Université Sorbonne Paris Cité

Thèse de doctorat

de l'Université Sorbonne Paris Cité

Préparée à l'Université Paris Diderot

Ecole doctorale HOB, ED561

INSERM UMR1170 / Equipe 2 « From HSC to megakaryocyte »

Physiopathological mechanisms of two congenital platelet disorders : Filaminopathy A and ANKRD26-related thrombocytopenia (THC2)

Par Alessandro Donada

Thèse de doctorat de Hématologie

Dirigée par Hana Raslova

Présentée et soutenue publiquement à Villejuif le 27 Septembre 2018

Président du jury : Pr. Ajzenberg Nadine

Rapporteur : Pr. Freson Kathleen

Rapporteur : Pr. Payrastre Bernard

Examinateur : Pr. Delhommeau François

Examinateur : Dr. Kauskot Alexandre

Directeur de thèse : Dr. Raslova Hana

“ O frati,” dissi, “che per cento milia
perigli siete giunti a l’occidente,
a questa tanto picciola vigilia

d’l ostril sensi ch’é del rimanente
non vogliate negar l’esperïenza,
di retro al sol, del mondo sanza gente.

Considerate la vostra semenza:
fatti non foste a viver come bruti,
ma per seguir virtute e canoscenza”.

Dante Alighieri, Divina Commedia

“Immigrants, we get the job done.”

Lin-Manuel Miranda, Hamilton: an American Musical

Acknowledgements

This thesis is the results of 4 years of work in an incredible lab, in many ways. I have received a special imprinting, as a scientist and a man, so I hope this manuscript could be valid as the training I have received.

A study claimed that on average, less than 1 individual is reading a Ph.D. manuscript from the beginning to the end. Therefore, I am very grateful to **Prof. Bernard Payrastre** and **Prof. Kathleen Freson** for their extremely complete and insightful reports. You clearly improved my perception of this work and helped me in the correction of the many typos and grammar blunders, and I am really looking forward to discuss with you those results during the Ph.D. defense.
I would also like to thank the other members of the committee, **Prof. Nadine Ajzenberg**, **Prof. François Delhommeau** and **Dr. Alexandre Kauskot** for their time and availability. I hope you will appreciate this thesis and I will try to answer to your questions at the best of my capabilities.

Hana, I do not think that I would be able to thank you enough for your mentorship. You believed in me, and I still wonder why you thought I was good enough to carry on with an internship in your lab, as I had no experience in haematology or cell culture. You have always being there for me, from an abstract to correct to a boot for a broken foot, I had the chance to have you supporting and believing in me. You tolerated my multiple injuries, you shared a laugh and you gave me a wonderful example of the scientist I would like to become. Thank you, from the deep of my heart.

William, the sharpness of your scientific mind amazes me everyday, since the first time I met you. You are an inspiration, I would feel extremely grateful to achieve just 1/10 of the scientific success you had in your career. Thank you for the time you dedicated me, and the enriching discussions we had during these last 4 years.

Thank you to the PIs who find the time to comment and help me out along the way: **Isabelle**, **Najet**, **Caroline** and **Thomas**, your feedback has always been very appreciated, so it has been your support.
At the same time, many many many thanks to all the people I have met during the last four years: **Laure**, **Virginie**, **Mira**, **Barbara**, **Larissa** (thank you so much for your help during the last FLNA experiments!), **Antonio**, the triade **Philippe-Cyril-Yann** (your teaching, help and patience has been invaluable!), **Rahma** and **Paule**, the Italian colony at Gustave Roussy **Francesco-Bojana-Giusy**,

the master of mice **Lamia** and everyone else that I have missed out purely because of space! Thank you to all!

Brahim, the Di Caprio of Berberie, what a fantastic lad! Thank you for your good humour and time spent listening to my non-sense without a single complains, even if I still can't feel the desert in a cup of black tea (sorry mate). I hope my cumbersome presence in the office did not bother you too much.

Ileana, the most amazing desk neighbor! You are a true legend, professionaly incredible and always smiling, teach me how to do it please! I really wish you all the best for your new career and I am sure you are going to rock it!!! Can't wait to meet the second little one…

To the young blood, **Rameez**, **Deshan**, **Graciela**, **Amandine**, **Camelia**, **Fabien**, remember to enjoy the ride as much as you can. It will not always be nice and cosy, quite the opposite, but it is worth the price! Keep calm and get your Ph.D. done!

Lise, you tried to teach me how to write my thesis in French, but that was probably too much for me. Nevertheless, I manage to get fluent enough to survive adequately in this city, and that is mostly because of you. Merci beaucoup for your example, you proved me what the word resilience really means.

Francesca, **Valentina** and **Manuel**, I must apologize: you probably expected a little bit more French or English for your internships, but eventually you managed to find me on your way. Sorry. I hope the experience was not too bad. From my side, you are exceptional: keep striving and mai molar, you are surely going to succeed, whatever you are planning to do!

I have to thank the man who introduced me to research and iPSCs, **Vladimir**. The stern gaze you gave me when I managed to aspire 4 out of 6 B100s of your precious cells is still with me, together with all the great advice and clever insights you shared. I hope you will enjoy this work as much as I have enjoyed your friendship.

Matt, this has been so hard, way more than what I expected. But you eventually made it very enjoyable. Every day, you left me with a satisfied smile on my face. You never pulled out, and that has been such a relief. I am so glad you came all the way for me, even when it was wet and tight.

Nathalie, I definitely cannot avoid to acknowledge you, together we have been probably the loudest, most complaining, definitely funniest duo of scientists who ever walked in the lab! As much as you were able to “trigger me”, you also taught me about the dedication for the job. We spent many hours, shoulder to shoulder, and it has been a blast! I sincerely wish you all the best for your future!

We started together, we end it up together. You are out for a real challenge, but you have the numbers for passing it with flying colours, even if the vWF is not specific enough and your hierarchy is just an approximation. You go sister, Stockholm is not that far in the end, and you still have to deliver many of the things you promised me! I will sorely miss you.

Je sais toujours pas qu'est ce qu'il y connait aux femmes Rick Hunter, but I know that you are one of the greatest asset of this lab. Your attitude is amazing, your professionality undeniable, your heart incredible. I am honoured to call you mon pote, but you are clearly more than that. You are a role model, and I am proud to be your friend.

I am in huge debt with all the lads of the **British Rugby Football Club of Paris**. The first thing I have done when I decided to move to Paris was looking for an internship, but the second was looking for a rugby club, as I knew that I would not be able to achieve much without some time on the pitch. I have been blessed with the most incredible bunch of brothers I could have ever imagined, and no, I would not sign your medical certificate for the licenses, I am not that type of doctor!

Last, but not least, the two who made all this possible: **mum** and **dad**, you made me the man I am. I really hope I have made you proud.

Titre: Physiopathological mechanisms of two congenital platelet disorders: Filaminopathy A and ANKRD26-related thrombocytopenia (THC2)

Résumé :

Les thrombopénies familiales sont une classe de maladies hématologiques congénitales affectant principalement la lignée mégacaryocytaire et s'accompagnant d'une diminution du nombre de plaquettes. Près de 50 gènes différents ont été associés à ces thrombopénies héréditaires, qui sont très hétérogènes sur les plans cliniques et physiopathologiques. Mes recherches ont porté sur deux thrombocytopénies congénitales différentes: la Filaminopathie A et la Thrombocytopénie 2. La première maladie est un syndrome lié à l'X, associé à des mutations dans le gène *FLNA* (Filamine A). Ces patients, présentent, au niveau hématologique, une macrothrombopénie associée à un risque hémorragique. La seconde maladie est de transmission autosomique dominante, et elle est due à des mutations dans la région 5' UTR du gène *ANKRD26*. Les patients présentent une dysmégacaryopoïèse, associée à une thrombopénie et un risque accru de développer des hémopathies myéloïdes.

Pour étudier la physiopathologie de ces deux maladies rares, j'ai exploité la technologie des cellules souches pluripotentes induites, afin de développer plusieurs lignées cellulaires spécifiques des patients. Ces outils expérimentaux ont permis de comprendre la physiopathologie de ces deux maladies, et de décrire les mécanismes moléculaires responsables du défaut de la formation de proplaquettes pour la Filaminopathie A, et de la prédisposition à la leucémie pour la Thrombocytopénie 2. Pour effectuer de telles études, nous avons mis au point un protocole de différenciation robuste, récapitulant efficacement les différentes étapes de la différenciation hématopoïétique et adapté à la différenciation *in vitro* de plusieurs lignées cellulaires. De plus, nous avons exploité une technique de modification du génome pour introduire efficacement différents mutants, afin de disséquer le rôle moléculaire de la filamine A dans la mégacaryopoïèse.

En ce qui concerne la filaminopathie A, nous avons pu décrire une relation originale entre une intégrine membranaire ($\alpha_{IIb}\beta_3$), la filamine A (FLNa) et une voie de signalisation cruciale (RhoA) pour la mégacaryopoïèse. Nos données supportent un modèle où l'absence de FLNa induit une activité anormale de la voie RhoA, en réponse à l'intégrine $\alpha_{IIb}\beta_3$ après liaison du fibrinogène. Concernant la thrombopénie 2, nous avons décrit un nouveau mécanisme associant l'augmentation d'expression d'*ANKRD26* à une activité dérégulée de la voie de signalisation dépendante du facteur de croissance G-CSF. Cette anomalie affecte la granulopoïèse et conduit à une amplification anormale de cette lignée myéloïde, augmentant potentiellement le risque d'acquisition d'autres mutations et de progression vers une leucémie aiguë myéloïde.

En conclusion, par ce travail, nous prouvons l'utilité des cellules souches pluripotentes induites pour la modélisation et l'étude de la physiopathologie de maladies. Nos résultats ouvrent la voie à d'autres études susceptibles de faire progresser la compréhension de la physiopathologie des troubles plaquettaires héréditaires.

Mots clefs : Mégacaryopoïèse, Pré-leucémie, iPSC, Cytosquelette, Signalisation

Title : Physiopathological mechanisms of two congenital platelet disorders : Filaminopathy A and ANKRD26-related thrombocytopenia (THC2)

Abstract

Inherited thrombocytopenias are a class of congenital haematological disorders affecting primarily the megakaryocytic lineage and accomunated by a decrease in platelet numbers. Almost 50 different genes have been associated to inherited platelet disorders, and huge differences exist between each disorder, in regard to clinical manifestation and pathobiology. My research interest have been focused on two different congenital thrombocytopenias: Filaminopathy A and Thrombocytopenia 2. The first disease is a X-linked syndrome associated to mutations in the gene *FLNA* (Filamin A), and patients display a mild to severe macrothrombocytopenia, associated with a lifelong bleeding tendency. The second disorder is an automal dominant condition caused by mutations in the 5' UTR of the *ANKRD26* gene. It is associated with dysmegakaryopoiesis, mild to severe thrombocytopenia and an increased risk to develop myeloid malignancies.

To study the physiopathology of those two rare diseases, we have exploited the induced pluripotent stem cell technology to develop several patient specific cell lines. Those experimental tools revealed invaluable for the understanding of the disease physiopathology, and allowed us to describe in great details the molecular mechanisms underlying the reduction in proplatelet formation for Filaminopathy A and the predisposition to leukemia for Thrombocytopenia 2. To perform such studies, we devised a robust differentiation protocol, recapitulating efficiently the haematopoietic differentiation and easily adapted to the *in vitro* differentiation of multiple cell lineages. Furthermore, we exploited a genome editing technique to introduce efficiently different protein mutants, in order to dissect the molecular role of Filamin A in megakaryopoiesis.

In regard of Filaminopathy A, we have been able to describe an original and novel relationship between a membrane integrin ($a_{IIb}b_3$), Filamin A and a crucial signalling pathway (RhoA) for megakaryopoiesis. Our data support a model where the absence of FLNa induces an abnormal

activity of the RhoA pathway, in response to the integrin $a_{IIb}b_3$ binding to fibrinogen. Concerning the thrombocytopenia 2, we described a novel mechanism that associated the increased expression of *ANKRD26* to a deregulated activity of the G-CSF-dependent signalling pathway. This anomaly impacts the normal granulopoiesis and lead to an abnormal amplification of this cell lineage, possibly increasing the risk of acquiring other mutational hits and progress towards a myeloid malignancy.

In conclusion, with this work we offer a proof of concept of the potentiality of disease modelling via induced pluripotent stem cells. Our results pave the way for further studies that could advance our understanding of the physiopathology of inherited platelet disorders.

Keywords : Megakaryocytes, Preleukemia, iPSC, Cytoskeleton, Signalling

TABLE OF CONTENTS

Abbreviations lists (listed alphabetically)

5-hMC: 5'-hydroxymethylcytosine

5' UTR: 5' untranslated region

αKG: alpha ketoglutarate

ABD: actin-binding domain

ACTN1: actinin 1

ADAM17: ADAM metallopeptidase domain 17

ADP: adenosine diphosphate

AEBP2: AE binding protein 2

AGM: aorta-gonado-mesonephros

ALL: acute lymphoblastic leukemia

AML: acute myeloid leukemia

ANKRD26: ankyrin-repeat domain 26

AP2: Adaptor protein 2

B-ALL: B-cell acute lymphoblastic leukemia

B: B-lymphocytes

BCAT1: branched chain aminoacid transaminase 1

BCL-XL: B cell lymphoma-extra large

BCR-ABL1: breakpoint cluster region-c-abl oncogene 1

bFGF: basic fibroblast growth factor

BFU-E: burst forming unit erythroid

BL-CFC: blast-like cell forming colony

BMI1: polycomb ring finger

BMP4: bone morphogenetic protein 4

BSS: Bernard Soulier syndrome

c-Myc: avian myelocytomatosis

cALL: childhood acute B-lymphoblastic leukemia

CALR: calreticulin

cAMP: cyclic adenosine monophosphate

CAMT: congenital amegakaryoblastic thrombocytopenia

CAR-T: chimeric antigen receptor T

CBL: Cbl proto-oncogene

CCNI2: cyclin I family member 2

CDC20: cell division cycle 20

CDC42: cell division cycle 42

CDK: cyclin-dependent kinase

CDK1: cyclin-dependent kinase 1

CDK5: cyclin dependent kinase 5

CDKN: cyclin-dependent kinase inhibitor

CDX: caudal type homeobox family

CFU-GM: colony-forming unit granulo-monocyte

CFU-MK: colony forming unit megakaryocyte

CLP: common lymphoid progenitor

CMP: common myeloid progenitor

CN-LOH: copy neutral-loss-of-heterozygosity

CNL: chronic neutrophilic leukemia

CNV: copy number variation

CREB: cAMP responsive element-binding protein

CRISPR/Cas9: clusterered regularly interspaced short palindromic repeats/CRISPR associated protein 9

CRM1: cytokine receptor homology module 1

CTCF: CCCTC-binding factor

DBA: Diamond-Blackfan anemia

DC: dendritic cells

DIAPH1: diaphanous related formin 1

DLL4: delta like canonical Notch ligand 4

DMS: demarcation membrane system

DNMT: DNA methyltransferase

DNMT3: DNA methyltransferase 3

DOCK7: dedicator of cytokinesis 7

DS B-ALL: Down syndrome B-Acute lymphoblastic leukemia

DS-AMKL: Down syndrome acute megakaryoblastic leukemia

EB: embryoid body

ECM: extracellular matrix

EED: embryonic ectoderm development

EHT: endothelial-to-haematopoietic transition

EJC: exon junction complex

EMP: erythro-myeloid progenitor

EpiSC: epiblast stem cell

EPO: erythropoietin

EPOR: erythropoietin receptor

ER: endoplasmic reticulum

ERG: v-ets avian erythroblastosisvirus E26 oncogene homolog

ERK1/2: mitogen activated kinase kinase kinase 1/2

ESC: embryonic stem cells

ESSRB: estrogen related receptor beta

ET: essential thrombocytemia

ETS: E26 transformation-specific

ETV6: ETS variant 6

EZH2: enhancer of Zeste 2

FACS: fluorescence activated cell sorting

FGF: fibroblast growth factor

FL: fetal liver

FLI1-RT: FLI1-related thrombocytopenia

FLI1: Friend leukemia integration 1

FLNA: filamin A (gene)

FLNa: filamin A (protein)

FLNB: filamin B

FLNC: filamin C

FOG-1: friend of GATA1

FOSB: Fos proto-oncogene B

FPD/AML: familial platelet disorder with predisposition to acute myeloid leukemia

FRET: fluorescence resonance energy transfer

G-CSF: granulocyte colony stimulating factor

G-CSFR: granulocyte colony stimulating factor receptor

GAP: GTPase activating protein

GATA1: GATA-binding factor 1

GATA1s: GATA1 short

GATA2: GATA-binding factor 2

GEF: guanine exchange factor

GFI1b: growth factor independence 1b

GMP: granulo-monocytic progenitor

GPA: glycophorin A

GPS: grey platelet syndrome

Gran: granulocytes

GSK3: glycogen synthase kinase 3

GT: Glanzmann thrombasthenia

GTPase: guanosine triphosphate hydrolase

H2AK119u1: monoubiquitination lysine 119 histone H2A

H3K27me3: trymethylation lysine 27 histone 3

H3K4me3: trymethylation lysine 4 histone 3

HAT: histone acetyltransferase

HDAC: histone deacetylase

HDAC1: histone deacetylase 1

HDAC2: histone deacetylase 2

HE: haemogenic endothelium

hESC: human embryonic stem cell

HLA: human leukocyte antigen

Hlf: hepatic leukemia factor

HMMS: hereditary myeloid malignancy syndromes

HOXA: homeobox A family

HOXA11: homeobox 11A

HOXA5: homeobox 5A

HOXA9: homeobox 9A

hPSC: human pluripotent stem cell

HSC: haematopoietic stem cell

IAHC: intra-aortic haematopoietic cluster

ICM: inner cell mass

IDH1/2: isocitrate dehydrogenase 1/2

IGF: insulin growth factor

IPD: inherited platelet disorder

IT: inherited thrombocytopenia

ITGA2B: integrin subunit alpha 2b

ITGB3: integrin subunit beta 3

JAK2: Janus kinase 2

JARID2: jumonji and AT-rich interaction domain containing 2

JBS: Jacobsen syndrome

JH1: janus homology 1 domain

KDR: kinase insert domain receptor

KLF1: kruppel like factor 1

KLF4: kruppel like factor 4

LCOR: ligand dependent nuclear receptor corepressor

LIF: leukemia inhibitor factor

LIMK: LIM kinase

LMO2: LIM domain only 2

LMPP: lymphoid-myeloid multipotent progenitors

LNK: SH2B adaptor protein 3

LSD1: lysine demethylase 1A

LYL1: lymphoblastic leukemia associated hematopoiesis regulator 1

MAPK: mitogen-activate protein kinase

MDS: myelodysplastic syndrome

Meis1: Meis homeobox 1

MEK1/2: mitogen activated kinase kinase 1/2

MEP: megakaryocytic-erythroid progenitor

mESC: mouse embryonic stem cell

MK-P: megakaryocyte progenitor

MK: megakaryocyte

MKL1: megakaryoblastic leukemia 1

MKpoiesis: megakaryopoiesis

MLC2: myosin regulatory light chain 2

MLL: mixed lineage leukemia

MMP9: matrix metallopeptidase 9

Mono: monocytes

MPL: myeloproliferative leukemia

MPN: myeloproliferative neoplasms

MPP: multipotent progenitor

MVB: multivesicular bodies

MYB: MYB proto-oncogene

MYH10: myosin IIB

MYH9-RD: MYH9-related disease

MYH9: myosin IIA

n-Myc: see c-Myc

NBEAL2: neurobeachin like 2

NF-kB: nuclear factor kappa B subunit

NuRD: nucleosome remodelling deacetylase

Oct4/POUF51: POU class 5 homeobox 1

OSKM: OCT4, SOX2, KLF4 and c-Myc

OxPHOS: oxidative phosphorylation

Pbx1: PBX homeobox 1

PF4: platelet factor 4

PH: pleckstrin homology

PI3K: phosphatydilinositol-4,5-biphosphate-3-kinase

PIM1: Pim-1 proto-oncogene

PIP_3: phospahtydilinostol-3,4,5-triphosphate

PKA: protein kinase A

PLT: platelet

PMF: primary myelofibrosis

PNH: periventricular nodular heterotopia

PPT: proplatelet

PRC: polycomb repressor complex

PRC1: polycomb repressor complex 1

PRC2: polycomb repressor complex 2

PRDM14: PR/SET domain 14

PRDM5: PR/SET domain 5

PRKACG: protein kinase cAMP-activated catalytic subunit gamma

PSC: pluripotent stem cell

PT-VWD: platelet type-von Willebrand disease

PV: polycythemia vera

Rac1: Rac family small GTPase 1

RAF1: Raf-1 proto-oncogene

RAS: Rous sarcoma

RBC: red blood cell

RBM15: RNA binding motif protein 15

RBM8B: RNA binding motif protein 8B

RhoA: ras homolog gene family, member A

ROCK1/2: RhoA-associated protein kinase 1/2

RORA: RAR related orphan receptor A

RUSAT: radio-ulnar synostosis associated thrombocytopenia

SAM: S-adenosylmethionine

Scl: see TAL1

SCN: severe congenital neutropenia

SF3B1: splicing factor 3b subunit 1

SH2: Src homology 2 domain

SMAD2/3: SMAD family member 2/3

SNCT: somatic nuclear cell transfer

SNP: single nucleotide polymorphisms

SOCS: suppressor of cytokine signalling

SOX2: SRY-box 2

SOX4: SRY-box 4

SPI1: spleen focus forming virus. Also called PU.1

SRC: SRC proto-oncogene

SRF: serum response factor

SRSF2: serine and arginine rich splicing factor 2

SSEA-1: stage specific embryonic antigen 1

STAT: signal transducer and activator of transcription

STAT1: signal transducer and activator of transcription1

STAT3: signal transducer and activator of transcription 3

STAT5: signal transducer and activator of transcription 5

SUZ12: suppressor of zeste 12 homolog

SYK: spleen associated tyrosine kinase

T: T-lymphocytes

TAD: topological associated domain

TAL1: T-cell acute lymphocytic leukemia 1

TAR: thrombocytopenia absent radii

TCA: tricyclic carboxylic acid cycle

TCR: T cell receptor

TET: ten-eleven translocation

TET2: ten-eleven translocation 2

TF: transcription factor

TGFβ: transforming growth factor β

THC2: thrombocytopenia type 2, also called ANKRD26-RT

THC5: thrombocytopenia type 5, also called ETV6-RT

TM: transmembrane

TPM4: tropomyosin 4

TPO: thrombopoietin

TUBB1: tubulin beta 1 class VI

TYK2: tyrosine kinase 2

VEGF: vascular endothelial growth factor

vWF: von Willebrand Factor

WAS: Wiskott-Aldrich syndrome

WASP: WAS protein

WNT: wingless

Xi: inactive X chromosome

ZFP37: zinc finger protein 37 homolog

Table of figures (as they appear in the text)

INTRODUCTION

CHAPTER I

1. The normal haematopoiesis

1.1. The haematopoietic system

1.1.1. The development of the haematopoietic system

If we consider the haematopoietic system as formed by freely moving cells, with structural and functional properties, then all the members of the animal kingdom display a haematopoietic system. The major distinction consist in the presence or absence of a circulatory system, that could be open or closed, an evolutionary feature that closely associate the ontogeny of both systems. The concept of a common origin between blood vessels and blood cells was experimentally defined by Sabin in 1917 in the chicken[1] and was confirmed for many other organisms, vertebrate and invertebrate alongside. At the cellular level, Murray in 1932 defined the haemangioblast: «The "angioblasts" are masses of cells or single cells generated from the mesenchyme, from which develop the endothelial vessels containing plasma and the blood islands; [...]. To "angioblast", however, I prefer the term "hæmangioblast". This expresses the fact that both endothelium and blood develop from the solid mass, whereas the term "angioblast" strictly refers only to the vessels, i.e., to the endothelium.»[2]. The intuition of a single oligopotent cellular precursor, able to pursue an haematopoietic fate or an endothelial one, was experimentally verified long time after the first, seminal, observations: in 2004, Gordon Keller and collaborators isolated a cellular population, originating in the posterior region of the primitive streak, that was able to migrate to the yolk sac and give rise to the blood islands, masses of cells responsible of the production of the first haematopoietic cells[3]. This result was largely based on the demonstration of the existence of an *in vitro* population, differentiated from mouse embryonic stem cells (mESCs), that was able to give rise to haematopoietic and endothelial cells[4]. This population, called Blast Like Cell Forming Colony (BL-CFC), expressed the marker *Kdr* (known also as *Flk1* or *Vegfr2*) and the mesodermal marker *Brachyury*. Nevertheless, Murray's definition was not specifically calling for a single bipotent progenitor, but was also leaving room for a subtler picture of the haematopoietic specification, with several, already poised, unipotent progenitors. In this regard, it is worth to notice that the BL-CFCs does not express some of the classic endothelial markers like CD31 or VE-cadherin, or some haematopoietic markers like CD41 or CD45. An alternative model, experimentally supported by Padrón-Barthe and colleagues, theorised a linear specification, from an independent epiblast population (so prior to gastrulation) to the early yolk sac blood island, through an intermediate state[5,6].

