID: 18497826.
18. Dooner GJ, Colvin GA, Dooner MS, Johnson KW, Quesenberry PJ. Gene expression fluctuations in murine hematopoietic stem cells with cell cycle progression. Journal of cellular physiology. 2008; 214(3):786–95. https://doi.org/10.1002/jcp.21273 PMID: 17894410; PubMed Central PMCID: PMC4286177.
19. Koga S, Yamaguchi N, Abe T, Minegishi M, Tsuchiya S, Yamamoto M, et al. Cell-cycle-dependent oscillation of GATA2 expression in hematopoietic cells. Blood. 2007; 109(10):4200–8. https://doi.org/10.1182/blood-2006-08-044149 PMID: 17255359.
20. Wilson NK, Kent DG, Buettner F, Shehata M, Macaulay IC, Calero-Nieto FJ, et al. Combined Single-Cell Functional and Gene Expression Analysis Resolves Heterogeneity within Stem Cell Populations. Cell stem cell. 2015; 16(6):712–24. https://doi.org/10.1016/j.stem.2015.04.004 PMID: 26004780; PubMed Central PMCID: PMC4460190.
21. Hoppe PS, Schwarzfischer M, Loeffler D, Kokkaliaris KD, Hilsenbeck O, Moritz N, et al. Early myeloid lineage choice is not initiated by random PU.1 to GATA1 protein ratios. Nature. 2016; 535(7611):299–302. https://doi.org/10.1038/nature18320 PMID: 27411635.
22. Cattoglio C, Pellin D, Rizzi E, Maruggi G, Corti G, Miselli F, et al. High-definition mapping of retroviral integration sites identifies active regulatory elements in human multipotent hematopoietic progenitors. Blood. 2010; 116(25):5507–17. https://doi.org/10.1182/blood-2010-05-283523 PMID: 20864581.
23. Livak KJ, Wills QF, Tipping AJ, Datta K, Mittal R, Goldson AJ, et al. Methods for qPCR gene expression profiling applied to 1440 lymphoblastoid single cells. Methods. 2013; 59(1):71–9. https://doi.org/10.1016/j.ymeth.2012.10.004 PMID: 23079396; PubMed Central PMCID: PMC3562442.
24. Tibshirani R, Guenther W, Hastie T. Estimating the number of clusters in a data set via gap statistics. Journal of Royal Statistical Society B. 2001; 63:411–23.
25. van der Maaten L, Hinton G. Data representation using t-SNE. Journal of Machine Learning Research. 2008; 9:2679–05.
26. Moignard V, Macaulay IC, Swiers G, Buettner F, Schutte J, Calero-Nieto FJ, et al. Characterization of transcriptional networks in blood stem and progenitor cells using high-throughput single-cell gene

expression analysis. Nature cell biology. 2013; 15(4):363–72. https://doi.org/10.1038/ncb2709 PMID: 23524953; PubMed Central PMCID: PMC3796878.

27. Bargaje R, Trachana K, Shelton MN, McGinnis CS, Zhou JX, Chadick C, et al. Cell population structure prior to bifurcation predicts efficiency of directed differentiation in human induced pluripotent cells. Proceedings of the National Academy of Sciences of the United States of America. 2017; 114(9):2271–6. https://doi.org/10.1073/pnas.1621412114 PMID: 28167799; PubMed Central PMCID: PMC5338498.
28. Giuliani A. Statistical mechanics of gene expression networks: Increasing connectivity as a response to stressful condition. Advances in Systems Biology. 2014; 3(1):3–6.
29. Mojtahedi M, Skupin A, Zhou J, Castano IG, Leong-Quong RY, Chang H, et al. Cell Fate Decision as High-Dimensional Critical State Transition. PLoS Biol. 2016; 14(12):e2000640. https://doi.org/10.1371/journal.pbio.2000640 PMID: 28027308; PubMed Central PMCID: PMC5189937.
30. Brummendorf TH, Dragowska W, Zijlmans J, Thornbury G, Lansdorp PM. Asymmetric cell divisions sustain long-term hematopoiesis from single-sorted human fetal liver cells. The Journal of experimental medicine. 1998; 188(6):1117–24. PMID: 9743530; PubMed Central PMCID: PMC2212541.
31. Gorgens A, Ludwig AK, Mollmann M, Krawczyk A, Durig J, Hanenberg H, et al. Multipotent hematopoietic progenitors divide asymmetrically to create progenitors of the lymphomyeloid and erythromyeloid lineages. Stem cell reports. 2014; 3(6):1058–72. https://doi.org/10.1016/j.stemcr.2014.09.016 PMID: 25448068; PubMed Central PMCID: PMC4263999.
32. Wagner W, Ansorge A, Wirkner U, Eckstein V, Schwager C, Blake J, et al. Molecular evidence for stem cell function of the slow-dividing fraction among human hematopoietic progenitor cells by genome-wide analysis. Blood. 2004; 104(3):675–86. https://doi.org/10.1182/blood-2003-10-3423 PMID: 15090461.
33. Bauer N, Wilsch-Brauninger M, Karbanova J, Fonseca AV, Strauss D, Freund D, et al. Haematopoietic stem cell differentiation promotes the release of prominin-1/CD133-containing membrane vesicles—a role of the endocytic-exocytic pathway. EMBO molecular medicine. 2011; 3(7):398–409. https://doi.org/10.1002/emmm.201100147 PMID: 21591261; PubMed Central PMCID: PMC3210830.
34. Pevny L, Simon MC, Robertson E, Klein WH, Tsai SF, D'Agati V, et al. Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1. Nature. 1991; 349(6306):257–60. https://doi.org/10.1038/349257a0 PMID: 1987478.
35. Graf T, Enver T. Forcing cells to change lineages. Nature. 2009; 462(7273):587–94. https://doi.org/10.1038/nature08533 PMID: 19956253.
36. Chaurasia P, Gajzer DC, Schaniel C, D'Souza S, Hoffman R. Epigenetic reprogramming induces the expansion of cord blood stem cells. The Journal of clinical investigation. 2014; 124(6):2378–95. https://doi.org/10.1172/JCI70313 PMID: 24762436; PubMed Central PMCID: PMC4038563.
37. De Felice L, Tatarelli C, Mascolo MG, Gregorj C, Agostini F, Fiorini R, et al. Histone deacetylase inhibitor valproic acid enhances the cytokine-induced expansion of human hematopoietic stem cells. Cancer research. 2005; 65(4):1505–13. https://doi.org/10.1158/0008-5472.CAN-04-3063 PMID: 15735039.
38. Borge OJ, Ramsfjell V, Cui L, Jacobsen SE. Ability of early acting cytokines to directly promote survival and suppress apoptosis of human primitive CD34+CD38- bone marrow cells with multilineage potential at the single-cell level: key role of thrombopoietin. Blood. 1997; 90(6):2282–92. PMID: 9310479.
39. Corre G, Stockholm D, Arnaud O, Kaneko G, Vinuelas J, Yamagata Y, et al. Stochastic fluctuations and distributed control of gene expression impact cellular memory. PLoS ONE. 2014; 9(12):e115574. https://doi.org/10.1371/journal.pone.0115574 PMID: 25531401; PubMed Central PMCID: PMC4274012.
40. Schwanhausser B, Wolf J, Selbach M, Busse D. Synthesis and degradation jointly determine the responsiveness of the cellular proteome. BioEssays: news and reviews in molecular, cellular and developmental biology. 2013; 35(7):597–601. https://doi.org/10.1002/bies.201300017 PMID: 23696377.
41. Vinuelas J, Kaneko G, Coulon A, Vallin E, Morin V, Mejia-Pous C, et al. Quantifying the contribution of chromatin dynamics to stochastic gene expression reveals long, locus-dependent periods between transcriptional bursts. BMC biology. 2013; 11:15. https://doi.org/10.1186/1741-7007-11-15 PMID: 23442824; PubMed Central PMCID: PMC3635915.
42. Kupiec JJ. A chance-selection model for cell differentiation. Cell death and differentiation. 1996; 3(4):385–90. PMID: 17180108.
43. Kupiec JJ. A Darwinian theory for the origin of cellular differentiation. Molecular & general genetics: MGG. 1997; 255(2):201–8. PMID: 9236778.
44. Paldi A. Stochastic gene expression during cell differentiation: order from disorder? Cellular and molecular life sciences: CMLS. 2003; 60(9):1775–8. https://doi.org/10.1007/s00018-003-23147-z PMID: 14523542.
45. Paldi A. What makes the cell differentiate? Progress in biophysics and molecular biology. 2012; 110(1):41–3. https://doi.org/10.1016/j.pbiomolbio.2012.04.003 PMID: 22543273.

46. Richard A, Boullu L, Herbach U, Bonnafoux A, Morin V, Vallin E, et al. Single-cell-based analysis highlights a surge in cell-to-cell molecular variability preceding irreversible commitment in a differentiation process. PLoS Biol. 2016;(14):e1002585. doi.org/10.1371/journal.pbio.1002585.

47. Huang S. Non-genetic heterogeneity of cells in development: more than just noise. Development. 2009; 136(23):3853–62. https://doi.org/10.1242/dev.035139 PMID: 19906852; PubMed Central PMCID: PMC2778736.

48. Furusawa C, Kaneko K. A dynamical-systems view of stem cell biology. Science. 2012; 338 (6104):215–7. https://doi.org/10.1126/science.1224311 PMID: 23066073.

49. Buganim Y, Faddah DA, Cheng AW, Itskovich E, Markoulaki S, Ganz K, et al. Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. Cell. 2012; 150(6):1209–22. https://doi.org/10.1016/j.cell.2012.08.023 PMID: 22980981; PubMed Central PMCID: PMC3457656.

50. Hayashi K, Lopes SM, Tang F, Surani MA. Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states. Cell stem cell. 2008; 3(4):391–401. https://doi.org/10.1016/j.stem.2008.07.027 PMID: 18940731; PubMed Central PMCID: PMC3847852.

51. Kobayashi T, Mizuno H, Imayoshi I, Furusawa C, Shirahige K, Kageyama R. The cyclic gene Hes1 contributes to diverse differentiation responses of embryonic stem cells. Genes & development. 2009; 23 (16):1870–5. https://doi.org/10.1101/gad.1823109 PMID: 19684110; PubMed Central PMCID: PMC2725939.

52. Zhao S, Guo Y, Sheng Q, Shyr Y. Heatmap3: An Improved Heatmap Package. https://CRANR-projectorg/package=heatmap3. 2015.

53. Le S, Josse J, Husson F. FactoMineR: An R Package for Multivariate Analysis. Journal of Statistical Software. 2008; 25:1–18.

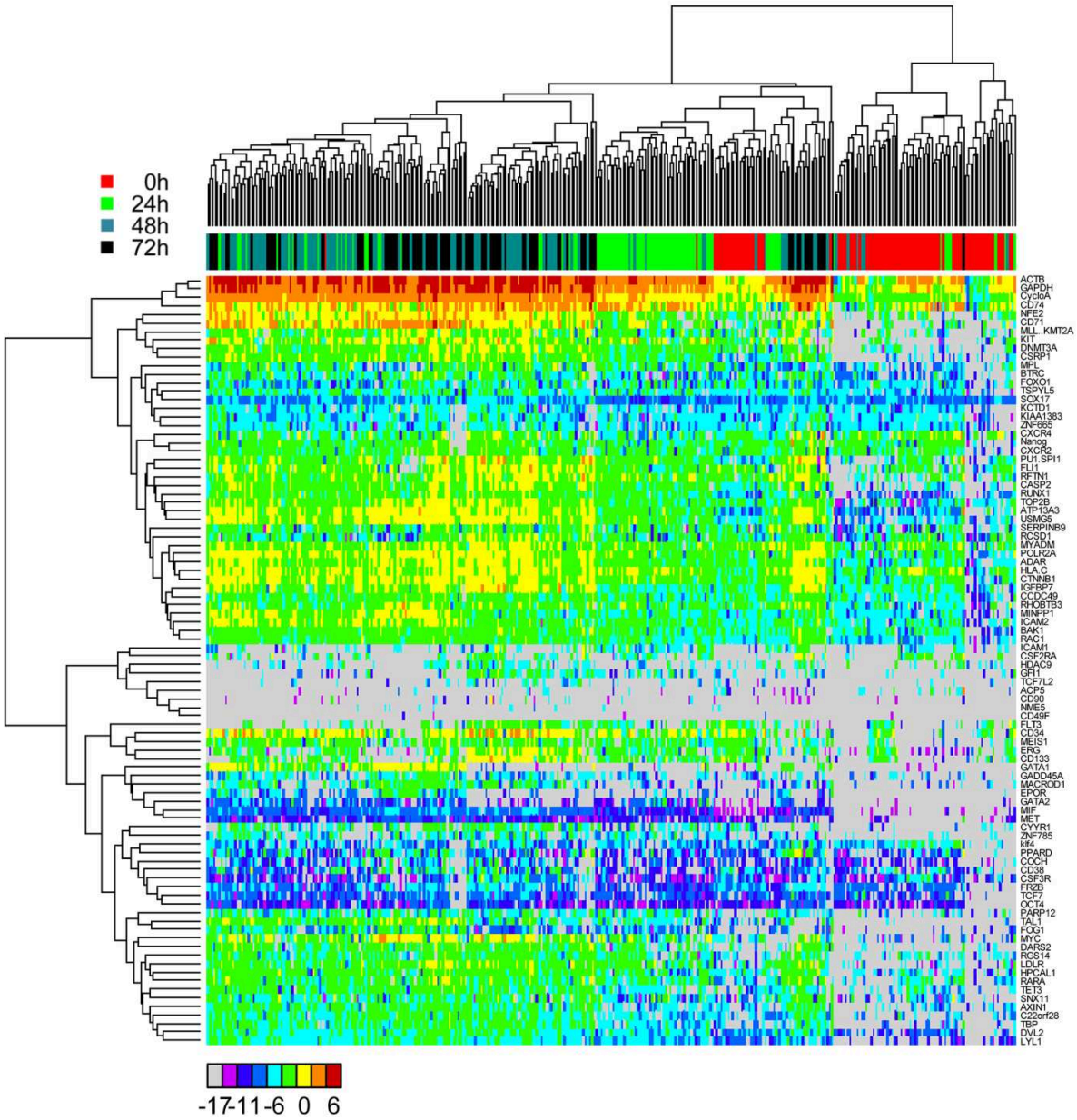
54. Wickham H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. 2009.

55. Grun D, Lyubimova A, Kester L, Wiebrands K, Basak O, Sasaki N, et al. Single-cell messenger RNA sequencing reveals rare intestinal cell types. Nature. 2015; 525(7568):251–5. https://doi.org/10.1038/nature14966 PMID: 26287467.

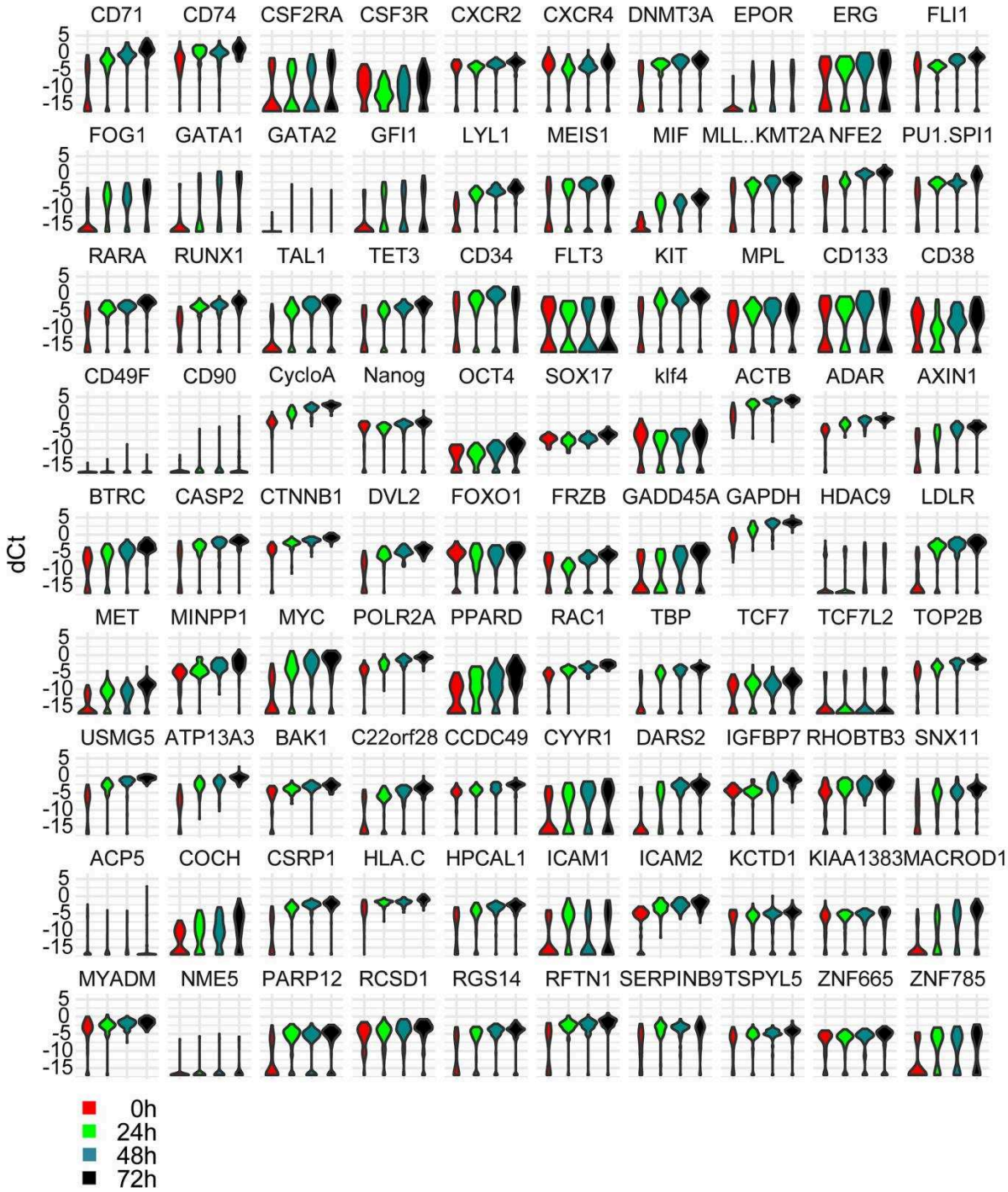
56. Cosette J, Moussy A, Onodi F, Auffret-Cariou A, Neildez-Nguyen TM, Paldi A, et al. Single Cell Dynamics Causes Pareto-Like Effect in Stimulated T Cell Populations. Scientific reports. 2015; 5:17756. https://doi.org/10.1038/srep17756 PMID: 26648396; PubMed Central PMCID: PMC4673432.

57. Neildez-Nguyen TM, Bigot J, Da Rocha S, Corre G, Boisgerault F, Paldi A, et al. Hypoxic culture conditions enhance the generation of regulatory T cells. Immunology. 2014; 144:431–43. https://doi.org/10.1111/imm.12388 PMID: 25243909; PubMed Central PMCID: PMC4557680.

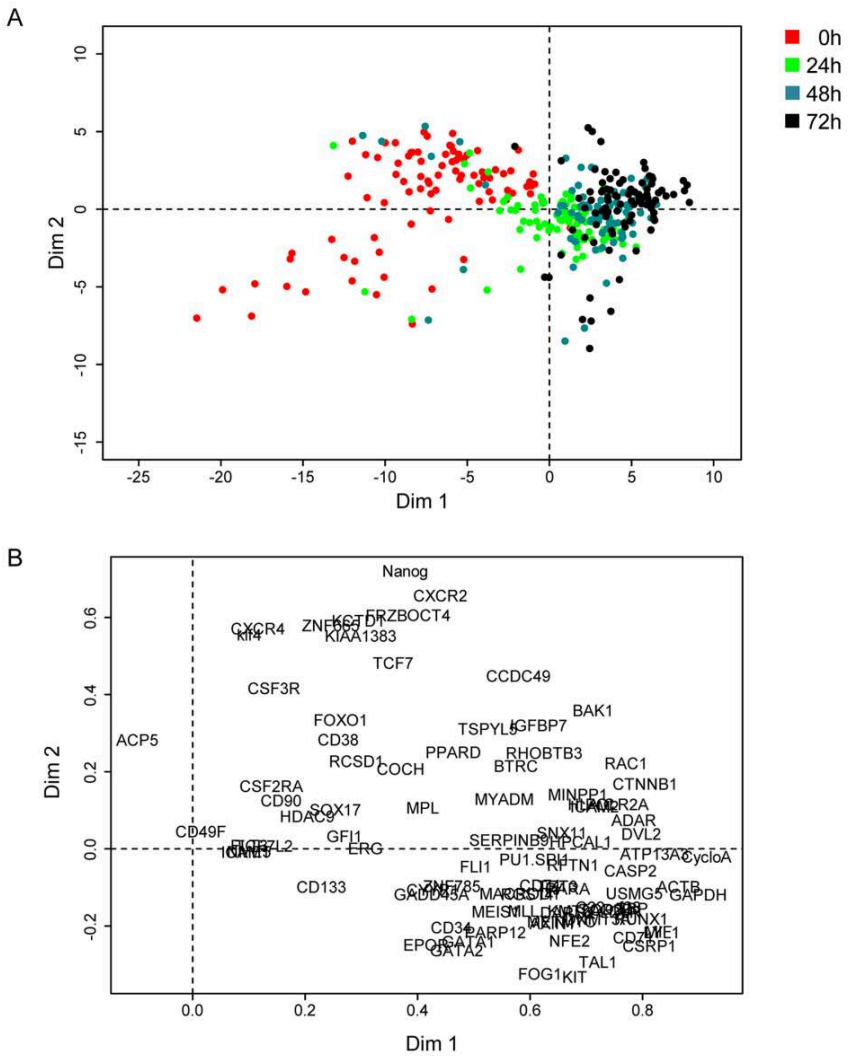
SUPPORTING INFORMATION



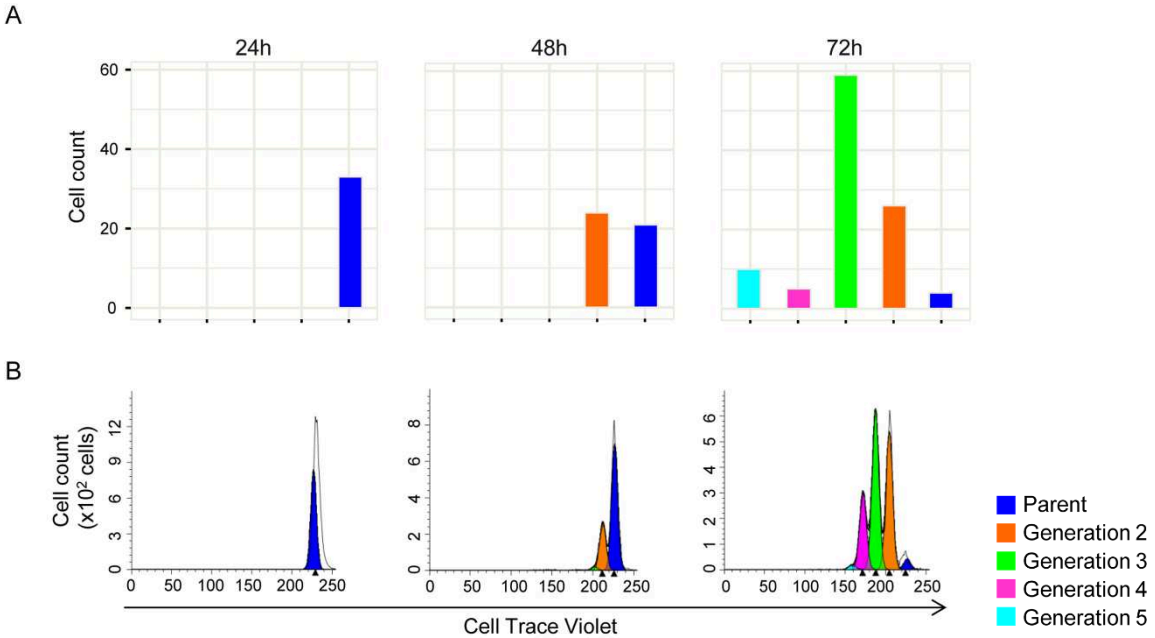
S1 Fig : Full set of gene expression data obtained using single-cell qRT-PCR in cord-blood CD34+ cells cultured in vivo with early-acting cytokines.