The chicken embryo was also instrumental for another major breakthrough in our understanding of the haematopoiesis development: in 1975, Dieterlen-Lievre demonstrated experimentally the existence of two different ontogenetic programs, originating in two different sites of the embryo[7]. The yolk sac-derived program consisted of a transient wave of functionally restricted haematopoietic cells, while the embryo-derived program was responsible of a persistent and functionally expanded wave of cells, that finally gave rise to the adult haematopoietic system. This second wave of haematopoiesis was called definitive, and included the generation of the long-standing pool of stem cells responsible of maintaining the entire blood system in the adult. Those cells are called haematopoietic stem cells (HSCs). This paradigm shift was confirmed in many other vertebrate models, including humans[8-10], thus proving a remarkable degree of conservation of the haematopoietic system ontogeny (Figure 1).

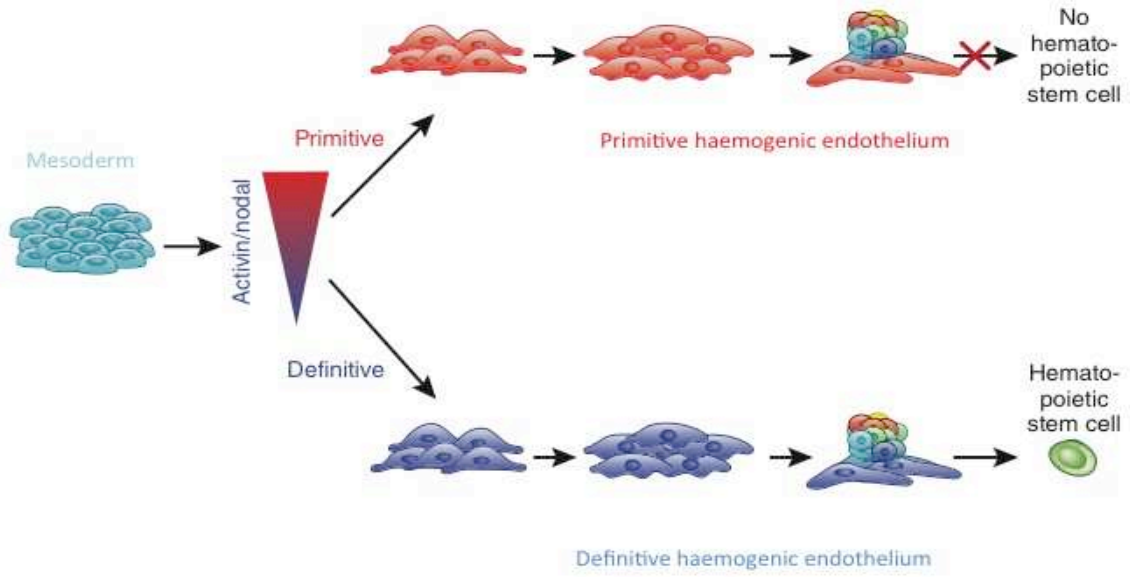


Figure 1: the endothelial to haematopoietic transition. Modified from Sturgeon et al 2013[29].

The primitive haematopoietic, as already said, is transient in nature and highly restricted, in functional terms: it includes only nucleated red blood cells, primitive megakaryocytes and macrophages. It emerges very early during the mouse development, approximately around E7. All those three cell types display molecular and functional differences with their definitive counterparts: the erythroblasts show different proliferative properties, do not undergo enucleation and express prevalently the ε-globin[11]. The megakaryocytes produce very few platelets and display a low degree of ploidy, and the macrophages mature without an intermediate monocytic stage[12,13]. As these cells emerge from the yolk sac before the development of a circulatory system[14,15], a deeper understanding of their functional roles in the context of the embryo development would be desirable, but very few information is available. It is nevertheless intriguing to speculate about how

these lineages are not directly generated from a clearly defined oligopotent progenitor: there are no evidences of such cellular intermediate and those cells exhibit clear differences in their differentiation and maturation, compared to the definitive counterparts. Is this an evidence of the dispensability of a multipotent progenitor cell for the generation of functional mature haematopoietic cells? What are the molecular determinants of this transition and to which degree are they different from the molecular determinants of the adult haematopoiesis?

The definitive haematopoietic program emerges after the onset of the primitive one and is firmly based on the existence of an endothelial progenitor able to originate cells of a haematopoietic nature. A cell with this differentiation potential is called hemogenic endothelium (HE), and endothelial-to-haematopoietic transition (EHT) is the name of the transition from one cell identity to the other. The typical example HE is the dorsal area of the aorta-gonado-mesonephros region (AGM region), a region described as the place where the first HSC emerges[16]. All the haematopoietic progenitors (HSCs included) that emerge from the AGM are organised in small clusters that bud from the ventral floor of the dorsal aorta, called intra-aortic haematopoietic clusters (IAHC). The endothelial nature of the progenitor cell is a consequence of the co-expression of haematopoietic and endothelial markers, both in mouse and human[17-19]. These HSCs, defined functionally as able to engraft successfully an adult recipient without any prior treatment, appear in the mouse at E10.

The AGM is not the only site of the embryo containing a HE, at least in the mouse: placenta[20], head[21], vitelline and umbilical arteries[22], and even the yolk sac[23] display the capacity to undergo EHT, at almost the same time as the AGM development. This means that, while different for their inception, the primitive and the definitive programs partially overlap. In fact, it was demonstrated that the hemangioblast can give rise directly to haematopoietic cells, but can also generate a HE[24,25]. Therefore, is this specific HE able to give rise to an HSC? There are evidences that the haematopoietic cells produced in the yolk sac from this haematopoietic program do not share the same feature of their primitive counterparts, but have different functional properties: for example, the erythrocytes express γ-globin together with the ε, and they follows a differentiation and maturation pattern involving the formation of burst forming unit (BFU-E), so a more intense proliferation of the erythroblasts[26]. In addition, it is possible to isolate other cell types not included in the primitive program: granulocytes, mast cells, T-lymphocytes and B-lymphocytes. The presence of the granulocytes could be related to a different macrophages maturation through a granulo-monocytic intermediate, so a hint of the presence of an at least bipotent haematopoietic progenitor[27]. More intriguing is the presence of lymphoid cells, in an otherwise completely myeloid scenario: is this the sign of the presence of a first multipotent progenitor, able to differentiate in all

the major haematopoietic lineages? Evidences suggest the overlapping of two different progenitor types, with a myeloid-exclusive one called erythro-myeloid progenitor (EMP), and a lymphoid-myeloid multipotent progenitor (LMPP)[28], although further work is required in this sense. These first multipotent progenitors do not show long-term stable engraftment in an adult recipient. The AGM-independent origin of these cells could be deducted thanks to the analysis of the emergence kynetics of these populations, as they are observed from E8 to E10, therefore prior to the development of a circulatory system. This feature avoids the possible contamination with AGM-dependent cells (HSCs derived from the AGM), as they can potentially colonise the yolk sac.
While the cells derived from the primitive wave display a somehow unclear functional role and are restricted to the early development, this pre-definitive haematopoiesis has been associated to several crucial events of the embryo development and give rise to a series of adult tissues, haematopoietic in origin, that retain some long term self-renewing properties. In particular the macrophages produced by EMPs and LMPPs are involved in the tissue remodelling during the neuronal and the vascular development[30], and originate different tissue-specific subsets of macrophages: the microglia (brain), the Kupffer cells (liver) and the Langherans cells (epidermis). About the lymphoid derivatives, some subsets of T and B-cells displays a certain persistence along the adult life: it is the case of the γδ subset, which resides in some peripheral tissue like the gut and bear a restricted repertoire of T-cell receptors[31]. About the B-cells, a subset called B-1 colonise the peritoneal and the pleural cavities[32]. These examples of HSC-independent haematopoiesis underlie some functional differences that deserve further investigations, and are possibly exploitable for therapeutic purposes.

There is another cell population, generated between the first primitive wave and the emergence of the first true HSC, that attracted some attention in the recent years. There is an entire body of work that prove the existence of nascent pre-HSCs, immature HSCs that originate in the yolk sac and the para-aortic splanchnopleura (the pre-AGM), after the primitive wave and alongside the EMPs and LMPPs[33]. They are defined as pre-HSCs, so not strictly called stem cell, because the earliest haematopoietic cell able to engraft a primary recipient emerges from the AGM at a later time point. Nevertheless, they can be converted into a true HSC, as they eventually gain the ability to engraft a secondary recipient, suggesting that there is a cell extrinsic component in the acquisition of the stemness. This notion of immature HSC gained traction recently, with the demonstration that it is possible to obtain the same maturation effect caused by the transplantation with the creation of an artificial niche *in vitro*, therefore recapitulating the fetal environment. This has been demonstrated only for the pre-HSCs derived from the pre-AGM[34-37]. A major requirement for this type of study is the ability to better define the pre-HSCs in phenotypic terms; as a consequence, the identification of

more univocal combinations of markers is underway, with the final goal of understanding more in details the transitions between these stages and uncovers any possible heterogeneity[38]. Finally, a comparable population of pre-HSCs, intended as progenitor cells able to engraft a recipient after a step of *in vitro* maturation, was recently described in the placenta[39]. Nevertheless, the physiological importance of this population has to been demonstrated yet, especially in terms of contribution to the successive haematopoiesis.

The understanding of the steps defining the HSC maturation from these progenitors would be instrumental in the demonstration of the existence of the same ontogenesis in the human species, and not only in the mouse (Figure 2).

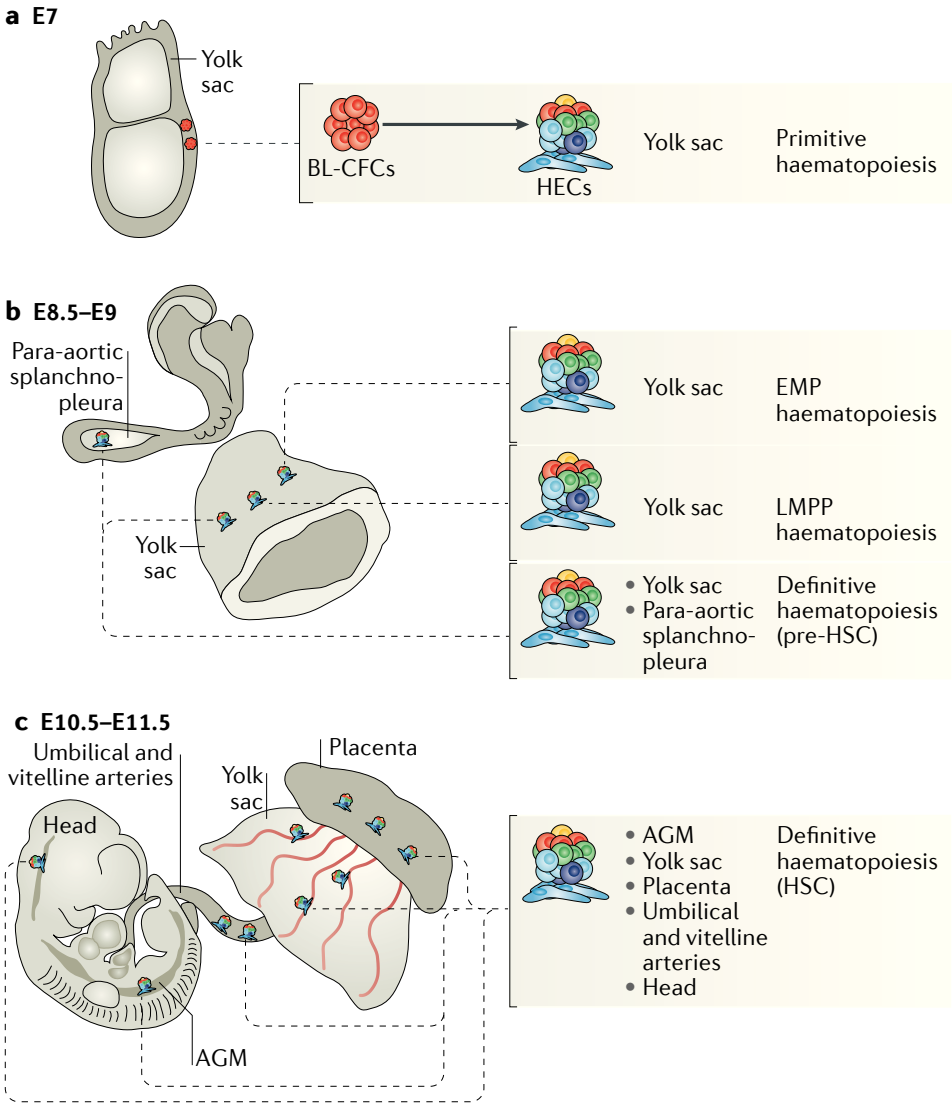


Figure 2: the different haematopoietic waves during embryo development. Readapted from Ditadi, Sturgeon and Keller 2016[40].

The number of HSCs produced by the AGM is minimal, compared to the needs of the growing embryo: while some of the required functionalities are assumed in first instance by the EMP/LMPP-derived haematopoiesis (the erythrocytes in particular are in high demand), there is clearly the need to amplify the stem and progenitor compartments. If the AGM is responsible for the emergence of the first stem cells, no actual haematopoietic differentiation is ongoing in that region. It is then necessary to have some haematopoietic sites that could act as "nurseries" for the nascent HSCs, microenvironment where they can amplify and differentiate. That is the role of the fetal liver (FL): colonised by progenitors of the pre-definitive haematopoiesis as well, it is the site where the haematopoiesis occurs during most of the fetal life[41]. From the FL, some progenitors migrate to the fetal thymus, where they will complete their differentiation and maturation towards the lymphoid fate. As the HSCs display a great proliferative potential, without losing at all their differentiation properties[42], it is quite surprising the difference with their adult life, where they spend most of their time in a quiescent state. Very few data are available about the molecular pathways involved, as there are some difficulties in the isolation of different progenitor population[43], a step required for a pairwise comparison with the adult equivalent, if they are equivalent (Figure 3). Finally, after birth, the stem and progenitor compartments migrate again, to their final niche, the bone marrow. Here they will spend the rest of their adult life, although some pathological anomalies could force them to leave it for other haematopoietic sites, like the spleen or the adult liver.

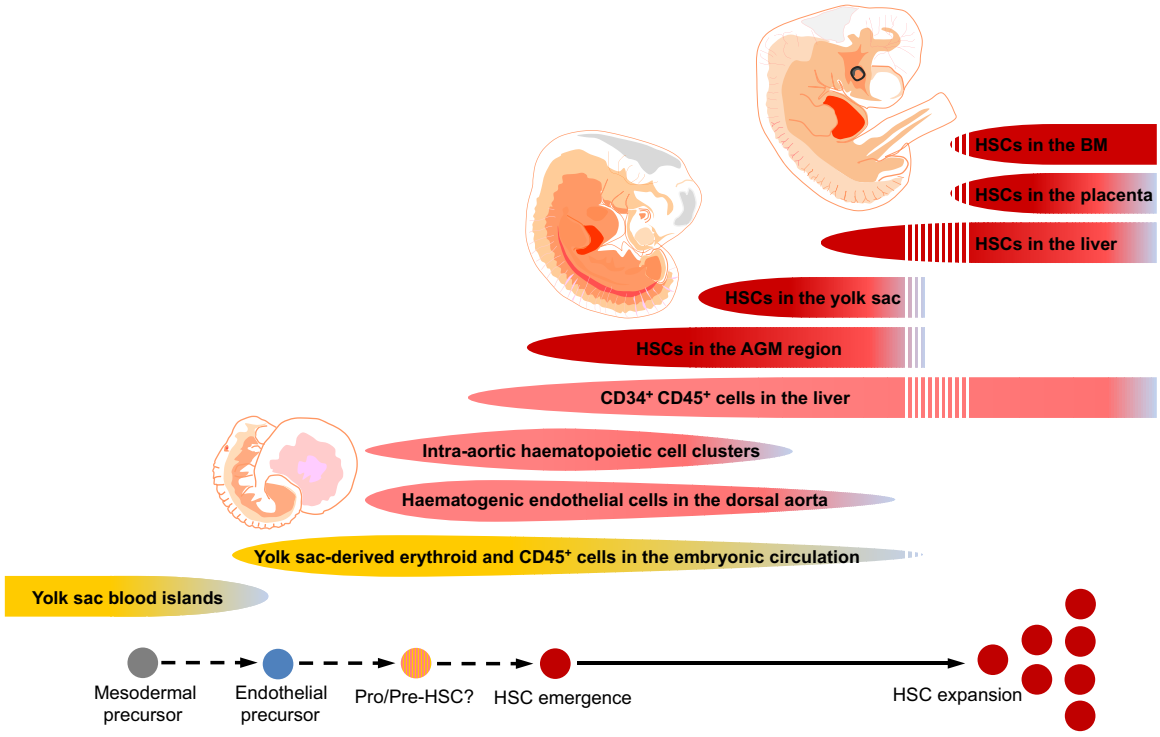


Figure 3: the HSCs emergence during development. The first HSC migrate from the AGM to other sites in the later stages of embryo development. Readapted from Ivanovs et al 2017[44].

1.1.2. The adult haematopoiesis

The haematopoietic system is the only liquid tissue found in complex animals. This peculiarity has attracted much attention from the biologists, and the understanding of the organization of this unique system has been a classic theme of stem cell research[45]. It is well rooted in the idea itself of stem cell a certain hierarchy, with the stem cell at the top and the differentiated cellular progeny at the bottom. The transition between these steps is unidirectional and top-bottom, reinforcing the centrality of stem cells in tissue organization. The basis for the understanding of the haematopoietic system could be traced back to the first experiments of bone marrow transplantation and the definition of a colony-forming unit: the clonal progeny of a single cell with a differentiation potential and the capacities of proliferation and self-renewal[46]. Another major contribution towards the definition of a robust model of the haematopoiesis has been the diffusion of the Fluorescence Activated Cell Sorting technology (FACS), as cell separation tool: cells are sorted accordingly to the expression of one or more surface proteins. This technology allowed Irving Weissman and colleagues, in a landmark publication in 1987, to isolate different haematopoietic populations from the mouse bone marrow, and test their functional potential in terms of colony formation and recipient reconstitution of the haematopoietic system[47]. This work paved the way towards the identification of different progenitor populations, each one progressively reduced in its ability to self-renew and/or differentiate. For example, a so-called common myeloid progenitor (CMP) is a cell able to differentiate only in myeloid cells, while a lymphoid progenitor would be specific for the lymphoid lineage only[48]. The consequent model has been widely accepted by the stem cell research community and consisted in the tree-like representation of the entire haematopoietic system: the HSCs at the top and the progressive branching of restricted progenitors, all the way down to the committed cells, at the base of the pyramid (Figure 4). The restriction is generally limited to two possibilities, a simplification that is functional to the model conceptualization and certainly true for some progenitors. The phenotypic definition of all these subsets of progenitors generated several identification codes, related to the expression of specific surface markers: an example of such identification scheme is the homogeneous expression of the members of the SLAM family[49], recognized as a valid way to discriminate between HSCs and progenitors, albeit not in a complete fashion.

All those definitions are strictly dependent on a single technical approach: the transplantation of a haematopoietic progenitor into a sublethally or otherwise conditioned bone marrow, in order to evaluate its capacity to regenerate the haematopoietic system. Two parameters are crucial for the validation of the engraftment: the presence, in the peripheral blood and the bone marrow, of all the haematopoietic lineages and the temporal stability of the transplant. Finally, it is common to

evaluate the reconstitution also in a recursive fashion, with serial engraftment experiments under a longer timeframe.

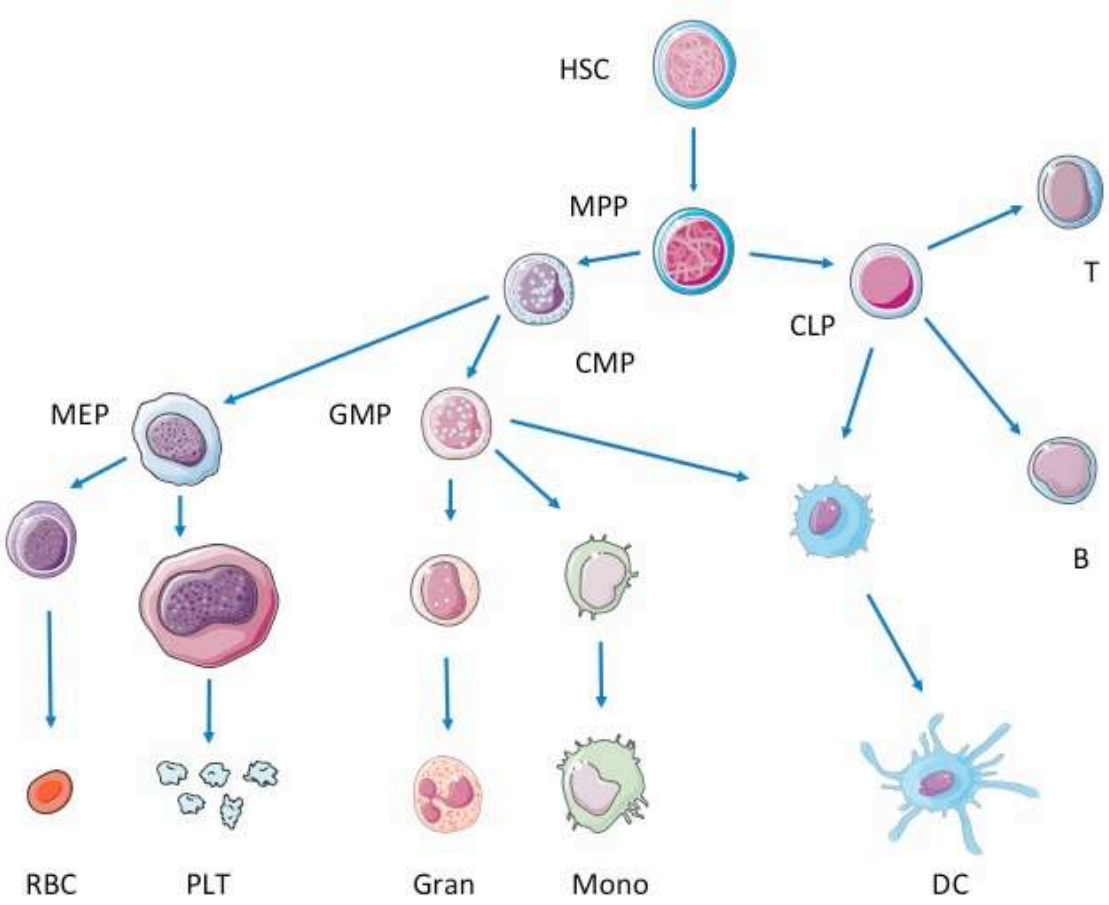


Figure 4: the haematopoietic differentiation tree in the early 2000s, largely based on the landmark paper from the Weissman lab[47]. **HSC**: Haematopoietic Stem Cell; **MPP**: MultiPotent Progenitor; **CMP**: Common Myeloid Progenitor; **CLP**: Common Lymphoid Progenitor; **MEP**: Megakaryocytic-Erythroid Progenitor; **GMP**: Granulo-Monocytic Progenitor; **RBC**: Red Blood Cell; **PLT**: Platelet; **Gran**: Granulocytes; **Mono**: Monocytes; **DC**: Dendritic Cells; **T**: T-lymphocytes; **B**: B-lymphocytes

Based on the time extension of the engraftment, a relative heterogeneity has been described inside otherwise homogeneous populations (accordingly to their immunophenotype). Many reasons have been provided for this heterogeneity, with the contrast between quiescent and actively cycling HSCs as one of the most studied. This dichotomy has a particular impact on the stemness attribute: phenotypically defined HSCs that are actively cycling resolve in shorter and less-robust engraftment, and only a very limited number of cell divisions allows the HSC to retain their stemness[50,51]. To arrange these differences, the distinction between long-term HSCs, short-term HSCs and multipotent progenitors has been introduced in the original model, formalizing a first heterogeneity in the HSC compartment in terms of self-renewal properties. In a similar way, not all HSCs possess the same multi-lineage differentiation potential, but most of them exhibit a bias

towards the myeloid lineage, progressively increasing along ageing[52,53]. The teams of John Dick and Sten Jacobsen proved convincingly the existence of lineage specific biases, describing an earlier branching of the progenitor compartment into a purely lymphoid progenitor and a lymphomyeloid progenitor[54,55]. Finally, several teams described a group of HSCs that is preferentially differentiating directly into the megakaryocytic (MK) lineage and is potentially deployed when higher levels of platelets are required, for example during inflammation[56-58]. This last result has attracted much attention, because this MK-biased HSCs retain their potential to reconstitute the full haematopoietic system in a transplantation setting, but they are essentially able to bypass the predicted differentiation pathway. The observed strong relationship between HSCs and MKs seems to be quite exclusive and it has not been formally described for other haematopoietic lineages[59]. Considered all this recent results, a new updated model describe a small subset of immunophenotypical HSCs as "true" stem cells, with the potential to reconstitute equally all the different lineages following transplantation. All the other HSCs subpopulations, while retaining most of the stem cell properties, show gene expression patterns that can be associated to the cell progeny toward which they are biased[60,61], therefore supporting the phenotypic observation (Figure 5).

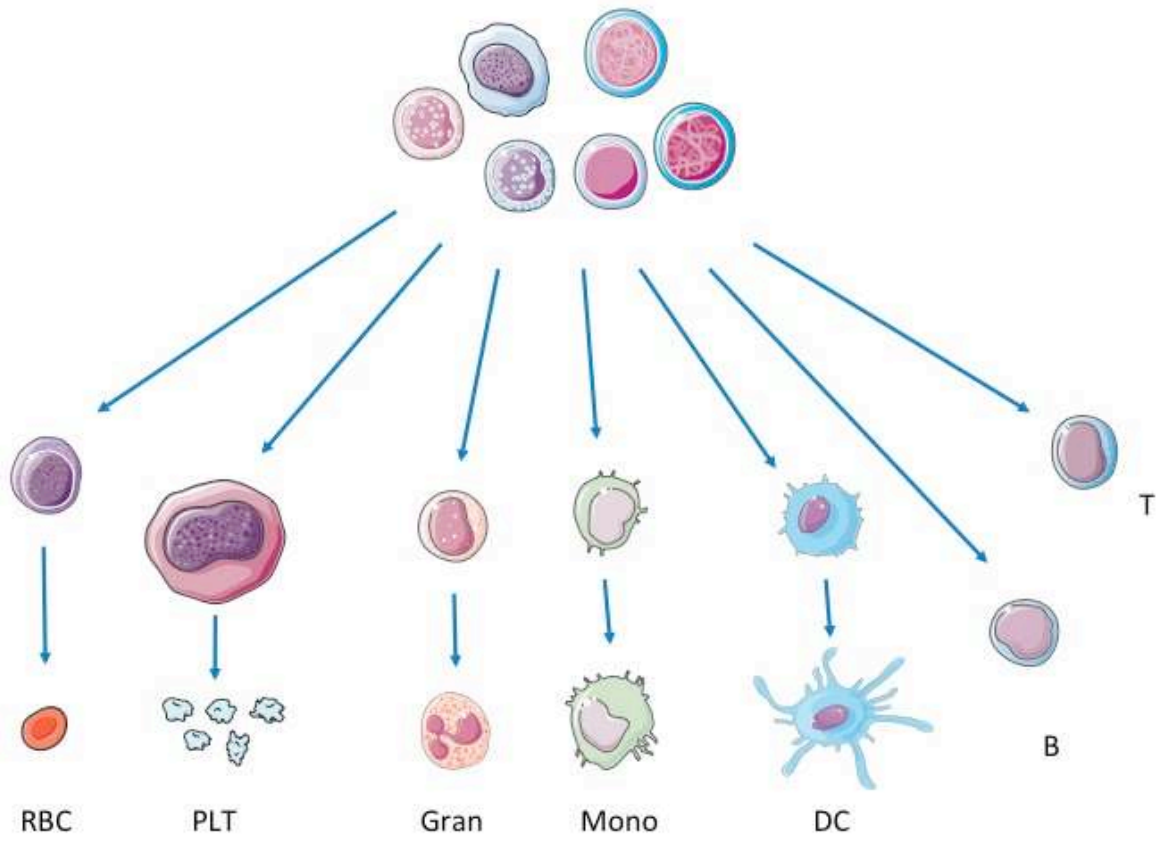


Figure 5: the revised haematopoietic differentiation hierarchy.

The definition of a model that takes into account the reported heterogeneity required a deep understanding of the molecular details responsible for cell specification. The classic bifurcation between two cell fates is largely supported by the strong cross-antagonism observed for certain transcription factors (TFs), like GATA1 and PU.1 (also called SPI1). Those genes are master regulators of different myeloid programs, and key blocks of their respective regulatory network. Both factors are positively regulating themselves and actively repressing each other. Therefore, at the progenitor level, those two TFs are in a dynamic equilibrium, where none of them is prevailing. In this way the cell retains the potential to lean towards the myelo-monocytic lineage or the erythro-megakaryocytic fate. When the balance shifts toward one of the two TF, for example GATA1, a progressive reduction of the levels of PU.1 should be expected, together with the increased expression of transcripts positively regulated by GATA1. The sum of this two type of events dictates effectively a higher probability for the progenitor to commit towards the erythro-megakaryocytic lineages (Figure 6).

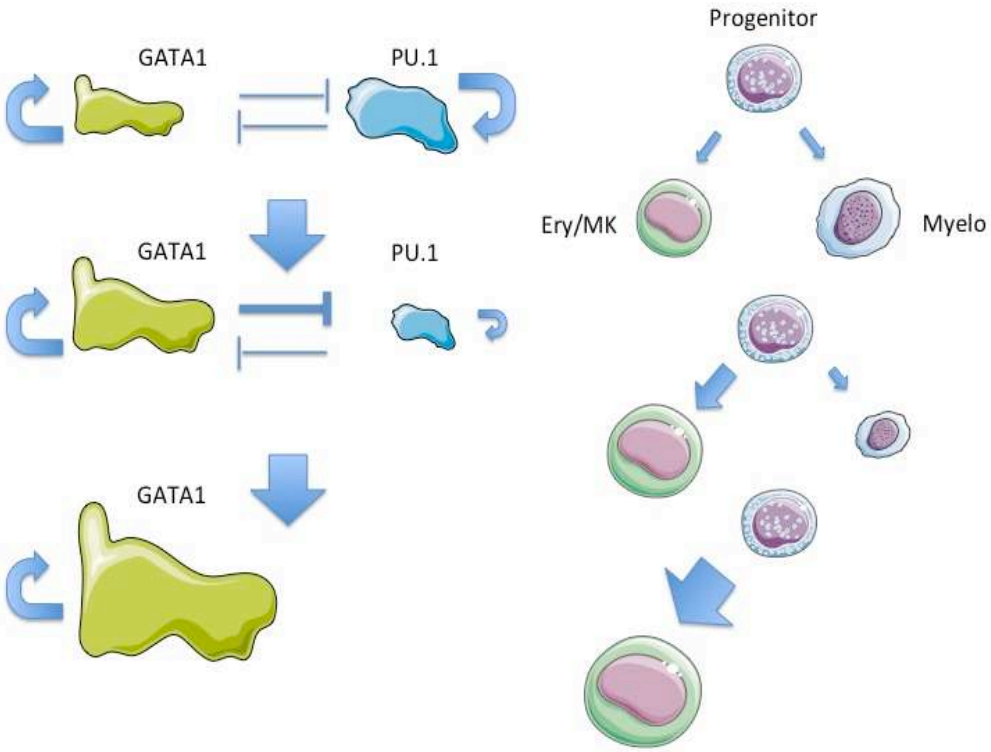


Figure 6: the cross antagonism between GATA1 and PU.1 and its phenotypic consequences on haematopoietic differentiation.

The cross antagonism fits well with the idea of a "choice between two options", but it is ill-equipped to justify the observed heterogeneity in the differentiation potential, where the same cell could be required to commit towards more than two cell fates. In the same way the FACS technology allowed the definition of the tree-like model, the diffusion of techniques like single cell

RNA sequencing allowed the definition of a whole new level of accuracy and the description of the transcriptional profile of single HSCs. This technological advance is one of the main contributor of the today definition of a third model: the stepping-stone of this new model is the observed phenotypical heterogeneity, and offer a shift from a step-wise, discrete vision of the haematopoiesis to one that is more dynamic and continuous[62]. Several observations point in this direction, thanks to the development of experimental tools able to assay the cell fate of HSCs without recurring to the transplantation[63,64]. In this model, most multipotent progenitors are primed towards a specific cell fate and, at the steady state, they generally commit towards it. If the organism requires a specific cell type, for example because of inflammation or infection, those primed cells still retain the ability to differentiate into other haematopoietic lineages than the one they are primed for. Another event that triggers this multipotent behavior is the transplantation: that explains why cells that have generally been associated to multipotency (because of the assay used for their characterization) are actually way more restricted in their differentiation potential[65,66]. The molecular landscape of these progenitors must be assayed in more stringent ways, but it could be postulated that they form a continuum of primed cells, at the top of a Waddington-like landscape: they oscillate towards a certain cell fate, but the requirements to switch towards a different one are fairly low, therefore in case of necessity they can transit to a different fate. This is experimentally represented (for now, but it could be not the only possibility) by the stress imposed by the transplantation, and the consequent repopulation of the recipient bone marrow. Nevertheless, at the steady state the cells are progressively descending towards a committed fate, and the transition to a different fate is consequently less likely. When the transition is too pronounced, the cells are irreversibly bound to that cell fate. This model is gaining traction and it describes the haematopoietic stem cell in terms of transitory state, instead of a discrete and isolate cell type[67] (Figure 7).

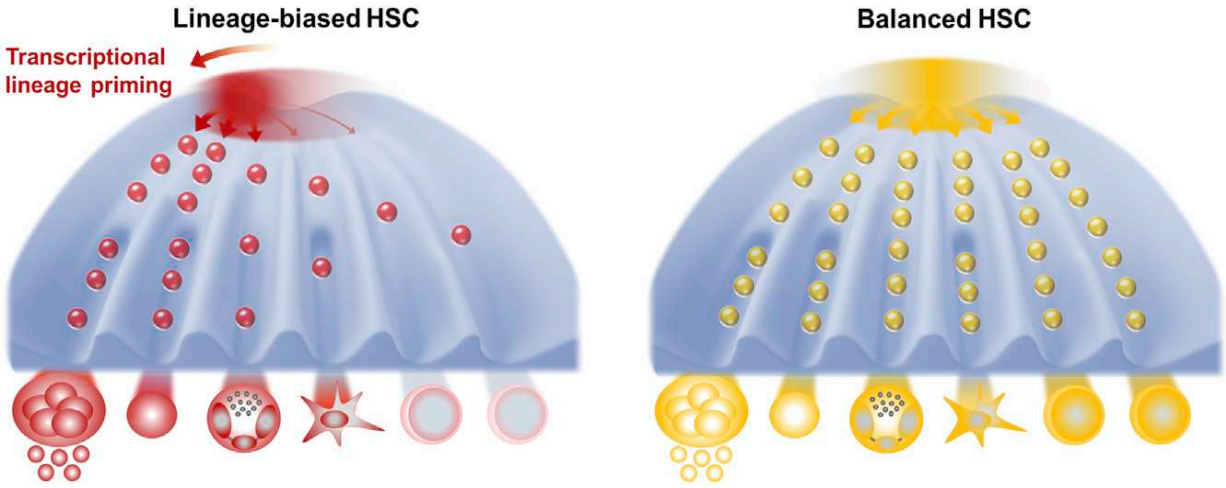


Figure 7: the Waddington-like haematopoietic tree, proposed by Haas, Trumpp and Milsom 2018[(62)]. This model describes the observed lineage-bias in HSC in probabilistic terms.

Most of the studies on the haematopoietic hierarchy have been performed in mouse models, therefore the question of the translatability of some aspects to the human haematopoiesis is very relevant. Taking advantage of the first clinical trials involving the lentiviral correction of a genetic disease, Aiuti and colleagues demonstrated a certain similarity between human and mouse, for what concerns the clonal dynamics of the engrafted populations. They were able to describe two waves of bone marrow reconstitution, and they explain this observation via the distinction between multipotent progenitors and long term HSCs, in order to justify the temporal shift between the two waves[68]. Nevertheless, the transplantation and the follow-up (three years) could be a confounding factor in the interpretation of the data. In this regard, the different longevities of human and mice constitute a major obstacle and must be taken into account when the clonal dynamics are analysed: it is reasonable to imagine that the human HSCs must be efficient for decades, compared to the short lifespan of a mouse; therefore, it is plausible to imagine major differences in quiescence and relative contributions to the steady-state haematopoiesis.

The idea of a highly heterogeneous haematopoiesis has been translated also to the developmental framework[69], although more work is required to define how much of this primitive heterogeneity is retained in the adult haematopoiesis. Other major questions refer to the mechanism behind the generation of the heterogeneity: is it acquired during the transition from fetal to adult life? Is it cell extrinsic or strictly cell intrinsic? Finally, the new theoretical framework is going to radically influence the conception of malignant haematopoiesis: the development of clonal haematopoiesis is possibly one of the more interesting test-bed of the continuous-like model, as otherwise healthy individuals are dependent on very few stem cells, that are forced to commit towards more fates than the ones they would do in a polyclonal setup[70]. Perhaps, it is the required increased output that induces a cellular stress, so creates the basis for the evolution of a true malignancy.

1.2. The megakaryopoiesis

1.2.1. Differentiation and maturation

The name megakaryocyte (MK) derives from the Greek and can be loosely translated in "Large Nucleus Cell". Indeed, the MK is one of the largest cells in the human body. It was first observed by James Homer Wright in 1906, who also made the observation that the MK is the cell of origin of platelets, discovered more than twenty years before by Giulio Bizzozero. The differentiation pathway from the HSC to the MK is called megakaryopoiesis (MKpoiesis) and could be broadly divided in two phases: an initial step of cellular commitment and moderate cell division (MK differentiation) and a latter, highly specific, stage of maturation (MK maturation). This entire process leads a single cell to the production of several thousand of platelets, as the human body requires an astounding number of these small anucleated fragments for an optimal haemostasis: between 1.5 and 4.5 x 10^8/mL. Combined with the short half-life of each platelet (around 10 days), this means that the human body must produce around 10^{11} platelets each day. Haemostasis is certainly the most important physiological process involving platelets, but it is not the only one: they have been described as central in inflammation, immune response and angiogenesis, not to mention their role in pathological phenomena like cancer metastasis and atherosclerosis.

The MK differentiation is not particularly dissimilar from other haematopoietic lineages, as it depends on the delicate interplay between intrinsic cues and extrinsic stimuli. In this phase, the progenitors are progressively rewiring their transcriptional landscape and committing to the MK lineage. The most evident consequence is the expression of certain markers, notably CD41 or integrin alpha 2b (*ITGA2B*): this surface protein is increasingly expressed along the MK differentiation and could be used in a functional assay for the identification of progenitor cells with a megakaryocytic differentiation potential. Those types of experiments must be performed on specific substrates supporting the growth and staining of the deriving colonies, as there are no evident morphological features that are specific of a megakaryocytic colony (CFU-MK). Due to the short timeframe of the MK differentiation, compared to other myeloid lineages, it is difficult to delineate a step-wise differentiation mechanism like the one observed in the erythropoiesis: the current view is that the different MK progenitors (MK-Ps) could be classified accordingly to the size of the colony and the expression of CD41. Beyond the stage of CFU-MK, the progenitor is irreversibly committed toward the MK lineage, and enters the second part of the MK differentiation, that consists essentially in the acquisition of a collection of highly specific integrins and surface markers. While the former stage is not distinguishable in cytology, in this phase of the MK differentiation it is possible to observe some early morphological features that are a trademark of

the MKs, notably the first α-granules and an initial membrane invagination, the precursor of the demarcation membrane system (DMS). In human MKpoiesis, a combination of just three surface markers allows a quite rigorous definition of several homogeneous populations, in terms of commitment and progression along the differentiation. The first marker is the transmembrane (TM) phosphoglycoprotein CD34, traditionally associated to HSCs, although it is worth noticing that could be expressed by several cell types outside the haematopoietic system. This marker is progressively lost during the MKpoiesis and is absent in mature MKs, so it is a reliable tool to evaluate the early commitment stage. The second marker, already mentioned, is CD41 (also called GPIIb), found in association with CD61 (GPIIIa). The deriving complex is also called integrin $\alpha_{IIb}\beta_3$, and constitutes a multipurpose receptor for extracellular proteins, with a particular affinity for the fibrinogen. Conversely to CD34, its expression increases during MKpoiesis, with the mature MKs expressing the highest levels of this surface marker. The third and final marker is CD42 (GPIb-GPIX-GPV), the receptor for the von Willebrand Factor (vWF). CD42 is expressed at the early onset of MK maturation and it is univocally expressed by committed MKs (Figure 8).

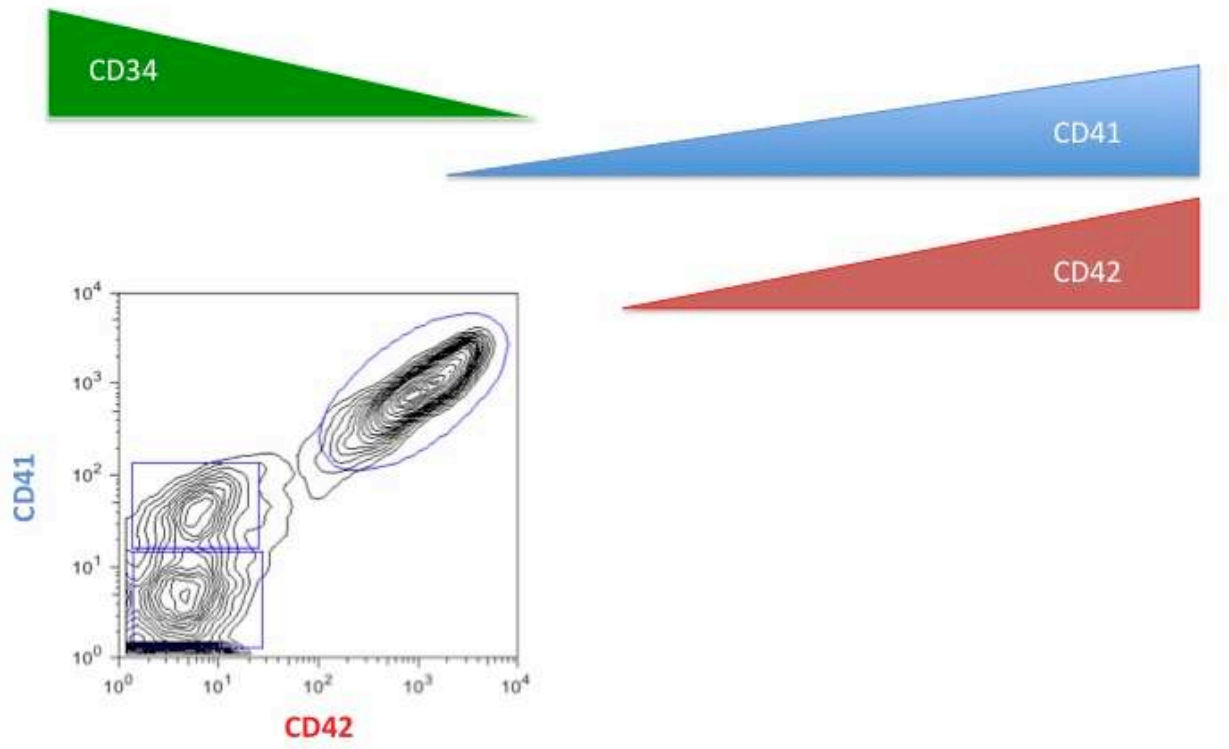


Figure 8: the MK differentiation analysis via flow cytometry. The displayed colours identify the expressed marker: CD34 in green, CD41 in blue and CD42 in red. Personal data.