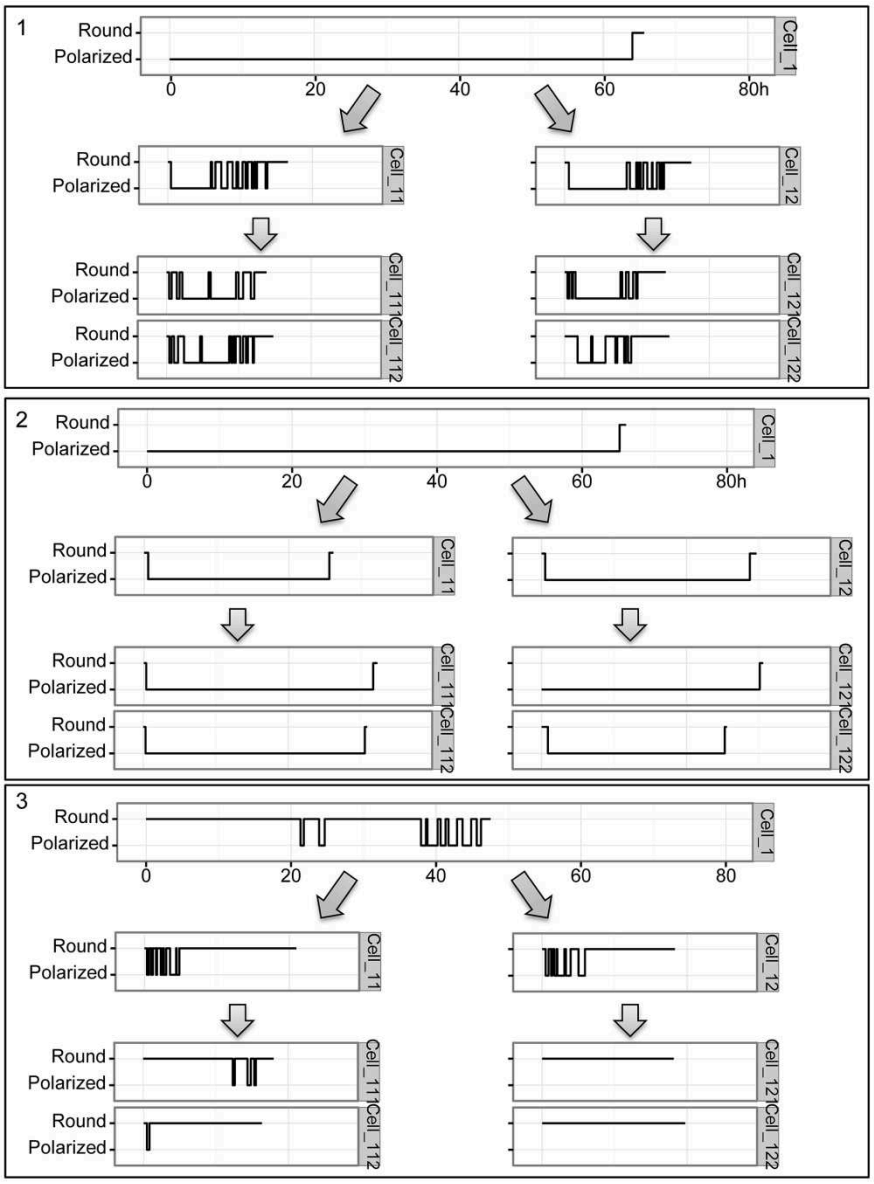
S2 Fig. Violin plot representation of individual gene expression levels at the four time points.



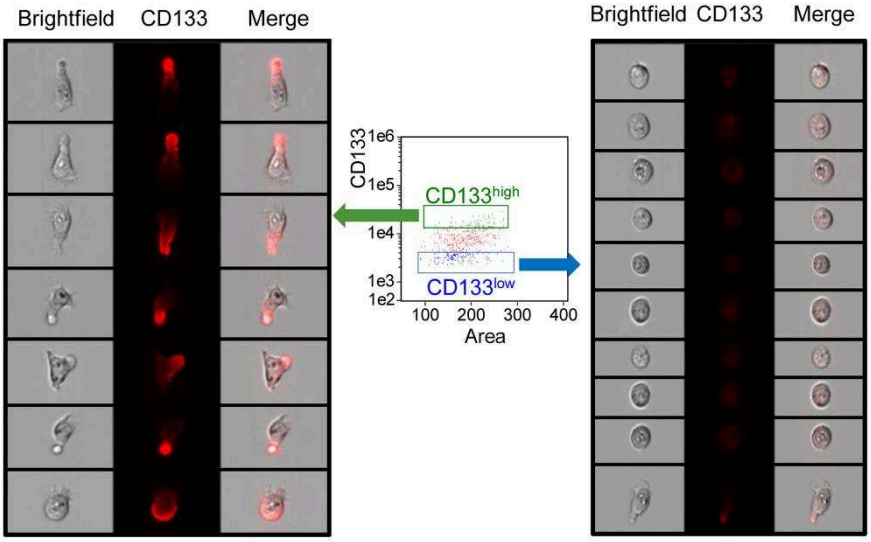
S3 Fig. Principal component analysis of single-cell expression profiles



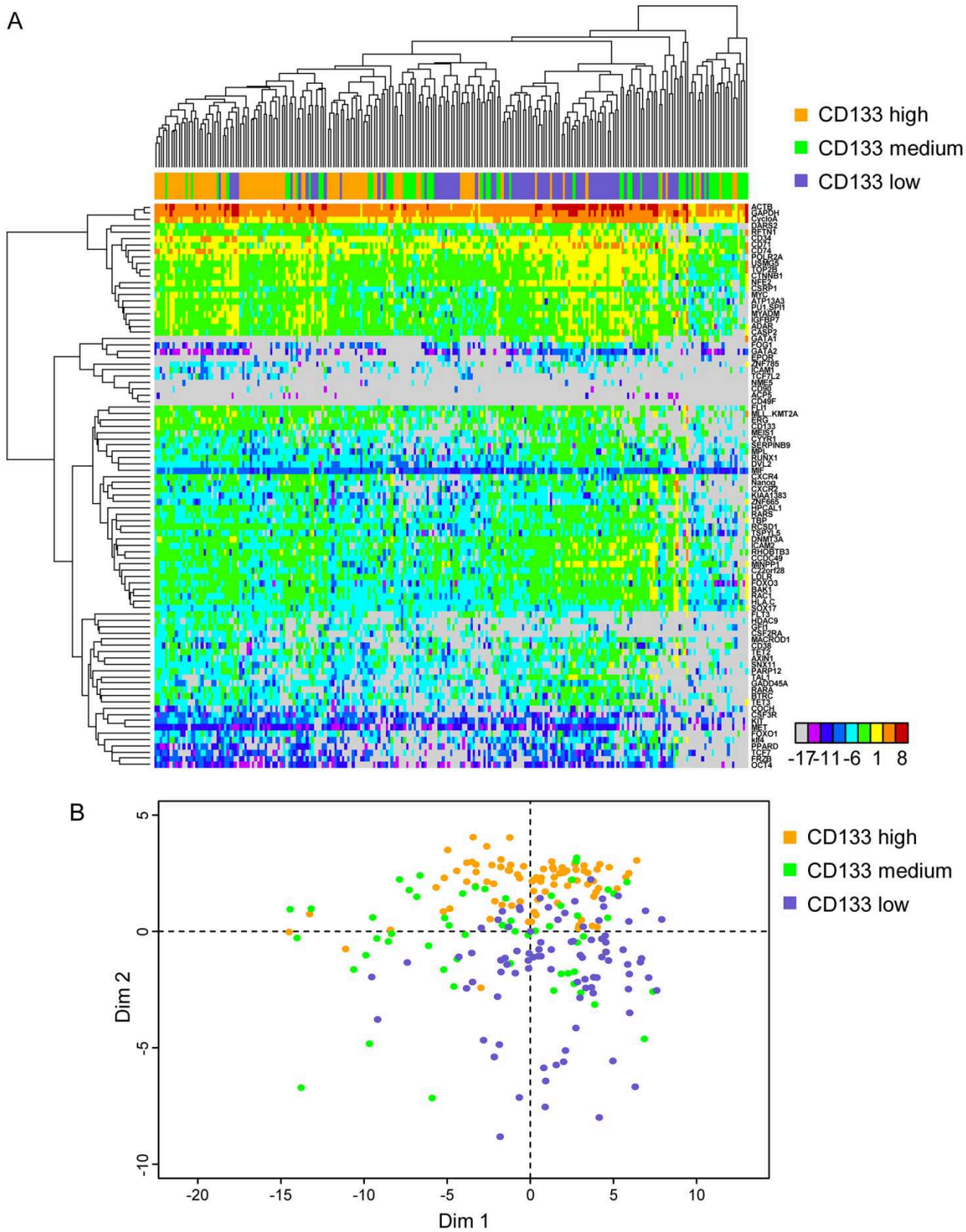
S4 Fig. Analysis of cell division rates



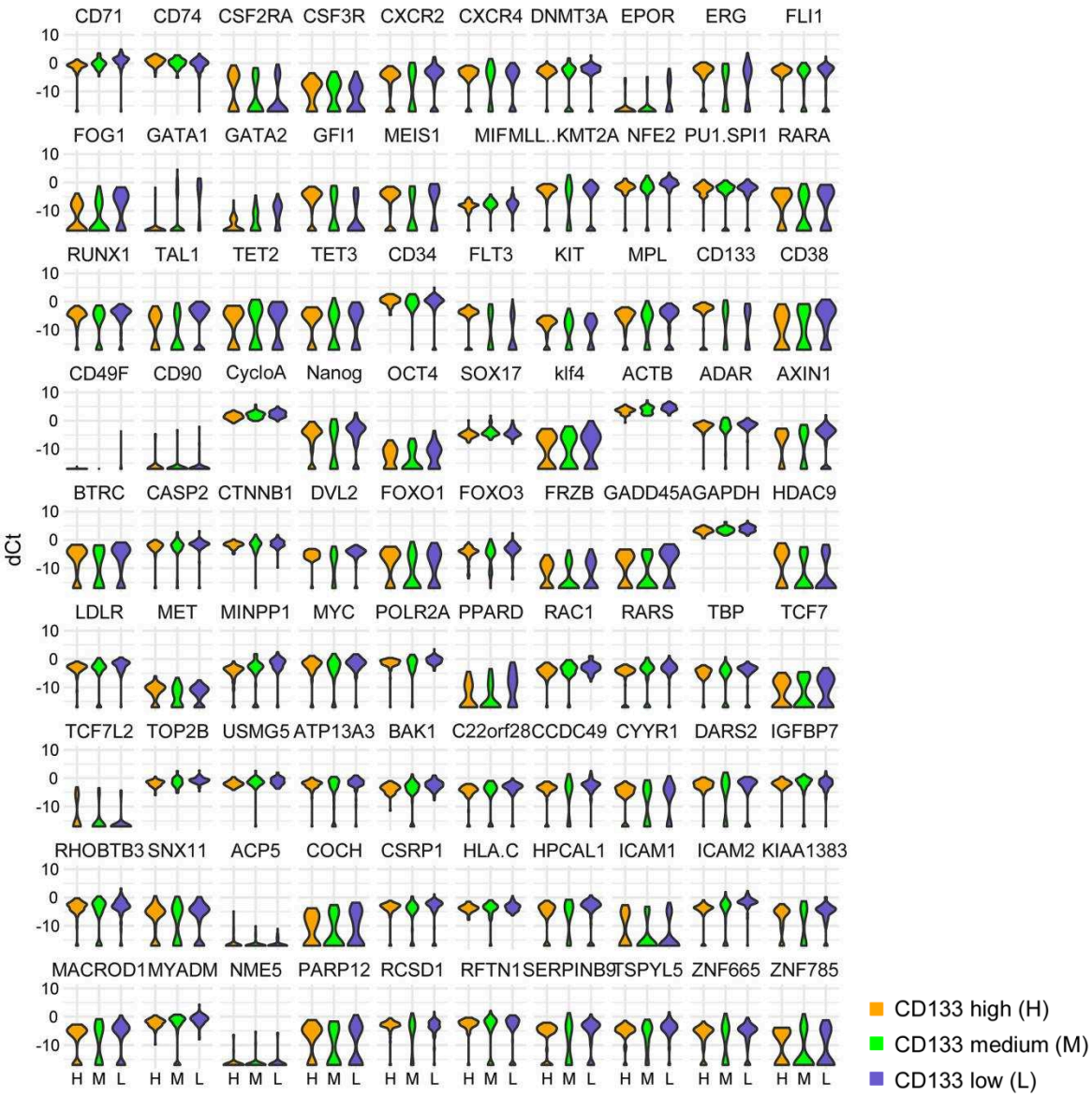
S5 Fig. Representations of morphological profiles of cells in three representative clones.



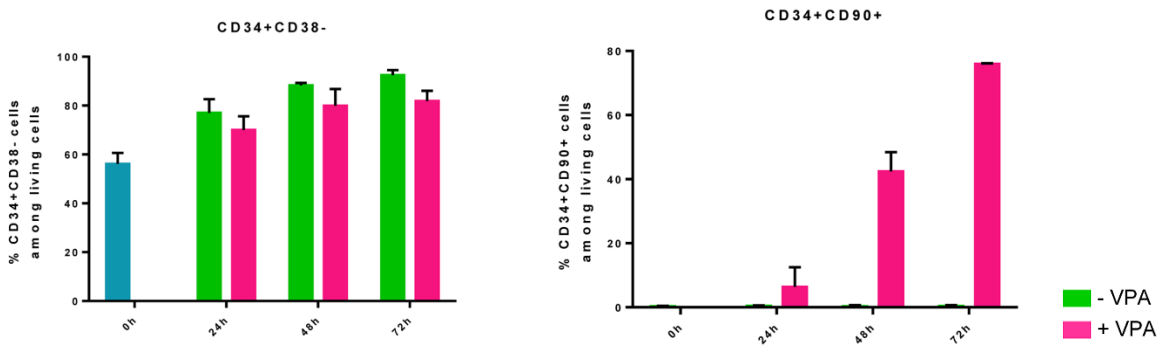
S6 Fig. Cell morphology and CD133 localisation.



S7 Fig. The full set of the gene expression data obtained on "high", "medium", and "low" CD133 expressing individual cells.



S8 Fig. Violin plot representation of individual gene expression levels in the "high", "medium", and "low" CD133 cells.



S9 Fig. Cytometry analysis of the effects of valproic acid on CD34+ cells.

Dans cette étude, notre but était d'identifier les étapes initiales, temporelles du destin cellulaire dans les cellules CD34+ humaines issues de sang de cordon. Sans stimulation par des cytokines, les cellules CD34+ restent quiescentes et meurent au bout de quelques jours de mise en culture. En revanche, l'activation par des cytokines permet aux cellules de survivre, d'être métaboliquement active et d'entrer dans le cycle cellulaire sans montrer, a priori, de signes de différenciation pendant les premiers jours.

Nous avons analysé les profils transcriptomiques des cellules CD34+ uniques à temps 0h, 24h, 48h et 72h. Nous avons montré que les changements transcriptomiques majeurs conduisant à un état « multi-primed » se produisaient très rapidement dans toutes les cellules avant la première division cellulaire. A partir de 48h, nous avons observé l'émergence de deux profils d'expression génique distincts statistiquement différents :

- plus de la moitié des cellules expriment des gènes plutôt caractéristiques des CMP comme GATA1 ou EPOR ;
- l'autre moitié exprime des gènes plutôt caractéristiques des MPP comme CD133, GFI1, KLF4 ou FLT3.

Les gènes spécifiques des cellules souches pluripotentes comme NANOG et OCT4 ont été exprimés à un niveau modéré dans les deux groupes de cellules observées. Les gènes sélectionnés de façon aléatoire étaient aussi de bons prédicteurs pour les deux groupes de cellules. Cette tendance observée à 48h a été renforcée dans les données obtenues à 72h.

Comme l'acide valproïque (VPA), inhibiteur des histones déacétylases (HDAC), est connu pour augmenter le pourcentage de chimérisme après transplantation des cellules CD34+ issues de sang de cordon et donc pour augmenter le pourcentage de MPP, nous l'avons utilisé comme contrôle pour confirmer ou infirmer le phénotype MPP observé (Bug et al., 2005; Chaurasia et al., 2014; Felice et al., 2005). Nous avons analysé à temps 0h, 24h, 48h et 72h, les profils transcriptomiques de ces cellules traitées avec VPA durant toute l'expérience. Grâce à la représentation en t-SNE des données avec et sans VPA, nous avons montré que l'un des deux groupes de cellules observés sans VPA présentait des caractéristiques transcriptomiques similaires que les cellules traitées avec VPA. Nous avons donc confirmé le phénotype MPP observé sans VPA à partir de 48h. De plus, nous avons observé que les cellules traitées avec VPA exprimaient, à un niveau modéré, des gènes de plusieurs lignées différentes, laissant ainsi suggérer que l'état multipotent serait un état transitoire où les cellules exprimeraient potentiellement tous les gènes.

Parallèlement, avec la microscopie en TL après isolement des cellules uniques, nous avons observé parmi les clones, trois types de phénotypes dynamiques :

- des cellules rondes stables tout au long de l'expérience ;
- des cellules polarisées stables ;
- des cellules qui fluctuent entre ces deux types de morphologies sur plusieurs générations.

Finalement, au cours de ces premiers jours de stimulation, nous avons observé trois types de phénotypes et deux profils d'expression génique des cellules CD34+. Des études préliminaires ont déjà montré que les cellules polarisées exprimaient fortement le CD133, expérience que nous avons validée avec la CFI (Görgens et al., 2014). Cela nous a laissé suggérer que les cellules polarisées stables, observées en TL auraient plutôt un profil MPP trouvé dans les analyses transcriptomiques alors que les cellules rondes stables auraient plutôt un profil CMP. Finalement, les cellules fluctuant entre polarisées et rondes oscilleraient entre les deux profils transcriptomiques.

Notre analyse a permis de mettre en évidence la nature dynamique et la cinétique variable de l'engagement du destin cellulaire dans les cellules hématopoïétiques et de relier le modèle d'expression génique à la morphologie cellulaire. Cela a permis d'identifier une nouvelle catégorie de cellules présentant des caractéristiques phénotypiques fluctuantes, démontrant ainsi la complexité du processus de prise de décision cellulaire, qui semble bien différent du modèle binaire habituellement proposé.

Notre hypothèse est la suivante : le choix du destin cellulaire serait accompagné par des fluctuations dans le profil de l'expression des gènes. Ce scénario dynamique relierait l'hypothèse d'un état « multi-primed » avec l'idée de transcriptomes fluctuants, en suggérant que la relaxation d'un état « multi-primed » métastable à un état stable impliquerait une phase incertaine, « hésitante », de longueur variable.

DISCUSSION ET PERSPECTIVES

I. Discussion

Bien que la différenciation soit un processus étudié depuis plusieurs décennies, la compréhension de la prise de décision des cellules n'est pas encore claire. A travers les résultats obtenus au cours de ma thèse à partir des deux modèles cellulaires, on remarque bien que la différenciation ne suit pas un processus linéaire, prédéterminé. En effet, malgré qu'elles soient toutes soumises aux mêmes stimuli, les cellules ne se comportent pas de la même manière.

I.A Hétérogénéité

Dans notre première étude, nous avons observé que les LT, pris individuellement, avaient des temps de divisions tous différents. Certains clones proliféraient plus vite que d'autres ; ils déterminaient ainsi la population finale (effet Pareto). Nous avons aussi vu que la différenciation des LT en LTreg n'était pas dépendante de la prolifération, qu'il s'agissait de deux processus totalement indépendants.
L'étude du comportement des cellules CD34+ a, elle, été plus loin en analysant le transcriptome de 90 gènes sur cellules uniques à différents temps. Nous n'avons vu aucune cellule identique. Comparé aux LTreg, nous avons aussi analysé la morphologie dynamique des cellules CD34+ en microscopie en TL. Bien qu'elles se trouvent dans le même environnement, nous avons observé deux formes différentes, rondes et polarisées, avec une dynamique propre à chacune.

On peut donc se poser la question suivante : est-ce que l'isolement des cellules a une influence sur leur comportement ? En effet, les cellules in vivo sont rarement seules, elles sont souvent entourées de cellules voisines qui peuvent jouer un rôle sur leur devenir (Isern and Méndez-Ferrer, 2011; Wozniak et al., 2004). Pour répondre à cette question, qui nous a souvent été posée, nous avons comparé le temps de prolifération en population et à l'échelle d'une cellule. Les résultats ont montré qu'il n'y avait pas de différence dans les deux expériences. Nous pouvons donc confirmer que l'hétérogénéité observée dans la prolifération des LTreg et des cellules CD34+ n'était pas due à leur isolement. Par contre, nous ne pouvons pas confirmer que l'isolement n'a aucune influence sur l'expression des gènes et donc sur le devenir des cellules. Selon moi, dans notre cas avec les microgrilles, l'isolement n'aurait pas d'impact sur le destin cellulaire car même si elles sont isolées, elles se trouvent dans le même environnement ; les cytokines et autres molécules qu'elles peuvent produire, peuvent donc être perçues par la cellule voisine.

Et si finalement, l'hétérogénéité des cellules était déjà présente, préexistante, avant de commencer l'expérience ? C'est une hypothèse à laquelle nous avons pensé et que nous n'avons pas écarté ; mais si tel est le cas alors d'où provient cette hétérogénéité ? Dans les gènes ? Les cellules seraient-elles « programmées » pour savoir à quel moment elles peuvent se diviser et se différencier ? C'est une hypothèse qui me parait peu probable. Surtout qu'il a déjà été démontré au sein de notre équipe qu'une population cellulaire isogénique analysée à l'échelle de la cellule unique, mise dans un environnement homogène, exprimait des protéines à des niveaux différents (Corre et al., 2014). Donc, même lorsque les cellules sont les plus « homogènes » possibles, elles ont des comportements hétérogènes. L'hétérogénéité préexistante ne peut donc pas tout expliquer, elle ne peut pas être l'unique explication de ce comportement observé. Nous sommes tous issus d'une seule cellule, et pourtant au cours de la vie de l'embryon, les cellules acquièrent des caractéristiques et des fonctions différentes.

Une explication possible à cette hétérogénéité individuelle serait le hasard. Comment est-il possible que, malgré l'hétérogénéité des comportements individuels, la différenciation cellulaire à l'échelle populationnelle se déroule de façon ordonnée et prédictible, si cela n'est pas due au hasard ? Bien qu'il soit souvent associé au mot imprévisible et désordonné, le hasard peut être organisé. Afin de mieux comprendre comment le hasard peut être ordonné, je vais illustrer mes propos en prenant un exemple, avec deux voies possibles, comme dans la métaphore du paysage de Waddington (Figure 41).

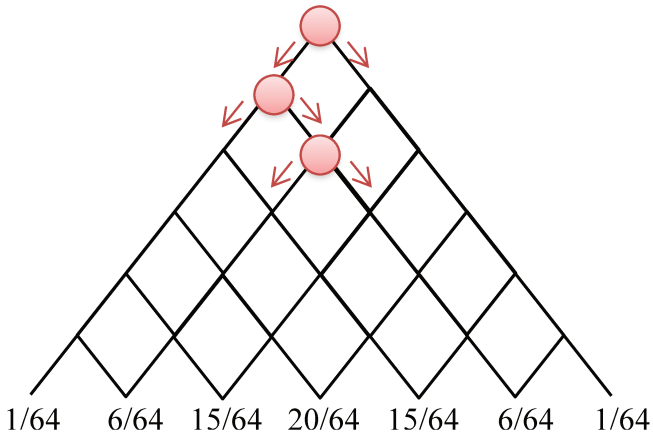


Figure 41 : Exemple de bifurcation aléatoire.

Dans cet exemple, la cellule n'a que deux possibilités : soit aller à droite (50% de chance), soit aller à gauche (50% de chance). On remarque que la cellule a plus de chance de se retrouver au milieu du paysage (31%) qu'aux extrémités (1,5%).

La cellule a le choix d'aller soit à droite, soit à gauche avec la même probabilité. A chaque étape, la cellule va « choisir » au hasard une voie. A la fin, on se rend compte que la probabilité que la cellule se retrouve au milieu du paysage est plus forte qu'aux extrémités. Cela s'explique facilement par le fait qu'il y ait plus de chemins possibles qui amènent au milieu du paysage qu'aux extrémités où il y en a qu'un seul. Si on augmente le nombre d'étapes entre le début et la fin de l'expérience alors la probabilité que la cellule aille au milieu augmentera, a contrario de la probabilité aux extrémités qui diminuera au point de devenir un évènement rare. On aura donc l'impression d'avoir un événement déterminé car la cellule se trouvera en majorité au milieu du paysage. Cette illustration peut totalement s'appliquer à la différenciation cellulaire avec des cellules qui ont le choix entre deux, voire plusieurs voies phénotypiques, et qui au final donnent une population prédictible. Dans le cas des CSH, les cellules prises individuellement ont des comportements hétérogènes qui donnent une population finale totalement prédictible avec des pourcentages de cellules sanguines identiques et connus tout au long de la vie.

Ce sont les raisons pour lesquelles je pense que le modèle stochastique est le modèle qui décrit le mieux le processus de différenciation comparé au modèle déterministe. D'autant plus qu'il n'y a pas un modèle consensus accepté pas toute la communauté scientifique et qu'aujourd'hui, on remet même en cause l'existence d'une population CMP (Notta et al., 2016; Paul et al., 2015). Une explication simple de la différenciation est que les cellules prises individuellement ont toutes une chance de s'engager dans une voie spécifique et qui, ensemble, forment une population prédictible.

I.B Choix du destin cellulaire

Les CSH, par leur capacité à s'auto-renouveler et à se différencier, permettent ainsi un renouvellement constant des cellules du sang tout au long de la vie, ce qui suscite beaucoup d'intérêt et de questions sur les mécanismes du maintien du nombre de CSH stables. Comment la cellule fait-elle son choix entre rester quiescente ou s'engager dans une voie de différenciation ?

L'une des hypothèses expliquant l'auto-renouvellement des CSH est la division asymétrique. Au moment de la division cellulaire, la cellule donnerait deux cellules filles différentes : l'une resterait quiescente, multipotente et l'autre s'engagerait dans une voie de différenciation. Le stock de cellules souches multipotentes serait donc toujours maintenu. Cette hypothèse a longtemps été controversée et l'est encore aujourd'hui. La pertinence fonctionnelle des héritages des protéines asymétriques pour le maintien de l'auto-renouvellement des CSH n'a pu être démontrée à ce jour, bien que le devenir des CSH filles issues de la division

asymétrique et la distribution polarisée des molécules ont déjà été décrits (Beckmann et al., 2007; Takano et al., 2004). Par exemple, Görgens et al. ont démontré grâce à l'imagerie en TL que lors de la division cellulaire, la protéine CD133 était répartie de façon inégale entre les cellules MPP filles dans 70% des cas. Ils en ont conclu que la séparation entre les progéniteurs lympho-myéloïdes et érythro-myéloïdes était due à cette division asymétrique (Görgens et al., 2014). Mais, ils n'ont pas démontré qu'elle pouvait maintenir le stock de CSH quiescentes. Selon moi, la division asymétrique ne peut pas tout expliquer. Au regard de nos résultats obtenus sur les cellules CD34+ humaines, nous avons montré que ces cellules n'avaient pas besoin de se diviser pour changer de comportement. Au cours des premières 48h de mise en culture, un tiers des cellules s'étaient divisées alors que nous avons observé des changements transcriptomiques dans toutes les cellules dès les premières 24h. A 48h, nous avons déjà vu deux groupes de cellules se former : les cellules avaient déjà pris leur décision sans se diviser.

De plus, dans les modèles de différenciation proposés, la cellule n'a pas la possibilité de revenir à son état initial, la réversibilité est impossible. Or, ce phénomène de dédifférenciation a déjà été observé chez les plantes. Pourquoi n'existerait-il pas dans la différenciation des CSH ? Ce phénomène peut être masqué à cause des analyses faites à des instants donnés et non en continu. En effet, en ne prenant que des instantanés, même à intervalles très courts, on rate des informations importantes ; par exemple, la fluctuation de l'expression des gènes, des protéines ou de la morphologie des cellules (Figure 42).

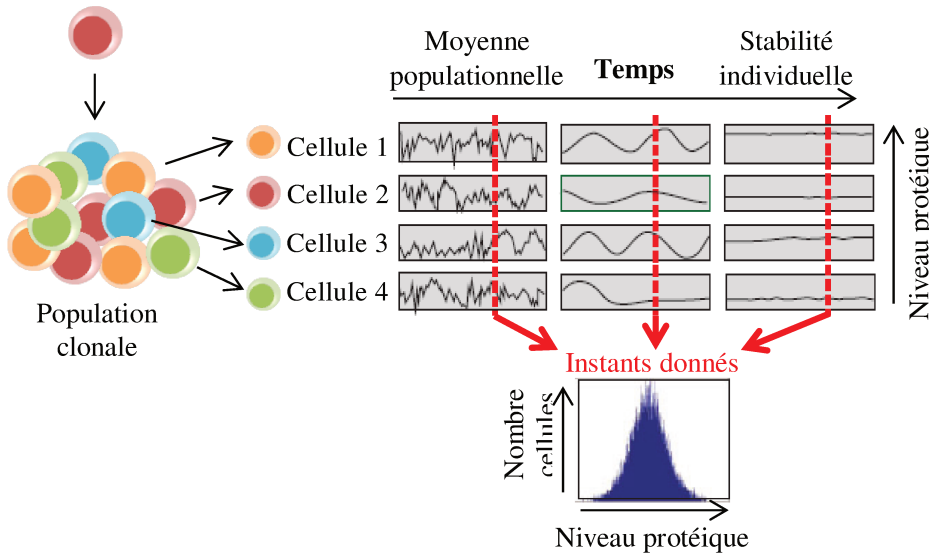


Figure 42 : Analyses instantanées vs dynamique phénotypique.

Des phénotypes cellulaires hétérogènes, analysés à des instants donnés, peuvent avoir le même profil protéique en cytométrie en flux alors que les cellules prises individuellement ont des comportements dynamiques totalement différents dans le temps. Schéma adapté de Huang, 2009.

Grâce à la microscopie en TL, nous avons observé trois phénotypes dans le comportement des cellules CD34+ humaines : des cellules rondes stables, des cellules polarisées stables et des cellules qui fluctuent entre ces deux types de morphologie. Cette dynamique morphologique observée peut être le reflet d'une fluctuation transcriptomique. Les cellules se trouveraient, à un moment donné, dans un état d'hésitation entre plusieurs phénotypes, plusieurs lignages (Huang et al., 2007).

On peut aussi se demander quand et à quelle vitesse se déroulent tous ces évènements ? Est-ce que la prise de décision des cellules est rapide ou est-ce un processus long ?
Grâce à la cinétique transcriptomique réalisée sur les cellules CD34+, nous avons pu montrer que les changements majeurs conduisant à un état « multi-primed » se produisaient très rapidement, pendant les premières 48h de mise en culture, dans toutes les cellules, avant la première division. A partir de 48h, nous avons observé l'émergence de deux profils d'expression génique distincts : des cellules exprimant des gènes plutôt caractéristiques des CMP et des cellules exprimant des gènes plutôt caractéristiques des MPP. Bien entendu, ces observations dépendent de l'environnement des cellules : lors de ma thèse, toutes les expériences ont été réalisées in vitro, avec un cocktail de cytokines spécifiques sensé ne pas favoriser la différenciation des cellules dans une voie particulière. Avec un autre système, le comportement de ces cellules aurait été différent. Cependant, il est à noter que le changement phénotypique observé est tout de même assez surprenant du point de vue de sa rapidité. Si ce changement phénotypique a été observé in vitro, alors il n'est pas impossible d'avoir le même phénomène in vivo.

I.C Proposition d'un modèle de différenciation stochastique

Grâce aux études réalisées au cours de ma thèse, nous proposons un nouveau modèle explicatif, stochastique de différenciation cellulaire (Figure 43) :

- Au départ, à temps 0h, les cellules sont non différenciées.
- Durant les premières 24 heures de mise en culture, les cellules changeraient d'état et se trouveraient dans un état « multi-primed », métastable, où elles exprimeraient plusieurs gènes de lignées différentes à un niveau modéré. Les cellules seraient capables d'exprimer à la fois des gènes des lignées myéloïdes et des gènes des lignées lymphoïdes.
- 24 heures plus tard, les cellules n'exprimeraient plus tous les gènes mais fluctueraient entre deux ou plusieurs profils transcriptomiques : c'est l'étape dite « hésitante », incertaine. La cellule n'aurait pas encore fait son choix entre les différents types de

phénotypes, elle se trouverait dans une phase d'instabilité phénotypique. C'est ce que l'on a observé dans la morphologie des cellules CD34+ : des cellules fluctuant entre le phénotype rond et polarisé.

– Une fois que la cellule a fait son choix de façon aléatoire, elle retrouverait un état stable.

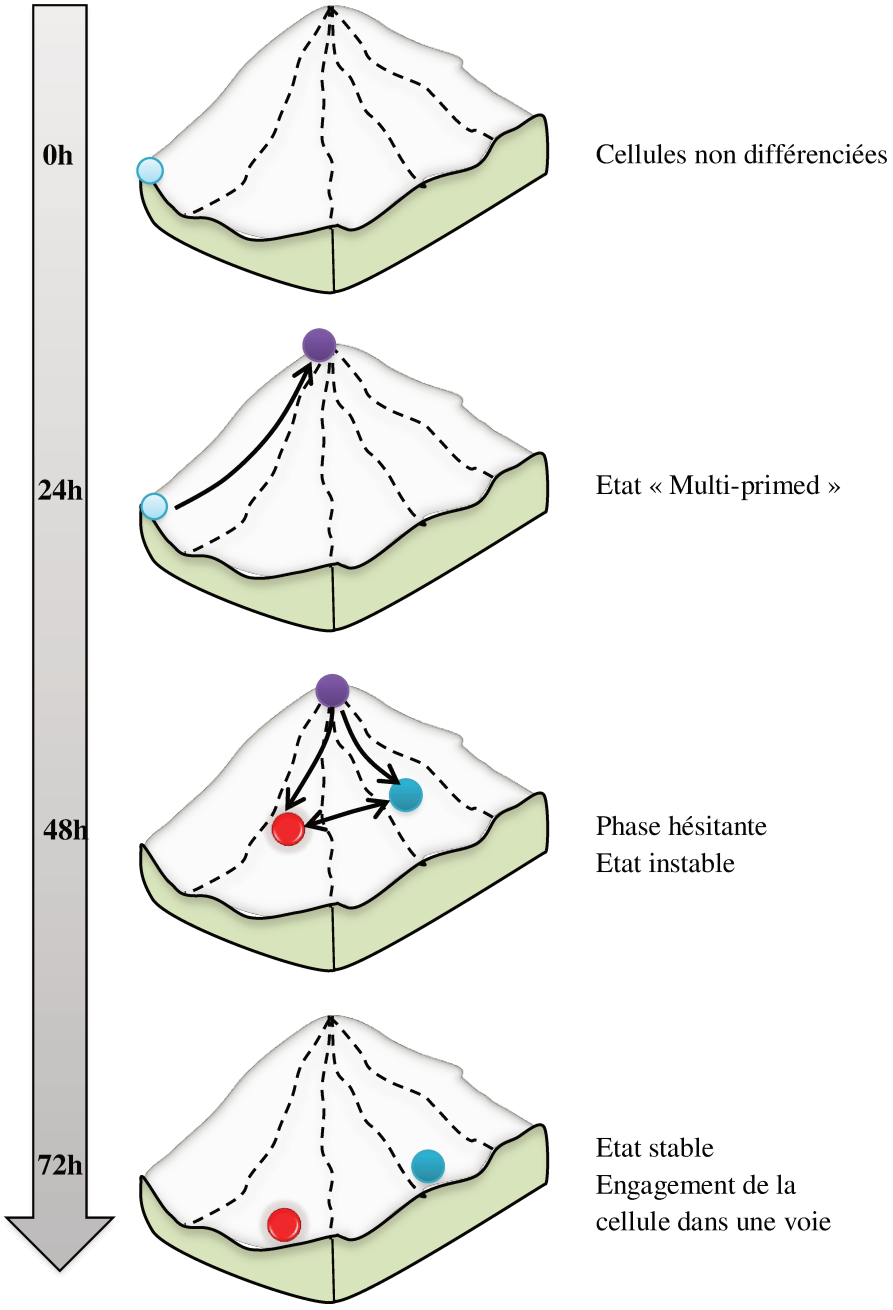


Figure 43 : Proposition d'un modèle stochastique de différenciation.

A temps 0h, les cellules sont non différenciées. Rapidement, à 24h, les cellules se trouvent dans un état dit « multi-primed » où elles expriment plusieurs gènes de lignées différentes, à un niveau modéré. A 48h, les cellules se trouvent dans un état instable où elles hésitent entre deux phénotypes. A 72h, elles ont « choisi » leur phénotype et se trouvent dans un état stable. Les différents événements se déroulent assez vite entre l'état non différencié et le choix du devenir cellulaire.

Au cours de ma thèse, nous avons pu identifier les premières phases de l'engagement du devenir des cellules CD34+ humaines issues de sang de cordon et nous avons pu définir une échelle de temps pour la différenciation cellulaire. Nous avons démontré que la prise de décision était un processus qui ne nécessitait pas de division cellulaire et qu'elle passait par un état transitoire, dynamique, de longueur variable pouvant durer sur plusieurs cycles cellulaires. Chaque cellule explore différentes possibilités en exprimant une grande variété de gènes avant de trouver une combinaison stable et de s'engager réellement dans une voie spécifique.

II. Perspectives

Grâce à l'étude de ces deux modèles de différenciation à l'échelle de la cellule unique, nous voyons bien que les cellules, prises individuellement, ont une certaine probabilité de s'engager dans une voie et que les événements ne sont pas figés ; une part aléatoire existe dans ce processus. Pour étudier le comportement des LT, nous avons créé un modèle mathématique basé sur nos observations. Cela nous a permis de mieux comprendre le comportement de ces cellules. L'avantage des modèles informatiques est que nous pouvons changer plusieurs paramètres, simultanément. Nous nous sommes d'abord basés sur un modèle « simple » de différenciation, les LT en LTreg où seulement trois paramètres ont été étudiés : la prolifération, la différenciation et la mort cellulaire. Dans ce cas, nous avons vu que seul le modèle stochastique pouvait expliquer leur comportement. Nous pourrions aller plus loin en étendant ce modèle aux cellules CD34+. La difficulté dans ce type cellulaire est que les cellules n'ont pas qu'une seule voie de différenciation mais elles ont plusieurs choix phénotypiques. Nous ne nous retrouvons plus dans un système Markovien mais dans un système beaucoup plus complexe avec une possibilité de réversibilité (fluctuation entre la forme ronde et polarisée). Dans les LT, nous n'avons jamais observé une dédifférenciation des LTreg en LT. De plus, dans le cas des cellules CD34+, la distribution de chaque événement change avec le temps : il faut le prendre en compte ce qui complexifie encore plus le modèle.

Dans le cas des cellules CD34+, nous avons analysé 90 gènes. Une expérience supplémentaire serait de faire un transcriptome total de tous les ARN de la cellule unique à des temps différents. Cela nous permettrait d'avoir une vue d'ensemble des gènes qui changent au cours du temps. En sélectionnant les gènes, nous savons que nous perdons de l'information. Néanmoins, le séquençage est une technologie qualitative alors que la RT-qPCR est quantitative, ce qui permet de mieux percevoir les variations d'expression génique. L'idéal aurait donc été de faire d'abord un séquençage des ARN totaux et ensuite de sélectionner les gènes qui diffèrent le plus dans le temps comme dans l'étude de Richard et al. mais en réalisant un séquençage sur cellule unique (Richard et al., 2016).

Dans les deux modèles cellulaires, nous avons étudié la prolifération, la différenciation et la mort cellulaire. Dans le cas des cellules CD34+, nous avons aussi étudié le transcriptome de 90 gènes à l'échelle de la cellule unique, leur dynamique morphologique et la cinétique des évènements. Suite aux résultats obtenus, on peut se demander quel évènement biologique

réagit en premier aux changements environnementaux. Est-ce d'abord la chromatine qui change en se réarrangeant et qui a une influence sur les autres événements ? Ou est-ce d'abord le métabolisme, le transcriptome, les protéines qui changent en premier ?

Des études ont montré que des propriétés métaboliques spécifiques étaient nécessaires pour maintenir l'auto-renouvellement et la multipotence des CSH (Folmes et al., 2012; Oburoglu et al., 2014; Suda et al., 2011). Les CSH quiescentes dépendraient principalement d'une glycolyse anaérobie pour la production d'énergie (Simsek et al., 2010; Takubo et al., 2013). La transition entre l'état souche et l'engagement cellulaire dans la différenciation s'accompagnerait d'un changement métabolique critique passant de la glycolyse à une phosphorylation oxydative (Piccoli et al., 2013; Piro et al., 2011). On peut alors se demander si le métabolisme a un rôle essentiel dans le processus de différenciation. Quelle est la part du métabolisme dans le devenir des cellules ? A-t-il une importance capitale dans le choix du devenir cellulaire ? Si oui, à quel moment intervient-il ? Précède-t-il tous les événements biologiques ? Est-il l'élément déclencheur de l'hypertranscription observée à 24h ? Et est ce qu'il a une influence sur la chromatine, la transcription, la morphologie dynamique ?
Afind de connaître son rôle dans la différenciation et le devenir cellulaire, il serait intéressant d'étudier l'effet de perturbateurs du métabolisme sur le transcriptome des cellules à l'échelle de la cellule unique. Neildez-Nguyen et al. ont étudié l'effet de l'oxygénation sur la prolifération des LTreg en comparant des cultures de LTreg primaires de souris sous 21% d'O_2 et sous 5% d'O_2 (Neildez-Nguyen et al., 2015). Ils ont démontré que lorsque les LT étaient soumis à des concentrations d'oxygène faibles, leur taux de prolifération augmentait. Une idée serait d'étudier l'influence de la concentration d'oxygène sur les cellules CD34+. On pourrait aussi utiliser, à des temps différents, des perturbateurs du métabolisme comme des inhibiteurs des voies glycolytiques ou des voies oxydatives tels que PS48, dinitrophénol ou l'acide bromo pyruvique. Ces molécules ont pour conséquence une diminution de la formation d'ATP et donc entraine un stress métabolique (Dunlop, 1934; Hindie et al., 2009; Sadowska-Bartosz and Bartosz, 2013).

De plus, la structure chromatinienne semble être un élément essentiel dans le processus de différenciation. Nos études de transcription réalisées sur les cellules CD34+ uniques traitées avec VPA ont montré que les cellules, dans ces conditions, sont bloquées dans un motif d'expression génique typique d'un état « multi-primed ». Pour rappel, VPA est un inhibiteur des HDAC c'est-à-dire qu'il permet de maintenir le groupement acétyle sur la chromatine et donc de garder la chromatine ouverte. Il a déjà été montré que les inhibiteurs des HDAC

augmentaient la stochasticité de l'expression des gènes en augmentant l'acétylation de la chromatine (Viñuelas et al., 2013).

Notre étude peut donc aller plus loin en effectuant des analyses cinétiques de la chromatine. Nous savons que les premiers changements s'opèrent dans les premières 48 heures et sans division cellulaire. Le but serait d'approfondir l'échelle de temps dans cette période, par exemple, en commençant par analyser les cellules toutes les 12h. La nouvelle technologie pour analyser l'état de la chromatine s'appelle ATAC-seq (assay for transposase-accessible chromatin with high throughput sequencing). Cette méthode détecte l'accessibilité de l'ADN avec une enzyme, la transposase Tn5 hyperactive, qui insère des adaptateurs de séquençage dans des régions accessibles de la chromatine (Buenrostro et al., 2001; Corces et al., 2016). Une étude récente a montré qu'il était possible d'analyser l'état chromatinien d'une seule cellule avec le système C1 de Fluidigm (Buenrostro et al., 2015). Il serait donc intéressant d'utiliser cette technologie à l'échelle de la cellule unique pour connaître l'hétérogénéité chromatinienne et faire le lien avec le transcriptome. On peut alors se demander si le réarrangement de la chromatine précède les changements observés dans l'expression des gènes.

De plus, dans l'étude sur les cellules CD34+, nous n'avons étudié qu'un seul type de perturbateur épigénétique, VPA. VPA est un acide aliphatique qui inhibe les HDAC de classe I et IIa et qui dégrade la protéine HDAC2, contrairement à d'autres molécules comme trichostatine A qui ne la dégradent pas mais inhibe son action (Krämer et al., 2003). Une étude intéressante serait de comparer l'effet de plusieurs perturbateurs sur le devenir des cellules CD34+. Aurait-on le même effet sur le devenir cellulaire avec un autre inhibiteur des HDAC que VPA ? Est-ce que l'effet de VPA sur le devenir cellulaire est dû à son mode d'action (dégrader la protéine HDAC2) ou à son effet final sur la chromatine (maintien de l'ouverture de la chromatine) ? On pourrait donc comparer l'effet d'autres molécules inhibitrices des HDAC comme trichostatine A ou d'autres molécules inhibitrices des HDAC de classe III (famille des sirtuines). On pourrait aussi étudier d'autres molécules qui ont le même effet sur la chromatine mais qui n'ont pas le même mécanisme d'action comme les inhibiteurs des méthyltransférases d'ADN (DNMT). Ces molécules, telles que la décitabine ou l'azacitidine, entrainent une hypométhylation et un remodelage de la chromatine et sont connues pour avoir un rôle sur le devenir des cellules CD34+ (Jackson et al., 2004; Milhem et al., 2004). Elles sont aussi utilisées pour obtenir des cellules souches pluripotentes induites à partir de cellules somatiques (Huangfu et al., 2008; Mikkelsen et al., 2008; Wang and Na, 2011). Les mécanismes épigénétiques comme les inhibiteurs des HDAC auraient donc une réelle influence sur le devenir des cellules.

Une autre étude intéressante serait de comparer le profil transcriptomique des cellules sœurs et des cellules cousines. Pourquoi les cellules sœurs peuvent-elles avoir des comportements différents alors qu'elles sont issues de la même cellule et qu'elles se trouvent dans le même environnement ? Cela permettrait de comprendre comment la cellule fait son choix pour s'engager dans une voie de différenciation. Si nous reprenons les analyses du comportement des LT, nous avons observé des clones hétérogènes c'est-à-dire que dans la population finale, il y avait des cellules différenciées (les LTreg) et des cellules non différenciées (les LT). Comment expliquer que les cellules sœurs ont des comportements différents ? Pourquoi une cellule décide-t-elle de s'engager dans une voie alors que sa sœur, pourtant soumise au même environnement, reste indifférente ?

Nous avons aussi observé ce phénomène dans les cellules CD34+ : une cellule pouvait avoir un phénotype stable rond ou polarisé, alors que sa sœur pouvait fluctuer rapidement entre les deux formes. Enfin, dans les deux modèles cellulaires, nous avons observé des temps de division différents entre les cellules sœurs et les cellules cousines.

Pour comprendre cette différence entre les cellules sœurs et les cellules cousines, une solution serait de filmer la cellule mère afin d'avoir l'information de sa dynamique morphologique. Une fois qu'elle se serait divisée, il faudrait séparer les deux cellules néoformées pour analyser leur transcriptome. La comparaison des temps de division des cellules sœurs n'est pas récente, des analyses dans les années 80 ont déjà démontré qu'il pouvait être différent (Nakahata and Ogawa, 1982; Suda et al., 1984).

Toute la difficulté dans cette démarche est de séparer les cellules issues d'un même clone. Dans notre système actuel avec les microgrilles, il est impossible de le faire car les puits sont beaucoup trop petits et on risquerait de perdre les cellules en bougeant la boite de culture. Pour contourner ce problème, une solution serait d'utiliser une puce microfluidique comme expliquée dans la Partie 3 I.B de l'introduction. On pourrait ainsi suivre la cellule mère en TL, récupérer ses cellules filles néoformées et faire du transcriptome sur quelques gènes ou sur plusieurs gènes.

Enfin, bien comprendre le processus de différenciation a un intérêt primordial dans les applications thérapeutiques telles que les cancers ou les maladies monogéniques. Actuellement, les patients atteints de maladies du sang sont soignés soit par thérapie cellulaire, soit par thérapie génique. Dans les deux cas, une fois les cellules extraites, la population des CSH est enrichie sur le critère de la protéine CD34 et amplifiée grâce à un cocktail de cytokines. Or, nous savons que le nombre de cellules souches dans cette population CD34+ est très faible et que plus on purifie la population, plus on perd le potentiel

multipotent des cellules (Quesenberry et al., 2014). Est-ce dû à leur manipulation ou aux critères phénotypiques peut être mal choisis ? Selon moi, il faudrait considérer la cellule souche comme un état dynamique et non comme une entité bien connue, bien définie. Par exemple, nous savons que le VPA augmente le nombre de cellules multipotentes en transformant les cellules de départ (et non en les sélectionnant) et améliore la transplantation (Bug et al., 2005; Chaurasia et al., 2014; Seet et al., 2009). Une solution alternative pour amplifier le nombre de cellules souches serait donc de transformer toute la population cellulaire et non de trier une sous population, qui ne contiendra, au final, que très peu de cellules souches. C'est une stratégie thérapeutique différente de celle proposée actuellement. Cette vision alternative considère donc la cellule souche plutôt comme un état dynamique avec plusieurs phénotypes possibles qu'une entité bien unique.

BIBLIOGRAPHIE

Abkowitz, J.L., Catlin, S.N., Guttorp, P., 1996. Evidence that hematopoiesis may be a stochastic process in vivo. Nat. Med. 2, 190–197.