The expression of CD42 indicates the entry into the last phase of MKpoiesis, called MK maturation. This is one of the most original differentiation processes in the human body, and consists in the controlled polyploidization of the MK, followed by the production of several cytoplasmic protrusion called proplatelets (PPTs) and the production of many proteins for platelet functionality. During this crucial step, the MK is migrating towards the sinusoidal vessels from the inner core of the bone marrow. The combination of cell intrinsic protrusion of the PPTs into the vessels and the flow shear stress of the blood circulation will allow the final step of the MKpoiesis, that is the progressive dispatch of the PPTs into the circulation. Those two steps are intertwined, but not mutually exclusive, as even diploid MKs (called micromegakaryocytes) can successfully produce platelets, albeit in very low number. Nevertheless, to produce a sufficient amount of platelets it is necessary to consider two parameters, notably the number and the size of MKs. As the platelets in mammals are anucleated, in contrast to other vertebrates, the increase in the MK cell size is a rapid and efficient strategy to produce large amounts of cell fragments in a short amount of time, compared to the simple cell proliferation. The size increase proceed parallel to an increase in the DNA content, as a higher number of alleles increase the mRNA transcription and, consequently, the protein synthesis necessary to sustain such an intense metabolic effort[71]. Two known mechanisms allow the synthesis of new DNA without cell division, and are the endomitosis and the endocycles. They differ only by the mitosis entry, present in the former case and absent in the latter (only rounds of $G_1/S/G_2$ phases). The absence of cytokinesis, feature specific for the endomitosis, is strictly dependent from the interplay between the actomyosin ring and the activation status of the small GTPase RhoA: in normal condition, active RhoA is responsible for the activation of the myosin machinery at the cleavage furrow, via recruitment of Myosin IIA (*MYH9*). Consequently, the progressive restriction of the contractile ring is induced, thus the effective cytokinesis (Figure 9). This holds true even for MK progenitors and diploid MKs, until the beginning of the endomitosis, where the activity of RhoA is suddenly reduced. One of the proposed mechanisms for the switch-off is the down-regulation of two guanine exchange factors (GEFs), GEF-H1 and ECT2, necessary for the activity of RhoA[72]. The other major non-exclusive mechanism responsible of the abortive cytokinesis is due to the absence of Myosin IIB (MYH10) at the cleavage furrow, due to the transcriptional silencing mediated by the TFs *RUNX1* and *FLI1*[73]. For the others cell cycle checkpoints, the MKs rely heavily on cyclin D1 and cyclin D3 among the other cyclins, both abundantly expressed in pre- and post-endomitotic MKs. The dependent interplay between cyclin-dependent kinases (CDKs) and kinases inhibitors (CDKNs) is extremely delicate and rich of redundancies: as an example, the ablation of those CDKs can rescue the deficiency in CDC20, an ubiquitin-ligase member involved in the degradation of cyclin B1, an activator of CDK1[76]

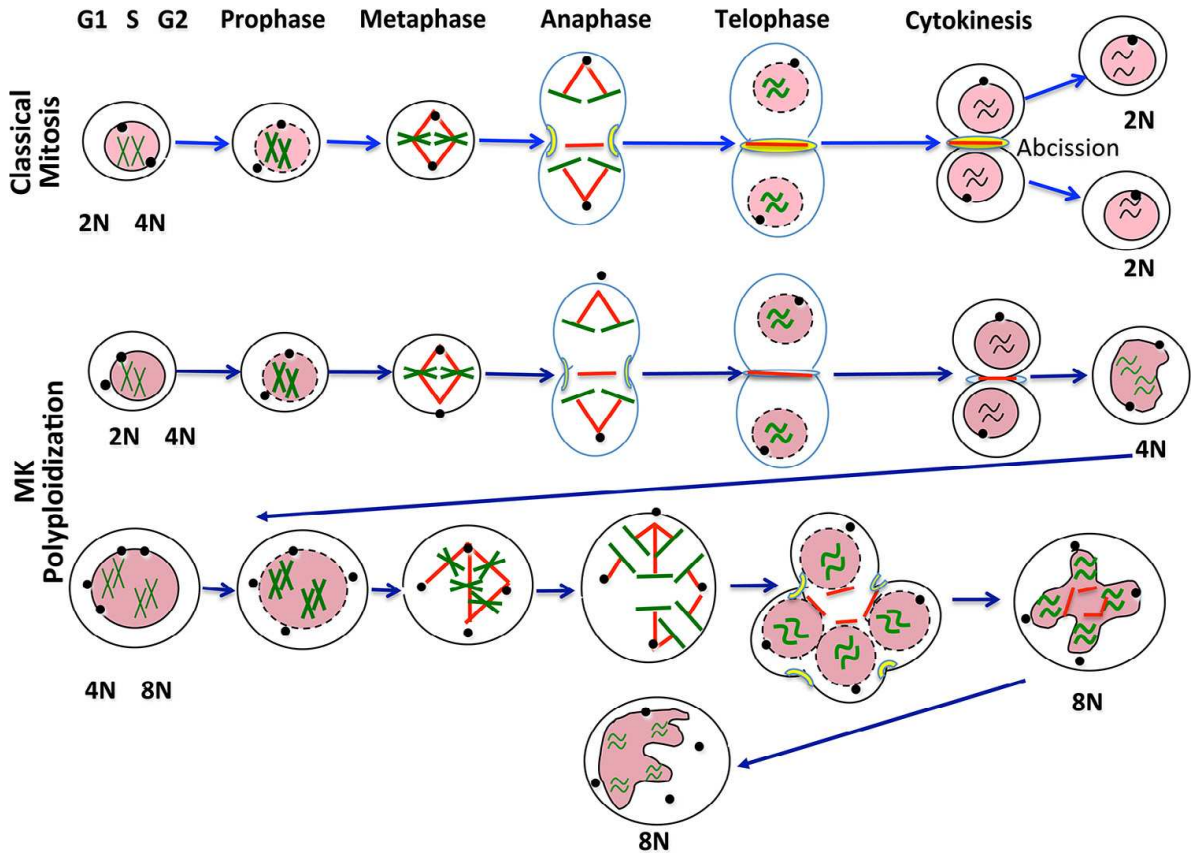


Figure 9: the endomitosis compared to the normal mitosis. It is possible, albeit less understood, for a MK 4n to perform cytokinesis (so a true mitotic cycle), therefore the early endomitosis do not exclude a mitosis[74]. Readapted from Mazzi and colleagues[75].

Nevertheless, it is worth noticing that more than 80% of polyploid MKs is the consequence of endomitosis instead of endocycles, raising the possibility that endomitosis allows a more efficient or safer polyploidization: for example the nucleus is assuming a multilobulated shape, due to the activity of the different mitotic spindles present in an endomitotic MK, while it is unclear if it is also the case for a MK undergoing endocycles. The simple ploidy analysis (typically performed by flow cytometry) cannot fully elucidate the differences between those two mechanisms, so it is tempting to speculate about the endocycles as a first step towards a defective MK maturation, in presence of specific genetic abnormalities.

Once the MK reaches its polyploid state, it enters a final, post-mitotic, phase, where the entire cellular machinery is dedicated to the large scale biogenesis of granules that stock several proteins crucial for haemostasis, angiogenesis and other platelets functions. At least three categories of granules have been described, the α-granules, the dense granules (or δ-granules) and the lysosomes, all differents in their content and origin. The biogenesis of the α-granules is the consequence of the

combination between endocytosis and secretion, at the level of the multi-vesicular bodies (MVBs), that acts as a molecular hub for the dispatch of the different proteins: while some are typically biosynthetic products of the MK itself (VEGF, P-selectin and PF4 among the others), others proteins are produced outside the MK and internalized by their respective cognate receptor. Classic examples of this pathway are fibrinogen and vWF, both internalized and packed in the α-granules together with their receptors CD41 and CD42. On the other side, δ-granules are only produced by the secretory pathway and share many biosynthetic aspects with the lysosomes; they contain mostly serotonin, ADP and other small molecules essential for haemostasis.

Simultaneously to the granules production, the MK is also undergoing a massive membrane rearrangement, with the development of a mature DMS, originating from the membrane invagination observed before the beginning of the polyploidization. Essentially, the DMS is a folded structure that provides the membranes necessary for the formation of the PPTs branches. The system is in continuity with the extracellular space, thus it is possible to observe small leukocytes trapped inside a MK, even if they are not really phagocyted by the MK (emperipolesis).

The formed DMS is progressively uncoiled by the microtubules, into the classic branched shape typical of a mature MK. At this point, the correct differentiation is strictly dependent from two machineries: the assembly of spectrin/actin/myosin cytoskeleton, that must be able to convey the microtubule forces towards the PPT fragments, generated by tubulin and the respective motor dynein. PPT elongation is crucial for the correct production of platelets in the blood circulation, as the MK is extending the fragment directly inside the sinusoidal vessel, where the shear stress detach the fragment, interrupting the membrane continuity with the rest of the cell. Further, the PPT fragment again into preplatelet, then into a barbell-like shape containing two individual platelets. Some of the molecular actors encountered during the polyploidization step are central again in proplatelet formation and elongation, like MYH9 and RhoA, together with effectors like the small GTPases CDC42 and Rac1. The case of the small GTPases is particularly interesting, as they are central in this maturation steps: RhoA activity must be silenced for a successful PPT formation, as it directly regulates three effectors that are crucial for the stability and correct function of the actomyosin cytoskeleton. The first one is the RhoA-associated protein kinase 1/2 (ROCK1/2), able to regulate the activity of the myosin regulatory light chain 2 (MLC2), thus the cytoskeleton contractility. A classic proxy for the activity of ROCK is the formation of stress fibers, and their excessive formation could affect the capacity of the MK to unwind the DMS, in order to form the PPT branches[77]. The second RhoA effector is DIAPH1, a formin involved in the stabilization of the microtubules and the actin filaments, and directly activated by the active form of RhoA[78]. As in the previous case, an abnormal stabilization of the F-actin polymerization could alter drastically the capacity of the microtubules to elongate the PPT branches[79]. The third effector of RhoA is LIM

kinase (LIMK),: acting together with cofilin, is responsible of the turnover of the actin filaments. In this case, LIMK activation is the consequence of a concerted regulation by CDC42 and Rac1, in addition to RhoA[80], and a striking example of how those two small GTPases are balancing the action of RhoA, acting in an opposite way and promoting a positive regulation of the PPT formation[81].

At the end of the proplatelet release, the MK role is completed and it is consequently eliminated by the bone marrow resident macrophages. Although there are no evidences that apoptosis is implicated directly in the release of PPTs, it is possible that some proteins implicated in the classic apoptosis are actively promoting the platelet release, especially in situation of acute need[82].

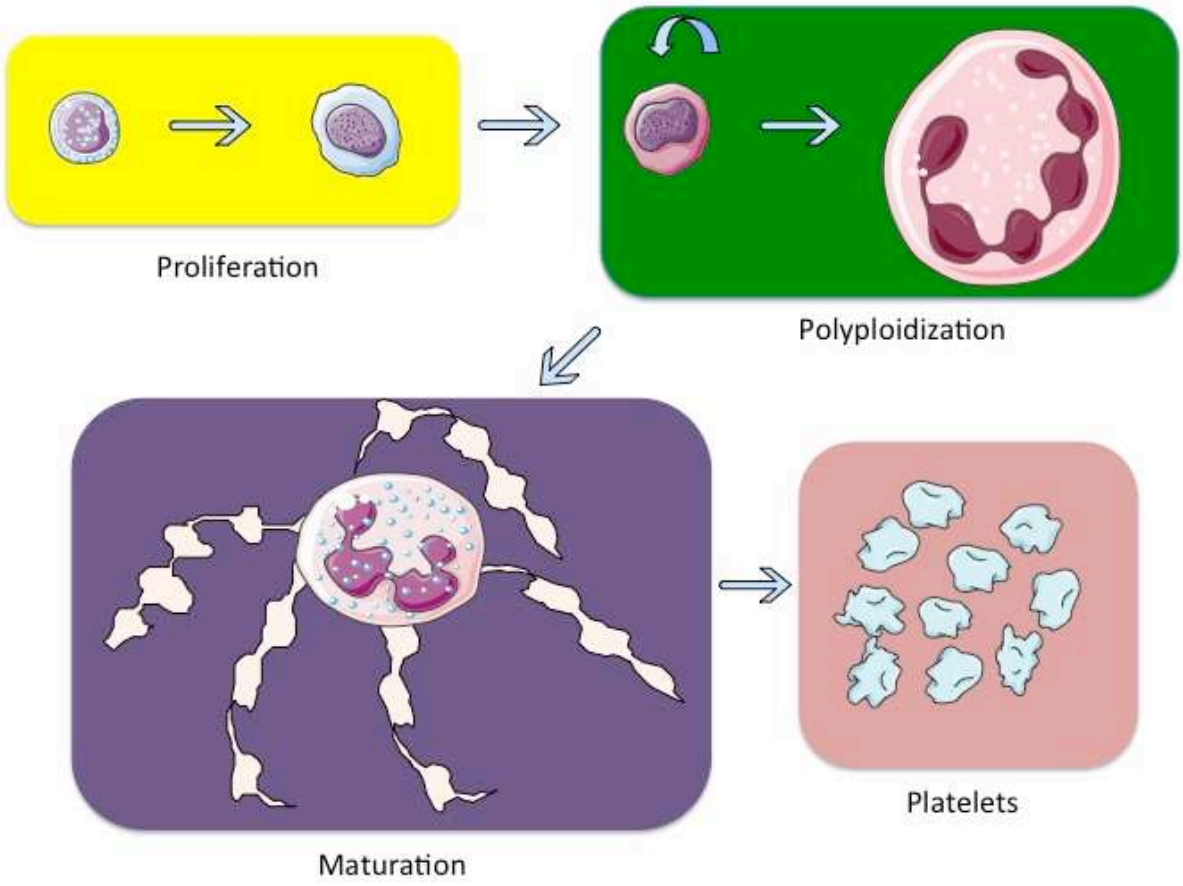


Figure 10: the megakaryopoiesis. In yellow the first step of MK differentiation (proliferation of MKs), in green the polyploidization step via endomitosis, in purple the proplatelet formation and finally the platelet release in pink.

1.2.2. Transcriptional regulation of megakaryopoiesis

The role of the transcriptional regulation during the haematopoietic differentiation is a classic example of the impact of transcription on lineage commitment and MKpoiesis is no exception: all the complex steps described in the former section are evidently the results of a strict regulation mediated by different TFs and others epigenetic actors. The central role of TFs in haematopoiesis and MKpoiesis can be related to the number of oncogenic lesions involving TFs, as they are essential in the definition of a cellular identity. Indeed, the first description of almost all the master TFs date to the analysis of leukemic chromosomal translocations or insertional mutagenesis[83]. It is peculiar how the activation/repression activity is not an intrinsic property of each factor, but it is the consequence of the interactions with other molecular partners. Typically, TFs activity is dependent from the binding to a genomic sequence, via specific domains that are able to recognize sequences more or less conserved across the genome. Therefore, the TF is usually containing some protein domains able to modulate the transcriptional machinery, in a positive or negative fashion. The consequences of a TF activity are essentially two: affecting the capacity of a cell to progress along the differentiation path, notably by increasing the expression of genes functionally involved in the MKpoiesis, and blocking the progression toward a different cell lineage, in particular during the early phases of MK differentiation.

As already mentioned, most of the TFs act as part of larger complexes involving several proteins[84] (Figure 11); therefore, the stoichiometry of each factor is crucial in ensuring the right activity of the entire complex: for example, GATA1 forms an axis with FOG-1 (Friend Of GATA1) to activate or repress its targets, and the lack of FOG1 phenocopies the lack of GATA1 for the defects in the erythroid lineage. But the lack of FOG1 leads also to the absence of any MK progenitor, while the lack of GATA1 expand greatly the pool of immature MKs. The explanation lies in the role of GATA2, a TF that is able to "fill-in" in the absence of GATA1, at the progenitor level, therefore partially rescuing the induced defect[85]. Another crucial partner of GATA1 is FLI1, together responsible of the expression of a subset of genes highly expressed along MKpoiesis and completely silenced in erythropoiesis. Those genes are expressed at very low level at the progenitor state, suggesting a development priming of the loci by a complex of at least seven proteins (heptad): RUNX1, FLI1, ERG, LYL1, TAL1, LMO2 and GATA2. The displacement of GATA2 by GATA1 and the recruitment of additional copies of FLI1 convert this priming into a full-blown expression, condition associated to the MK fate. Conversely, if no additional FLI1 copies are recruited, for example because of the antagonistic effect of an erythroid specific TF (KLF1), the developmental priming became a repression of the MK-specific subset[86]. Another intriguing possibility for the modulation of the cell identity is the expression of different isoforms of the same TF, therefore

changing the relative functional properties. This is the case of RUNX1 that could be expressed in different isoforms accordingly to the ontogeny and the differentiation, so acting differentially in the same transcriptional context[87]. Another example, from the pathological haematopoiesis, is the expression of a shorter isoform of GATA1, named GATA1 short (GATA1s), in presence of loss-of-function mutations in the exon 1. Those mutations are well described in the case of Down Syndrome Acute Megakaryoblastic Leukemia (DS-AMKL), an acute leukemia that affects specifically individual carrying a trisomy 21 and that is always associated with a GATA1s mutation in the founding clone[88]. This shorter isoform lacks the N-terminal domain, but the functional consequences are still unclear and matter of investigation.

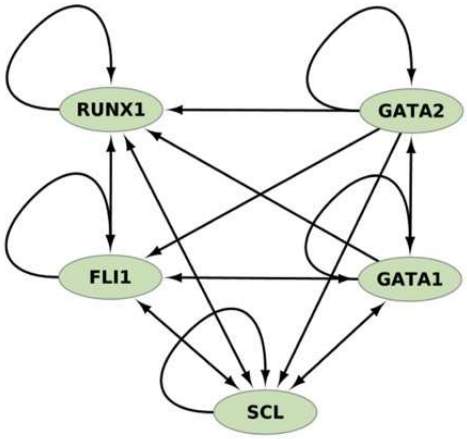


Figure 11: an example of the transcriptional network operating in MKs, presenting the different feedback and feedforward loops. Readapted from Tijssen and colleagues[84].

The TFs induce the expression of their targets by recruitment of the transcription machinery based on the RNA polymerases, but this is not the only functional consequence of their interaction with the genome: a TF can induce indirectly an extensive remodelling of the chromatin, by the interaction with different types of epigenetic regulators. The specificity of those molecular complexes is the consequence of the pioneering activity of the TF that are interacting with, as they are generally devoid of any capacity to recognize a specific DNA sequence. Between the many regulators, some of the most cited in the regulation of the MKpoiesis are the members of the histone deacetylases (HDACs) class: those enzymes remove acetyl groups from the histone tails, removing negative charges that reduce the affinity of the nucleosomes for the DNA. Therefore, the DNA tends to be condensed and more compact, with the consequence of a reduced accessibility to external factors. In an opposite way, histone acetyltransferases (HATs) mediate the adding of acetyl group to the histone tails, favouring an open chromatin state. The relevance for the MKpoiesis of the HDAC class members is due to the well-documented thrombocytopenia associated to the clinical use of HDAC-specific chemical inhibitors in oncology. HDAC1 and HDAC2, while

included in the NuRD complex, interact directly with FOG1; conseuquently, their inhibition disrupt massively the early stages of MKpoiesis and leads to an increase in double-strand breaks formation[89]. Nevertheless, HDAC proteins can also modify other proteins than histones, notably cytoskeleton-associated proteins[90] and TFs, although this last intriguing link has not been proven definitely yet. The other major regulators of chromatin accessibility are the two Polycomb Repressor Complexes (PRCs), PRC1 and PRC2 (Figure 12): instead of adding or removing acetyl groups, those two complexes catalyze ubiquitination and methylation of lysine residues on specific histones. They can act in tandem or independently, but in both cases they reduce the accessibility to the underlying sequences, promoting a compact chromatin. The classic sequence of events consists in the initial trimethylation of the lysine 27 of the histone 3 (H3K27me3), by the catalytic subunit EZH2 of the PRC2. Successively, the PRC1 is recruited on the same region, via its domain CBX, and the catalytic subunit BMI1 introduces the mono-ubiquitination of the lysine 119 of the histone H2A, causing a more robust silencing of the underlying genomic sequence. This combined action converts a relatively weak silencing (the trimethylation) into a more stable chromatin modification, therefore highlighting the differences between those two complexes. Indeed, the H3K27me3 is more typically associated to a locus poised to be repressed than an effectively repressed one, and is not unusual to find the same region to be marked with another trimethylation mark, this time on the lysine 4 of the histone 3 (H3K4me3), typically catalyzed by another class of methyltransferases, the family of the KMT2 or MLL. The PRC1 recruitment can also be recruited to the chromatin in a PRC2-independent manner, by direct interaction with RUNX1 and the core binding factor β[91]. For the recruitment of PRC2 to the chromatin, the complex exploit the subunit EED, the only essential subunit together with SUZ12 and the catalytic domain; nevertheless, this minimal form of the complex displays a very poor interaction with the nucleosomes, and requires more subunits to effectively recognize its targets, for example JARID2 and AEBP2. Even more, JARID2 is also implicated in the crosstalk between PRC1 and PRC2, as it has been described to be able to recognize specifically the H2AK119u1 modification catalyzed by the PRC1[92], in order to ensure a stable silencing.

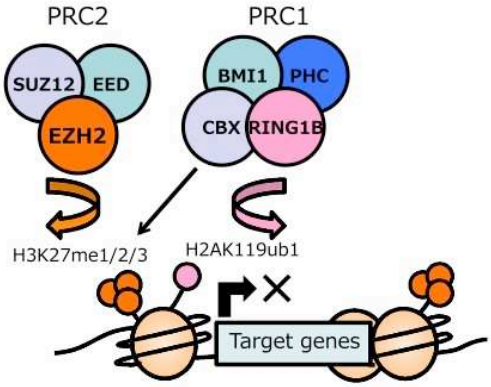


Figure 12: the action mechanisms of PRC1 and PRC2 complexes. Readapted from Sashida and Iwama[93].