Acuff, N.V., Linden, J., 2017. Using Visualization of t-Distributed Stochastic Neighbor Embedding To Identify Immune Cell Subsets in Mouse Tumors. J. Immunol. 198, 4539–4546. https://doi.org/10.4049/jimmunol.1602077

Adolfsson, J., Månsson, R., Buza-Vidas, N., Hultquist, A., Liuba, K., Jensen, C.T., Bryder, D., Yang, L., Borge, O.-J., Thoren, L.A.M., Anderson, K., Sitnicka, E., Sasaki, Y., Sigvardsson, M., Jacobsen, S.E.W., 2005. Identification of Flt3+ Lympho-Myeloid Stem Cells Lacking Erythro-Megakaryocytic Potential. Cell 121, 295–306. https://doi.org/10.1016/j.cell.2005.02.013

Aiuti, A., Biasco, L., Scaramuzza, S., Ferrua, F., Cicalese, M.P., Baricordi, C., Dionisio, F., Calabria, A., Giannelli, S., Castiello, M.C., Bosticardo, M., Evangelio, C., Assanelli, A., Casiraghi, M., Nunzio, S.D., Callegaro, L., Benati, C., Rizzardi, P., Pellin, D., Serio, C.D., Schmidt, M., Kalle, C.V., Gardner, J., Mehta, N., Neduva, V., Dow, D.J., Galy, A., Miniero, R., Finocchi, A., Metin, A., Banerjee, P.P., Orange, J.S., Galimberti, S., Valsecchi, M.G., Biffi, A., Montini, E., Villa, A., Ciceri, F., Roncarolo, M.G., Naldini, L., 2013. Lentiviral Hematopoietic Stem Cell Gene Therapy in Patients with Wiskott-Aldrich Syndrome. Science 341, 1233151. https://doi.org/10.1126/science.1233151

Akashi, K., Traver, D., Miyamoto, T., Weissman, I.L., 2000. A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. Nature 404, 193–197. https://doi.org/10.1038/35004599

Akel, S., Kolialexi, A., Mavrou, A., Metaxotou, C., Loukopoulos, D., Yataganas, X., 2000. Evaluation at single cell level of residual Philadelphia negative hemopoietic stem cells in chronic phase CML patients. Cancer Genet. Cytogenet. 122, 93–100. https://doi.org/10.1016/S0165-4608(00)00280-6

Allen, R.D., Zacharski, L.R., Widirstky, S.T., Rosenstein, R., Zaitlin, L.M., Burgess, D.R., 1979. Transformation and motility of human platelets: details of the shape change and release reaction observed by optical and electron microscopy. J. Cell Biol. 83, 126–142. https://doi.org/10.1083/jcb.83.1.126

Amir, E.D., Davis, K.L., Tadmor, M.D., Simonds, E.F., Levine, J.H., Bendall, S.C., Shenfeld, D.K., Krishnaswamy, S., Nolan, G.P., Pe'er, D., 2013. viSNE enables visualization of high dimensional single-cell data and reveals phenotypic heterogeneity of leukemia. Nat. Biotechnol. 31, 545–552. https://doi.org/10.1038/nbt.2594

Angelo, M., Bendall, S.C., Finck, R., Hale, M.B., Hitzman, C., Borowsky, A.D., Levenson, R.M., Lowe, J.B., Liu, S.D., Zhao, S., Natkunam, Y., Nolan, G.P., 2014. Multiplexed ion beam imaging of human breast tumors. Nat. Med. 20, 436–442. https://doi.org/10.1038/nm.3488

Angermueller, C., Clark, S.J., Lee, H.J., Macaulay, I.C., Teng, M.J., Hu, T.X., Krueger, F., Smallwood, S.A., Ponting, C.P., Voet, T., Kelsey, G., Stegle, O., Reik, W., 2016. Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. Nat. Methods 13, 229–232. https://doi.org/10.1038/nmeth.3728

Aranyi, T., Stockholm, D., Yao, R., Poinsignon, C., Wiart, T., Corre, G., Touleimat, N., Tost, J., Galy, A., Paldi, A., 2016. Systemic epigenetic response to recombinant lentiviral vectors independent of proviral integration. Epigenetics Chromatin 9, 29. https://doi.org/10.1186/s13072-016-0077-1

Artavanis-Tsakonas, S., Rand, M.D., Lake, R.J., 1999. Notch signaling: cell fate control and signal integration in development. Science 284, 770–776.

Banchereau, J., Steinman, R.M., 1998. Dendritic cells and the control of immunity. Nature 392, 245–252. https://doi.org/10.1038/32588

Baum, C., Düllmann, J., Li, Z., Fehse, B., Meyer, J., Williams, D.A., Kalle, C. von, 2003. Side effects of retroviral gene transfer into hematopoietic stem cells. Blood 101, 2099–2113. https://doi.org/10.1182/blood-2002-07-2314

Baum, C.M., Weissman, I.L., Tsukamoto, A.S., Buckle, A.M., Peault, B., 1992. Isolation of a candidate human hematopoietic stem-cell population. Proc. Natl. Acad. Sci. U. S. A. 89, 2804–2808.

Becher, B., Schlitzer, A., Chen, J., Mair, F., Sumatoh, H.R., Teng, K.W.W., Low, D., Ruedl, C., Riccardi-Castagnoli, P., Poidinger, M., Greter, M., Ginhoux, F., Newell, E.W., 2014. High-dimensional analysis of the murine myeloid cell system. Nat. Immunol. 15, 1181–1189. https://doi.org/10.1038/ni.3006

Beckmann, J., Scheitza, S., Wernet, P., Fischer, J.C., Giebel, B., 2007. Asymmetric cell division within the human hematopoietic stem and progenitor cell compartment: identification of asymmetrically segregating proteins. Blood 109, 5494–5501. https://doi.org/10.1182/blood-2006-11-055921

Behbehani, G.K., Bendall, S.C., Clutter, M.R., Fantl, W.J., Nolan, G.P., 2012. Single-cell mass cytometry adapted to measurements of the cell cycle. Cytometry A 81A, 552–566. https://doi.org/10.1002/cyto.a.22075

Benavente-Babace, A., Gallego-Pérez, D., Hansford, D.J., Arana, S., Pérez-Lorenzo, E., Mujika, M., 2014. Single-cell trapping and selective treatment via co-flow within a microfluidic platform. Biosens. Bioelectron. 61, 298–305. https://doi.org/10.1016/j.bios.2014.05.036

Bendall, S.C., Davis, K.L., Amir, E.D., Tadmor, M.D., Simonds, E.F., Chen, T.J., Shenfeld, D.K., Nolan, G.P., Pe'er, D., 2014. Single-Cell Trajectory Detection Uncovers Progression and Regulatory Coordination in Human B Cell Development. Cell 157, 714–725. https://doi.org/10.1016/j.cell.2014.04.005

Bendall, S.C., Simonds, E.F., Qiu, P., Amir, E.D., Krutzik, P.O., Finck, R., Bruggner, R.V., Melamed, R., Trejo, A., Ornatsky, O.I., Balderas, R.S., Plevritis, S.K., Sachs, K., Pe'er, D., Tanner, S.D., Nolan, G.P., 2011. Single-Cell Mass Cytometry of Differential Immune and Drug Responses Across a Human Hematopoietic Continuum. Science 332, 687–696. https://doi.org/10.1126/science.1198704

Bergmann, C.E., Hoefer, I.E., Meder, B., Roth, H., Royen, N. van, Breit, S.M., Jost, M.M., Aharinejad, S., Hartmann, S., Buschmann, I.R., 2006. Arteriogenesis depends on circulating monocytes and macrophage accumulation and is severely depressed in op/op mice. J. Leukoc. Biol. 80, 59–65. https://doi.org/10.1189/jlb.0206087

Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., Jaenisch, R., Wagschal, A., Feil, R., Schreiber, S.L., Lander, E.S., 2006. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell 125, 315–326. https://doi.org/10.1016/j.cell.2006.02.041

Bertrand, J.Y., Chi, N.C., Santoso, B., Teng, S., Stainier, D.Y.R., Traver, D., 2010. Haematopoietic stem cells derive directly from aortic endothelium during development. Nature 464, 108–111. https://doi.org/10.1038/nature08738

Bilate, A.M., Lafaille, J.J., 2012. Induced CD4+Foxp3+ regulatory T cells in immune tolerance. Annu. Rev. Immunol. 30, 733–758. https://doi.org/10.1146/annurev-immunol-020711-075043

Blaese, R.M., 1993. Development of gene therapy for immunodeficiency: adenosine deaminase deficiency. Pediatr. Res. 33, S49-53; discussion S53-55. https://doi.org/10.1203/00006450-199305001-00278

Blau, H.M., Brazelton, T.R., Weimann, J.M., 2001. The Evolving Concept of a Stem Cell. Cell 105, 829–841. https://doi.org/10.1016/S0092-8674(01)00409-3

Bloom, W., Bartelmez, G.W., 1940. Hematopoiesis in young human embryos. Am. J. Anat. 67, 21–53. https://doi.org/10.1002/aja.1000670103

Bloom, W., Fawcett, D.W., Maximow, A.A., 1975. A textbook of histology, 10th ed. ed. Saunders, Philadelphia.

Bodenmiller, B., Zunder, E.R., Finck, R., Chen, T.J., Savig, E.S., Bruggner, R.V., Simonds, E.F., Bendall, S.C., Sachs, K., Krutzik, P.O., Nolan, G.P., 2012. Multiplexed mass cytometry profiling of cellular states perturbed by small-molecule regulators. Nat. Biotechnol. 30, 858–867. https://doi.org/10.1038/nbt.2317

Bogdan, P., Deasy, B.M., Gharaibeh, B., Roehrs, T., Marculescu, R., 2014. Heterogeneous Structure of Stem Cells Dynamics: Statistical Models and Quantitative Predictions. Sci. Rep. 4. https://doi.org/10.1038/srep04826

Boisset, J.-C., van Cappellen, W., Andrieu-Soler, C., Galjart, N., Dzierzak, E., Robin, C., 2010. In vivo imaging of haematopoietic cells emerging from the mouse aortic endothelium. Nature 464, 116–120. https://doi.org/10.1038/nature08764

Bortin, M.M., 1970. A compendium of reported human bone marrow transplants. Transplantation 9, 571–587.

Broudy, V.C., Kaushansky, K., 1995. Thrombopoietin, the c-mpl ligand, is a major regulator of platelet production. J. Leukoc. Biol. 57, 719–725.

Brown, G., Bunce, C.M., Guy, G.R., 1985. Sequential determination of lineage potentials during haemopoiesis. Br. J. Cancer 52, 681–686.

Brown, G., Hughes, P.J., Michell, R.H., Rolink, A.G., Ceredig, R., 2007. The sequential determination model of hematopoiesis. Trends Immunol. 28, 442–448. https://doi.org/10.1016/j.it.2007.07.007

Brown, S., Couchy, I., Fraisier, V., Satiat-Jeunemaitre, B., 2001. Microscopies, marquages et imageries de la cellule végétale : vers une analyse structurale et fonctionnelle du protéome.

Broxmeyer, H.E., Douglas, G.W., Hangoc, G., Cooper, S., Bard, J., English, D., Arny, M., Thomas, L., Boyse, E.A., 1989. Human umbilical cord blood as a potential source of transplantable hematopoietic stem/progenitor cells. Proc. Natl. Acad. Sci. U. S. A. 86, 3828–3832.

Bruggner, R.V., Bodenmiller, B., Dill, D.L., Tibshirani, R.J., Nolan, G.P., 2014. Automated identification of stratifying signatures in cellular subpopulations. Proc. Natl. Acad. Sci. 111, E2770–E2777. https://doi.org/10.1073/pnas.1408792111

Buenrostro, J.D., Wu, B., Chang, H.Y., Greenleaf, W.J., 2001. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide, in: Current Protocols in Molecular Biology. John Wiley & Sons, Inc.

Buenrostro, J.D., Wu, B., Litzenburger, U.M., Ruff, D., Gonzales, M.L., Snyder, M.P., Chang, H.Y., Greenleaf, W.J., 2015. Single-cell chromatin accessibility reveals principles of regulatory variation. Nature 523, 486–490. https://doi.org/10.1038/nature14590

Buettner, F., Natarajan, K.N., Casale, F.P., Proserpio, V., Scialdone, A., Theis, F.J., Teichmann, S.A., Marioni, J.C., Stegle, O., 2015. Computational analysis of cell-to-cell heterogeneity in single-cell RNA-sequencing data reveals hidden subpopulations of cells. Nat. Biotechnol. 33, 155–160. https://doi.org/10.1038/nbt.3102

Bug, G., Gül, H., Schwarz, K., Pfeifer, H., Kampfmann, M., Zheng, X., Beissert, T., Boehrer, S., Hoelzer, D., Ottmann, O.G., Ruthardt, M., 2005. Valproic Acid Stimulates Proliferation and Self-renewal of Hematopoietic Stem Cells. Cancer Res. 65, 2537–2541. https://doi.org/10.1158/0008-5472.CAN-04-3011

Buggenthin, F., Marr, C., Schwarzfischer, M., Hoppe, P.S., Hilsenbeck, O., Schroeder, T., Theis, F.J., 2013. An automatic method for robust and fast cell detection in bright field images from high-throughput microscopy. BMC Bioinformatics 14, 297. https://doi.org/10.1186/1471-2105-14-297

Carlo, D.D., Wu, L.Y., Lee, L.P., 2006. Dynamic single cell culture array. Lab. Chip 6, 1445–1449. https://doi.org/10.1039/B605937F

Carr, K., Carr, I., 1970. How cells settle on glass: A study by light and scanning electron microscopy of some properties of normal and stimulated macrophages. Z. Für Zellforsch. Mikrosk. Anat. 105, 234–241. https://doi.org/10.1007/BF00335473

Cartier, N., Hacein-Bey-Abina, S., Bartholomae, C.C., Bougnères, P., Schmidt, M., Kalle, C.V., Fischer, A., Cavazzana-Calvo, M., Aubourg, P., 2012. Lentiviral hematopoietic cell gene therapy for X-linked adrenoleukodystrophy. Methods Enzymol. 507, 187–198. https://doi.org/10.1016/B978-0-12-386509-0.00010-7

Cartier, N., Hacein-Bey-Abina, S., Bartholomae, C.C., Veres, G., Schmidt, M., Kutschera, I., Vidaud, M., Abel, U., Dal-Cortivo, L., Caccavelli, L., Mahlaoui, N., Kiermer, V., Mittelstaedt, D., Bellesme, C., Lahlou, N., Lefrère, F., Blanche, S., Audit, M., Payen, E., Leboulch, P., l'Homme, B., Bougnères, P., Von Kalle, C., Fischer, A., Cavazzana-Calvo, M., Aubourg, P., 2009. Hematopoietic stem cell gene therapy with a lentiviral vector in X-linked adrenoleukodystrophy. Science 326, 818–823. https://doi.org/10.1126/science.1171242

Catlin, S.N., Busque, L., Gale, R.E., Guttorp, P., Abkowitz, J.L., 2011. The replication rate of human hematopoietic stem cells in vivo. Blood 117, 4460–4466. https://doi.org/10.1182/blood-2010-08-303537

Cattoglio, C., Pellin, D., Rizzi, E., Maruggi, G., Corti, G., Miselli, F., Sartori, D., Guffanti, A., Serio, C.D., Ambrosi, A., Bellis, G.D., Mavilio, F., 2010. High-definition mapping of retroviral integration sites identifies active regulatory elements in human multipotent hematopoietic progenitors. Blood 116, 5507–5517. https://doi.org/10.1182/blood-2010-05-283523

Cavazzana-Calvo, M., Hacein-Bey, S., Basile, G. de S., Gross, F., Yvon, E., Nusbaum, P., Selz, F., Hue, C., Certain, S., Casanova, J.-L., Bousso, P., Deist, F.L., Fischer, A., 2000. Gene Therapy of Human Severe Combined Immunodeficiency (SCID)-X1 Disease. Science 288, 669–672. https://doi.org/10.1126/science.288.5466.669

Cavazzana-Calvo, M., Payen, E., Negre, O., Wang, G., Hehir, K., Fusil, F., Down, J., Denaro, M., Brady, T., Westerman, K., Cavallesco, R., Gillet-Legrand, B., Caccavelli, L., Sgarra, R., Maouche-Chrétien, L., Bernaudin, F., Girot, R., Dorazio, R., Mulder, G.-J., Polack, A., Bank, A., Soulier, J., Larghero, J., Kabbara, N., Dalle, B., Gourmel, B., Socie, G., Chrétien, S., Cartier, N., Aubourg, P., Fischer, A., Cornetta, K., Galacteros, F., Beuzard, Y., Gluckman, E., Bushman, F., Hacein-Bey-Abina, S., Leboulch, P., 2010. Transfusion independence and HMGA2 activation after gene therapy of human β-thalassaemia. Nature 467, 318–322. https://doi.org/10.1038/nature09328

Chakravorty, R., Rawlinson, D., Zhang, A., Markham, J., Dowling, M.R., Wellard, C., Zhou, J.H.S., Hodgkin, P.D., 2014. Labour-Efficient In Vitro Lymphocyte Population Tracking and Fate Prediction Using Automation and Manual Review. PLOS ONE 9, e83251. https://doi.org/10.1371/journal.pone.0083251

Chang, H.H., Hemberg, M., Barahona, M., Ingber, D.E., Huang, S., 2008. Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. Nature 453, 544–547. https://doi.org/10.1038/nature06965

Charrier, S., Dupré, L., Scaramuzza, S., Jeanson-Leh, L., Blundell, M.P., Danos, O., Cattaneo, F., Aiuti, A., Eckenberg, R., Thrasher, A.J., Roncarolo, M.G., Galy, A., 2007. Lentiviral vectors targeting WASp expression to hematopoietic cells, efficiently transduce and correct cells from WAS patients. Gene Ther. 14, 415–428. https://doi.org/10.1038/sj.gt.3302863

Chaurasia, P., Gajzer, D.C., Schaniel, C., D'Souza, S., Hoffman, R., 2014. Epigenetic reprogramming induces the expansion of cord blood stem cells. J. Clin. Invest. 124, 2378–2395. https://doi.org/10.1172/JCI70313

Chen, H., Sun, J., Wolvetang, E., Cooper-White, J., 2015. High-throughput, deterministic single cell trapping and long-term clonal cell culture in microfluidic devices. Lab. Chip 15, 1072–1083. https://doi.org/10.1039/C4LC01176G

Chen, M.J., Yokomizo, T., Zeigler, B.M., Dzierzak, E., Speck, N.A., 2009. Runx1 is required for the endothelial to haematopoietic cell transition but not thereafter. Nature 457, 887–891. https://doi.org/10.1038/nature07619

Chester, C., Maecker, H.T., 2015. Algorithmic Tools for Mining High-Dimensional Cytometry Data. J. Immunol. 195, 773–779. https://doi.org/10.4049/jimmunol.1500633

Chirumbolo, S., 2012. State-of-the-art review about basophil research in immunology and allergy: is the time right to treat these cells with the respect they deserve? Blood Transfus. 10, 148–168. https://doi.org/10.2450/2011.0020-11

Choi, K., Kennedy, M., Kazarov, A., Papadimitriou, J.C., Keller, G., 1998. A common precursor for hematopoietic and endothelial cells. Development 125, 725–732.

Chotinantakul, K., Leeanansaksiri, W., 2012. Hematopoietic Stem Cell Development, Niches, and Signaling Pathways. Bone Marrow Res. 2012. https://doi.org/10.1155/2012/270425

Civin, C.I., Strauss, L.C., Brovall, C., Fackler, M.J., Schwartz, J.F., Shaper, J.H., 1984. Antigenic analysis of hematopoiesis. III. A hematopoietic progenitor cell surface antigen defined by a monoclonal antibody raised against KG-1a cells. J. Immunol. 133, 157–165.

Clausell-Tormos, J., Lieber, D., Baret, J.-C., El-Harrak, A., Miller, O.J., Frenz, L., Blouwolff, J., Humphry, K.J., Köster, S., Duan, H., Holtze, C., Weitz, D.A., Griffiths, A.D., Merten, C.A., 2008. Droplet-Based Microfluidic Platforms for the Encapsulation and Screening of Mammalian Cells and Multicellular Organisms. Chem. Biol. 15, 427–437. https://doi.org/10.1016/j.chembiol.2008.04.004

Coifman, R.R., Lafon, S., Lee, A.B., Maggioni, M., Nadler, B., Warner, F., Zucker, S.W., 2005. Geometric diffusions as a tool for harmonic analysis and structure definition of data: diffusion maps. Proc. Natl. Acad. Sci. U. S. A. 102, 7426–7431. https://doi.org/10.1073/pnas.0500334102

Corces, M.R., Buenrostro, J.D., Wu, B., Greenside, P.G., Chan, S.M., Koenig, J.L., Snyder, M.P., Pritchard, J.K., Kundaje, A., Greenleaf, W.J., Majeti, R., Chang, H.Y., 2016. Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution. Nat. Genet. 48, 1193–1203. https://doi.org/10.1038/ng.3646

Corre, G., Stockholm, D., Arnaud, O., Kaneko, G., Viñuelas, J., Yamagata, Y., Neildez-Nguyen, T.M.A., Kupiec, J.-J., Beslon, G., Gandrillon, O., Paldi, A., 2014. Stochastic Fluctuations and Distributed Control of Gene Expression Impact Cellular Memory. PLOS ONE 9, e115574. https://doi.org/10.1371/journal.pone.0115574

Cosette, J., Moussy, A., Onodi, F., Auffret-Cariou, A., Neildez-Nguyen, T.M.A., Paldi, A., Stockholm, D., 2015. Single Cell Dynamics Causes Pareto-Like Effect in Stimulated T Cell Populations. Sci. Rep. 5, 17756. https://doi.org/10.1038/srep17756

Costa, M.R., Ortega, F., Brill, M.S., Beckervordersandforth, R., Petrone, C., Schroeder, T., Götz, M., Berninger, B., 2011. Continuous live imaging of adult neural stem cell division and lineage progression in vitro. Development 138, 1057–1068. https://doi.org/10.1242/dev.061663

Coutu, D.L., Schroeder, T., 2013. Probing cellular processes by long-term live imaging – historic problems and current solutions. J Cell Sci 126, 3805–3815. https://doi.org/10.1242/jcs.118349

Cross, J.C., 2005. How to make a placenta: mechanisms of trophoblast cell differentiation in mice--a review. Placenta 26 Suppl A, S3-9. https://doi.org/10.1016/j.placenta.2005.01.015

Cross, M.A., Enver, T., 1997. The lineage commitment of haemopoietic progenitor cells. Curr. Opin. Genet. Dev. 7, 609–613. https://doi.org/10.1016/S0959-437X(97)80007-X

Cui, C., Shu, W., Li, P., 2016. Fluorescence In situ Hybridization: Cell-Based Genetic Diagnostic and Research Applications. Front. Cell Dev. Biol. 4. https://doi.org/10.3389/fcell.2016.00089

Curry, J.L., J, J.T., Wolf, N., 1967. Hemopoieitc Spleen Colony Studies. J. Exp. Med. 125, 703–720. https://doi.org/10.1084/jem.125.4.703

Day, D., Pham, K., Ludford-Menting, M.J., Oliaro, J., Izon, D., Russell, S.M., Gu, M., 2008. A method for prolonged imaging of motile lymphocytes. Immunol. Cell Biol. 87, 154–158. https://doi.org/10.1038/icb.2008.79

Debili, N., Wendling, F., Cosman, D., Titeux, M., Florindo, C., Dusanter-Fourt, I., Schooley, K., Methia, N., Charon, M., Nador, R., 1995. The Mpl receptor is expressed in the megakaryocytic lineage from late progenitors to platelets. Blood 85, 391–401.

Deeg, H.J., Storb, R., Thomas, E.D., Appelbaum, F., Buckner, C.D., Clift, R.A., Doney, K., Johnson, L., Sanders, J.E., Stewart, P., Sullivan, K.M., Witherspoon, R.P., 1983. Fanconi's anemia treated by allogeneic marrow transplantation. Blood 61, 954–959.

Delhommeau, F., Dupont, S., Della Valle, V., James, C., Trannoy, S., Massé, A., Kosmider, O., Le Couedic, J.-P., Robert, F., Alberdi, A., Lécluse, Y., Plo, I., Dreyfus, F.J., Marzac, C., Casadevall, N., Lacombe, C., Romana, S.P., Dessen, P., Soulier, J., Viguié, F., Fontenay, M., Vainchenker, W., Bernard, O.A., 2009. Mutation in TET2 in myeloid cancers. N. Engl. J. Med. 360, 2289–2301. https://doi.org/10.1056/NEJMoa0810069

Dieterlen-Lièvre, F., Beaupain, D., Martin, C., 1976. Origin of erythropoietic stem cells in avian development: shift from the yolk sac to an intraembryonic site. Ann. Immunol. 127, 857–863.