Another route to modify, in a reversible fashion, the transcriptional properties of the genome is the DNA methylation, particularly of the cytosines residues of the CpG dinucleotides. Commonly associated to promoters, it is considered as a repressive mark. Two families of proteins are responsible of the establishment and maintenance of the methylation patterns: the DNA methyltransferases DNMTs and the methylcytosine dioxygenase of the Ten-Eleven Translocation family TETs, with the former adding the marks (5mC) and the latter removing them (5hmC, successfully converted in the unmodified cytosine via dehydroxylation). The two most important members of these two families for haematopoietic cells are DNMT3a/b and TET2, and they are frequently found mutated in acute leukemia. The absence of these two proteins affects in opposite ways the DNA methylation status (hypomethylation for *DNMT3*$^{-/-}$, hypermethylation for *TET2*$^{-/-}$), and they are associated to anomalies in the progenitor and stem cell compartments. In particular *TET2*$^{-/-}$ displays a differentiation blockage, coupled with an amplification of the progenitor compartment: scarce evidences are available on the link between *TET2*$^{-/-}$-dependent hypermethylation of multiple promoters and a differentiation block specific for MKs[94]. A study of the erythropoiesis in zebrafish suggests that the reduction in 5-hMC levels reduces the expression of lineage-specific TFs, impacting the ability of the HSCs to commit to the erythroid fate[95]. Similar effects to the TET2 loss of function has been observed in the case of isocytrate dehydrogenase (*IDH1/2*) mutations for acute leukemia, as this TCA enzyme, when mutated, produce a potent inhibitor of the alpha-ketoglutarate (αKG), O_2-dependent dioxygenases like TET2[96,97]. In a conceptually similar way, an enzyme involved in the metabolism of branched aminoacids (BCAT1) could deplete the levels of αKG when over-active, therefore phenocopying the effects of IDH1/2 and TET2 mutations: accordingly, *BCAT1* overexpression results in increased DNA methylation[98].

The delicate interplay of all the various epigenetic regulators is only one of the possible layers of epigenetic control of the chromatin. The 3D accessibility of the genome is crucial for the pioneering activity of the TF, as it leads to the remodelling of the chromatin via the epigenetic regulators activity. The genome accessibility is therefore crucial for the activation of the desired gene expression program or, viceversa, the repression of the undesired one. This is the role of the cohesins (Figure 13), a four-core complex ring-shaped that allows, together with the protein CTCF, the definition of a topological associating domain (TAD): the DNA sequences inside a TAD interact more frequently between themselves than with sequences outside the TAD. This results is achieved via a loop-extrusion mechanism: the cohesin ring allows the DNA to pass, forming a loop. As the CTCF proteins recognize and tightly bind the DNA, the moment when the complex CTCF-DNA reach the cohesin ring induce a block of the DNA extrusion, as the CTCF protein cannot pass

through the ring. It means that a fully formed loop contains two CTCF domains at the extremities, just outside the cohesin ring. When this mechanism is deregulated, for example because of mutations in one of the core subunits, the accessibility to specific DNA regions is altered, impacting the global transcriptomic activity of that genomic area[99].

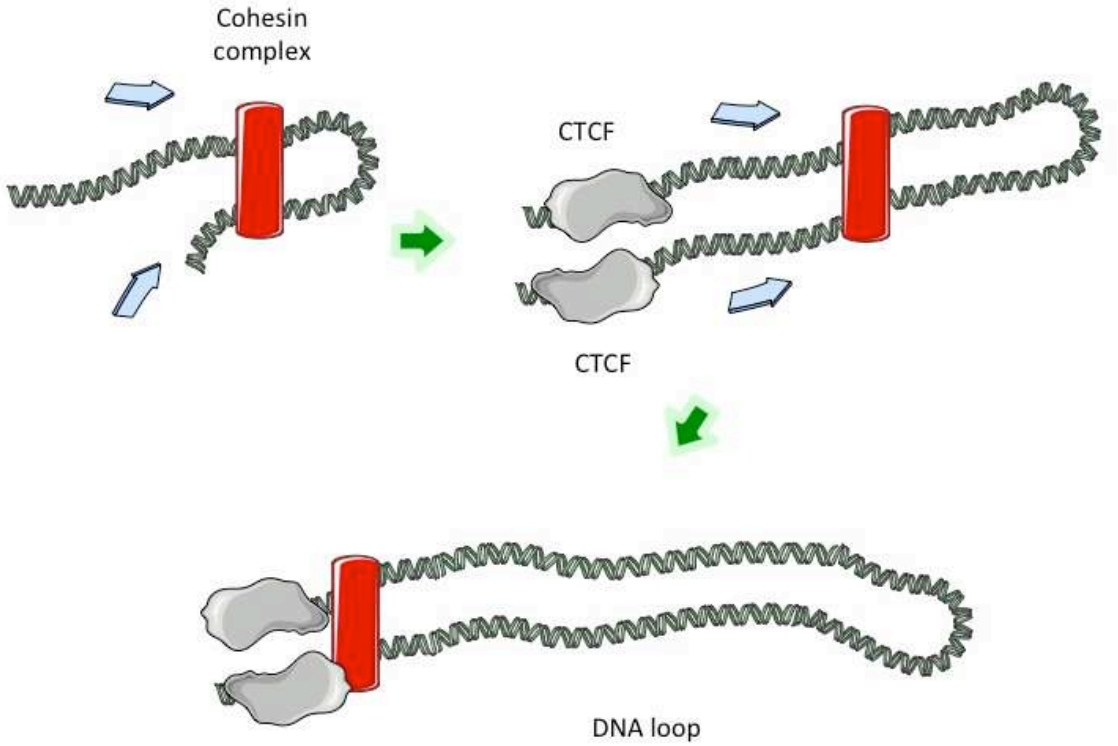


Figure 13: the cohesin complex activity. The creation of the DNA loop allows the definition of specific DNA regions that are exposed to gene regulation, therefore a spatial regulation of gene expression.

Another layer of epigenetic regulation, not directly related to the genomic activity, is ensured by the splicing factors, as they are frequently found mutated in the pathological haematopoiesis (myeloproliferative neoplasms and myelodysplastic syndromes), with *SRSF2* and *SF3B1* as the most mutated genes[100]. Another protein able to greatly affect the gene expression in MKs is RBM15: it is a RNA-binding protein, able to bind the intronic regions of the pre-mRNA of MK-associated genes like *GATA1*, *RUNX1*, *TAL1* and *MPL*. The binding of RBM15 allows the recruitment of SF3B1, linking the two proteins.

Finally, there is an entire layer of gene expression regulation that is mediated by different RNAs, notably long non-coding RNAs and microRNAs: both categories are relatively poorly understood, in particular the long non-coding RNAs, but they are quite crucial for the control of the cell identity, in particular at the progenitor level.

1.2.3. MPL-mediated signalling and its importance for megakaryopoiesis

The MKpoiesis and its regulation depends on the cardinal axis formed by the soluble factor thrombopoietin (TPO), the cognate receptor myeloproliferative leukemia (MPL, also called TPO receptor or TPOR) and the tyrosine kinase JAnus Kinase 2 (JAK2). Large parts of the processes previously described depend heavily on these three proteins and their equilibrium. The TPO receptor belongs to the family of the homodimeric type I cytokine receptors, the same of EPOR and G-CSFR. The three receptors are fundamental for the successful differentiation and maturation of three different myeloid lineages, respectively MKs (MPL), erythrocytes (EPOR) and granulocytes (G-CSFR) (Figure 14).

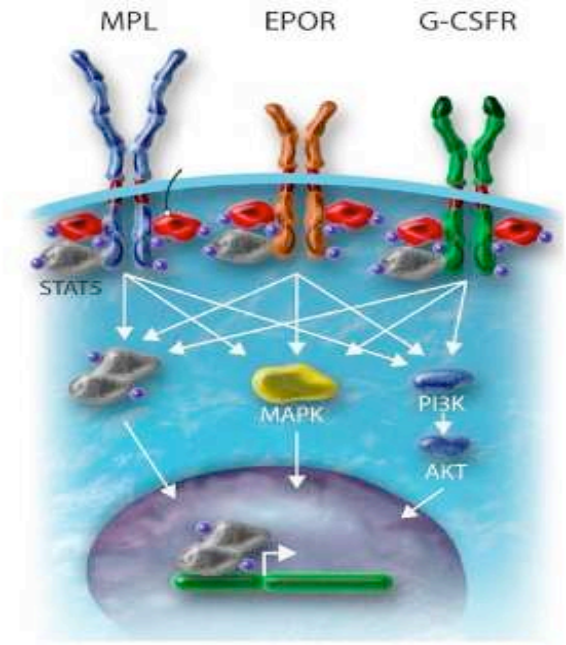


Figure 14: The three main receptors for myeloid cells diffentiation and survival. Readapted from Vainchenker and Kralovics[101].

Encoded by the gene *MPL*, the TPO receptor consists in a polypeptide of 634 aminoacids: 490 in the extracellular domain, 122 in the intracellular portion and a TM helix of 22 aminoacids. It is a heavily glycosylated protein, in particular on four extracellular asparagine residues, with a final molecular weight of approximately 90 kDa. As a homodimer, it interacts with a single molecule of TPO, in a highly exclusive and complementary way. Mutations in some crucial residues in the cytokine receptor homology module CRM1, one of the two modules constituting the extracellular domain, disrupt completely the TPO binding and lead to the complete lacks of MPL signalling and the absence of the MK lineage[102,103]. On the other side, the deletion of the entire CRM1 leads to a ligand-independent activation of the receptor[104]. Even in a more extreme way, it is possible to

activate the MPL receptor by the simple addition of a FLAG motif in the extracellular domain, thus treating the cells with an antibody anti-FLAG; this triggers a similar cascade as the one mediated by TPO binding[105]. These examples underlie the huge impact that the structural features of the receptor have on its functional effects.

As a member of the homodimeric type I receptor family, MPL lacks any intrinsic kinase activity, therefore it must rely on the presence of some sort of activator, in order to trigger the signalling cascade. This role is essentially played by the tyrosine kinase JAK2, with a non-negligible contribution from another member of the same family, TYK2[106]. JAK2 and MPL associate in the endoplasmic reticulum, in a non-covalent fashion and through a specific box of the intracellular portion of the receptor enriched in proline residues, called Box1[107]. This box interacts with the FERM domain of JAK2 and it is mandatory for the activation of the receptor. Nevertheless, a second interaction between the receptor and the kinase is necessary for the signalling cascade, and it is between the Box2 (MPL) and the Src Homology 2 domain (SH2)-like domain (JAK2). At the C-terminus of the JAK2 protein (the C-terminus), two major kinase domains are involved in the catalytic activity and its relative regulation: the more distal, called JH1, retains the active site of the enzyme, while the more proximal is almost completely devoid of any catalytic activity, but it acts as a negative regulator, keeping the kinase in an inactive state in absence of any stimulation, effectively preventing any self-activation. The mechanism involves the JH2-mediated phosphorylation of two residues, Ser^{523} and Tyr^{570}, in a sequential way; this cascade prevents the JH1 domain to assume an active conformation[108]. For this reason, alterations in the pseudokinase domain destabilize the normal JAK2 activity: the valine-to-phenylalanine substitution in the codon 617, also called JAK2V617F, is a classic example of it, and it has been described in the BCR-ABL1-negative MPNs[109-112]. The mutation gives rise to an hyperproliferative phenotype affecting the EPOR signalling (Polycytemia Vera or PV, mutation found in more than 90% of cases) or the MPL one (Essential Thrombocythemia or ET and Primary MyeloFibrosis or PMF, around 60% of cases), but it has never been described as implicated in the G-CSFR-dependent signalling[101]. While the V617F mutation induce a constitutive activation of JAK2, it still requires the presence of a homodimeric receptor, therefore it is different from other gain-of-function mutation of JAK2, like the one found in Down Syndrome B-Acute Lymphoid Leukemia (DS B-ALL) cases[113].

To recapitulate the sequence of events, once the TPO bind the MPL extracellular domain CRM1, the entire dimer change conformation, assuming a tighter organization and bringing the two monomers closer to each other. The spatial proximity induces the dimerization of the two TM helices, therefore the juxtaposition of the two intracellular portions, both associated with a monomer JAK2. Now, the two monomers are close enough to trans-phosphorylate each other, effectively switching to the active form and triggering the phosphorylation cascade.

Several mechanisms are in place, in absence of the ligand, to prevent the activation of the receptor: for example the His^{499} residue, that is located in the TM domain, is able to prevent the receptor dimerization, a highly human-specific mechanism that underlies subtle differences between species[114]. Another example that is unique for MPL and play a major role in the receptor activation dynamics is the cytoplasmic amphiphatic motif of five aminoacids, RWQFP, localized right after the TM domain. This domain contains a crucial tryptophan (W515) that, if substituted by any residue (except proline and cysteine), leads to a constitutive activation of the receptor[115]. This tryptophan is found mutated in a small percentage of ET and PMF[116]. Also the deletion of the entire motif leads to the same type of defect[117], therefore this small region has been interpreted as a negative regulator of the MPL activation, as it blocks, by steric hindrance, the juxtaposition of the two monomers, via the dimerization of the two TM helices. Due to the combination of these two features, the MPL receptor could be described as intrinsically poised to be active, and the TPO binding "releases the brake" and finally allows the activation of the signalling cascade (Figure 15).
It is also worth to mention that anomalies in the MPL partners could also lead to the inappropriate activation of the receptor. The most famous case is the CALR-MPL interaction: as a chaperon, the calreticulin is responsible of the correct folding of the fledging MPL in the ER, and ensures the retaining (and degradation) of any misfolded receptor. Mutations that destroy the CALR ER-retention peptide[118] modify the protein role, and transform it into a rogue chaperon able, once the receptor is on the cell surface, to activate the signalling cascade in absence of TPO, whatever is the maturation status of MPL[119,120].

The transphosphorylation of the two JAK2 molecules is the real trigger of the cascade: the first targets of their catalytic activity are several tyrosine residues in the intracellular domain of the receptor. While these residues have positive or negative effects on the signalling, they all contribute to the definition of a recruitment platform for the different protein scaffolds and intracellular mediators. There are essentially three major routes: STATs protein, the PI3K pathway and the MAPK cascade[121]. The first effectors that are recruited are the members of the STAT family, as they are directly phosphorylated by JAK2 and could be recruited on the MPL intracellular domain, through their SH2 domain. The most important STAT member for the MPL/JAK2 axis is STAT5, as demonstrated by different mouse genetic studies[122,123], followed by STAT1 and STAT3. The JAK2-mediated phosphorylation is necessary for the dimerization of two compatible STATs, as they can form also heterodimers. The formation of the dimer has a major impact on the nuclear translocation of these proteins, where they can bind efficiently the promoter regions of proliferative and anti-apoptotic genes, notably *BCL-XL*, *PIM1*, *CyclinD* and *SOCS*. Nevertheless, not all STAT members require phosphorylation for the nuclear import[124], and there is even an alternative model

that is claiming that it is the unphosphorylated (uSTAT5) that is actively binding the genome and repressing several MK genes, and the phosphorylation is actually relieving this inhibition[125].

In chronological order, the second pathway to be activated is the PI3K cascade. It is a lipid kinase, triggered indirectly by MPL and able to catalyze the production of the membrane-bound mediator phosphatidylinositol (3,4,5)-trisphosphate (PIP_3). This lipid recruits different proteins containing a pleckstrin homology (PH) domain, for example AKT. This enzyme is central in several survival pathways, but also in the activation of the mTOR complex, classically involved in cell proliferation and MK maturation; differently from STATs and MAPK, the PI3K pathway does not involved directly a TF. Finally, the last pathway to be activated temporally is the MAPK cascade, due to the larger number of mediators involved: RAS, RAF1, MEK1/2 and finally ERK1/2, these ones able to translocate in the nucleus in their phosphorylated form actively modifying TFs, notably RUNX1. While it is a major pro-proliferative pathway, the MAPK cascade is particularly involved in the MK differentiation and maturation: the chemical and genetic inhibition of this pathway leads to a higher ploidy level and a more efficient PPT formation[126,127]. This last observation highlights how an effective MKpoiesis requires a strong activation of MPL in the early steps of differentiation, but during the maturation a general switch-off of the signal is necessary, similarly to what is happening during the cellular senescence[128]. Conversely, other pathways, like the one activated by the surface integrins, augment their impact on the cellular physiology and become prominent during MK maturation.

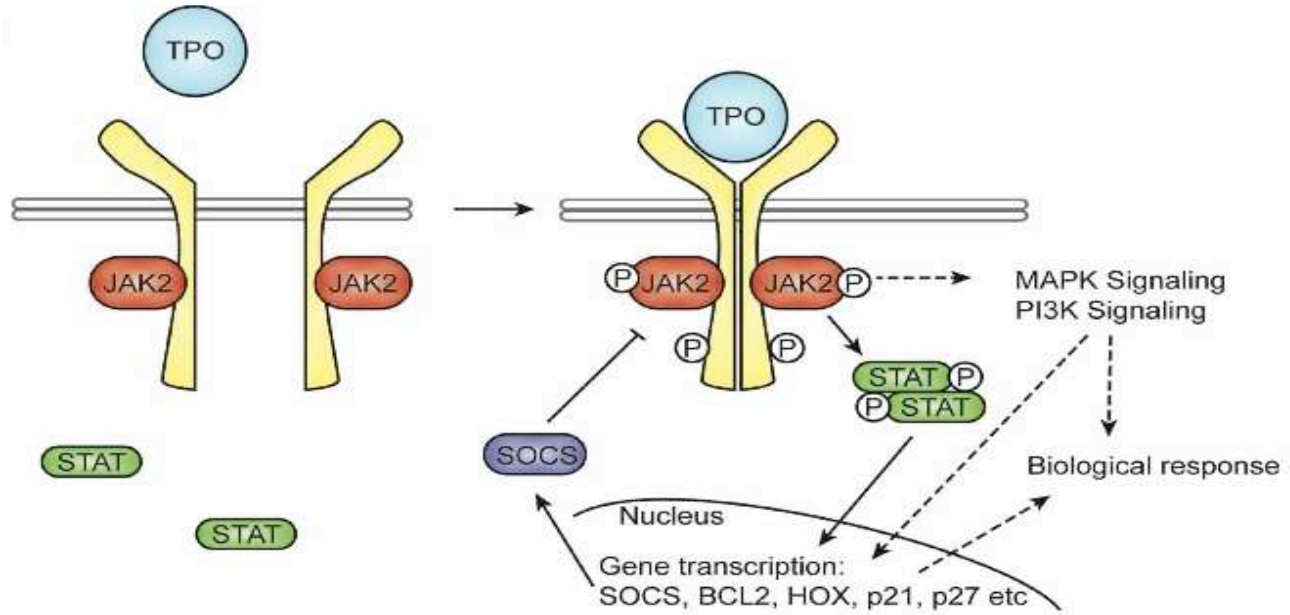


Figure 15: The schematic representation of the MPL activation via TPO binding. Readapted from De Graaf and Metcalf[129].

As every signalling pathway has a transient nature, also the MPL signal is supposed to be terminated in a highly regulated and timely fashion. The JAK2 phosphorylation of the MPL receptor, while a trigger for the cascade, also induces the activation and recruitment of several negative regulators. Nevertheless, the most important mechanism for the down-regulation of the TPO signal consists in the removal of the activated receptor from the cell surface, via clathrin-mediated internalization. This mechanism allows not only to switch-off the cascade, but also to remove from the extracellular space the TPO, contributing in an essential way to the regulation of the physiological level of the cytokine in the blood, a model proposed in the second part of the 1990s by different teams[130,131]. In this model, all cells expressing MPL, but mostly MKs and platelets, contribute to the clearance of the TPO from the plasma (Figure 16). Therefore, exists an inverse relationship between the TPO levels and the amount of MPL able to interact with TPO, a concept that could be associated to the number of MKs and platelet[129], but only in presence of a fully functional MPL: the mutation P106L leads to a partial trafficking defect, consisting in a diminished but not completely abolished expression of the receptor at the surface. The consequence is that the MK progenitors are able to differentiate and mature, but they are not particularly efficient in the TPO clearance, thus the individual carrying this mutation presents the paradoxical situation of a thrombocytosis with high TPO plasma levels[132]. A similar result has been obtained by the expression of reduced levels of the receptor in the *mpl*$^{-/-}$ context[133]. In conclusion, the regulation of TPO levels, unlike the EPO levels, is not directly dependent from the gene transcription at the sites of TPO production (liver, kidney, spleen); conversely, the TPO levels are inversely proportional to the number of MPL-expressing cells, in particularly MKs and platelets[134].

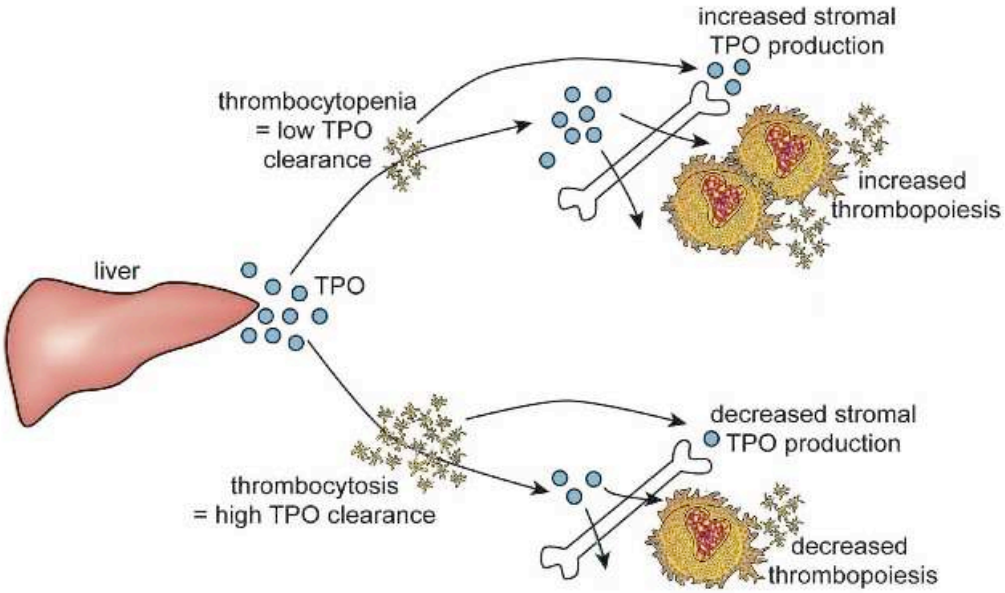


Figure 16: the TPO clearance, via MPL. The MPL expressing cells ensures the correct regulation of the cytokine levels and impacts the MKpoiesis and platelets production (thrombopoiesis). Readapted from De Graaf and Metcalf[129].

The MPL internalization occurs via a clathrin-mediated endocytosis, and it requires the assembly of vesicles via the indirect interaction between the adaptors AP2 and a specific hydrophobic motive in the intracellular domains, containing one of the tyrosine residues phosphorylated by JAK2 in the early stage of MPL activation[135]. Once inside the endocytic vesicles, MPL could be degraded, via ubiquitination or lysomal routing, or rapidly recycled at the surface: this last option in particular seems particularly suitable for the TPO clearance, as it allows the same MPL molecule to bind several TPO molecules in a short period of time[136]. In conclusion, it is evident the link between the interaction TPO-MPL and the immediate down-regulation of the signal. Nevertheless, the signalling has a long-lasting nature, as it activates multiple pathways at the same time, so several other systems are required to tune out the signals. As already mentioned, one possibility is the degradation of the signalling actors, for example by ubiquitination: several E3 ubiquitin ligases are apt to the task, the more famous is CBL, recruited via JAK2 to monoubiquitinate MPL[137]. Another entire family of E3 ubiquitin ligases that plays a crucial role is the Suppressor Of CytokineS (SOCS), recruited on phosphorylated proteins via the SH2 domain. They are of particular interest because they are transcriptionally regulated by STATs, therefore involved in a clearly negative feedback loop that could be easily disrupted in MPNs[138]. Other regulators, not related to the ubiquitin-mediated degradation, are LNK, an adaptor protein that could be also found mutated in MPNs[139], and several phosphatases.

It is not surprising that anomalies in the TPO-MPL-JAK2 axis lead to a disrupted MKpoiesis. Therefore many efforts have been directed towards the development of therapeutic strategies acting on it, mostly to battle against thrombocytopenia and anemia. Currently, two MPL agonists are approved for clinical use. The first one is called romiplostim and it is a peptide mimicking the TPO effect by binding with high affinity the same site described for the cytokine. The second one is called eltrombopag and it is a small-molecule interacting with the juxtamembrane end of the extracellular domain, able to induce the dimerization of the receptor, thus its activation. While both molecules promote a platelet increase via MKpoiesis stimulation, differences could be expected in the activation of the signalling pathways, as a direct consequence of the different activation mechanisms[140,141]. Less advanced is the development of MPL antagonists, in order to decrease the MPL activity, as the most preferred strategy is to target directly the kinase downstream. Nevertheless, a synthetic peptide able to bind in a similar manner to TPO was able to reduce the myelofibrosis and inhibit JAK2V617F-mutated cell proliferation[142].

CHAPTER II

2. Pathological haematopoiesis: Inherited platelet disorders

"Nature is nowhere accustomed more openly to display her secret mysteries than in cases where she shows traces of her workings apart from the beaten path; nor is there any better way to advance the proper practice of medicine than to give our minds to the discovery of the usual law of nature, by the careful investigation of cases of rarer forms of disease. For it has been found in almost all things, that what they contain of useful or of applicable, is hardly perceived unless we are deprived of them, or they become deranged in some way."

William Harvey, 1657 Letter to Dr. John Vlackveld

The definition of inherited platelet disorder (IPD) could be used in the description of all conditions that are congenital and affects the number and/or the function of platelets. Consequently, anomalies in the MKpoiesis and the thrombopoiesis (included the elimination of platelets from the circulation) could lead to a platelet disorder. An increase in the number of platelets is called thrombocytosis, a decrease is called thrombocytopenia, and an alteration in the platelet function is called thrombastenia. While only a minority of platelet disorders has a strong genetic component, it has nonetheless been at the forefront of the haematological research, because it has historically been a source of insights into the process of MKpoiesis and platelet production: the first description of a causal protein for a platelet disorder dates back to the early 1970s[143] and since then, at least more than 50 different genes have been described as responsible of a platelet disorder[144]. Combining this increased knowledge of the IPDs etiology with the larger accessibility to molecular diagnostic is giving encouraging results for the diagnostic of these pathologies[145]. Nevertheless, almost half of the thrombocytopenias suspected of a congenital origin are, to this date, still not associated to any genetic anomalies. Focusing the attention on the disorders showing a reduced number of platelet, different classification can be provided for inherited thrombocytopenias (ITs). Historically, the first one was based on the size of the platelets and distinguished between giants and macroplatelets, normal size platelets and microplatelets[146]. An alternative classification could be based on the stage of MKpoiesis that is principally affected by the mutated gene, therefore on the molecular etiology and not on the pathological manifestation. In this case the distinction could be operated between genes affecting the MK differentiation and genes affecting the MK maturation and platelet formation[147]. For the sake of clarity, I will introduce the different conditions accordingly to the differentiation stage where they have the largest impact.

2.1. MK Differentiation

In this group of diseases, the pathological mechanisms involves the early stages of MK commitment and differentiation, therefore it is possible to find many TFs as the responsible players. The consequence is the frequent presence of different clinical manifestations that encompass the thrombocytopenia and the diagnosis could be complicated, especially when other haematopoietic lineages are affected.

Congenital Amegakaryoblastic Thrombocytopenia (CAMT)

Due to the major role of the TPO/MPL axis, is not unexpected that alterations in this signalling pathway could lead to thrombocytopenia. The name of this severe disorder is quite self-explanatory, as there is an almost complete absence of MKs in patient bone marrow, followed by a tri-lineage anemia and bone marrow failure. Mutations that induce a MPL loss-of-function are typically associated to CAMT. When the anomaly completely abolish the expression of the receptor, the CAMT manifests very rapidly (in the first year from birth), while milder phenotypes could be observed for mutations that induce relatively smaller functional consequences on the TPO/MPL signalling axis[148]. The absence of a mouse model closely mimicking the human phenotype[149] prompted the development of alternative experimental strategies, like the generation of induced pluripotent stem cell lines from patient cells. This strategy has proven effective in recapitulating the major phenotypes associated to the disease, providing a valuable tool for further physiopathology studies[150]. The only curative possibility for these patients is the bone marrow transplantation, although the use of TPO agonist could be a viable option for the milder form of CAMT, in order to prevent the bone marrow aplasia. Recently, two teams independently described different mutations affecting the TPO and linked to CAMT[151-153]; in both cases, the use of TPO agonists improved the patient conditions, as the MPL protein function was not.

Radio-Ulnar Synostosis associated Thrombocytopenia (RUSAT) and Thrombocytopenia with Absent Radii (TAR)

Similarly to CAMT, there are two other ITs that incept very early on in life. These forms are also associated to two different bone development defects that are absent in CAMT: the bilateral absence of the radius bone (TAR) and the formation of a radio-ulnar synostosis (RUSAT). Moreover, these patients do not evolve towards a pancytopenic clinical picture, and the TAR patients even recover from the thrombocytopenia. While the genetic etiology of RUSAT has been assigned to mutations in the *HOXA11* gene, the understanding of the genetic causes of TAR has been more complicated. A seminal paper in 2012 proved that the reduced expression of the Exon Junction Complex (EJC)

subunit Y14 is associated to more than 95% cases of TAR; the particularity is that this haploinsufficiency is the consequence of the compound inheritance of a loss-of function mutation on one allele of the gene *RBM8B* and a SNP affecting the gene expression on the other allele. This hypomorphic mechanism significantly reduces the level of Y14 protein in affected individuals, compared to healthy non-affected familiars, and leads to the clinical manifestation of the TAR syndrome[154]. Nevertheless, a paucity of experimental data on the thrombocytopenia etiology for both syndromes, also due to the lack of valid experimental models, hampers the understanding of the pathology.

Paris Trousseau thrombocytopenia, Jacobsen syndrome (JBS) and FLI1-RT

The Jacobsen syndrome has been first described in 1973 as a congenital disorder with several clinical features, including a macrothrombocytopenia[155]. Twenty years later, a French team described a variant of this syndrome (Paris Trousseau thrombocytopenia), affecting the same region of the chromosome 11 deleted in the JBS, although a reduced extension of the mutation in Paris Trousseau cases[156]. Nevertheless, both syndromes share the macrothrombocytopenia, with the curious feature of the existence of two platelet populations, a normal size one and a population of large platelets displaying giant α-granules and deficient in dense granules, a feature that had a diagnostic impact[157]. Moreover, patients bone marrow smears show the existence of two MK populations, one composed of normal MKs and one of small hypoploid MKs. The deletions on chromosome 11 encompass *FLI1*, a well-known regulator of haematopoiesis and specifically MKpoiesis, as it acts in tandem with RUNX1 in the down-regulation of several genes crucial for the correct MK differentiation, like *MYH10*. The development in the year 2000 of a heterozygous knock out of *Fli1* recapitulated the dysmegakaryopoiesis observed in JBS and Paris Trousseau thrombocytopenia patients[158]. Finally, different point mutations in *FLI1* have been described as responsible of the macrothrombocytopenia, without the other syndromic clinical manifestations[159,160]. To resume, the causal reason of all these syndromes is the hemizygous loss of *FLI1*.

The consequences on the functional level has been investigated in the early 2000s and have been proved particularly interesting, as a report in 2004 described a monoallelic expression of *FLI1* during the early MKpoiesis ($CD41^+CD42^-$), with the switch from monoallelic to biallelic expression at the transition to the more mature $CD41^+CD42^+$ stage. Therefore, in *FLI1*-mutated cells it is possible that a cell find itself in the transitory state of a complete loss of FLI1, when the active transcription switch from the wild type to the mutated allele. Due to the particular impact of FLI1 during the late MKpoiesis, it is then clear that those cells would not be able to mature correctly and they will give rise to the hypoploid MKs observed in the patient bone marrow. Conversely, the cells

that were expressing the wild-type allele are able to terminate correctly the MK maturation[161]. This provocative model explains the existence of two MK populations; it would be interesting to investigate wheter this unusual gene expression mechanism exists also for other genes.

FPD/AML – THC5 – THC2

Many ITs are associated to other clinical manifestations, and are typically part of more complex syndromes. If the CAMT can also affects the rest of the haematopoiesis and leads to bone marrow aplasia, this trio of ITs with normal platelet volume shares a subtler haematological phenotype: all the three are associated with an increased risk of developing blood malignancies, both for the myeloid (FPD/AML and THC2) and the lymphoid lineage (THC5). Clinical signs other than the thrombocytopenia includes dysmegakaryopoiesis and the reduced presence of α-granules in the mature MKs.

FPD/AML

Familial Platelet Disorder with predisposition to Acute Myeloid Leukemia (FPD/AML) has been the first IT to be associated to a predisposition to myeloid malignancies (myelodysplastic syndrome or acute myeloid leukemia), and more rarely to lymphoid malignancies. First described in the 1980s, the autosomal dominant disorder has been finally associated, in 1999, to mutations in the TF *RUNX1*[162], a member of the core binding factor. Mutations are not particularly localized to certain hotspots and could also include intragenic deletions conducing to haploinsufficiency. These mutations do not seems to be specific for FPD/AML, but they overlap with the somatic mutations found in many cases of MDS and AML. Nevertheless, they could be classified in two loose categories: haploinsufficiency-like and dominant-negative-like mutations[163]. This classification could be particularly important in light of the predisposition associated to germinal RUNX1 mutations: around 35% of FPD/AML cases develop a myeloid malignancy and the epidemiology of this pathology has always been hampered by its extreme rarity, although the increased awareness is revealing more cases of this congenital disorder[164]. In the process of myeloid leukemia evolution the most frequent event during leukemogenesis is the acquisition of a somatic mutation on the intact allele, leading to the total loss of its activity and the emergence of a full-blown malignancy[165].

RUNX1 is one of the most important TF in haematopoiesis. Crucial in the development of the first HSCs[166], the loss of this TF in the context of the adult haematopoiesis is not fundamentally harming the entire system, but cause numerous defects in multiple lineages, including a defective platelet production associated to a presence of hypoploid MKs in mouse[167,168]. The defect has been confirmed in primary patients cells and induced pluripotent stem cells, confirming its role also in

human haematopoiesis[169-172]. In the MKpoiesis RUNX1, in partnership with FLI1 and other MK-specific TFs, regulates different genes that are involved in the most distinguished steps of the MK differentiation. Indeed, heterozygous mutations for RUNX1 impair the normal MK differentiation, polyploidization and PPT formation[173], in the same way as the reduction of RUNX1 levels via short hairpin RNA inhibition. The effect is direct, for example the retained expression of the *MYH10* protein, responsible for the reduced endomitosis[73], or indirect, as in the case of the reduced expression of the different members of the $\alpha_{IIb}\beta_3$ complex, that can successively regulate the MKpoiesis. Finally, it is worth noticing that several targets of RUNX1 are themselves involved in other ITs (e.g. THC2), therefore FPD/AML could share some similarities with other forms.

THC2 or ANKRD26-RT

The second congenital platelet disorder described to be associated to a leukemic predisposition is the THC2 (Thrombocytopenia type 2), characterized by autosomal dominant inheritance pattern, mild to severe reduction in platelet levels and the absence of any evident platelet defect. This disease has been associated to mutations in the regulatory region of the gene ANKRD26, a member of a primate specific gene family counting several members and characterized by the presence of multiple ankyrin-repeat domains. The detailed description of this IT could be found in a review article that I have co-authored (Annex).

THC5 or ETV6-RT

The latest addition to the subcategory of IT predisposing to leukemia is THC5 (thrombocytopenia type 5). *ETV6* mutations have been described for several pedigrees affected by an autosomal dominant bleeding disorder, with reduced platelet counts and normal platelet volume. No evident defects in platelet function have been described, although a mild reduction in response to arachidonic acid could be observed for some patients. Moreover, an increased incidence of acute lymphoblastic (ALL), AML and MDS have been described, associated to increased number of circulating haematopoietic progenitors ($CD34^+$ cells in the peripheral blood[174]). The leukemic penetrance is variable but it could be estimated that around 30% of germline carriers develops some kind of haematologic malignancy. Mutations in the ETV6 gene have been previously described as cancer-related and appear in some mutations catalog like COSMIC. Nevertheless, the classic mutation linking *ETV6* with leukemia is the chromosomal translocation responsible of the fusion protein ETV6-RUNX1, described in pediatric B-ALL[175].

The gene *ETV6* encodes for a TF of the ETS family, as it displays a C-terminal domain able to recognize the ETS genomic consensus sequence. The central portion of the protein is responsible for the regulation of the TF activity, while the N-terminal domain is involved in the dimerization

and nuclear localization of the protein; in fact, this TF can be active only in a homo- or heterodimer, where it acts as a repressor[176]. The mutational landscape gives a picture not dissimilar to the FPD/AML case, where mutations lead to a dominant-negative effect of the mutated *ETV6* allele. Indeed, the disruption of the correct interaction with the transcriptional partner, or the mislocalization of the protein due to the alteration of the localization signal, are all responsible for a reduced activity of the intact ETV6. The consequence is the increased expression of genes normally down regulated, again in a similar fashion to *RUNX1* mutations. Most mutations occurs in the ETS domain or in the central regulatory one[177], therefore it is surprising to observe a mislocalization of the protein, as the nuclear signal does not seems to be altered[178]. To this date, a true haploinsufficiency has not been reported.

The analysis of the MKpoiesis in ETV6 patient cells revealed a paradoxical increase in the number of MK progenitors and a larger number of $CD41^{+}CD42^{+}$ cells, a notable difference compared to the FPD/AML and ANKRD26-RT forms, where the MK numbers were reduced or unmodified. Nevertheless, *ETV6*-mutated MKs displayed a defective PPT formation, with an abnormal cytoskeletal organization, associated to a reduced expression of key genes involved in the final maturation, like the small GTPases RhoA and CDC42. No clear association between these observations has been proposed yet. It is curious to notice that the reduced CDC42 expression could also explain the increased number of circulating progenitors, as the inhibition of this small GTPases improved the progenitors mobilization.

SRC

There is only one IT associated to a tyrosine kinase and it has been recently identified in nine cases of autosomal dominant bone pathologies associated to haematological symptoms like myelofibrosis, trilineage dysplasia with a moderate myeloid hyperproliferation and bleeding[179]. *SRC* is historically the first oncogene that has been described in the 1970s, but it has never been associated to a genetic disease. The only mutation described to this date (E527K) consists in a gain-of-function, linked to the loss of the self-inhibitory mechanism that keeps the kinase inactive. Patients MKs are dysplastic and unable to mature correctly, and the same results are obtained when the mutated protein is overexpressed in normal MKs. The defective maturation seems to be linked to the increased formation of podosomes on fibrinogen, direct consequence of the SRC increased activity. Moreover, a PPT formation defect could be observed, reversed by the use of a SRC inhibitor. A clear link between the podosomes formation and the PPT formation defect remains to be proved, but some data generated in another IT, the Wiskott-Aldrich syndrome (WAS), demonstrate that a reduced podosome formation induce an anticipated PPT formation and the release of platelets in the bone marrow and not in the sinusoid vessels.

GATA1

As already mentioned, GATA1 is a crucial TF for the normal and pathological myeloid haematopoiesis. While mutations in the first exons induce a truncated, short isoforms associated to myeloid leukemia, anomalies in the rest of the protein typically leads to macrothrombocytopenia and anemia, albeit these two phenotypes encompass a large phenotypic variability. This is not particularly surprising, as GATA1 is actively regulating several genes that are expressed specifically in MKs only or in erythrocytes only[180]. For the MK lineage, it is possible to observe small dysplastic MKs, highly proliferating and unable to mature correctly, because of the extensive defects in all the crucial steps of the MKpoiesis, like the endomitosis, the DMS formation and the PPTs formation[181]. Moreover, the few macroplatelets produced displays fewer granules and are highly defective in their functional properties, because of decreased levels of all the major glycoproteins and surface integrins. The conventional distinction between mutations affecting the DNA binding properties and mutations affecting the GATA1-FOG1 axis is not fully explaining the huge phenotypic variability for the erythrocytic lineage[182].

GFI1B

Growth factor independence 1b (GFI1B) is a TF involved in different stages of the haematopoiesis, both for the myeloid and the lymphoid lineages. It generally acts as a repressor, recruiting different chromatin modifiers like LSD1[183]. Mutations in this gene has been reported in different cases of autosomal dominant macrothrombocytopenia with a disrupted cytoskeleton in platelets, plus defects affecting both α-and δ-granules. Moreover, a consistent feature among patients is the increase in the CD34 expression on platelets[184]. Other signs of this IT are the presence of dysplastic erythroid cells (anisopoikilocytosis) and mild myelofibrosis[185]. Mutation in the gene seems to cause dysmegakaryopoiesis linked to an increased number of MKs in the bone marrow; nevertheless, the impact of this TF on human MKpoiesis remains poorly understood. Mouse model knock out for the murine *Gfi1b* show a defective erythropoiesis and MKpoiesis only when homozygous, supporting a dominant negative impact of GFI1B[186].

2.2. MK maturation and platelet release

This group of disease present many cases of macrothrombocytopenias, as mutations in the proteins involved in the PPT formation are invariably altering the size of the produced platelets. It is unusual, compared to the first group, to find other haematological defects in those patients; nevertheless, a syndromic picture could be observed for several cases, especially when the affected proteins are largely expressed in other tissues.

Grey Platelet Syndrome (GPS)

The Grey Platelet Syndrome is one of the earliest descriptions of non-immune thrombocytopenia[187]. The name derives from the characteristic grey colour of platelets in blood smears, caused by the absence of α-granules. Moreover, those platelets tend to have a larger size than the usual and defective aggregation properties, mostly because of the lack of α-granules content. Other clinical signs are bone marrow fibrosis, splenomegaly and emperipolesis[188]. The inheritance pattern is typically autosomal recessive and the mapping revealed biallelic missense, nonsense or frameshift mutations in the *NBEAL2* gene as likely responsible of this IT[189-191]. Healthy carriers displays no clinical signs whatsoever, therefore a complete loss of function is necessary for the pathology to manifest, as confirmed by *NBEAL2* expression in GPS platelets. In the same way, *Nbeal2*$^{+/-}$ mice do not display any haematological abnormalities. Conversely, homozygous knock out mice are severely thrombocytopenic, with splenomegaly and profound lack of α-granules. *Nbeal2*$^{-/-}$ MKs are producing fewer proplatelets and are less polyploid than the wild type, although a role of NBEAL2 in the process of endomitosis it is not confirmed[192]. This murine data have been confirmed in human, via *in vitro* culture of patient progenitors[193]. Nevertheless, conflicting reports have been issued in regard of the PPT formation and the ploidy state[194,195], leaving a certain uncertainty on this matter. The reported emperipolesis (presence of living leukocytes inside GPS MKs) is a curious feature that could be linked to a defective DMS formation and maintenance, although further data is necessary to support this hypothesis. It has also been reported that *Nbeal2*$^{-/-}$ mice display an increased incidence of infection and neutrophil defects[196], a feature that has been sparsely found in patients too. A recent report tried to understand the mechanistic basis of the GPS and identified several partners for NBEAL2 that could be linked to a defective PPT formation; in particular DOCK7 seems to be a promising candidate, as it appears to be downregulated in patients platelets[197].

Most of GPS families present an autosomal recessive inheritance, but several pedigrees displaying grey platelets have been described with an autosomal-dominant inheritance pattern. Successively, they have been associated to *GFI1B* and *GATA1* mutations, and their many similarities may prompt

a debate about considering all those ITs as a single entity, at least for what concern the MK lineage. In particular, the presence in all these disorders of a reduced amount of α-granules and emperipolesis support a model where those three genes are functionally linked. Recently, Freson and colleagues proved that GATA1 could regulate the *NBEAL2* expression via a long-range enhancer, therefore partially explaining the similar phenotypes between those two macrothrombocytopenia[198]. Further studies are needed to confirm if GATA1-GFI1B co-transcriptional complexes could regulate *NBEAL2* expression during MKpoiesis.

Wiskott-Aldrich Syndrome (WAS)

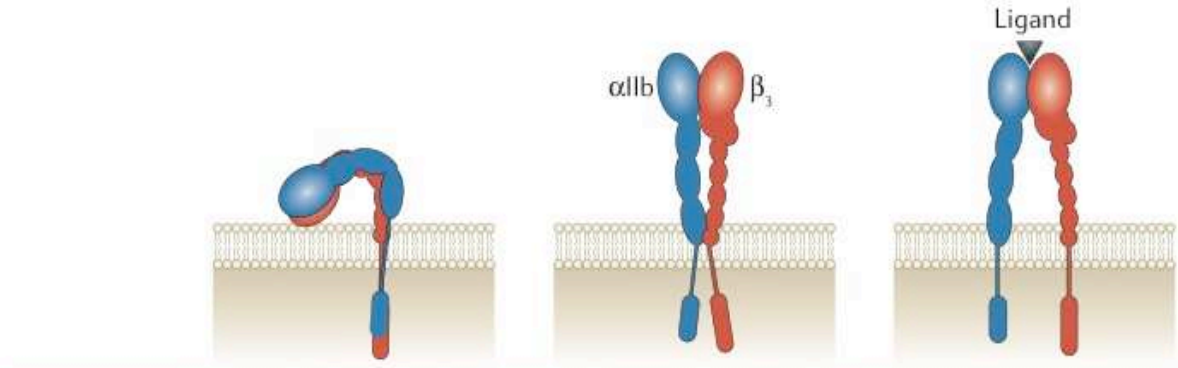
This X-linked complex syndrome, due to the defective production or activity of the WASP protein, is associated to a major immunodeficiency when there is a complete absence of the protein. It is associated to platelets of smaller size called microplatelets[199] that are quickly cleared by the macrophages; consequently, a splenectomy could in some cases improve the thrombocytopenia and normalizes the platelet count. Nevertheless, a peculiar defect in PPT formation has been observed in some mouse models, and consists in the precocious proplatelet extension and platelet release in the bone marrow space. This feature has to be proven in the human settings, as the only results obtained ex vivo from human cells do not display a defective PPT formation[200]. It is worth to notice that WAS is the only IT that has been treated with gene therapy to this day, via a lentiviral correction of the WAS deficiency and consequent bone marrow transplantation[201].