Dominguez, M.H., Chattopadhyay, P.K., Ma, S., Lamoreaux, L., McDavid, A., Finak, G., Gottardo, R., Koup, R.A., Roederer, M., 2013. Highly multiplexed quantitation of gene expression on single cells. J. Immunol. Methods 391, 133–145. https://doi.org/10.1016/j.jim.2013.03.002

Doulatov, S., Notta, F., Laurenti, E., Dick, J.E., 2012. Hematopoiesis: A Human Perspective. Cell Stem Cell 10, 120–136. https://doi.org/10.1016/j.stem.2012.01.006

Duffy, K.R., Wellard, C.J., Markham, J.F., Zhou, J.H.S., Holmberg, R., Hawkins, E.D., Hasbold, J., Dowling, M.R., Hodgkin, P.D., 2012. Activation-induced B cell fates are selected by intracellular stochastic competition. Science 335, 338–341. https://doi.org/10.1126/science.1213230

Dumoulin, J.C., Land, J.A., Van Montfoort, A.P., Nelissen, E.C., Coonen, E., Derhaag, J.G., Schreurs, I.L., Dunselman, G.A., Kester, A.D., Geraedts, J.P., Evers, J.L., 2010. Effect of in vitro culture of human embryos on birthweight of newborns. Hum. Reprod. Oxf. Engl. 25, 605–612. https://doi.org/10.1093/humrep/dep456

Dunlop, D.M., 1934. The use of 2:4-dinitrophenol as a metabolic stimulant. Br. Med. J. 1, 524–527.

Durdik, M., Kosik, P., Kruzliakova, J., Jakl, L., Markova, E., Belyaev, I., Durdik, M., Kosik, P., Kruzliakova, J., Jakl, L., Markova, E., Belyaev, I., 2017. Hematopoietic stem/progenitor cells are less prone to undergo apoptosis than lymphocytes despite similar DNA damage response. Oncotarget 5. https://doi.org/10.18632/oncotarget.16455

Dvorak, A.M., Estrella, P., Ishizaka, T., 1994. Vesicular transport of peroxidase in human eosinophilic myelocytes. Clin. Exp. Allergy J. Br. Soc. Allergy Clin. Immunol. 24, 10–18.

Dzierzak, E., Speck, N.A., 2008. Of lineage and legacy: the development of mammalian hematopoietic stem cells. Nat. Immunol. 9, 129–136. https://doi.org/10.1038/ni1560

Edd, J.F., Di Carlo, D., Humphry, K.J., Köster, S., Irimia, D., Weitz, D.A., Toner, M., 2008. Controlled encapsulation of single-cells into monodisperse picolitre drops. Lab. Chip 8, 1262–1264. https://doi.org/10.1039/b805456h

Eilken, H.M., Nishikawa, S.-I., Schroeder, T., 2009. Continuous single-cell imaging of blood generation from haemogenic endothelium. Nature 457, 896–900. https://doi.org/10.1038/nature07760

Eldar, A., Elowitz, M.B., 2010. Functional roles for noise in genetic circuits. Nature 467, 167–173. https://doi.org/10.1038/nature09326

Elowitz, M.B., Levine, A.J., Siggia, E.D., Swain, P.S., 2002. Stochastic Gene Expression in a Single Cell. Science 297, 1183–1186. https://doi.org/10.1126/science.1070919

Enger, J., Goksör, M., Ramser, K., Hagberg, P., Hanstorp, D., 2004. Optical tweezers applied to a microfluidic system. Lab. Chip 4, 196–200. https://doi.org/10.1039/B307960K

Enver, T., Greaves, M., 1998. Loops, Lineage, and Leukemia. Cell 94, 9–12. https://doi.org/10.1016/S0092-8674(00)81215-5

Enver, T., Pera, M., Peterson, C., Andrews, P.W., 2009. Stem Cell States, Fates, and the Rules of Attraction. Cell Stem Cell 4, 387–397. https://doi.org/10.1016/j.stem.2009.04.011

Epstein, R.B., Graham, T.C., Buckner, C.D., Bryant, J., Thomas, E.D., 1966. Allogeneic Marrow Engraftment by Cross Circulation in Lethally Irradiated Dogs. Blood 28, 692–707.

Etzrodt, M., Endele, M., Schroeder, T., 2014. Quantitative Single-Cell Approaches to Stem Cell Research. Cell Stem Cell 15, 546–558. https://doi.org/10.1016/j.stem.2014.10.015

Fehling, H.J., Lacaud, G., Kubo, A., Kennedy, M., Robertson, S., Keller, G., Kouskoff, V., 2003. Tracking mesoderm induction and its specification to the hemangioblast during embryonic stem cell differentiation. Dev. Camb. Engl. 130, 4217–4227.

Felice, L.D., Tatarelli, C., Mascolo, M.G., Gregorj, C., Agostini, F., Fiorini, R., Gelmetti, V., Pascale, S., Padula, F., Petrucci, M.T., Arcese, W., Nervi, C., 2005. Histone Deacetylase Inhibitor Valproic Acid Enhances the Cytokine-Induced Expansion of Human Hematopoietic Stem Cells. Cancer Res. 65, 1505–1513. https://doi.org/10.1158/0008-5472.CAN-04-3063

Fernández-Gonzalez, R., Ramirez, M.A., Bilbao, A., De Fonseca, F.R., Gutiérrez-Adán, A., 2007. Suboptimal in vitro culture conditions: an epigenetic origin of long-term health effects. Mol. Reprod. Dev. 74, 1149–1156. https://doi.org/10.1002/mrd.20746

Ferrell, J.E., 2012. Bistability, Bifurcations, and Waddington's Epigenetic Landscape. Curr. Biol. 22, R458–R466. https://doi.org/10.1016/j.cub.2012.03.045

Fischer, A., Landais, P., Friedrich, W., Gerritsen, B., Fasth, A., Porta, F., Vellodi, A., Benkerrou, M., Jais, J.P., Cavazzana-Calvo, M., 1994. Bone marrow transplantation (BMT) in Europe for primary immunodeficiencies other than severe combined immunodeficiency: a report from the European Group for BMT and the European Group for Immunodeficiency. Blood 83, 1149–1154.

Fisson, S., Vendomèle, J., 2017. Bases immunologiques générales, in: Uvéite. Elsevier Masson, pp. 5–12.

Folmes, C.D.L., Dzeja, P.P., Nelson, T.J., Terzic, A., 2012. Metabolic plasticity in stem cell homeostasis and differentiation. Cell Stem Cell 11, 596–606. https://doi.org/10.1016/j.stem.2012.10.002

Ford, C.E., Hamerton, J.L., Barnes, D.W., Loutit, J.F., 1956. Cytological identification of radiation-chimaeras. Nature 177, 452–454.

Fulwyler, M.J., 1980. Flow cytometry and cell sorting. Blood Cells 6, 173–184.

Fulwyler, M.J., 1965. Electronic separation of biological cells by volume. Science 150, 910–911.

Gaspar, H.B., Thrasher, A.J., 2005. Gene therapy for severe combined immunodeficiencies. Expert Opin. Biol. Ther. 5, 1175–1182. https://doi.org/10.1517/14712598.5.9.1175

Gaudilliere, B., Fragiadakis, G.K., Bruggner, R.V., Nicolau, M., Finck, R., Tingle, M., Silva, J., Ganio, E.A., Yeh, C.G., Maloney, W.J., Huddleston, J.I., Goodman, S.B., Davis,

M.M., Bendall, S.C., Fantl, W.J., Angst, M.S., Nolan, G.P., 2014. Clinical recovery from surgery correlates with single-cell immune signatures. Sci. Transl. Med. 6, 255ra131. https://doi.org/10.1126/scitranslmed.3009701

Gawaz, M., Langer, H., May, A.E., 2005. Platelets in inflammation and atherogenesis. J. Clin. Invest. 115, 3378–3384. https://doi.org/10.1172/JCI27196

Geering, B., Stoeckle, C., Conus, S., Simon, H.-U., 2013. Living and dying for inflammation: neutrophils, eosinophils, basophils. Trends Immunol. 34, 398–409. https://doi.org/10.1016/j.it.2013.04.002

Geissmann, F., Manz, M.G., Jung, S., Sieweke, M.H., Merad, M., Ley, K., 2010. Development of Monocytes, Macrophages, and Dendritic Cells. Science 327, 656–661. https://doi.org/10.1126/science.1178331

Gekas, C., Dieterlen-Lièvre, F., Orkin, S.H., Mikkola, H.K.A., 2005. The placenta is a niche for hematopoietic stem cells. Dev. Cell 8, 365–375. https://doi.org/10.1016/j.devcel.2004.12.016

Georgopoulos, K., Bigby, M., Wang, J.-H., Molnar, A., Wu, P., Winandy, S., Sharpe, A., 1994. The ikaros gene is required for the development of all lymphoid lineages. Cell 79, 143–156. https://doi.org/10.1016/0092-8674(94)90407-3

Ghosh, S., Thrasher, A.J., Gaspar, H.B., 2015. Gene therapy for monogenic disorders of the bone marrow. Br. J. Haematol. https://doi.org/10.1111/bjh.13520

Giesen, C., Wang, H.A.O., Schapiro, D., Zivanovic, N., Jacobs, A., Hattendorf, B., Schüffler, P.J., Grolimund, D., Buhmann, J.M., Brandt, S., Varga, Z., Wild, P.J., Günther, D., Bodenmiller, B., 2014. Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. Nat. Methods 11, 417–422. https://doi.org/10.1038/nmeth.2869

Gilner, J.B., Walton, W.G., Gush, K., Kirby, S.L., 2007. Antibodies to Stem Cell Marker Antigens Reduce Engraftment of Hematopoietic Stem Cells. STEM CELLS 25, 279–288. https://doi.org/10.1634/stemcells.2006-0076

Girbl, T., Lunzer, V., Greil, R., Namberger, K., Hartmann, T.N., 2014. The CXCR4 and adhesion molecule expression of CD34+ hematopoietic cells mobilized by "on-demand" addition of plerixafor to granulocyte–colony-stimulating factor. Transfusion (Paris) 54, 2325–2335. https://doi.org/10.1111/trf.12632

Gluckman, E., Devergié, A., Bourdeau-Esperou, H., Thierry, D., Traineau, R., Auerbach, A., Broxmeyer, H.E., 1990. Transplantation of umbilical cord blood in Fanconi's anemia. Nouv. Rev. Fr. Hematol. 32, 423–425.

Gluckman, E., Devergie, A., Schaison, G., Bussel, A., Berger, R., Sohier, J., Bernard, J., 1980. Bone marrow transplantation in Fanconi anaemia. Br. J. Haematol. 45, 557–564.

Gluckman, E., Rocha, V., 2005. History of the clinical use of umbilical cord blood hematopoietic cells. Cytotherapy 7, 219–227. https://doi.org/10.1080/14653240510027136

Goldberg, L.R., Dooner, M.S., Johnson, K.W., Papa, E.F., Pereira, M.G., Del Tatto, M., Adler, D.M., Aliotta, J.M., Quesenberry, P.J., 2014. The murine long-term multi-lineage renewal marrow stem cell is a cycling cell. Leukemia 28, 813–822. https://doi.org/10.1038/leu.2013.252

Goodman, J.W., Hodgson, G.S., 1962. Evidence for Stem Cells in the Peripheral Blood of Mice. Blood 19, 702–714.

Görgens, A., Ludwig, A.-K., Möllmann, M., Krawczyk, A., Dürig, J., Hanenberg, H., Horn, P.A., Giebel, B., 2014. Multipotent Hematopoietic Progenitors Divide Asymmetrically to Create Progenitors of the Lymphomyeloid and Erythromyeloid Lineages. Stem Cell Rep. 3, 1058–1072. https://doi.org/10.1016/j.stemcr.2014.09.016

Görgens, A., Radtke, S., Horn, P.A., Giebel, B., 2013a. New relationships of human hematopoietic lineages facilitate detection of multipotent hematopoietic stem and

progenitor cells. Cell Cycle Georget. Tex 12, 3478–3482. https://doi.org/10.4161/cc.26900
Görgens, A., Radtke, S., Möllmann, M., Cross, M., Dürig, J., Horn, P.A., Giebel, B., 2013b. Revision of the Human Hematopoietic Tree: Granulocyte Subtypes Derive from Distinct Hematopoietic Lineages. Cell Rep. 3, 1539–1552. https://doi.org/10.1016/j.celrep.2013.04.025
Grégori, G., Rajwa, B., Patsekin, V., Jones, J., Furuki, M., Yamamoto, M., Paul Robinson, J., 2014. Hyperspectral cytometry. Curr. Top. Microbiol. Immunol. 377, 191–210. https://doi.org/10.1007/82_2013_359
Grover, A., Mancini, E., Moore, S., Mead, A.J., Atkinson, D., Rasmussen, K.D., O'Carroll, D., Jacobsen, S.E.W., Nerlov, C., 2014. Erythropoietin guides multipotent hematopoietic progenitor cells toward an erythroid fate. J. Exp. Med. 211, 181–188. https://doi.org/10.1084/jem.20131189
Grün, D., Lyubimova, A., Kester, L., Wiebrands, K., Basak, O., Sasaki, N., Clevers, H., van Oudenaarden, A., 2015. Single-cell messenger RNA sequencing reveals rare intestinal cell types. Nature 525, 251–255. https://doi.org/10.1038/nature14966
Gucker, F.T., O'Konski, C.T., Pickard, H.B., Pitts, J.N., 1947. A Photoelectronic Counter for Colloidal Particles1. J. Am. Chem. Soc. 69, 2422–2431. https://doi.org/10.1021/ja01202a053
Guo, G., Huss, M., Tong, G.Q., Wang, C., Li Sun, L., Clarke, N.D., Robson, P., 2010. Resolution of Cell Fate Decisions Revealed by Single-Cell Gene Expression Analysis from Zygote to Blastocyst. Dev. Cell 18, 675–685. https://doi.org/10.1016/j.devcel.2010.02.012
Guo, G., Luc, S., Marco, E., Lin, T.-W., Peng, C., Kerenyi, M.A., Beyaz, S., Kim, W., Xu, J., Das, P.P., Neff, T., Zou, K., Yuan, G.-C., Orkin, S.H., 2013. Mapping Cellular Hierarchy by Single-Cell Analysis of the Cell Surface Repertoire. Cell Stem Cell 13, 492–505. https://doi.org/10.1016/j.stem.2013.07.017
Hacein-Bey-Abina, S., Gaspar, H.B., Blondeau, J., Caccavelli, L., Charrier, S., Buckland, K., Picard, C., Six, E., Himoudi, N., Gilmour, K., McNicol, A.-M., Hara, H., Xu-Bayford, J., Rivat, C., Touzot, F., Mavilio, F., Lim, A., Treluyer, J.-M., Héritier, S., Lefrère, F., Magalon, J., Pengue-Koyi, I., Honnet, G., Blanche, S., Sherman, E.A., Male, F., Berry, C., Malani, N., Bushman, F.D., Fischer, A., Thrasher, A.J., Galy, A., Cavazzana, M., 2015. Outcomes Following Gene Therapy in Patients With Severe Wiskott-Aldrich Syndrome. JAMA 313, 1550–1563. https://doi.org/10.1001/jama.2015.3253
Hacein-Bey-Abina, S., Von Kalle, C., Schmidt, M., McCormack, M.P., Wulffraat, N., Leboulch, P., Lim, A., Osborne, C.S., Pawliuk, R., Morillon, E., Sorensen, R., Forster, A., Fraser, P., Cohen, J.I., de Saint Basile, G., Alexander, I., Wintergerst, U., Frebourg, T., Aurias, A., Stoppa-Lyonnet, D., Romana, S., Radford-Weiss, I., Gross, F., Valensi, F., Delabesse, E., Macintyre, E., Sigaux, F., Soulier, J., Leiva, L.E., Wissler, M., Prinz, C., Rabbitts, T.H., Le Deist, F., Fischer, A., Cavazzana-Calvo, M., 2003. LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. Science 302, 415–419. https://doi.org/10.1126/science.1088547
Haghverdi, L., Buettner, F., Theis, F.J., 2015. Diffusion maps for high-dimensional single-cell analysis of differentiation data. Bioinformatics btv325. https://doi.org/10.1093/bioinformatics/btv325
Hamasaki, N., Yamamoto, M., 2000. Red Blood Cell Function and Blood Storage. Vox Sang. 79, 191–197. https://doi.org/10.1046/j.1423-0410.2000.7940191.x
Han, Y., Gu, Y., Zhang, A.C., Lo, Y.-H., 2016. Review: imaging technologies for flow cytometry. Lab. Chip 16, 4639–4647. https://doi.org/10.1039/c6lc01063f
Hansmann, L., Blum, L., Ju, C.-H., Liedtke, M., Robinson, W.H., Davis, M.M., 2015. Mass Cytometry Analysis Shows That a Novel Memory Phenotype B Cell Is Expanded in

Multiple Myeloma. Cancer Immunol. Res. 3, 650–660. https://doi.org/10.1158/2326-6066.CIR-14-0236-T

Hashimshony, T., Wagner, F., Sher, N., Yanai, I., 2012. CEL-Seq: Single-Cell RNA-Seq by Multiplexed Linear Amplification. Cell Rep. 2, 666–673. https://doi.org/10.1016/j.celrep.2012.08.003

Heissig, B., Hattori, K., Dias, S., Friedrich, M., Ferris, B., Hackett, N.R., Crystal, R.G., Besmer, P., Lyden, D., Moore, M.A.S., Werb, Z., Rafii, S., 2002. Recruitment of stem and progenitor cells from the bone marrow niche requires MMP-9 mediated release of kit-ligand. Cell 109, 625–637.

Hindie, V., Stroba, A., Zhang, H., Lopez-Garcia, L.A., Idrissova, L., Zeuzem, S., Hirschberg, D., Schaeffer, F., Jørgensen, T.J.D., Engel, M., Alzari, P.M., Biondi, R.M., 2009. Structure and allosteric effects of low-molecular-weight activators on the protein kinase PDK1. Nat. Chem. Biol. 5, 758–764. https://doi.org/10.1038/nchembio.208

Hippen, K.L., Merkel, S.C., Schirm, D.K., Nelson, C., Tennis, N.C., Riley, J.L., June, C.H., Miller, J.S., Wagner, J.E., Blazar, B.R., 2011. Generation and large-scale expansion of human inducible regulatory T cells that suppress graft-versus-host disease. Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg. 11, 1148–1157. https://doi.org/10.1111/j.1600-6143.2011.03558.x

Hobbs, J.R., 1981. Bone marrow transplantation for inborn errors. Lancet Lond. Engl. 2, 735–739.

Hoppe, P.S., Coutu, D.L., Schroeder, T., 2014. Single-cell technologies sharpen up mammalian stem cell research. Nat. Cell Biol. 16, 919–927. https://doi.org/10.1038/ncb3042

Hoppe, P.S., Schwarzfischer, M., Loeffler, D., Kokkaliaris, K.D., Hilsenbeck, O., Moritz, N., Endele, M., Filipczyk, A., Gambardella, A., Ahmed, N., Etzrodt, M., Coutu, D.L., Rieger, M.A., Marr, C., Strasser, M.K., Schauberger, B., Burtscher, I., Ermakova, O., Bürger, A., Lickert, H., Nerlov, C., Theis, F.J., Schroeder, T., 2016. Early myeloid lineage choice is not initiated by random PU.1 to GATA1 protein ratios. Nature 535, 299–302. https://doi.org/10.1038/nature18320

Horii, T., Yanagisawa, E., Kimura, M., Morita, S., Hatada, I., 2010. Epigenetic differences between embryonic stem cells generated from blastocysts developed in vitro and in vivo. Cell. Reprogramming 12, 551–563. https://doi.org/10.1089/cell.2009.0104

Hosic, S., Murthy, S.K., Koppes, A.N., 2016. Microfluidic Sample Preparation for Single Cell Analysis. Anal. Chem. 88, 354. https://doi.org/10.1021/acs.analchem.5b04077

Hough, S.R., Laslett, A.L., Grimmond, S.B., Kolle, G., Pera, M.F., 2009. A Continuum of Cell States Spans Pluripotency and Lineage Commitment in Human Embryonic Stem Cells. PLoS ONE 4, e7708. https://doi.org/10.1371/journal.pone.0007708

Hu, M., Krause, D., Greaves, M., Sharkis, S., Dexter, M., Heyworth, C., Enver, T., 1997. Multilineage gene expression precedes commitment in the hemopoietic system. Genes Dev. 11, 774–785. https://doi.org/10.1101/gad.11.6.774

Hu, P., Zhang, W., Xin, H., Deng, G., 2016. Single Cell Isolation and Analysis. Front. Cell Dev. Biol. 4, 116. https://doi.org/10.3389/fcell.2016.00116

Hu, Y., Huang, Y., Yi, Y., Wang, H., Liu, B., Yu, J., Wang, D., 2017. Single-cell RNA sequencing highlights transcription activity of autophagy-related genes during hematopoietic stem cell formation in mouse embryos. Autophagy 13, 770–771. https://doi.org/10.1080/15548627.2016.1278093

Huang, S., 2012. The molecular and mathematical basis of Waddington's epigenetic landscape: a framework for post-Darwinian biology? BioEssays News Rev. Mol. Cell. Dev. Biol. 34, 149–157. https://doi.org/10.1002/bies.201100031

Huang, S., 2009. Non-genetic heterogeneity of cells in development: more than just noise. Development 136, 3853–3862. https://doi.org/10.1242/dev.035139

Huang, S., Eichler, G., Bar-Yam, Y., Ingber, D.E., 2005. Cell fates as high-dimensional attractor states of a complex gene regulatory network. Phys. Rev. Lett. 94, 128701. https://doi.org/10.1103/PhysRevLett.94.128701

Huang, S., Guo, Y.-P., May, G., Enver, T., 2007. Bifurcation dynamics in lineage-commitment in bipotent progenitor cells. Dev. Biol. 305, 695–713. https://doi.org/10.1016/j.ydbio.2007.02.036

Huang, S., Terstappen, L.W., 1994. Lymphoid and myeloid differentiation of single human CD34+, HLA-DR+, CD38- hematopoietic stem cells. Blood 83, 1515–1526.

Huangfu, D., Maehr, R., Guo, W., Eijkelenboom, A., Snitow, M., Chen, A.E., Melton, D.A., 2008. Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. Nat. Biotechnol. 26, 795–797. https://doi.org/10.1038/nbt1418

Huber, T.L., Kouskoff, V., Fehling, H.J., Palis, J., Keller, G., 2004. Haemangioblast commitment is initiated in the primitive streak of the mouse embryo. Nature 432, 625–630. https://doi.org/10.1038/nature03122

Huyhn, A., Dommergues, M., Izac, B., Croisille, L., Katz, A., Vainchenker, W., Coulombel, L., 1995. Characterization of hematopoietic progenitors from human yolk sacs and embryos. Blood 86, 4474–4485.

Isern, J., Méndez-Ferrer, S., 2011. Stem cell interactions in a bone marrow niche. Curr. Osteoporos. Rep. 9, 210–218. https://doi.org/10.1007/s11914-011-0075-y

Islam, S., Kjällquist, U., Moliner, A., Zajac, P., Fan, J.-B., Lönnerberg, P., Linnarsson, S., 2011. Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. Genome Res. 21, 1160–1167. https://doi.org/10.1101/gr.110882.110

Islam, S., Zeisel, A., Joost, S., La Manno, G., Zajac, P., Kasper, M., Lönnerberg, P., Linnarsson, S., 2014. Quantitative single-cell RNA-seq with unique molecular identifiers. Nat. Methods 11, 163–166. https://doi.org/10.1038/nmeth.2772

Jackson, M., Krassowska, A., Gilbert, N., Chevassut, T., Forrester, L., Ansell, J., Ramsahoye, B., 2004. Severe global DNA hypomethylation blocks differentiation and induces histone hyperacetylation in embryonic stem cells. Mol. Cell. Biol. 24, 8862–8871. https://doi.org/10.1128/MCB.24.20.8862-8871.2004

Jacobson, L.O., Simmons, E.L., Marks, E.K., Eldredge, J.H., 1951. Recovery from Radiation Injury. Science 113, 510–511. https://doi.org/10.1126/science.113.2940.510

Jaffredo, T., Gautier, R., Eichmann, A., Dieterlen-Lievre, F., 1998. Intraaortic hemopoietic cells are derived from endothelial cells during ontogeny. Development 125, 4575–4583.

Jobin, C., Cloutier, M., Simard, C., Néron, S., 2015. Heterogeneity of in vitro–cultured CD34+ cells isolated from peripheral blood. Cytotherapy 17, 1472–1484. https://doi.org/10.1016/j.jcyt.2015.05.006

Johnson, F.L., Look, A.T., Gockerman, J., Ruggiero, M.R., Dalla-Pozza, L., Billings, F.T., 1984. Bone-marrow transplantation in a patient with sickle-cell anemia. N. Engl. J. Med. 311, 780–783. https://doi.org/10.1056/NEJM198409203111207

Jubery, T.Z., Srivastava, S.K., Dutta, P., 2014. Dielectrophoretic separation of bioparticles in microdevices: a review. Electrophoresis 35, 691–713. https://doi.org/10.1002/elps.201300424

Judson, G., Jones, A., Kellogg, R., Buckner, D., Eisel, R., Perry, S., Greenough, W., 1968. Closed continuous-flow centrifuge. Nature 217, 816–818.

Julien, E., El Omar, R., Tavian, M., 2016. Origin of the hematopoietic system in the human embryo. FEBS Lett. 590, 3987–4001. https://doi.org/10.1002/1873-3468.12389

Juttner, C.A., To, L.B., Haylock, D.N., Dyson, P.G., 1990. Peripheral blood stem cell selection, collection and auto-transplantation. Prog. Clin. Biol. Res. 333, 447–459; discussion 460.

Kalina, T., Stuchlý, J., Janda, A., Hrušák, O., Růžičková, Š., Šedivá, A., Litzman, J., Vlková, M., 2009. Profiling of polychromatic flow cytometry data on B-cells reveals patients' clusters in common variable immunodeficiency. Cytometry A 75A, 902–909. https://doi.org/10.1002/cyto.a.20801

Kanter, I., Kalisky, T., 2015. Single Cell Transcriptomics: Methods and Applications. Front. Oncol. 5. https://doi.org/10.3389/fonc.2015.00053

Karasuyama, H., Obata, K., Wada, T., Tsujimura, Y., Mukai, K., 2011. Newly appreciated roles for basophils in allergy and protective immunity. Allergy 66, 1133–1141. https://doi.org/10.1111/j.1398-9995.2011.02613.x

Karponi, G., Psatha, N., Lederer, C.W., Adair, J.E., Zervou, F., Zogas, N., Kleanthous, M., Tsatalas, C., Anagnostopoulos, A., Sadelain, M., Rivière, I., Stamatoyannopoulos, G., Yannaki, E., 2015. Plerixafor+G-CSF–mobilized CD34+ cells represent an optimal graft source for thalassemia gene therapy. Blood 126, 616–619. https://doi.org/10.1182/blood-2015-03-629618

Kauffman, S., 1969. Homeostasis and differentiation in random genetic control networks. Nature 224, 177–178.

Kauffman, S.A., 1993. The Origins of Order: Self-organization and Selection in Evolution. Oxford University Press.

Kaufmann, K.B., Büning, H., Galy, A., Schambach, A., Grez, M., 2013. Gene therapy on the move. EMBO Mol. Med. 5, 1642–1661. https://doi.org/10.1002/emmm.201202287

Kawamoto, H., Katsura, Y., 2009. A new paradigm for hematopoietic cell lineages: revision of the classical concept of the myeloid–lymphoid dichotomy. Trends Immunol. 30, 193–200. https://doi.org/10.1016/j.it.2009.03.001

Kawamoto, H., Ohmura, K., Fujimoto, S., Katsura, Y., 1999. Emergence of T Cell Progenitors Without B Cell or Myeloid Differentiation Potential at the Earliest Stage of Hematopoiesis in the Murine Fetal Liver. J. Immunol. 162, 2725–2731.

Kawamoto, H., Ohmura, K., Katsura, Y., 1997. Direct evidence for the commitment of hematopoietic stem cells to T, B and myeloid lineages in murine fetal liver. Int. Immunol. 9, 1011–1019.

Kawamoto, H., Wada, H., Katsura, Y., 2010. A revised scheme for developmental pathways of hematopoietic cells: the myeloid-based model. Int. Immunol. 22, 65–70. https://doi.org/10.1093/intimm/dxp125

Kay, M.A., Manno, C.S., Ragni, M.V., Larson, P.J., Couto, L.B., McClelland, A., Glader, B., Chew, A.J., Tai, S.J., Herzog, R.W., Arruda, V., Johnson, F., Scallan, C., Skarsgard, E., Flake, A.W., High, K.A., 2000. Evidence for gene transfer and expression of factor IX in haemophilia B patients treated with an AAV vector. Nat. Genet. 24, 257–261. https://doi.org/10.1038/73464

Kessinger, A., Armitage, J.O., Landmark, J.D., Smith, D.M., Weisenburger, D.D., 1988. Autologous peripheral hematopoietic stem cell transplantation restores hematopoietic function following marrow ablative therapy. Blood 71, 723–727.

Khattri, R., Cox, T., Yasayko, S.-A., Ramsdell, F., 2003. An essential role for Scurfin in CD4+CD25+ T regulatory cells. Nat. Immunol. 4, 337–342. https://doi.org/10.1038/ni909

Kissa, K., Herbomel, P., 2010. Blood stem cells emerge from aortic endothelium by a novel type of cell transition. Nature 464, 112–115. https://doi.org/10.1038/nature08761

Klein, A.M., Mazutis, L., Akartuna, I., Tallapragada, N., Veres, A., Li, V., Peshkin, L., Weitz, D.A., Kirschner, M.W., 2015. Droplet Barcoding for Single-Cell Transcriptomics Applied to Embryonic Stem Cells. Cell 161, 1187–1201. https://doi.org/10.1016/j.cell.2015.04.044

Klinger, M.H.F., Jelkmann, W., 2002. Role of blood platelets in infection and inflammation. J. Interferon Cytokine Res. Off. J. Int. Soc. Interferon Cytokine Res. 22, 913–922. https://doi.org/10.1089/10799900260286623

Kobel, S., Valero, A., Latt, J., Renaud, P., Lutolf, M., 2010. Optimization of microfluidic single cell trapping for long-term on-chip culture. Lab. Chip 10, 857–863. https://doi.org/10.1039/B918055A

Kokkaliaris, K.D., Loeffler, D., Schroeder, T., 2012. Advances in tracking hematopoiesis at the single-cell level. Curr. Opin. Hematol. 19, 243–249. https://doi.org/10.1097/MOH.0b013e32835421de

Kolaczkowska, E., Kubes, P., 2013. Neutrophil recruitment and function in health and inflammation. Nat. Rev. Immunol. 13, 159–175. https://doi.org/10.1038/nri3399

Kondo, M., Weissman, I.L., Akashi, K., 1997. Identification of clonogenic common lymphoid progenitors in mouse bone marrow. Cell 91, 661–672.

Köster, S., Angilè, F.E., Duan, H., Agresti, J.J., Wintner, A., Schmitz, C., Rowat, A.C., Merten, C.A., Pisignano, D., Griffiths, A.D., Weitz, D.A., 2008. Drop-based microfluidic devices for encapsulation of single cells. Lab. Chip 8, 1110–1115. https://doi.org/10.1039/b802941e

Kosugi, Y., Sato, R., Genka, S., Shitara, N., Takakura, K., 1988. An interactive multivariate analysis of FCM data. Cytometry 9, 405–408. https://doi.org/10.1002/cyto.990090419

Krämer, O.H., Zhu, P., Ostendorff, H.P., Golebiewski, M., Tiefenbach, J., Peters, M.A., Brill, B., Groner, B., Bach, I., Heinzel, T., Göttlicher, M., 2003. The histone deacetylase inhibitor valproic acid selectively induces proteasomal degradation of HDAC2. EMBO J. 22, 3411–3420. https://doi.org/10.1093/emboj/cdg315

Kupiec, J.-J., 1983. A probabilist theory for cell differentiation, embryonic mortality and DNA C-value paradox. Specul. Sci. Technol. 6, 471–478.

Kurimoto, K., Yabuta, Y., Ohinata, Y., Saitou, M., 2007. Global single-cell cDNA amplification to provide a template for representative high-density oligonucleotide microarray analysis. Nat. Protoc. 2, 739–752. https://doi.org/10.1038/nprot.2007.79

Lacaud, G., Kouskoff, V., 2017. Hemangioblast, hemogenic endothelium, and primitive versus definitive hematopoiesis. Exp. Hematol. 49, 19–24. https://doi.org/10.1016/j.exphem.2016.12.009

Laiosa, C.V., Stadtfeld, M., Graf, T., 2006. Determinants of lymphoid-myeloid lineage diversification. Annu. Rev. Immunol. 24, 705–738. https://doi.org/10.1146/annurev.immunol.24.021605.090742

Lancrin, C., Sroczynska, P., Stephenson, C., Allen, T., Kouskoff, V., Lacaud, G., 2009. The haemangioblast generates haematopoietic cells through a haemogenic endothelium stage. Nature 457, 892–895. https://doi.org/10.1038/nature07679

Landaw, S.A., 1988. Factors that accelerate or retard red blood cell senescence. Blood Cells 14, 47–67.

Landecker, H., 2006. Microcinematography and the History of Science and Film. Isis 97, 121–132. https://doi.org/10.1086/501105

Lansdorp, P.M., Sutherland, H.J., Eaves, C.J., 1990. Selective expression of CD45 isoforms on functional subpopulations of CD34+ hemopoietic cells from human bone marrow. J. Exp. Med. 172, 363–366.

Laughlin, M.J., Eapen, M., Rubinstein, P., Wagner, J.E., Zhang, M.-J., Champlin, R.E., Stevens, C., Barker, J.N., Gale, R.P., Lazarus, H.M., Marks, D.I., van Rood, J.J., Scaradavou, A., Horowitz, M.M., 2004. Outcomes after transplantation of cord blood or bone marrow from unrelated donors in adults with leukemia. N. Engl. J. Med. 351, 2265–2275. https://doi.org/10.1056/NEJMoa041276

LeBien, T.W., Tedder, T.F., 2008. B lymphocytes: how they develop and function. Blood 112, 1570–1580. https://doi.org/10.1182/blood-2008-02-078071

Lee, H., Ruane, D., Law, K., Ho, Y., Garg, A., Rahman, A., Esterházy, D., Cheong, C., Goljo, E., Sikora, A.G., Mucida, D., Chen, B.K., Govindraj, S., Breton, G., Mehandru, S., 2015. Phenotype and function of nasal dendritic cells. Mucosal Immunol. 8, 1083–1098. https://doi.org/10.1038/mi.2014.135

Levine, J.H., Simonds, E.F., Bendall, S.C., Davis, K.L., Amir, E.D., Tadmor, M.D., Litvin, O., Fienberg, H.G., Jager, A., Zunder, E.R., Finck, R., Gedman, A.L., Radtke, I., Downing, J.R., Pe'er, D., Nolan, G.P., 2015. Data-Driven Phenotypic Dissection of AML Reveals Progenitor-like Cells that Correlate with Prognosis. Cell 162, 184–197. https://doi.org/10.1016/j.cell.2015.05.047

Levsky, J.M., Shenoy, S.M., Pezo, R.C., Singer, R.H., 2002. Single-Cell Gene Expression Profiling. Science 297, 836–840. https://doi.org/10.1126/science.1072241

Lewis, J.P., Trobaugh, F.E., 1964. Haematopoietic stem cells. Nature 204, 589–590.

Liang, J., Cai, W., Sun, Z., 2014. Single-Cell Sequencing Technologies: Current and Future. J. Genet. Genomics 41, 513–528. https://doi.org/10.1016/j.jgg.2014.09.005

Lindström, S., Andersson-Svahn, H., 2010. Overview of single-cell analyses: microdevices and applications. Lab. Chip 10, 3363–3372. https://doi.org/10.1039/C0LC00150C

Liu, M., Gao, W., van Velkinburgh, J.C., Wu, Y., Ni, B., Tian, Y., 2016. Role of Ets Proteins in Development, Differentiation, and Function of T-Cell Subsets. Med. Res. Rev. 36, 193–220. https://doi.org/10.1002/med.21361

Liu, S., Trapnell, C., 2016. Single-cell transcriptome sequencing: recent advances and remaining challenges. F1000Research 5. https://doi.org/10.12688/f1000research.7223.1

Livak, K.J., Wills, Q.F., Tipping, A.J., Datta, K., Mittal, R., Goldson, A.J., Sexton, D.W., Holmes, C.C., 2013. Methods for qPCR gene expression profiling applied to 1440 lymphoblastoid single cells. Methods San Diego Calif 59, 71–79. https://doi.org/10.1016/j.ymeth.2012.10.004

Lo, S.-J., Yao, D.-J., 2015. Get to Understand More from Single-Cells: Current Studies of Microfluidic-Based Techniques for Single-Cell Analysis. Int. J. Mol. Sci. 16, 16763–16777. https://doi.org/10.3390/ijms160816763

Locatelli, F., Rocha, V., Reed, W., Bernaudin, F., Ertem, M., Grafakos, S., Brichard, B., Li, X., Nagler, A., Giorgiani, G., Haut, P.R., Brochstein, J.A., Nugent, D.J., Blatt, J., Woodard, P., Kurtzberg, J., Rubin, C.M., Miniero, R., Lutz, P., Raja, T., Roberts, I., Will, A.M., Yaniv, I., Vermylen, C., Tannoia, N., Garnier, F., Ionescu, I., Walters, M.C., Lubin, B.H., Gluckman, E., 2003. Related umbilical cord blood transplantation in patients with thalassemia and sickle cell disease. Blood 101, 2137–2143. https://doi.org/10.1182/blood-2002-07-2090

Lord, B.I., Mori, K.J., Wright, E.G., Lajtha, L.G., 1976. Inhibitor of stem cell proliferation in normal bone marrow. Br. J. Haematol. 34, 441–445.

Lucarelli, G., Polchi, P., Izzi, T., Manna, M., Agostinelli, F., Delfini, C., Porcellini, A., Galimberti, M., Moretti, L., Manna, A., 1984. Allogeneic marrow transplantation for thalassemia. Exp. Hematol. 12, 676–681.

Luckett, W.P., 1978. Origin and differentiation of the yolk sac and extraembryonic mesoderm in presomite human and rhesus monkey embryos. Am. J. Anat. 152, 59–97. https://doi.org/10.1002/aja.1001520106

Ludin, A., Itkin, T., Gur-Cohen, S., Mildner, A., Shezen, E., Golan, K., Kollet, O., Kalinkovich, A., Porat, Z., D'Uva, G., Schajnovitz, A., Voronov, E., Brenner, D.A., Apte, R.N., Jung, S., Lapidot, T., 2012. Monocytes-macrophages that express α-smooth muscle actin preserve primitive hematopoietic cells in the bone marrow. Nat. Immunol. 13, 1072–1082. https://doi.org/10.1038/ni.2408

Lugli, E., Roederer, M., Cossarizza, A., 2010. Data Analysis in Flow Cytometry: The Future Just Started. Cytom. Part J. Int. Soc. Anal. Cytol. 77, 705–713. https://doi.org/10.1002/cyto.a.20901

Lunger, I., Fawaz, M., Rieger, M.A., 2017. Single cell analyses to reveal hematopoietic stem cell fate decisions. FEBS Lett. n/a-n/a. https://doi.org/10.1002/1873-3468.12712

Macaulay, I.C., Svensson, V., Labalette, C., Ferreira, L., Hamey, F., Voet, T., Teichmann, S.A., Cvejic, A., 2016. Single-Cell RNA-Sequencing Reveals a Continuous Spectrum

of Differentiation in Hematopoietic Cells. Cell Rep. 14, 966–977. https://doi.org/10.1016/j.celrep.2015.12.082

Macosko, E.Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A.R., Kamitaki, N., Martersteck, E.M., Trombetta, J.J., Weitz, D.A., Sanes, J.R., Shalek, A.K., Regev, A., McCarroll, S.A., 2015. Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell 161, 1202–1214. https://doi.org/10.1016/j.cell.2015.05.002

Maillard, I., Koch, U., Dumortier, A., Shestova, O., Xu, L., Sai, H., Pross, S.E., Aster, J.C., Bhandoola, A., Radtke, F., Pear, W.S., 2008. Canonical notch signaling is dispensable for the maintenance of adult hematopoietic stem cells. Cell Stem Cell 2, 356–366. https://doi.org/10.1016/j.stem.2008.02.011

Mair, F., Hartmann, F.J., Mrdjen, D., Tosevski, V., Krieg, C., Becher, B., 2016. The end of gating? An introduction to automated analysis of high dimensional cytometry data. Eur. J. Immunol. 46, 34–43. https://doi.org/10.1002/eji.201545774

Majeti, R., Park, C.Y., Weissman, I.L., 2007. Identification of a hierarchy of multipotent hematopoietic progenitors in human cord blood. Cell Stem Cell 1, 635–645. https://doi.org/10.1016/j.stem.2007.10.001

Mancini, S.J.C., Mantei, N., Dumortier, A., Suter, U., MacDonald, H.R., Radtke, F., 2005. Jagged1-dependent Notch signaling is dispensable for hematopoietic stem cell self-renewal and differentiation. Blood 105, 2340–2342. https://doi.org/10.1182/blood-2004-08-3207

Mann, R.C., Popp, D.M., Hand, R.E., 1984. The use of projections for dimensionality reduction of flow cytometric data. Cytometry 5, 304–307. https://doi.org/10.1002/cyto.990050314

Manno, C.S., Pierce, G.F., Arruda, V.R., Glader, B., Ragni, M., Rasko, J.J., Rasko, J., Ozelo, M.C., Hoots, K., Blatt, P., Konkle, B., Dake, M., Kaye, R., Razavi, M., Zajko, A., Zehnder, J., Rustagi, P.K., Nakai, H., Chew, A., Leonard, D., Wright, J.F., Lessard, R.R., Sommer, J.M., Tigges, M., Sabatino, D., Luk, A., Jiang, H., Mingozzi, F., Couto, L., Ertl, H.C., High, K.A., Kay, M.A., 2006. Successful transduction of liver in hemophilia by AAV-Factor IX and limitations imposed by the host immune response. Nat. Med. 12, 342–347. https://doi.org/10.1038/nm1358

Månsson, R., Hultquist, A., Luc, S., Yang, L., Anderson, K., Kharazi, S., Al-Hashmi, S., Liuba, K., Thorén, L., Adolfsson, J., Buza-Vidas, N., Qian, H., Soneji, S., Enver, T., Sigvardsson, M., Jacobsen, S.E.W., 2007. Molecular Evidence for Hierarchical Transcriptional Lineage Priming in Fetal and Adult Stem Cells and Multipotent Progenitors. Immunity 26, 407–419. https://doi.org/10.1016/j.immuni.2007.02.013

Market-Velker, B.A., Fernandes, A.D., Mann, M.R.W., 2010. Side-by-side comparison of five commercial media systems in a mouse model: suboptimal in vitro culture interferes with imprint maintenance. Biol. Reprod. 83, 938–950. https://doi.org/10.1095/biolreprod.110.085480

Martin, P., Aulagner, P.G., 2009. La greffe de cellules souches hématopoïétiques. Actual. Pharm. Hosp., Greffe de cellules souches hématopoïétiques et prévention de la maladie du greffon contre l'hôte 5, 16–28. https://doi.org/10.1016/S1769-7344(09)70205-7

Matsuzaki, Y., Kinjo, K., Mulligan, R.C., Okano, H., 2004. Unexpectedly efficient homing capacity of purified murine hematopoietic stem cells. Immunity 20, 87–93.

Mazurier, F., Doedens, M., Gan, O.I., Dick, J.E., 2003. Rapid myeloerythroid repopulation after intrafemoral transplantation of NOD-SCID mice reveals a new class of human stem cells. Nat. Med. 9, 959–963. https://doi.org/10.1038/nm886

Mazutis, L., Gilbert, J., Ung, W.L., Weitz, D.A., Griffiths, A.D., Heyman, J.A., 2013. Single-cell analysis and sorting using droplet-based microfluidics. Nat. Protoc. 8, 870–891. https://doi.org/10.1038/nprot.2013.046

McCredie, K.B., Hersh, E.M., Freireich, E.J., 1971. Cells Capable of Colony Formation in the Peripheral Blood of Man. Science 171, 293–294. https://doi.org/10.1126/science.171.3968.293

McDonald, L.W., Hayes, T.L., 1969. Correlation of scanning electron microscope and light microscope images of individual cells in human blood and blood clots. Exp. Mol. Pathol. 10, 186–198.

McGrath, K.E., 2015. Utilization of imaging flow cytometry to define intermediates of megakaryopoiesis in vivo and in vitro. J. Immunol. Methods 423, 45–51. https://doi.org/10.1016/j.jim.2015.03.002

Medvinsky, A., Dzierzak, E., 1996. Definitive Hematopoiesis Is Autonomously Initiated by the AGM Region. Cell 86, 897–906. https://doi.org/10.1016/S0092-8674(00)80165-8

Mestas, J., Hughes, C.C.W., 2004. Of Mice and Not Men: Differences between Mouse and Human Immunology. J. Immunol. 172, 2731–2738. https://doi.org/10.4049/jimmunol.172.5.2731

Metcalf, D., 1990. The colony stimulating factors discovery, development, and clinical applications. Cancer 65, 2185–2195. https://doi.org/10.1002/1097-0142(19900515)65:10<2185::AID-CNCR2820651005>3.0.CO;2-4

Mikkelsen, T.S., Hanna, J., Zhang, X., Ku, M., Wernig, M., Schorderet, P., Bernstein, B.E., Jaenisch, R., Lander, E.S., Meissner, A., 2008. Dissecting direct reprogramming through integrative genomic analysis. Nature 454, 49–55. https://doi.org/10.1038/nature07056

Milhem, M., Mahmud, N., Lavelle, D., Araki, H., DeSimone, J., Saunthararajah, Y., Hoffman, R., 2004. Modification of hematopoietic stem cell fate by 5aza 2'deoxycytidine and trichostatin A. Blood 103, 4102–4110. https://doi.org/10.1182/blood-2003-07-2431

Miller, J.S., McCullar, V., Punzel, M., Lemischka, I.R., Moore, K.A., 1999. Single adult human CD34(+)/Lin-/CD38(-) progenitors give rise to natural killer cells, B-lineage cells, dendritic cells, and myeloid cells. Blood 93, 96–106.

Miller, L.W., Cornish, V.W., 2005. Selective chemical labeling of proteins in living cells. Curr. Opin. Chem. Biol., Proteomics and genomics 9, 56–61. https://doi.org/10.1016/j.cbpa.2004.12.007

Misra, S., 2013. Human gene therapy: a brief overview of the genetic revolution. J. Assoc. Physicians India 61, 127–133.

Mohrin, M., Bourke, E., Alexander, D., Warr, M.R., Barry-Holson, K., Le Beau, M.M., Morrison, C.G., Passegué, E., 2010. Hematopoietic stem cell quiescence promotes error-prone DNA repair and mutagenesis. Cell Stem Cell 7, 174–185. https://doi.org/10.1016/j.stem.2010.06.014

Moignard, V., Macaulay, I.C., Swiers, G., Buettner, F., Schütte, J., Calero-Nieto, F.J., Kinston, S., Joshi, A., Hannah, R., Theis, F.J., Jacobsen, S.E., de Bruijn, M., Göttgens, B., 2013. Characterisation of transcriptional networks in blood stem and progenitor cells using high-throughput single cell gene expression analysis. Nat. Cell Biol. 15, 363–372. https://doi.org/10.1038/ncb2709

Moignard, V., Woodhouse, S., Haghverdi, L., Lilly, A.J., Tanaka, Y., Wilkinson, A.C., Buettner, F., Macaulay, I.C., Jawaid, W., Diamanti, E., Nishikawa, S.-I., Piterman, N., Kouskoff, V., Theis, F.J., Fisher, J., Göttgens, B., 2015. Decoding the regulatory network of early blood development from single-cell gene expression measurements. Nat. Biotechnol. 33, 269–276. https://doi.org/10.1038/nbt.3154

Mojtahedi, M., Skupin, A., Zhou, J., Castaño, I.G., Leong-Quong, R.Y.Y., Chang, H., Trachana, K., Giuliani, A., Huang, S., 2016. Cell Fate Decision as High-Dimensional Critical State Transition. PLoS Biol. 14, e2000640. https://doi.org/10.1371/journal.pbio.2000640

Moldavan, A., 1934. Photo-electric technique for the counting of microscopical cells. Science 80, 188–189. https://doi.org/10.1126/science.80.2069.188

Moore, M.A., Metcalf, D., 1970. Ontogeny of the haemopoietic system: yolk sac origin of in vivo and in vitro colony forming cells in the developing mouse embryo. Br. J. Haematol. 18, 279–296.

Moris, N., Pina, C., Arias, A.M., 2016. Transition states and cell fate decisions in epigenetic landscapes. Nat. Rev. Genet. 17, 693–703. https://doi.org/10.1038/nrg.2016.98

Morrison, S.J., Wandycz, A.M., Hemmati, H.D., Wright, D.E., Weissman, I.L., 1997. Identification of a lineage of multipotent hematopoietic progenitors. Dev. Camb. Engl. 124, 1929–1939.