Glanzmann thrombastenia

The Glanzmann thrombasthenia (GT) stood out from the other platelet disorders, as it is not clearly associated to a reduced number of platelets, but it is mostly associated to an abnormal fixation of fibrinogen, a key component of the extracellular matrix and constituent of the bone marrow. This unusual phenotype has been associated to a deep decrease of the surface expression of the complex GPIIb/IIIa, also known as integrin $\alpha_{IIb}\beta_3$. This defective expression is caused by mutations in the genes *ITGA2B* and *ITGB3*, but the pathology manifest only when they are affecting both alleles, according to a classic autosomal recessive inheritance pattern[202]. A diagnosis of GT is usually performed in absence of platelet aggregation with physiological agonists and a decrease in the expression of the $\alpha_{IIb}\beta_3$ integrin by flow cytometry. Nevertheless, it is possible to observe normal levels of $\alpha_{IIb}\beta_3$ integrin and defective aggregation properties, therefore patients with classic GT could be segregated into three categories, according to absence (<5%), reduction (<25%) and no differential expression of this integrin[203]. The defective activation for a large range of agonists is a consequence of the ability of $\alpha_{IIb}\beta_3$ to interact also with other ligands than fibrinogen.

A wide range of mutations has been reported for GT, encompassing the entire genes. As GPIIb and

GPIIIa (respectively α_{IIb} and β_3, also identified as CD41 and CD61) are TM proteins, so they can be subdivided into three portions, with an extracellular domain responsible of the fibrinogen interaction, a TM fragment and the important intracellular domain, necessary for the regulation of the complex formation and activation. The two subunits are synthesized separately, but assembled as a complex quite early, already in the ER. α_{IIb} is produced as a precursor and successively cleaved in the Golgi, while β_3 as a non-glycosylated chain. Therefore, mutations that influence these early steps are usually associated to a complete absence of the complex at the surface[204], like in the case of the V298F substitution, that cause the retention of the complex in the Golgi, but does not alter the formation of the complex in the ER[205]. A trickier phenotype is observed when the mutations are not affecting the maturation of the complex, but its functional properties: for example a triplet of mutations in the extracellular domain (D145N, R240Q and W substitutions in the βA portion) affect the ability of the β_3 chain to interact with the fibrinogen after activation, but does not alter its expression or its stability. This last feature is particularly unusual, as the integrin needs to be "pre-activated", in order to be able to interact with the fibrinogen (Figure 17). Mechanistically, the integrin is active after a conformational change from a bent to an extended structure, therefore allowing the binding of the platelet to the substrate[206]. This activation is due to the signalling inside-out, induced by the activity of two intracellular proteins, talin and kindlin, that are able to interact with the intracellular portion of the β_3 and prepare the complex to its full activation. The interaction with the fibrinogen (or other ligand) triggers the second signalling cascade, named outside-in, and normally mediated by SRC and SYK kinases[207]. Several mutations have been described as able to disrupt all these steps[208].



Affinity	Low	Intermediate	High
Conformation	Bent	Extended	Open
Outside-in signalling	No	No	Yes

Figure 17: the transition from the inactive to active conformation for the integrin $\alpha_{IIb}\beta_3$. This highly conserved mechanism allows a finer control of the signalling dependent from this integrin. Readapted from Ley and colleagues[209].

The classic manifestation of a GT does not includes a decrease in platelet levels; nevertheless, the $\alpha_{IIb}\beta_3$ complex is involved in the PPT formation and the final steps of MK maturation, and some groups reported such defects in different models of GT[210,211]. Those mutations do not share the same inheritance pathway of most GT-causing lesions, but they segregate in an autosomal dominant pattern and induce a gain-of-function in *ITGA2B* and *ITGB3*, associated to different degrees of macrothrombocytopenia, but milder platelet defects than the classic GT cases. In most cases, the mutations are located in the cytoplasmic part of the protein and induce such conformational changes to the complex that the integrin results already active also in resting condition. Patient cells retain a normal differentiation, but show an aberrant PPT formation that has been linked to a defective re-organization of the acto-myosin cytoskeleton[212].

Finally, as the β_3 component of the $\alpha_{IIb}\beta_3$ integrin is prominent for other cell types than MKs and platelets, it is surprising that GT-mutations in *ITGB3* does not present other symptoms than bleeding, due to the importance of the integrin for cancer, cardiovascular diseases and inflammation[209]. In a mouse model, the loss of β_3 cause an enhanced tumor growth, osteosclerosis and a marked defect in the placental development, as this integrin is part of the vitronectin receptor $\alpha_v\beta_3$[213].

Bernard Soulier thrombocytopenia, PT-VWD and Di George syndrome

The Bernard Soulier syndrome (BSS) is one of the first IT that has been described, in the late 1940s[214]. Patient platelets are bigger than the reference volume (giant platelets) and clearly defective in their aggregation properties under different stimulation. The thrombocytopenia could be very variable, ranging from mild to very severe, but the global bleeding tendency is particularly high, due to the combination of reduced platelet levels and impaired platelet aggregation. Patients with BSS have been differentiated from GT cases because they were lacking the expression of the glycoprotein GPIb, but they were expressing normal levels of $\alpha_{IIb}\beta_3$ integrin[215]. This glycoprotein has been successively associated to an essential complex for haemostasis, the GPIb-IX-V complex, able to bind to the vWF. This surface complex is the results of the highly regulated assembly of four subunits: GPIbα and GPIbβ (forming the GPIb unit), GPIX and GPV, assembled in the stoichiometric ratio of 2:4:2:1 (Figure 18). These four subunits are also identified as different cluster of differentiation, and the entire receptor is usually called CD42. Although its primary ligand is the vWF, this receptor is also able to bind other molecules like thrombin, P-selectin and even some types of snake venom proteins[216].

Gene	Protein	CD (in human)	Stochiometry (for a single complex)
GP1BA	GPIbα	CD42b	2
GP1BB	GPIbβ	CD42C	4
GP9	GPIX	CD42a	2
GP5	GPV	CD42d	1

Table 2: the GPIb-IX-V different subunits, with the respective names and sizes. The stoichiometry is to be intended for a single receptor molecule.

All these subunits are TM proteins type I, although they are vastly different in their length and conformation. As member of the type I family, they are devoid of any enzymatic activity. The most important member of the complex is GPIbα, as it is the subunit containing all the binding sites for the extracellular ligands. Most of the protein is extracellular and contains, other than the ligand binding sites, some crucial residues for the correct assembly of the entire complex. Its cytoplasmic tail is well known for the interaction with Filamin A (FLNA), PI3K and 14-3-3 proteins. The other part of the GPIb complex, the β subunit, is much smaller than the α one, but it is fundamental for the assembly of the entire complex, as the interaction between the GPIb subunits is the first event in the assembly of the entire CD42 complex. GPIX is essential for the stabilization of the GPIb subunits, therefore the expression of each subunit could be used as an indicator for the entire complex assembly. All these three subunits interact with each other, in three different ways: 1) disulfide bonds between the three GPIb subunits, 2) interactions between the TM domains and 3) the GPIbβ and GPIX extracellular domains. The disulfide bonds are essential in connecting one GPIbα with two GPIbβ, via the Cys^{484} and Cys^{485} on the α subunit and the Cys^{122} on the β subunit[217]. The correct formation of these disulfide bonds is also ensured by the correct juxtaposition between the TM domains of the subunits involved, as the TM domains replacement decreased the binding efficiency, therefore the entire complex expression at the surface[218]. In a similar fashion, replacing the TM domain of GPIX induce the same destabilization in disulfide bonding, demonstrating the central role of GPIX in the complex stabilization[219]. At the core of this juxtaposition is a combination of polar residues interaction and leucine zipper motifs. The final mechanism of assembly of the CD42 receptor is the interaction between the extracellular domains of GPIbβ and GPIX: this is highly required, as GPIX is not stable in absence of this interaction, so the entire complex is unstable[220]. Finally, the role of GPV is less clear than the other subunits: it is weakly associated with GPIb, although that could be just a consequence of the different protein half-lives[221]. Nevertheless, this protein does not seem to be essential for the complex: mice lacking this gene do not show any platelet anomaly or haemostasic alterations[222].

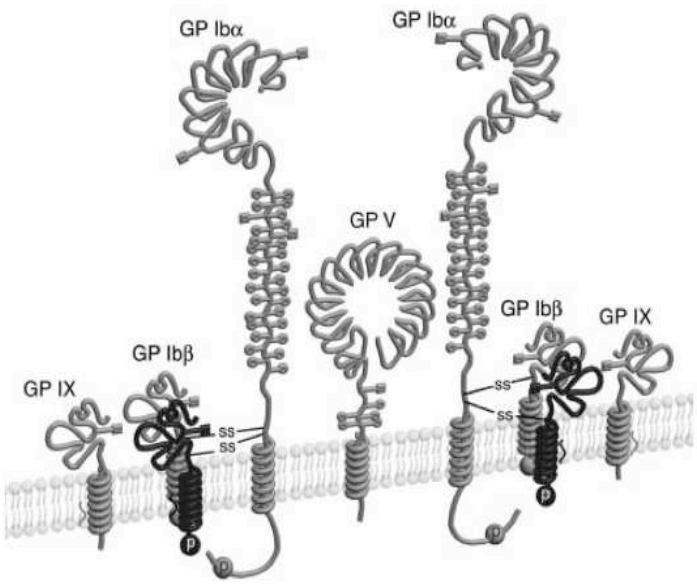


Figure 18: the GPIb-IX-V complex. In evidence the disulphide bonds between GPIbα and the two GPIbβ molecules (in light grey and black). Readapted from Luo and colleagues[217].

While the first mutations have been identified in the *GP1BA* gene (encoding for the GPIbα), this is not the only causal gene for BSS[223]. Mutations in *GP9*[224] (encoding for GPIX) and *GP1BB*[225] (encoding for GPIbβ) have also been reported, to a current total of approximately 50 mutations described as causing a biallelic BSS. No mutations have been reported to this date in the *GP5* gene (encoding for GPV). Most of these mutations affect the assembly of the complex in the endoplasmic reticulum, therefore their expression levels at the membrane. They tend to be concentrated in the extracellular and TM regions of the protein, and they affect aminoacids residues that are not exposed, but buried in the protein. Nevertheless, there are also mutations able to induce a milder form of BSS and transmitted according to an autosomal dominant pathway: the most famous is the A156V, also called Bolzano mutation, as it has been described for the first time in Northern Italy[226]. This monoallelic mutation induces a macrothrombocytopenia with almost normal levels of CD42, but a marked disruption in platelets ability to bind vWF. The MK differentiation for biallelic and monoallelic BSS patients does not appear to be disrupted, except for the altered CD42 expression. However, those MK are unable to extend proplatelets *in vitro*, suggesting an essential role of this complex for the correct MK maturation[227-229]. The defect appears more evident in biallelic than monoallelic patients, another sign of a correlation between the thrombocytopenia degree and the mutation severity. This phenotype is confirmed also in different animal models of BSS[230,231].

GP1BA mutations are not only responsible for BSS, but they have been associated also to a different form of congenital platelet disorder, named Platelet Type von Willebrand disease (PT-

VWD). In this case, the *GP1BA* mutations cause a gain-of-function, with an excessive binding of GPIbα to the vWF, therefore an increased platelet clearance, reduced level of vWF and a bleeding phenotype[232]. The causing mutations are all localized in the extracellular patch responsible of the vWF binding, so surface exposed, and they act as stabilizer of the binding, favoring a higher receptor affinity. Finally, the hemizygous deletion of portions of the chromosome 22, where the *GP1BB* gene is located, is associated to the Di George syndrome. The platelet levels for these patients remind of a monoallelic BSS[233].

MYH9-RD

The disorder called MYH9-related thrombocytopenia is, at the moment, the most diffused IPD worldwide. It included 4 different syndromes, affecting not only the platelet lineage but also several other tissues; including the kidney and the ear. As for the ANKRD26-RT, I have co-authored a review that describes the disease in more detail (Annex).

DIAPH1

This gene has been very recently associated to macrothrombocytopenia and hearing loss, a feature similar to the MYH9-RD, as a result of a large international collaboration[234]. In these patients, a mutation gain-of-function do not altered MK differentiation, but caused a defective PPT formation, possibly linked to the observed increased microtubules stability in platelets. As a member of the formin family, the protein DIAPH1 is strictly associated to the cytoskeleton remodelling and microtubules elongation, as a classic RhoA GTPase effector. The gene has been associated to the final stages of human MKpoiesis by an earlier report that highlighted the importance of this gene in the correct PPT formation process. In this case, a constitutively active protein induced a defective PPT formation and an increase in stress fibers formation, suggesting a role as a negative regulator of the late maturation[235].

ACTN1

Actinin 1 (ACTN1) is one of four homologous protein mainly involved in the cross-linking of filamentous actin (F-actin) during the cytoskeleton reorganization. Moreover, ACTN1 interacts with the $\alpha_{IIb}\beta_3$ complex in platelets, via the β_3 portion of the integrin, possibly regulating the inside-out signalling[236]. To allow the interaction with the actin cytoskeleton, ACTN1 must form a homodimer, a feature that it is calcium-dependent and directly modify the dimer affinity for actin, via conformation change. Heterozygous mutations in this gene have been associated to an inherited macrothrombocytopenia with no syndromic features and no evident platelets functional defects, probably because of the tissue specific expression of this gene or a generic compensation effects

due to the other homologous genes[237]. The molecular etiology is currently unknown and warrants further investigations.

Filaminopathies A (FLNA) – PRKACG

The FLNa protein, encoded by the *FLNA* gene, is an actin-binding protein that has been described in platelets for more than thirty years[238]. Mutations in this gene have been associated to a large spectrum of rare diseases affecting several organs, possibly a consequence of the ubiquitous nature of this protein. The most characterized condition is the Periventricular Nodular Heterotopia (PNH), a neurodevelopmental disorder affecting the migration of neuronal progenitors[239]. Other conditions include skeletal dysplasia, cardiac valvular defects and intestinal obstruction[240]. The filamins family includes two other isoforms (*FLNB* and *FLNC*) that shares a high degree of homology with *FLNA* and they can even form heterodimers, although they do not seems to be particularly active in MKs[241]. Due to its association with GPIbα in platelets and MKs, these two components have always been strictly associated, and their interaction has been described as crucial for a correct activity of the GPIb-GPIX-GPV complex. Nevertheless, FLNa has been proved to interact also with the β_3 family, therefore it plays a role in the activity and regulation of $\alpha_{IIb}\beta_3$[242]. Several mouse models have been generated for studying the role of this protein in MKs and platelets[243-245], and although some conflicting reports due to the model differences, the general consensus is that FLNa deficiency induce the formation of larger and more fragile platelets, that are cleared rapidly from the circulation, therefore the macrothrombocytopenia. This feature of macroplatelets has been accordingly described in patients carrying mutations in the *FLNA* gene and affected by PNH: the patients display both intragenic deletions and point mutations[246]. Nevertheless, two platelets populations, with different adhesive properties and sizes, have been observed in $FLNA^{mut}$ carriers[247]. The *in vitro* MKpoiesis appears unperturbed, but very few data is available on human cells, as the patient heterogeneity complicates the molecular characterization.

Strictly related to Filaminopathies A is the role of *PRKACG*: this gene encode for the catalytic subunit of the PKA, a kinase that is crucial for the FLNa stability, as it is able to phosphorylate a serine residue (Ser^{2152}) that blocks the protein degradation[248]. Biallelic mutation in PRKACG abrogates the catalytic activity of PKA and in MKs leads to a sensible reduction of the FLNa levels. Individuals carrying this biallelic mutations on *PRKACG* are severely thrombocytopenic and their platelets are larger than usual, sharing similarities with the macroplatelet population found in Filaminopathy A patients[249].

TUBB1

TUBB1 is the primary β-tubulin isoform present in MKs, so it is not surprising that mutations in this essential microtubules component leads to thrombocytopenia. Heterozygous mutations of the *TUBB1* gene have been described recently in an autosomal dominant macrothrombocytopenia with normal bone marrow, and they lead to an outright decrease in protein levels and a major defect in PPT formation[250]. Interestingly, while the murine heterozygous knock-out is not recapitulating the macrothrombocytopenia, a high prevalence of giant platelets could be found in the Cavalier King Charles Spaniel dogs, as they carries a mutation in the canine *TUBB1*[251].

TPM4

Tropomyosin 4 (*TPM4*) is another gene identified by the same consortium that described DIAPH1. In this case, the gene haploinsufficiency conduces to an autologous dominant macrothrombocytopenia, with mild platelet defects. In a murine model, a dose-dependent effect on PPT formation and platelet biogenesis can be observed, as a consequence of a non-redundant effect on the acto-myosin complex. An increased degradation of some constituents of the cytoskeleton has been observed, themselves responsible of other forms of macrothrombocytopenias (FLNA, ACTN1, cofilin), but further elucidation of the mechanisms is required[252].

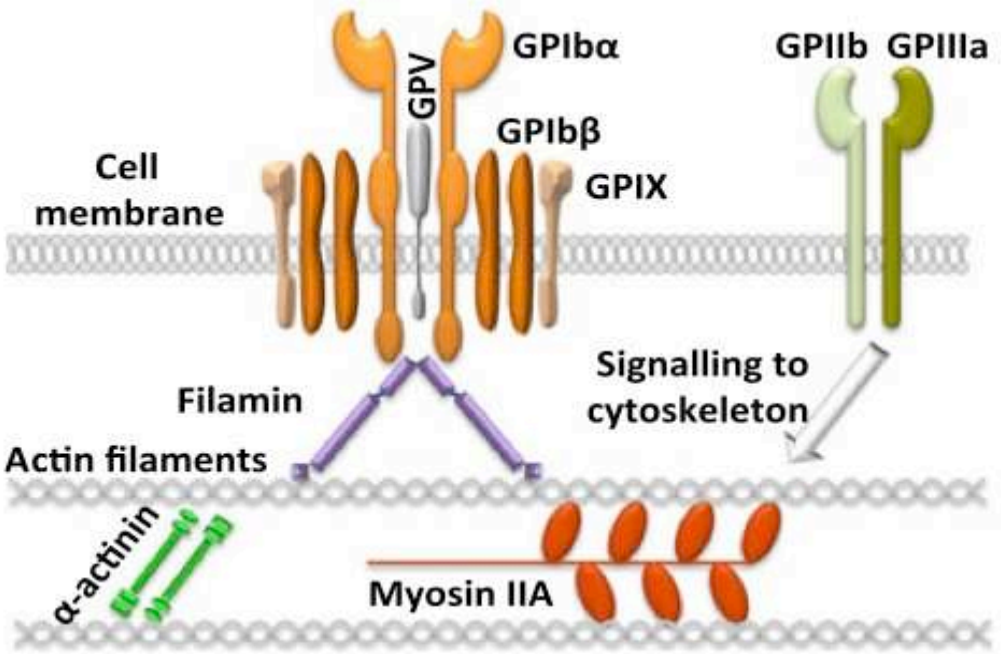


Figure 19: a general schematic representation of some of the proteins involved in macrothrombocytopenias. Readapted from Pecci and Balduini[253].

Inherited Thrombocytopenia (IT) class	Disease Name	Gene	Platelet Size	Syndromic form	Haematological associated phenotype
MK differentiation	Congenital Amegakaryoblastic Thrombocytopenia (CAMT)	*MPL/TPO*	Normal	No	Bone marrow failure
	Radio Ulnar Synostosis associated Thrombocytopenia (RUSAT)	*HOXA11*	Normal	Yes	-
	Thrombocytopenia with Absent Radii (TAR)	*RBM8B*	Normal	Yes	-
	Jacobsen syndrome	*FLI1*	Macro	Yes	-
	Paris-Trousseau thrombocytopenia	*FLI1*	Macro	Yes	-
	FLI1-related thrombocytopenia	*FLI1*	Macro	No	-
	Familial Platelet Disorder with predisposition to Acute Myeloid Leukemia (FPD/AML)	*RUNX1*	Normal	No	Predisposition to MDS/AML
	ANKRD26-related thrombocytopenia (THC2 or ANKRD26-RT)	*ANKRD26*	Normal	No	Predisposition to MDS/AML
	ETV6-related thrombocytopenia (THC5 or ETV6-RT)	*ETV6*	Normal	No	Predisposition to ALL/MDS/AML
	SRC-related thrombocytopenia (SRC-RT)	*SRC*	Normal	Yes	Myelofibrosis, MDS (trilineage dysplasia)
	GATA1-related thrombocytopenia (GATA1-RT)	*GATA1*	Macro	No	Anemia
	GFI1B-related thrombocytopenia (GFI1B-RT)	*GFI1B*	Macro	No	Myelofibrosis, anisopoikilocytosis
MK maturation	Grey Platelet Syndrome (GPS)	*NBEAL2*	Macro	No	Myelofibrosis, splenomegaly
	Wiskott-Aldrich Syndrome (WAS)	*WAS*	Micro	Yes	Immunodeficiency
	Glanzmann Thrombastenia (GT)	*ITGA2B/ITGB3*	Normal/Macro	No	-
	Bernard Soulier	*GP1BA/GPIBB/GP9*	Macro	No	-
	Platelet Type von Willebrand Factor Disease (PT-VWD)	*GP1BA*	Normal	No	-
	Di George syndrome	*GP1BA*	Macro	Yes	-
	MYH9-related disease (MYH9-RD)	*MYH9*	Macro	Yes	-
	DIAPH1-related thrombocytopenia (DIAPH1-RT)	*DIAPH1*	Macro	Yes	-
	ACTN1-related thrombocytopenia (ACTN1-RT)	*ACTN1*	Macro	No	-
	Filaminopathy A (FLNA-RT)	*FLNA*	Macro	Yes	-
	PRKACG-related thrombocytopenia (PRKACG-RT)	*PRKACG*	Macro	No	
	TUBB1-related thrombocytopenia (TUBB1-RT)	*TUBB1*	Macro	No	-
	TPM4-related thrombocytopenia (TPM4-RT)	*TPM4*	Macro	No	-

Table 1: list of the most common congenital platelet disorders, organized accordingly to the impacted phase of MKpoiesis.