Mossadegh-Keller, N., Sarrazin, S., Kandalla, P.K., Espinosa, L., Stanley, E.R., Nutt, S.L., Moore, J., Sieweke, M.H., 2013. M-CSF instructs myeloid lineage fate in single haematopoietic stem cells. Nature 497, 239–243. https://doi.org/10.1038/nature12026

Mullard, A., 2011. Single-cell profiling sheds new light. Nat. Rev. Drug Discov. 10, 477–478. https://doi.org/10.1038/nrd3487

Nakahata, T., Ogawa, M., 1982. Hemopoietic colony-forming cells in umbilical cord blood with extensive capability to generate mono- and multipotential hemopoietic progenitors. J. Clin. Invest. 70, 1324–1328.

Narsinh, K.H., Sun, N., Sanchez-Freire, V., Lee, A.S., Almeida, P., Hu, S., Jan, T., Wilson, K.D., Leong, D., Rosenberg, J., Yao, M., Robbins, R.C., Wu, J.C., 2011. Single cell transcriptional profiling reveals heterogeneity of human induced pluripotent stem cells. J. Clin. Invest. 121, 1217–1221. https://doi.org/10.1172/JCI44635

Nathwani, A.C., Tuddenham, E.G.D., Rangarajan, S., Rosales, C., McIntosh, J., Linch, D.C., Chowdary, P., Riddell, A., Pie, A.J., Harrington, C., O'Beirne, J., Smith, K., Pasi, J., Glader, B., Rustagi, P., Ng, C.Y.C., Kay, M.A., Zhou, J., Spence, Y., Morton, C.L., Allay, J., Coleman, J., Sleep, S., Cunningham, J.M., Srivastava, D., Basner-Tschakarjan, E., Mingozzi, F., High, K.A., Gray, J.T., Reiss, U.M., Nienhuis, A.W., Davidoff, A.M., 2011. Adenovirus-associated virus vector-mediated gene transfer in hemophilia B. N. Engl. J. Med. 365, 2357–2365. https://doi.org/10.1056/NEJMoa1108046

Neildez-Nguyen, T.M.A., Bigot, J., Da Rocha, S., Corre, G., Boisgerault, F., Paldi, A., Galy, A., 2015. Hypoxic culture conditions enhance the generation of regulatory T cells. Immunology 144, 431–443. https://doi.org/10.1111/imm.12388

Nelissen, E.C., Van Montfoort, A.P., Coonen, E., Derhaag, J.G., Geraedts, J.P., Smits, L.J., Land, J.A., Evers, J.L., Dumoulin, J.C., 2012. Further evidence that culture media affect perinatal outcome: findings after transfer of fresh and cryopreserved embryos. Hum. Reprod. Oxf. Engl. 27, 1966–1976. https://doi.org/10.1093/humrep/des145

Nestorowa, S., Hamey, F.K., Sala, B.P., Diamanti, E., Shepherd, M., Laurenti, E., Wilson, N.K., Kent, D.G., Göttgens, B., 2016. A single-cell resolution map of mouse hematopoietic stem and progenitor cell differentiation. Blood 128, e20–e31. https://doi.org/10.1182/blood-2016-05-716480

Newell, E.W., Sigal, N., Bendall, S.C., Nolan, G.P., Davis, M.M., 2012. Cytometry by Time-of-Flight Shows Combinatorial Cytokine Expression and Virus-Specific Cell Niches within a Continuum of CD8+ T Cell Phenotypes. Immunity 36, 142–152. https://doi.org/10.1016/j.immuni.2012.01.002

Nie, Y., Han, Y.-C., Zou, Y.-R., 2008. CXCR4 is required for the quiescence of primitive hematopoietic cells. J. Exp. Med. 205, 777–783. https://doi.org/10.1084/jem.20072513

Notta, F., Doulatov, S., Laurenti, E., Poeppl, A., Jurisica, I., Dick, J.E., 2011. Isolation of Single Human Hematopoietic Stem Cells Capable of Long-Term Multilineage Engraftment. Science 333, 218–221. https://doi.org/10.1126/science.1201219

Notta, F., Zandi, S., Takayama, N., Dobson, S., Gan, O.I., Wilson, G., Kaufmann, K.B., McLeod, J., Laurenti, E., Dunant, C.F., McPherson, J.D., Stein, L.D., Dror, Y., Dick,

J.E., 2016. Distinct routes of lineage development reshape the human blood hierarchy across ontogeny. Science 351, aab2116. https://doi.org/10.1126/science.aab2116
Oberlin, E., Tavian, M., Blazsek, I., Péault, B., 2002. Blood-forming potential of vascular endothelium in the human embryo. Development 129, 4147–4157.
Oburoglu, L., Tardito, S., Fritz, V., de Barros, S.C., Merida, P., Craveiro, M., Mamede, J., Cretenet, G., Mongellaz, C., An, X., Klysz, D., Touhami, J., Boyer-Clavel, M., Battini, J.-L., Dardalhon, V., Zimmermann, V.S., Mohandas, N., Gottlieb, E., Sitbon, M., Kinet, S., Taylor, N., 2014. Glucose and glutamine metabolism regulate human hematopoietic stem cell lineage specification. Cell Stem Cell 15, 169–184. https://doi.org/10.1016/j.stem.2014.06.002
Ogawa, M., Porter, P.N., Nakahata, T., 1983. Renewal and commitment to differentiation of hemopoietic stem cells (an interpretive review). Blood 61, 823–829.
Ohishi, K., Varnum-Finney, B., Bernstein, I.D., 2002. Delta-1 enhances marrow and thymus repopulating ability of human CD34(+)CD38(-) cord blood cells. J. Clin. Invest. 110, 1165–1174. https://doi.org/10.1172/JCI16167
Olsson, A., Venkatasubramanian, M., Chaudhri, V.K., Aronow, B.J., Salomonis, N., Singh, H., Grimes, H.L., 2016. Single-cell analysis of mixed-lineage states leading to a binary cell fate choice. Nature 537, 698–702. https://doi.org/10.1038/nature19348
Ortega, F., Costa, M.R., Simon-Ebert, T., Schroeder, T., Götz, M., Berninger, B., 2011. Using an adherent cell culture of the mouse subependymal zone to study the behavior of adult neural stem cells on a single-cell level. Nat. Protoc. 6, 1847–1859. https://doi.org/10.1038/nprot.2011.404
Ott, M.G., Schmidt, M., Schwarzwaelder, K., Stein, S., Siler, U., Koehl, U., Glimm, H., Kühlcke, K., Schilz, A., Kunkel, H., Naundorf, S., Brinkmann, A., Deichmann, A., Fischer, M., Ball, C., Pilz, I., Dunbar, C., Du, Y., Jenkins, N.A., Copeland, N.G., Lüthi, U., Hassan, M., Thrasher, A.J., Hoelzer, D., von Kalle, C., Seger, R., Grez, M., 2006. Correction of X-linked chronic granulomatous disease by gene therapy, augmented by insertional activation of MDS1-EVI1, PRDM16 or SETBP1. Nat. Med. 12, 401–409. https://doi.org/10.1038/nm1393
Ottersbach, K., Dzierzak, E., 2005. The murine placenta contains hematopoietic stem cells within the vascular labyrinth region. Dev. Cell 8, 377–387. https://doi.org/10.1016/j.devcel.2005.02.001
Paldi, A., 2013. Effects of the in vitro manipulation of stem cells: epigenetic mechanisms as mediators of induced metabolic fluctuations. Epigenomics 5, 429–437. https://doi.org/10.2217/epi.13.35
Paldi, A., 2003. Stochastic gene expression during cell differentiation: order from disorder? Cell. Mol. Life Sci. CMLS 60, 1775–1778. https://doi.org/10.1007/s00018-003-23147-z
Panch, S.R., Szymanski, J., Savani, B.N., Stroncek, D.F., 2017. Sources of Hematopoietic Stem and Progenitor Cells and Methods to Optimize Yields for Clinical Cell Therapy. Biol. Blood Marrow Transplant. https://doi.org/10.1016/j.bbmt.2017.05.003
Papayannopoulos, V., Zychlinsky, A., 2009. NETs: a new strategy for using old weapons. Trends Immunol. 30, 513–521. https://doi.org/10.1016/j.it.2009.07.011
Pappenheim, A., 1896. Ueber Entwickelung und Ausbildung der Erythroblasten. Arch. Für Pathol. Anat. Physiol. Für Klin. Med. 145, 587–643. https://doi.org/10.1007/BF01969901
Pardali, E., Waltenberger, J., 2012. Monocyte function and trafficking in cardiovascular disease. Thromb. Haemost. 108, 804–811. https://doi.org/10.1160/TH12-04-0276
Park, S., Wong, D.J., Ooi, C.C., Kurtz, D.M., Vermesh, O., Aalipour, A., Suh, S., Pian, K.L., Chabon, J.J., Lee, S.H., Jamali, M., Say, C., Carter, J.N., Lee, L.P., Kuschner, W.G., Schwartz, E.J., Shrager, J.B., Neal, J.W., Wakelee, H.A., Diehn, M., Nair, V.S., Wang, S.X., Gambhir, S.S., 2016. Molecular profiling of single circulating tumor cells

from lung cancer patients. Proc. Natl. Acad. Sci. 113, E8379–E8386. https://doi.org/10.1073/pnas.1608461113

Parks, D.R., Bryan, V.M., Oi, V.T., Herzenberg, L.A., 1979. Antigen-specific identification and cloning of hybridomas with a fluorescence-activated cell sorter. Proc. Natl. Acad. Sci. U. S. A. 76, 1962.

Passegué, E., Wagers, A.J., Giuriato, S., Anderson, W.C., Weissman, I.L., 2005. Global analysis of proliferation and cell cycle gene expression in the regulation of hematopoietic stem and progenitor cell fates. J. Exp. Med. 202, 1599–1611. https://doi.org/10.1084/jem.20050967

Paul, F., Arkin, Y., Giladi, A., Jaitin, D.A., Kenigsberg, E., Keren-Shaul, H., Winter, D., Lara-Astiaso, D., Gury, M., Weiner, A., David, E., Cohen, N., Lauridsen, F.K.B., Haas, S., Schlitzer, A., Mildner, A., Ginhoux, F., Jung, S., Trumpp, A., Porse, B.T., Tanay, A., Amit, I., 2015. Transcriptional Heterogeneity and Lineage Commitment in Myeloid Progenitors. Cell 163, 1663–1677. https://doi.org/10.1016/j.cell.2015.11.013

Péault, B., Coltey, M., Le Douarin, N.M., 1988. Ontogenic emergence of a quail leukocyte/endothelium cell surface antigen. Cell Differ. 23, 165–174.

Péault, B., Weissman, I.L., Buckle, A.M., Tsukamoto, A., Baum, C., 1993. Thy-1-expressing CD34+ human cells express multiple hematopoietic potentialities in vitro and in SCID-hu mice. Nouv. Rev. Fr. Hematol. 35, 91–93.

Perié, L., Duffy, K.R., 2016. Retracing the in vivo haematopoietic tree using single-cell methods. FEBS Lett. 590, 4068–4083. https://doi.org/10.1002/1873-3468.12299

Perié, L., Duffy, K.R., Kok, L., de Boer, R.J., Schumacher, T.N., 2015. The Branching Point in Erythro-Myeloid Differentiation. Cell 163, 1655–1662. https://doi.org/10.1016/j.cell.2015.11.059

Piccoli, C., Agriesti, F., Scrima, R., Falzetti, F., Di Ianni, M., Capitanio, N., 2013. To breathe or not to breathe: the haematopoietic stem/progenitor cells dilemma. Br. J. Pharmacol. 169, 1652–1671. https://doi.org/10.1111/bph.12253

Picelli, S., Björklund, Å.K., Faridani, O.R., Sagasser, S., Winberg, G., Sandberg, R., 2013. Smart-seq2 for sensitive full-length transcriptome profiling in single cells. Nat. Methods 10, 1096–1098. https://doi.org/10.1038/nmeth.2639

Picelli, S., Faridani, O.R., Björklund, Å.K., Winberg, G., Sagasser, S., Sandberg, R., 2014. Full-length RNA-seq from single cells using Smart-seq2. Nat. Protoc. 9, 171–181. https://doi.org/10.1038/nprot.2014.006

Pillay, J., Braber, I. den, Vrisekoop, N., Kwast, L.M., Boer, R.J. de, Borghans, J.A.M., Tesselaar, K., Koenderman, L., 2010. In vivo labeling with 2H2O reveals a human neutrophil lifespan of 5.4 days. Blood 116, 625–627. https://doi.org/10.1182/blood-2010-01-259028

Pina, C., Fugazza, C., Tipping, A.J., Brown, J., Soneji, S., Teles, J., Peterson, C., Enver, T., 2012. Inferring rules of lineage commitment in haematopoiesis. Nat. Cell Biol. 14, 287–294. https://doi.org/10.1038/ncb2442

Piro, D., Piccoli, C., Guerra, L., Sassone, F., D'Aprile, A., Favia, M., Castellani, S., Di Gioia, S., Lepore, S., Garavaglia, M.L., Trotta, T., Maffione, A.B., Casavola, V., Meyer, G., Capitanio, N., Conese, M., 2011. Hematopoietic Stem/Progenitor Cells Express Functional Mitochondrial Energy-Dependent Cystic Fibrosis Transmembrane Conductance Regulator. Stem Cells Dev. 21, 634–646. https://doi.org/10.1089/scd.2011.0041

Piyasena, M.E., Graves, S.W., 2014. The intersection of flow cytometry with microfluidics and microfabrication. Lab. Chip 14, 1044–1059. https://doi.org/10.1039/c3lc51152a

Polliack, A., Lampen, N., Clarkson, B.D., de Harven, E., Bentwich, Z., Siegal, F.P., Kunkel, H.G., 1973. Identification of human B and T lymphocytes by scanning electron microscopy. J. Exp. Med. 138, 607–624.

Porter, J.R., Telford, W.G., Batchelor, E., 2017. Single-cell Gene Expression Profiling Using FACS and qPCR with Internal Standards. J. Vis. Exp. JoVE. https://doi.org/10.3791/55219

Qian, H., Buza-Vidas, N., Hyland, C.D., Jensen, C.T., Antonchuk, J., Månsson, R., Thoren, L.A., Ekblom, M., Alexander, W.S., Jacobsen, S.E.W., 2007. Critical role of thrombopoietin in maintaining adult quiescent hematopoietic stem cells. Cell Stem Cell 1, 671–684. https://doi.org/10.1016/j.stem.2007.10.008

Qiu, P., Simonds, E.F., Bendall, S.C., Gibbs, K.D., Bruggner, R.V., Linderman, M.D., Sachs, K., Nolan, G.P., Plevritis, S.K., 2011. Extracting a cellular hierarchy from high-dimensional cytometry data with SPADE. Nat. Biotechnol. 29, 886–891. https://doi.org/10.1038/nbt.1991

Quesenberry, P.J., Goldberg, L., Dooner, M., Aliotta, J., 2014. Marrow hematopoietic stem cells revisited: they exist in a continuum and are not defined by standard purification approaches; then there are the microvesicles. Cancer Genet. 4, 56. https://doi.org/10.3389/fonc.2014.00056

Quesenberry, P.J., Goldberg, L.R., Dooner, M.S., 2015. A stem cell apostasy: a tale of four H words. STEM CELLS 33, 15–20. https://doi.org/10.1002/stem.1829

Raj, A., van Oudenaarden, A., 2008. Nature, Nurture, or Chance: Stochastic Gene Expression and Its Consequences. Cell 135, 216–226. https://doi.org/10.1016/j.cell.2008.09.050

Ramalho-Santos, M., Willenbring, H., 2007. On the Origin of the Term "Stem Cell." Cell Stem Cell 1, 35–38. https://doi.org/10.1016/j.stem.2007.05.013

Ramser, K., Hanstorp, D., 2010. Optical manipulation for single-cell studies. J. Biophotonics 3, 187–206. https://doi.org/10.1002/jbio.200910050

Ramsköld, D., Luo, S., Wang, Y.-C., Li, R., Deng, Q., Faridani, O.R., Daniels, G.A., Khrebtukova, I., Loring, J.F., Laurent, L.C., Schroth, G.P., Sandberg, R., 2012. Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. Nat. Biotechnol. 30, 777–782. https://doi.org/10.1038/nbt.2282

Ren, K., Zhou, J., Wu, H., 2013. Materials for Microfluidic Chip Fabrication. Acc. Chem. Res. 46, 2396–2406. https://doi.org/10.1021/ar300314s

Rendu, F., Brohard-Bohn, B., 2001. The platelet release reaction: granules' constituents, secretion and functions. Platelets 12, 261–273. https://doi.org/10.1080/09537100120068170

Ribeil, J.-A., Hacein-Bey-Abina, S., Payen, E., Magnani, A., Semeraro, M., Magrin, E., Caccavelli, L., Neven, B., Bourget, P., El Nemer, W., Bartolucci, P., Weber, L., Puy, H., Meritet, J.-F., Grevent, D., Beuzard, Y., Chrétien, S., Lefebvre, T., Ross, R.W., Negre, O., Veres, G., Sandler, L., Soni, S., de Montalembert, M., Blanche, S., Leboulch, P., Cavazzana, M., 2017. Gene Therapy in a Patient with Sickle Cell Disease. N. Engl. J. Med. 376, 848–855. https://doi.org/10.1056/NEJMoa1609677

Richard, A., Boullu, L., Herbach, U., Bonnafoux, A., Morin, V., Vallin, E., Guillemin, A., Papili Gao, N., Gunawan, R., Cosette, J., Arnaud, O., Kupiec, J.-J., Espinasse, T., Gonin-Giraud, S., Gandrillon, O., 2016. Single-Cell-Based Analysis Highlights a Surge in Cell-to-Cell Molecular Variability Preceding Irreversible Commitment in a Differentiation Process. PLoS Biol. 14, e1002585. https://doi.org/10.1371/journal.pbio.1002585

Rieger, M.A., Hoppe, P.S., Smejkal, B.M., Eitelhuber, A.C., Schroeder, T., 2009. Hematopoietic Cytokines Can Instruct Lineage Choice. Science 325, 217–218. https://doi.org/10.1126/science.1171461

Rieger, M.A., Schroeder, T., 2008. Exploring Hematopoiesis at Single Cell Resolution. Cells Tissues Organs 188, 139–149. https://doi.org/10.1159/000114540

Rizo, A., Dontje, B., Vellenga, E., de Haan, G., Schuringa, J.J., 2008. Long-term maintenance of human hematopoietic stem/progenitor cells by expression of BMI1. Blood 111, 2621–2630. https://doi.org/10.1182/blood-2007-08-106666

Rizo, A., Olthof, S., Han, L., Vellenga, E., de Haan, G., Schuringa, J.J., 2009. Repression of BMI1 in normal and leukemic human CD34(+) cells impairs self-renewal and induces apoptosis. Blood 114, 1498–1505. https://doi.org/10.1182/blood-2009-03-209734

Robin, C., Bollerot, K., Mendes, S., Haak, E., Crisan, M., Cerisoli, F., Lauw, I., Kaimakis, P., Jorna, R., Vermeulen, M., Kayser, M., van der Linden, R., Imanirad, P., Verstegen, M., Nawaz-Yousaf, H., Papazian, N., Steegers, E., Cupedo, T., Dzierzak, E., 2009. Human placenta is a potent hematopoietic niche containing hematopoietic stem and progenitor cells throughout development. Cell Stem Cell 5, 385–395. https://doi.org/10.1016/j.stem.2009.08.020

Rocha, V., 2016. Umbilical cord blood cells from unrelated donor as an alternative source of hematopoietic stem cells for transplantation in children and adults. Semin. Hematol. 53, 237–245. https://doi.org/10.1053/j.seminhematol.2016.08.002

Rocha, V., Labopin, M., Sanz, G., Arcese, W., Schwerdtfeger, R., Bosi, A., Jacobsen, N., Ruutu, T., de Lima, M., Finke, J., Frassoni, F., Gluckman, E., Acute Leukemia Working Party of European Blood and Marrow Transplant Group, Eurocord-Netcord Registry, 2004. Transplants of umbilical-cord blood or bone marrow from unrelated donors in adults with acute leukemia. N. Engl. J. Med. 351, 2276–2285. https://doi.org/10.1056/NEJMoa041469

Rosenberg, H.F., Dyer, K.D., Foster, P.S., 2013. Eosinophils: changing perspectives in health and disease. Nat. Rev. Immunol. 13, 9–22. https://doi.org/10.1038/nri3341

Rossaint, J., Zarbock, A., 2015. Platelets in leucocyte recruitment and function. Cardiovasc. Res. 107, 386–395. https://doi.org/10.1093/cvr/cvv048

Saadatpour, A., Guo, G., Orkin, S.H., Yuan, G.-C., 2014. Characterizing heterogeneity in leukemic cells using single-cell gene expression analysis. Genome Biol. 15, 525. https://doi.org/10.1186/s13059-014-0525-9

Sabin, F., 1920. Studies on the origin of blood vessels and of red corpuscles as seen in the living blastoderm of the chick during the second day of incubation. Contrib. Embryol. 9, 213–262.

Sadowska-Bartosz, I., Bartosz, G., 2013. Effect of 3-bromopyruvic acid on human erythrocyte antioxidant defense system. Cell Biol. Int. 37, 1285–1290. https://doi.org/10.1002/cbin.10160

Sakaguchi, S., Sakaguchi, N., Shimizu, J., Yamazaki, S., Sakihama, T., Itoh, M., Kuniyasu, Y., Nomura, T., Toda, M., Takahashi, T., 2001. Immunologic tolerance maintained by CD25+ CD4+ regulatory T cells: their common role in controlling autoimmunity, tumor immunity, and transplantation tolerance. Immunol. Rev. 182, 18–32.

Sasagawa, Y., Nikaido, I., Hayashi, T., Danno, H., Uno, K.D., Imai, T., Ueda, H.R., 2013. Quartz-Seq: a highly reproducible and sensitive single-cell RNA sequencing method, reveals non-genetic gene-expression heterogeneity. Genome Biol. 14, 3097. https://doi.org/10.1186/gb-2013-14-4-r31

Schmitz, N., Dreger, P., Suttorp, M., Rohwedder, E.B., Haferlach, T., Loffler, H., Hunter, A., Russell, N.H., 1995. Primary transplantation of allogeneic peripheral blood progenitor cells mobilized by filgrastim (granulocyte colony-stimulating factor). Blood 85, 1666–1672.

Schneider, H., Chaovapong, W., Matthews, D.J., Karkaria, C., Cass, R.T., Zhan, H., Boyle, M., Lorenzini, T., Elliott, S.G., Giebel, L.B., 1997. Homodimerization of erythropoietin receptor by a bivalent monoclonal antibody triggers cell proliferation and differentiation of erythroid precursors. Blood 89, 473–482.

Schofield, R., 1978. The relationship between the spleen colony-forming cell and the haemopoietic stem cell. Blood Cells 4, 7–25.

Schröter, C., Rué, P., Mackenzie, J.P., Martinez Arias, A., 2015. FGF/MAPK signaling sets the switching threshold of a bistable circuit controlling cell fate decisions in

embryonic stem cells. Dev. Camb. Engl. 142, 4205–4216. https://doi.org/10.1242/dev.127530
Scott, E.W., Simon, M.C., Anastasi, J., Singh, H., 1994. Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages. Science 265, 1573–1577. https://doi.org/10.1126/science.8079170
Seet, L.-F., Teng, E., Lai, Y.-S., Laning, J., Kraus, M., Wnendt, S., Merchav, S., Chan, S.L., 2009. Valproic acid enhances the engraftability of human umbilical cord blood hematopoietic stem cells expanded under serum-free conditions*. Eur. J. Haematol. 82, 124–132. https://doi.org/10.1111/j.1600-0609.2008.01169.x
Sekhavati, F., Endele, M., Rappl, S., Marel, A.-K., Schroeder, T., Rädler, J.O., 2015. Marker-free detection of progenitor cell differentiation by analysis of Brownian motion in micro-wells. Integr. Biol. Quant. Biosci. Nano Macro 7, 178–183. https://doi.org/10.1039/c4ib00158c
Sekulovic, S., Gasparetto, M., Lecault, V., Hoesli, C.A., Kent, D.G., Rosten, P., Wan, A., Brookes, C., Hansen, C.L., Piret, J.M., Smith, C., Eaves, C.J., Humphries, R.K., 2011. Ontogeny stage-independent and high-level clonal expansion in vitro of mouse hematopoietic stem cells stimulated by an engineered NUP98-HOX fusion transcription factor. Blood 118, 4366–4376. https://doi.org/10.1182/blood-2011-04-350066
Serbina, N.V., Jia, T., Hohl, T.M., Pamer, E.G., 2008. Monocyte-Mediated Defense Against Microbial Pathogens. Annu. Rev. Immunol. 26, 421–452. https://doi.org/10.1146/annurev.immunol.26.021607.090326
Shekhar, K., Brodin, P., Davis, M.M., Chakraborty, A.K., 2014. Automatic Classification of Cellular Expression by Nonlinear Stochastic Embedding (ACCENSE). Proc. Natl. Acad. Sci. 111, 202–207. https://doi.org/10.1073/pnas.1321405111
Shepherd, B.E., Guttorp, P., Lansdorp, P.M., Abkowitz, J.L., 2004. Estimating human hematopoietic stem cell kinetics using granulocyte telomere lengths. Exp. Hematol. 32, 1040–1050. https://doi.org/10.1016/j.exphem.2004.07.023
Shi, C., Pamer, E.G., 2011. Monocyte recruitment during infection and inflammation. Nat. Rev. Immunol. 11, 762–774. https://doi.org/10.1038/nri3070
Siminovitch, L., Mcculloch, E.A., Till, J.E., 1963. The distribution of colony-forming cells among spleen colonies. J. Cell. Comp. Physiol. 62, 327–336.
Simsek, T., Kocabas, F., Zheng, J., Deberardinis, R.J., Mahmoud, A.I., Olson, E.N., Schneider, J.W., Zhang, C.C., Sadek, H.A., 2010. The distinct metabolic profile of hematopoietic stem cells reflects their location in a hypoxic niche. Cell Stem Cell 7, 380–390. https://doi.org/10.1016/j.stem.2010.07.011
Singh, H., 1996. Gene targeting reveals a hierarchy of transcription factors regulating specification of lymphoid cell fates. Curr. Opin. Immunol. 8, 160–165. https://doi.org/10.1016/S0952-7915(96)80053-7
Song, Y., Bahnson, A., Hall, N., Yu, H., Shen, H., Koebler, D., Houck, R., Xie, Y., Cheng, T., 2010. Stem cell traits in long-term co-culture revealed by time-lapse imaging. Leukemia 24, 153–161. https://doi.org/10.1038/leu.2009.191
Spain, L.M., Guerriero, A., Kunjibettu, S., Scott, E.W., 1999. T Cell Development in PU.1-Deficient Mice. J. Immunol. 163, 2681–2687.
Spitzer, M.H., Nolan, G.P., 2016. Mass Cytometry: Single Cells, Many Features. Cell 165, 780–791. https://doi.org/10.1016/j.cell.2016.04.019
Stegle, O., Teichmann, S.A., Marioni, J.C., 2015. Computational and analytical challenges in single-cell transcriptomics. Nat. Rev. Genet. 16, 133–145. https://doi.org/10.1038/nrg3833
Steinbach, K.H., Schick, P., Trepel, F., Raffler, H., Döhrmann, J., Heilgeist, G., Heltzel, W., Li, K., Past, W., van der Woerd-de Lange, J.A., Theml, H., Fliedner, T.M., Begemann,

H., 1979. Estimation of kinetic parameters of neutrophilic, eosinophilic, and basophilic granulocytes in human blood. Blut 39, 27–38.

Stier, S., Cheng, T., Dombkowski, D., Carlesso, N., Scadden, D.T., 2002. Notch1 activation increases hematopoietic stem cell self-renewal in vivo and favors lymphoid over myeloid lineage outcome. Blood 99, 2369–2378.

Suda, T., Suda, J., Ogawa, M., 1984. Disparate differentiation in mouse hemopoietic colonies derived from paired progenitors. Proc. Natl. Acad. Sci. U. S. A. 81, 2520–2524.

Suda, T., Takubo, K., Semenza, G.L., 2011. Metabolic regulation of hematopoietic stem cells in the hypoxic niche. Cell Stem Cell 9, 298–310. https://doi.org/10.1016/j.stem.2011.09.010

Sugiyama, D., Ogawa, M., Hirose, I., Jaffredo, T., Arai, K., Tsuji, K., 2003. Erythropoiesis from acetyl LDL incorporating endothelial cells at the preliver stage. Blood 101, 4733–4738. https://doi.org/10.1182/blood-2002-09-2799

Sugiyama, T., Kohara, H., Noda, M., Nagasawa, T., 2006. Maintenance of the hematopoietic stem cell pool by CXCL12-CXCR4 chemokine signaling in bone marrow stromal cell niches. Immunity 25, 977–988. https://doi.org/10.1016/j.immuni.2006.10.016

Swiers, G., Patient, R., Loose, M., 2006. Genetic regulatory networks programming hematopoietic stem cells and erythroid lineage specification. Dev. Biol. 294, 525–540. https://doi.org/10.1016/j.ydbio.2006.02.051

Swirski, F.K., Nahrendorf, M., Etzrodt, M., Wildgruber, M., Cortez-Retamozo, V., Panizzi, P., Figueiredo, J.-L., Kohler, R.H., Chudnovskiy, A., Waterman, P., Aikawa, E., Mempel, T.R., Libby, P., Weissleder, R., Pittet, M.J., 2009. Identification of splenic reservoir monocytes and their deployment to inflammatory sites. Science 325, 612–616. https://doi.org/10.1126/science.1175202

Sykes, S.M., Scadden, D.T., 2013. Modeling Human Hematopoietic Stem Cell Biology in the Mouse. Semin. Hematol., Mouse Medicine and Human Disease 50, 92–100. https://doi.org/10.1053/j.seminhematol.2013.03.029

Takano, H., Ema, H., Sudo, K., Nakauchi, H., 2004. Asymmetric Division and Lineage Commitment at the Level of Hematopoietic Stem Cells. J. Exp. Med. 199, 295–302. https://doi.org/10.1084/jem.20030929

Takubo, K., Nagamatsu, G., Kobayashi, C.I., Nakamura-Ishizu, A., Kobayashi, H., Ikeda, E., Goda, N., Rahimi, Y., Johnson, R.S., Soga, T., Hirao, A., Suematsu, M., Suda, T., 2013. Regulation of glycolysis by Pdk functions as a metabolic checkpoint for cell cycle quiescence in hematopoietic stem cells. Cell Stem Cell 12, 49–61. https://doi.org/10.1016/j.stem.2012.10.011

Tan, W.-H., Takeuchi, S., 2007. A trap-and-release integrated microfluidic system for dynamic microarray applications. Proc. Natl. Acad. Sci. 104, 1146–1151. https://doi.org/10.1073/pnas.0606625104

Tanaka, E.M., 2003. Regeneration: if they can do it, why can't we? Cell 113, 559–562. https://doi.org/10.1016/S0092-8674(03)00395-7

Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., Wang, X., Bodeau, J., Tuch, B.B., Siddiqui, A., Lao, K., Surani, M.A., 2009. mRNA-Seq whole-transcriptome analysis of a single cell. Nat. Methods 6, 377–382. https://doi.org/10.1038/nmeth.1315

Tavian, M., Hallais, M.F., Peault, B., 1999. Emergence of intraembryonic hematopoietic precursors in the pre-liver human embryo. Development 126, 793–803.

Tavian, M., Robin, C., Coulombel, L., Péault, B., 2001. The Human Embryo, but Not Its Yolk Sac, Generates Lympho-Myeloid Stem Cells: Mapping Multipotent Hematopoietic Cell Fate in Intraembryonic Mesoderm. Immunity 15, 487–495. https://doi.org/10.1016/S1074-7613(01)00193-5

Thaker, Y.R., Schneider, H., Rudd, C.E., 2015. TCR and CD28 activate the transcription factor NF-κB in T-cells via distinct adaptor signaling complexes. Immunol. Lett. 163, 113–119. https://doi.org/10.1016/j.imlet.2014.10.020

Thomas, E.D., Buckner, C.D., Sanders, J.E., Papayannopoulou, T., Borgna-Pignatti, C., De Stefano, P., Sullivan, K.M., Clift, R.A., Storb, R., 1982. Marrow transplantation for thalassaemia. Lancet Lond. Engl. 2, 227–229.

Thomas, E.D., Lochte, H.L., Cannon, J.H., Sahler, O.D., Ferrebee, J.W., 1959. Supralethal whole body irradiation and isologous marrow transplantation in man. J. Clin. Invest. 38, 1709–1716. https://doi.org/10.1172/JCI103949

Thomas, E.D., Lochte, H.L.J., Lu, W.C., Ferrebee, J.W., 1957. Intravenous Infusion of Bone Marrow in Patients Receiving Radiation and Chemotherapy. N. Engl. J. Med. 257, 491–496. https://doi.org/10.1056/NEJM195709122571102

Till, J.E., McCulloch, E.A., 1980. Hemopoietic stem cell differentiation. Biochim. Biophys. Acta 605, 431–459.

Till, J.E., McCulloch, E.A., 1961. A direct measurement of the radiation sensitivity of normal mouse bone marrow cells. Radiat. Res. 14, 213–222.

Till, J.E., McCulloch, E.A., Siminovitch, L., 1964. A stochastic model of stem cell proliferation, based on the growth of spleen colony-forming cells. Proc. Natl. Acad. Sci. U. S. A. 51, 29–36.

Trott, J., Martinez Arias, A., 2013. Single cell lineage analysis of mouse embryonic stem cells at the exit from pluripotency. Biol. Open 2, 1049–1056. https://doi.org/10.1242/bio.20135934

Tsang, J.C.H., Yu, Y., Burke, S., Buettner, F., Wang, C., Kolodziejczyk, A.A., Teichmann, S.A., Lu, L., Liu, P., 2015. Single-cell transcriptomic reconstruction reveals cell cycle and multi-lineage differentiation defects in Bcl11a -deficient hematopoietic stem cells. Genome Biol. 16, 178. https://doi.org/10.1186/s13059-015-0739-5

Turrini, F., Arese, P., Yuan, J., Low, P.S., 1991. Clustering of integral membrane proteins of the human erythrocyte membrane stimulates autologous IgG binding, complement deposition, and phagocytosis. J. Biol. Chem. 266, 23611–23617.

Tzeng, Y.-S., Li, H., Kang, Y.-L., Chen, W.-C., Cheng, W.-C., Lai, D.-M., 2011. Loss of Cxcl12/Sdf-1 in adult mice decreases the quiescent state of hematopoietic stem/progenitor cells and alters the pattern of hematopoietic regeneration after myelosuppression. Blood 117, 429–439. https://doi.org/10.1182/blood-2010-01-266833

Van der Maaten, L., Hinton, G., 2008. Visualizing Data using t-SNE. J. Mach. Learn. Res. 9, 2579–2605.

Van Nguyen, A., Kirchmair, M., Fürhapter, C., Romani, N., Sepp, N., 2004. Adhesive interactions between CD34(+)-derived dendritic cell precursors and dermal microvascular endothelial cells studied by scanning electron microscopy. Cell Tissue Res. 315, 139–143. https://doi.org/10.1007/s00441-003-0780-7

Varnum-Finney, B., Brashem-Stein, C., Bernstein, I.D., 2003. Combined effects of Notch signaling and cytokines induce a multiple log increase in precursors with lymphoid and myeloid reconstituting ability. Blood 101, 1784–1789. https://doi.org/10.1182/blood-2002-06-1862

Varnum-Finney, B., Halasz, L.M., Sun, M., Gridley, T., Radtke, F., Bernstein, I.D., 2011. Notch2 governs the rate of generation of mouse long- and short-term repopulating stem cells. J. Clin. Invest. 121, 1207–1216. https://doi.org/10.1172/JCI43868

Velten, L., Haas, S.F., Raffel, S., Blaszkiewicz, S., Islam, S., Hennig, B.P., Hirche, C., Lutz, C., Buss, E.C., Nowak, D., Boch, T., Hofmann, W.-K., Ho, A.D., Huber, W., Trumpp, A., Essers, M.A.G., Steinmetz, L.M., 2017. Human haematopoietic stem cell lineage commitment is a continuous process. Nat. Cell Biol. 19, 271–281. https://doi.org/10.1038/ncb3493

Verd, B., Crombach, A., Jaeger, J., 2014. Classification of transient behaviours in a time-dependent toggle switch model. BMC Syst. Biol. 8, 43. https://doi.org/10.1186/1752-0509-8-43

Viñuelas, J., Kaneko, G., Coulon, A., Vallin, E., Morin, V., Mejia-Pous, C., Kupiec, J.-J., Beslon, G., Gandrillon, O., 2013. Quantifying the contribution of chromatin dynamics to stochastic gene expression reveals long, locus-dependent periods between transcriptional bursts. BMC Biol. 11, 15. https://doi.org/10.1186/1741-7007-11-15

Visconti, R.P., Ebihara, Y., LaRue, A.C., Fleming, P.A., McQuinn, T.C., Masuya, M., Minamiguchi, H., Markwald, R.R., Ogawa, M., Drake, C.J., 2006. An In Vivo Analysis of Hematopoietic Stem Cell Potential. Circ. Res. 98, 690–696. https://doi.org/10.1161/01.RES.0000207384.81818.d4

Vorburger, S.A., Hunt, K.K., 2002. Adenoviral Gene Therapy. The Oncologist 7, 46–59. https://doi.org/10.1634/theoncologist.7-1-46

Waddington, C., 1957. The Strategy of the Genes: A Discussion of Some Aspects of Theoretical Biology. London: George Allen & Unwin, Ltd.

Wagner, J.E., Barker, J.N., DeFor, T.E., Baker, K.S., Blazar, B.R., Eide, C., Goldman, A., Kersey, J., Krivit, W., MacMillan, M.L., Orchard, P.J., Peters, C., Weisdorf, D.J., Ramsay, N.K.C., Davies, S.M., 2002. Transplantation of unrelated donor umbilical cord blood in 102 patients with malignant and nonmalignant diseases: influence of CD34 cell dose and HLA disparity on treatment-related mortality and survival. Blood 100, 1611–1618. https://doi.org/10.1182/blood-2002-01-0294

Wagner, P.D., 1977. Diffusion and chemical reaction in pulmonary gas exchange. Physiol. Rev. 57, 257–312.

Walther, W., Stein, U., 2000. Viral vectors for gene transfer: a review of their use in the treatment of human diseases. Drugs 60, 249–271.

Wang, J., Xu, L., Wang, E., Huang, S., 2010. The Potential Landscape of Genetic Circuits Imposes the Arrow of Time in Stem Cell Differentiation. Biophys. J. 99, 29–39. https://doi.org/10.1016/j.bpj.2010.03.058

Wang, J.C.Y., Doedens, M., Dick, J.E., 1997. Primitive Human Hematopoietic Cells Are Enriched in Cord Blood Compared With Adult Bone Marrow or Mobilized Peripheral Blood as Measured by the Quantitative In Vivo SCID-Repopulating Cell Assay. Blood 89, 3919–3924.

Wang, J.-H., Nichogiannopoulou, A., Wu, L., Sun, L., Sharpe, A.H., Bigby, M., Georgopoulos, K., 1996. Selective Defects in the Development of the Fetal and Adult Lymphoid System in Mice with an Ikaros Null Mutation. Immunity 5, 537–549. https://doi.org/10.1016/S1074-7613(00)80269-1

Wang, P., Na, J., 2011. Mechanism and methods to induce pluripotency. Protein Cell 2, 792–799. https://doi.org/10.1007/s13238-011-1107-1

Weigel, D., Jürgens, G., 2002. Stem cells that make stems. Nature 415, 751–754. https://doi.org/10.1038/415751a

Weissman, I.L., Anderson, D.J., Gage, F., 2001. Stem and progenitor cells: origins, phenotypes, lineage commitments, and transdifferentiations. Annu. Rev. Cell Dev. Biol. 17, 387–403. https://doi.org/10.1146/annurev.cellbio.17.1.387

Wen, L., Tang, F., 2016. Single-cell sequencing in stem cell biology. Genome Biol. 17. https://doi.org/10.1186/s13059-016-0941-0

Whitesides, G.M., 2006. The origins and the future of microfluidics. Nature 442, 368–373. https://doi.org/10.1038/nature05058

Wieth, J.O., Andersen, O.S., Brahm, J., Bjerrum, P.J., Borders, C.L., 1982. Chloride--bicarbonate exchange in red blood cells: physiology of transport and chemical modification of binding sites. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 299, 383–399.

Wilkins, R.C., Rodrigues, M.A., Beaton-Green, L.A., 2017. The Application of Imaging Flow Cytometry to High-Throughput Biodosimetry. Genome Integr. 8. https://doi.org/10.4103/2041-9414.198912

Wilson, N.K., Kent, D.G., Buettner, F., Shehata, M., Macaulay, I.C., Calero-Nieto, F.J., Sánchez Castillo, M., Oedekoven, C.A., Diamanti, E., Schulte, R., Ponting, C.P., Voet, T., Caldas, C., Stingl, J., Green, A.R., Theis, F.J., Göttgens, B., 2015. Combined Single-Cell Functional and Gene Expression Analysis Resolves Heterogeneity within Stem Cell Populations. Cell Stem Cell 16, 712–724. https://doi.org/10.1016/j.stem.2015.04.004

Wozniak, M.A., Modzelewska, K., Kwong, L., Keely, P.J., 2004. Focal adhesion regulation of cell behavior. Biochim. Biophys. Acta BBA - Mol. Cell Res., Cell Adhesion and Signalling 1692, 103–119. https://doi.org/10.1016/j.bbamcr.2004.04.007

Wu, A.M., Till, J.E., Siminovitch, L., McCulloch, E.A., 1967. A cytological study of the capacity for differentiation of normal hemopoietic colony-forming cells. J. Cell. Physiol. 69, 177–184. https://doi.org/10.1002/jcp.1040690208

Wu, J., Tzanakakis, E.S., 2013. Deconstructing stem cell population heterogeneity: Single-cell analysis and modeling approaches. Biotechnol. Adv. 31, 1047. https://doi.org/10.1016/j.biotechadv.2013.09.001

Wu, M., Singh, A.K., 2012. Single-cell protein analysis. Curr. Opin. Biotechnol., Analytical biotechnology 23, 83–88. https://doi.org/10.1016/j.copbio.2011.11.023

Yamagata, Y., Parietti, V., Stockholm, D., Corre, G., Poinsignon, C., Touleimat, N., Delafoy, D., Besse, C., Tost, J., Galy, A., Paldi, A., 2012. Lentiviral Transduction of CD34+ Cells Induces Genome-Wide Epigenetic Modifications. PLoS ONE 7, e48943. https://doi.org/10.1371/journal.pone.0048943

Yamamoto, R., Morita, Y., Ooehara, J., Hamanaka, S., Onodera, M., Rudolph, K.L., Ema, H., Nakauchi, H., 2013. Clonal Analysis Unveils Self-Renewing Lineage-Restricted Progenitors Generated Directly from Hematopoietic Stem Cells. Cell 154, 1112–1126. https://doi.org/10.1016/j.cell.2013.08.007

Ye, F., Huang, W., Guo, G., 2017. Studying hematopoiesis using single-cell technologies. J. Hematol. Oncol.J Hematol Oncol 10, 27. https://doi.org/10.1186/s13045-017-0401-7

Yeaman, M.R., 1997. The role of platelets in antimicrobial host defense. Clin. Infect. Dis. Off. Publ. Infect. Dis. Soc. Am. 25, 951-968; quiz 969-970.

Yoshihara, H., Arai, F., Hosokawa, K., Hagiwara, T., Takubo, K., Nakamura, Y., Gomei, Y., Iwasaki, H., Matsuoka, S., Miyamoto, K., Miyazaki, H., Takahashi, T., Suda, T., 2007. Thrombopoietin/MPL signaling regulates hematopoietic stem cell quiescence and interaction with the osteoblastic niche. Cell Stem Cell 1, 685–697. https://doi.org/10.1016/j.stem.2007.10.020

Yvernogeau, L., Gautier, R., Khoury, H., Menegatti, S., Schmidt, M., Gilles, J.-F., Jaffredo, T., 2016. An in vitro model of hemogenic endothelium commitment and hematopoietic production. Development 143, 1302–1312. https://doi.org/10.1242/dev.126714

Zaretsky, I., Polonsky, M., Shifrut, E., Reich-Zeliger, S., Antebi, Y., Aidelberg, G., Waysbort, N., Friedman, N., 2012. Monitoring the dynamics of primary T cell activation and differentiation using long term live cell imaging in microwell arrays. Lab. Chip 12, 5007–5015. https://doi.org/10.1039/C2LC40808B

Zeng, F., Chen, M., Katsumata, M., Huang, W., Gong, Z., Hu, W., Qian, H., Xiao, Y., Ren, Z., Huang, S., 2005. Identification and characterization of engrafted human cells in human/goat xenogeneic transplantation chimerism. DNA Cell Biol. 24, 403–409. https://doi.org/10.1089/dna.2005.24.403

Zhang, K., Han, X., Li, Y., Li, S.Y., Zu, Y., Wang, Z., Qin, L., 2014. Hand-Held and Integrated Single-Cell Pipettes. J. Am. Chem. Soc. 136, 10858–10861. https://doi.org/10.1021/ja5053279

Zhou, F., Li, X., Wang, W., Zhu, P., Zhou, J., He, W., Ding, M., Xiong, F., Zheng, X., Li, Z., Ni, Y., Mu, X., Wen, L., Cheng, T., Lan, Y., Yuan, W., Tang, F., Liu, B., 2016. Tracing haematopoietic stem cell formation at single-cell resolution. Nature 533, 487–492. https://doi.org/10.1038/nature17997

Zhu, Y.Y., Machleder, E.M., Chenchik, A., Li, R., Siebert, P.D., 2001. Reverse transcriptase template switching: a SMART approach for full-length cDNA library construction. BioTechniques 30, 892–897.

Zipori, D., 2004. The nature of stem cells: state rather than entity. Nat. Rev. Genet. 5, 873–878. https://doi.org/10.1038/nrg1475

Zopf, C.J., Quinn, K., Zeidman, J., Maheshri, N., 2013. Cell-Cycle Dependence of Transcription Dominates Noise in Gene Expression. PLOS Comput. Biol. 9, e1003161. https://doi.org/10.1371/journal.pcbi.1003161

Zovein, A.C., Hofmann, J.J., Lynch, M., French, W.J., Turlo, K.A., Yang, Y., Becker, M.S., Zanetta, L., Dejana, E., Gasson, J.C., Tallquist, M.D., Iruela-Arispe, M.L., 2008. Fate tracing reveals the endothelial origin of hematopoietic stem cells. Cell Stem Cell 3, 625–636. https://doi.org/10.1016/j.stem.2008.09.018

Résumé

Bien que largement étudiés, les mécanismes fondamentaux de prise de décision dans les processus de différenciation cellulaire restent mal compris. Les théories déterministes, souvent basées sur des études populationnelles, atteignent rapidement leur limite lorsqu'il s'agit d'expliquer les différences de choix individuels de cellules, pourtant exposées au même environnement.
L'objectif de ma thèse est donc d'étudier les premières étapes de la différenciation des cellules hématopoïétiques à l'échelle de la cellule unique, par des analyses transcriptomiques, protéomiques et morphologiques. Ce travail a été effectué sur deux modèles de différenciation : les lymphocytes T régulateurs et les cellules CD34+ humaines issues de sang de cordon. Nous avons observé le comportement de ces cellules uniques après stimulation.
Grâce à la combinaison de la microscopie en time lapse et des analyses moléculaires réalisées à l'échelle de la cellule individuelle, nous avons pu démontrer que le choix du devenir cellulaire n'était pas unique, programmé. La cellule passe d'abord par un état dit « multi-primed », métastable où elle exprime des gènes de plusieurs lignées différentes, puis elle passe par une phase dite « incertaine », instable où elle hésite entre deux phénotypes avant de se stabiliser dans un état fixe. Nos observations sont cohérentes avec une explication stochastique de la prise de décision.
La différenciation serait donc un processus spontané, dynamique, fluctuant et non un processus prédéterminé. Les décisions du destin cellulaire sont prises séparément par les cellules individuelles.

Abstract

Despite intensively studies, the fundamental mechanisms of cell fate decision during cellular differentiation still remain unclear. The deterministic mechanisms, often based on studies of large cell populations, cannot explain the difference between individual cell fates choices placed in the same environment.
The aim of my thesis work is to study the first steps of hematopoietic cell differentiation at the single cell level thanks to transcriptomic, proteomic and morphological analyses. Two differentiation models have been used: T regulatory lymphocytes and human cord blood-derived CD34+ cells. The behavior of individual cells following stimulation has been analyzed.
Using time-lapse microscopy coupled to single cell molecular analyses, we could demonstrate that the cell fate choice is not a unique, programmed event. First, the cell reaches a metastable "multi-primed" state, which is characterized by a mixed lineage gene expression pattern. After transition through an "uncertain", unstable state, characterized by fluctuations between two phenotypes, the cell reaches a stable state. Our observations are coherent with a stochastic model of cell fate decision. The differentiation is likely to be a spontaneous, dynamic, fluctuating and not a deterministic process. The cell fate decisions are taken by individual cells.

Mots Clés

Cellules hématopoïétiques, cellules uniques, destin cellulaire, dynamique morphologique et moléculaire, comportement hétérogène, stochasticité.

Keywords

Hematopoietic cells, single cell, cell fate, morphological and molecular dynamic, heterogeneous behavior, stochasticity.