Chapter III

3. Pluripotent Stem Cells, disease modelling and therapeutic applications in haematology

3.1 Pluripotency and induced pluripotency

3.1.1 Pluripotency and its molecular determinants

3.1.1.1. The murine pluripotency

Pluripotency is classically defined as the ability to self-renew indefinitely and to differentiate into the three germ layers (ectoderm, mesoderm and endoderm) and all the derivative tissues[254]. Pluripotent stem cells (PSCs) are typically found in the inner cell mass of an embryo (ICM), so they are not involved in the generation of most of the extra-embryonic tissues. *In vivo,* pluripotency is retained until the gastrulation, when early somatic and germ cells finally emerge from the pluripotent pseudo-stratified epithelium of the epiblast. On the contrary, pluripotency can be captured *in vitro* and maintained for a long period of time (theoretically indefinitely), by isolation and culture of cells derived from the ICM in the short temporal range between the early blastocyst and the gastrulation. These cells are all generically called Embryonic Stem Cells (ESCs). The first ESCs were derived in 1981 from the mouse embryo[255] and were initially propagated on a feeder of inactivated fibroblasts, in presence of fetal calf serum. These requirements have been progressively abandoned for more controlled and specific cell culture conditions: the feeder layer has been substituted by the leukemia inhibitor factor (LIF), while the role of serum replaced by just two molecular inhibitors, added to the culture medium (2i). The LIF is primarily activating STAT3, a TF actively implicated in the self-renewal, while the two inhibitors are acting on the WNT/β-catenin pathway (CHIR99021, an inhibitor of the GSK3 family enzymes) and the Erk signalling cascade (PD0325901, an inhibitor of MEK). Mouse ESCs cultivated in 2i and LIF are less heterogeneous than cells cultivated in serum: this specific setup activates the signalling pathways that regulate precisely the pluripotency-associated factors and any lineage-specific genes, giving a more uniform picture[256] (Figure 20). From the functional point of view, the mESCs are able to integrate a tetraploid blastocyst and contribute effectively to the generation of a healthy and viable offspring, without incurring in any tumorigenesis, a valuable proof that the cells are retaining the control of a correct developmental program. For this reason, they have been called naive or ground state ESCs, to underscore their "blank" condition that could be directed towards any developmental fate. In the absence of any differentiating cues, naive ESCs are a homogeneous entity able to self-renew extensively.

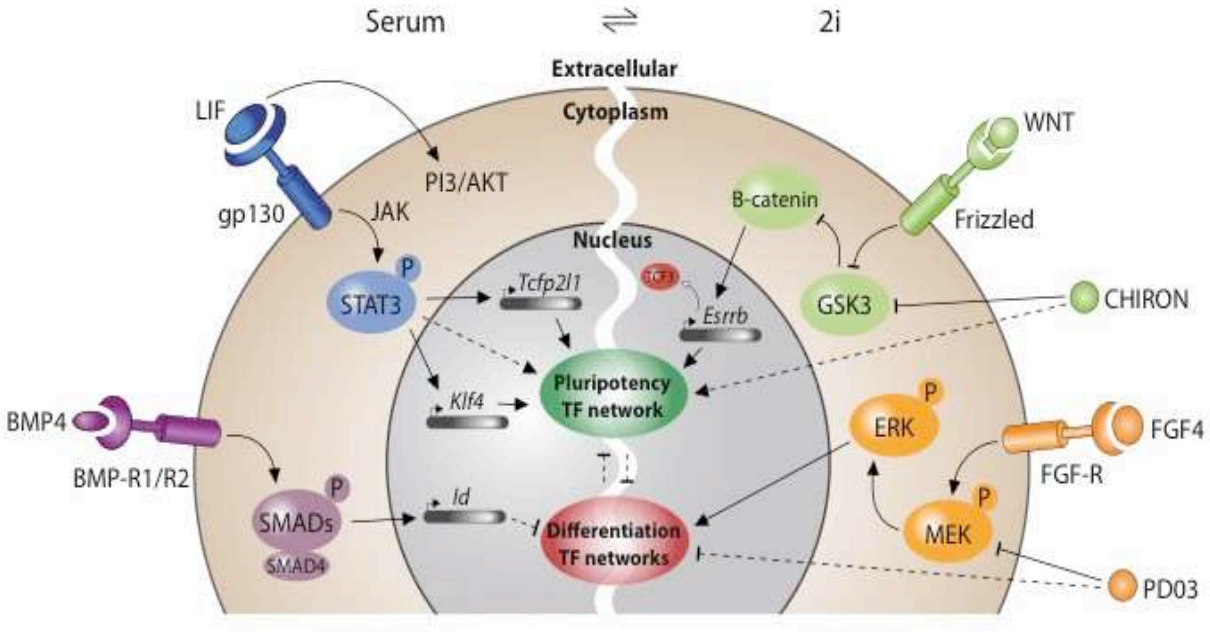


Figure 20: the principal signalling pathways activated during the *in vitro* culture of mouse PSCs. Readapted from Hackett and Surani[257].

Since the pluripotency *in vivo* is restricted to a certain time window, it is not unreasonable to imagine that cells isolated at different time points could display different degrees of pluripotency, notably in their differentiation potentials. In 2007, two teams have been able to isolate and maintain in culture PSCs derived from the post-implantation epiblast, just before the gastrulation[258,259]. The derived cells were called Epiblast Stem Cells (EpiSCs) and were unable to survive in the classic culture setup used for the mESCs, a strong evidence of the intrinsic differences between the two cell types. Moreover, EpiSCs were not able to integrate a developing blastocyst as the ESCs (no healthy and fertile chimaeras), proving a functional restriction in their pluripotency. Nevertheless, mEpiSCs were described as PSCs, as they were still retaining a clear differentiation potential towards the three germ layers, both *in vitro* and *in vivo*. To mark the differences between those two types of pluripotency, the mEpiSCs have been called also primed ESCs, to highlight their restricted pluripotency. While naive mESCs can be converted into primed mESCs by simple modification of the culture conditions, the opposite transition is possible only with the introduction of transgenic TFs univocally associated to the naive state (Nanog amongst others); this sort of unidirectional transition recapitulate the temporal order of appearance of the two states (pre-implantation versus post-implantation[260]). These two discrete pluripotent states are separated by a “continuum of pluripotencies”, where different cell identities can ideally swing towards one of the two states. The differences in the molecular characteristics of naive and primed cells are related to their cellular metabolism, their epigenetic landscape and their transcriptomic profile[257] (Figure 21). The naive state displays a reduced DNA methylation degree and a general "relaxed" chromatin organization

(exemplified by the expression of both X chromosome in XX cells), compared to the primed state where many lineage-associated genes are poised (bivalent chromatin) and there is only one active X chromosome. A textbook example of the different epigenetic status between the two states is the usage of *Oct4* (a classic pluripotent master factor) enhancer: naive cells express the gene through controls of the activity of the distal enhancer, while the primed cells through the proximal enhancer. Minor differences can be observed when comparing the transcriptomic profiles: some TFs are differentially expressed, mostly as a consequence of the culture conditions (LIF+2i versus FGF/Activin for the primed PSCs). Quite interestingly, the metabolism of the two cell types seems to play a crucial role in the definition of the two different profiles[261]. Primed PSCs have a metabolism that evokes the Warburg effect: they are highly proliferative and rely heavily on glycolysis, instead of the more catabolically efficient oxidative phosphorylation (OxPHOS), even in presence of oxygen. On the other side, naive PSCs display a strong mitochondrial activity, a sign of a flourishing OxPHOS. One of the consequences of a strong glycolysis is a larger anabolic activity, with the allocation of metabolite intermediates to biosynthetic pathways for nucleotides, amino acids and lipids. Some of these intermediates are also crucial cofactors for the activity of many epigenetic enzymes, like the HDACs and the DNMTs. The first depends directly on the acetyl-CoA that is produced by the oxidation of the pyruvate produced at the end of the glycolysis (and not used for fueling the trycyclic carboxylic acid cycle or TCA), while the activity of the methyltransferases depends on the availability of S-adenosylmethionine (SAM), a classic product of the amino acid biosynthesis. So different metabolite availability could well influence the state of the chromatin, and primed cells needs a more decorated chromatin than naive cells.

Globally, the transition from one state to the other seems predominantly dependent from the epigenetic status: naive cells are well guarded against differentiation, as their cell fate seems not committed towards any fate different from the pluripotency itself. Advancing towards a more mature developmental condition impose a poised state, where there is still a general pluripotency ongoing, but the cell is more prone to restrain its potential towards a less potent state, transitioning finally towards a true somatic or germ cell fate. This transition could be described as a formative phase, where the PSCs acquire a relative sensitivity to the differentiation cues[262].

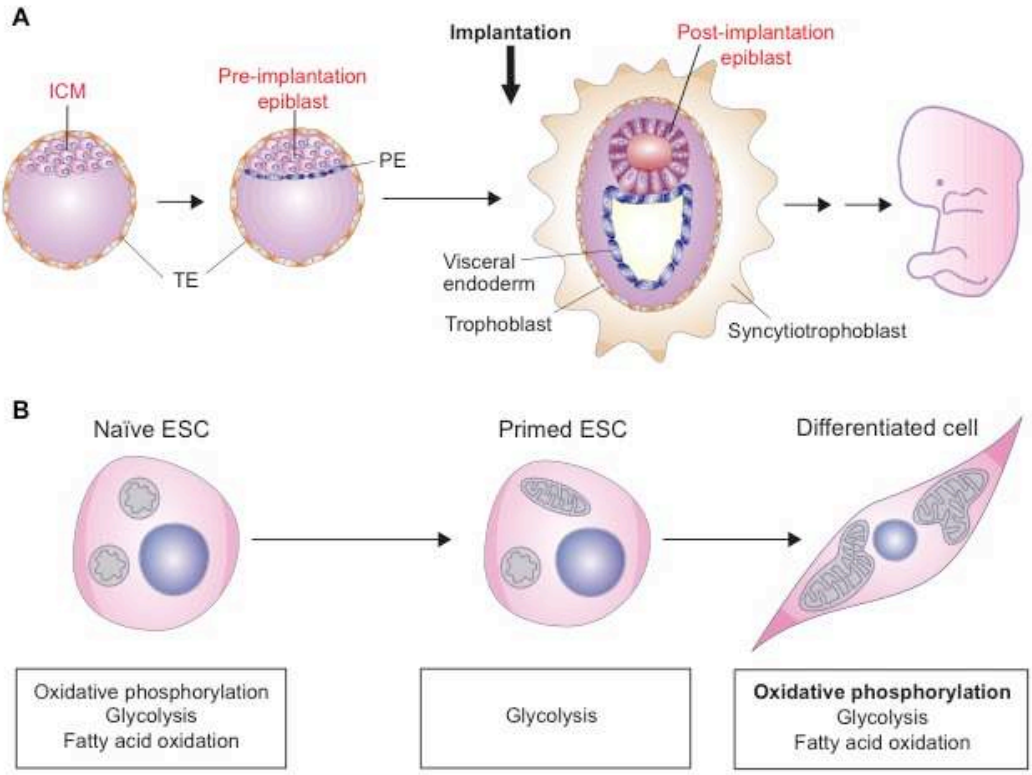


Figure 21: the comparison between naïve pluripotency, primed pluripotency and the more committed state. Readapted from Mathieu and Ruohola-Baker[261].

3.1.1.2. The human pluripotency

The first human ESCs have been isolate, by Thomson and colleagues[263], from the same blastocyst stage as the mESC. Nevertheless, the hESCs do not share the same cell culture requirement of mESCs. Even if both can be maintained on a feeder layer, adding of human LIF cannot replace efficiently the inactivated fibroblasts, because hESCs display a weaker STAT3 activity compared to the murine counterparts. Moreover, the inhibition of GSK3 does not increase the expression of the human *ESSRB* (that it is actually absent in the human naive epiblast *in vivo*) and, in general, does not particularly affect the pluripotency network. On the contrary, a prolonged inhibition of the β-catenin pathway conduces to a self-renewal loss. Surprisingly, the hESCs culture requirements are closer to the mEpiSCs ones, with a certain reliance on the FGF-dependent signalling pathway. The molecular mechanisms associated to the human pluripotency are generally conserved, albeit some differences are retained between the human and mouse. The core transcriptional network, directly responsible of the maintenance of the pluripotency transcriptional program, is fundamentally the same, with the main actors OCT4, SOX2 and NANOG (OSN) as the core essential factors of the

entire auto-regulatory loop. Together, they sustain the self-renewal and block the differentiation programs, mostly by co-localization of the entire triade on several genomic regions. Half of the genes, bound by OSN at their respective promoters, are actively transcribed in hESCs, while the other half (enriched in lineage-specific regulators) seems to be negatively regulated. Those regions in particular carry a bivalent chromatin state of active (H3K4me3) and repressive (H3K27me3) marks, reminiscent of the poised state of the mEpiSCs. In fact, OSN and the PRC2 targets overlap in a significant way; moreover, the PRC2 is proved to interact directly with PRDM14, a TF that regulate directly the expression of *OCT4*, via its proximal enhancer. The activity of the core pluripotency network is directly dependent from the FGF signalling, via induction of the *NANOG* expression. Also involved through its downstream effectors SMAD2/3, the activity of the transforming growth factor β (TGFβ)/Activin/Nodal cascade seems to converge on the regulation of *NANOG,* amongst other pluripotency factors. Indeed, inhibition of the TGFβ pathway rapidly induces neuro-ectodermal differentiation, while it does not affect greatly the mESC naive state[264].

These and other evidences, pointing to the similarities between the hESCs and the mEpiSCs, suggested that the hESCs are in a primed state, prompting the search for an *in vitro* human equivalent of the mESCs naive. This new type of hPSC should display a profoundly different metabolic profile, different functional capacities and finally, a reduced priming of the epigenetic landscape, exemplified by the expression of both X chromosomes. Several groups tried recently to create this transition, by transgenic induction combined with different cell culture conditions[265], but one of the major difficulties is linked to the way the different groups choose to assess the degree of pluripotency of their cells. The characterization could be performed at both molecular and functional levels. Molecular markers of pluripotency include the X inactivation status, the levels of DNA methylation, the expression of a panel of univocally associated TF and the *OCT4* enhancer usage. Taking in account the inevitable differences between human and mouse (notably for the markers expression), some reports confirmed a different molecular profile for the so-called naive hESCs, a profile that could be loosely associated to the naive mESCs. Nevertheless, only the functional tests could establish the true developmental potential of a cell: a range of different assays can be deployed, in a progressively more stringent manner. The less restrictive assay is the *in vitro* differentiation: the culture conditions used for the pluripotency maintaining are replaced by differentiation-inducing cues and the expression of molecular markers associated to the three germ layers is subsequently assessed. After this first test, the cells are evaluated for the *in vivo* generation of a teratoma, an embryonic tumour containing cells derived from the three germ layers. While these two strategies are quite effective for the routine pluripotency evaluation, there are at least three other tests that are even more stringent, but they demand the effective contribution of the

PSCs to the generation of a chimaera, both in its somatic and germ component, a task that require a certain cross-species compatibility and it is subject to a high technical variability that reduced greatly the robustness of the assays. In conclusion, to this date, there is no functional proof of a *bona fide* naive hPSC.

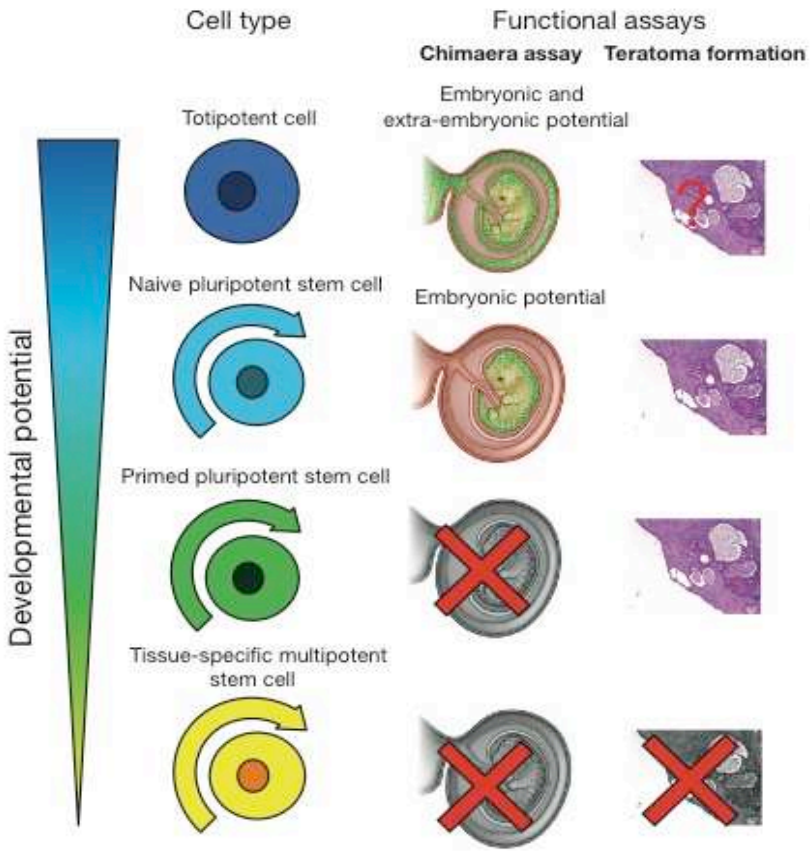


Figure 22: a brief resume of the classic requirements to assess the pluripotency status. Readapted from De Los Angeles and colleagues[254].

3.1.2. The induced pluripotency

The derivation of stable pluripotent cell lines from an embryo it is not the only way to generate PSCs. Scientists developed different strategies for the artificial derivation of PSCs (Figure 23). The more direct alternative consists in the cellular fusion of an already pluripotent stem cell with a non-pluripotent stem cells, in order to create a hybrid cell, albeit with a high degree of instability[266]. This type of demonstration proves that the PSCs have the ability to induce a modification of the differentiation potential of the PSCs, although it is unclear by which mechanism. In a more convincing way, the somatic cell nuclear transfer (SNCT), pioneered by the Nobel Prize John Gurdon in the late 1950s, demonstrate univocally that the nucleus of a somatic cell (the genomic content) could be reprogrammed to a higher potency state (notably a totipotent state) by the cellular content of a higher potency cell, e.g. a egg cell[267]. This seminal body of work culminated in the cloning of entire mammalian organisms[268] and the derivation of embryonic stem cells[269], supporting the notion that development is a sequence of reversible epigenetic modifications, while the genome itself retains its totipotency.

The third alternative approach to the generation of PSCs is the *in vitro* reprogramming by defined factors: a non-pluripotent stem cells could be induced to become pluripotent by the imposition of a genetic program, notably by the exogenous expression of a defined combination of genetic factors. The idea of inducing a new cellular identity by rewiring of the transcriptional landscape stems from the work of Lassar and Graf, who both proved in different systems (fibroblasts and lymphocytes) that the overexpression of a single TF could induce a cellular transition toward a developmentally related or unrelated cell type[270]. Building on this theoretical framework, the research for a single or multiple factors able to establish a pluripotent transcriptional identity led to the seminal discovery of the induced pluripotent stem cells (iPSCs) by Yamanaka and Takahashi[271]. The painstaking research of the minimal combination of TFs able to induce the conversion from a somatic identity to a pluripotent one resolved in the identification of four TFs: Oct4, Sox2, Klf4 and c-Myc, also defined as OSKM. The same combination, described first in mouse, was proved to be able to generate iPSCs also from other mammalian cells, including human cells[272]. Since this landmark discovery, other combinations of TFs and/or miRNAs have been described, but they all rely on the reactivation of a fundamental pluripotent circuitry that include Oct4 and Sox2, not very dissimilar from the fundamental transcriptional network identified in the embryonic stem cells. Originally, the reprogramming factors were delivered by retroviral vector-mediated over-expression and successively silenced by DNA and histone methylation. This system display a very low efficiency (between 0,1 and 0,01%) and could fail to induce the endogenous expression of the pluripotency

factor, leading to an incomplete reprogramming and forcing the cells to rely on the indefinite exogenous expression of the OSKM factors. In a more efficient and safer way, the use of non-integrating expression systems has progressively replaced the retrovirus, as technological platform for the reprogramming: the cited systems included single-strand RNA virus and mRNA encoding for the OSKM factors. Reports of small molecules-mediated reprogramming[273] and the use of activating antibodies in substitution of certain TFs[274] opened new possibilities for pluripotency reprogramming, although those systems need to be further validated.

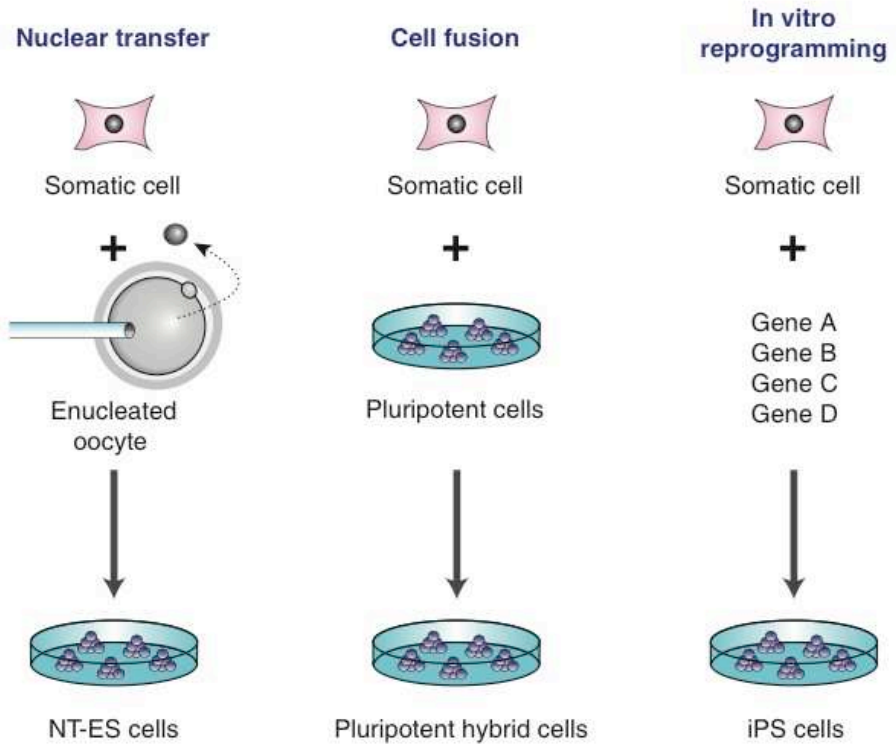


Figure 23: the three different approaches to artificially generate a PSCs. Readapted from Hochedlinger and Jaenisch[272].

The entire process of reprogramming relies on the reactivation of endogenous loci that are plausibly in an unfavourable state, as somatic committed cells do not express transcription regulators associated to pluripotency. Then, it is remarkable how the simple ectopic over-expression allows the cellular conversion in such robust and reproducible way. The pulse gave by the TFs leads to the down-regulation of the molecular markers associated to the cell type of origin, in the context of a general loss of somatic characters. Successively, the cells acquire a series of markers loosely associated to pluripotency (notably SSEA-1), to culminate in the expression of endogenous, bona-fide pluripotency genes, like *OCT4*, *NANOG* and, in the late stages, *SOX2*. The transition to pluripotency progresses through successive bottlenecks, with several molecular roadblocks, and it is at the origin of the very low reprogramming efficiency. The inability to overcome those blockages could be irreversible (failed reprogramming), requiring a longer latency or the presence of

additional factors (TFs, epigenetic regulators or miRNAs) to the OSKM combination. Successive boosts of OSKM could work too, but they also increase the risk of a persistent dependency of the reprogrammed cells to the exogenous factors (incomplete reprogramming). It is also very plausible that exogenous OSKM expression must be turned off to obtain a successful reprogramming, accordingly to the observed epigenetic silencing observed in the retroviral-based reprogramming. Those experimental evidences support a model of a stochastic nature of the reprogramming process, opposed to the idea of an "elite" population of rare cells closer to the pluripotency (Figure 24).

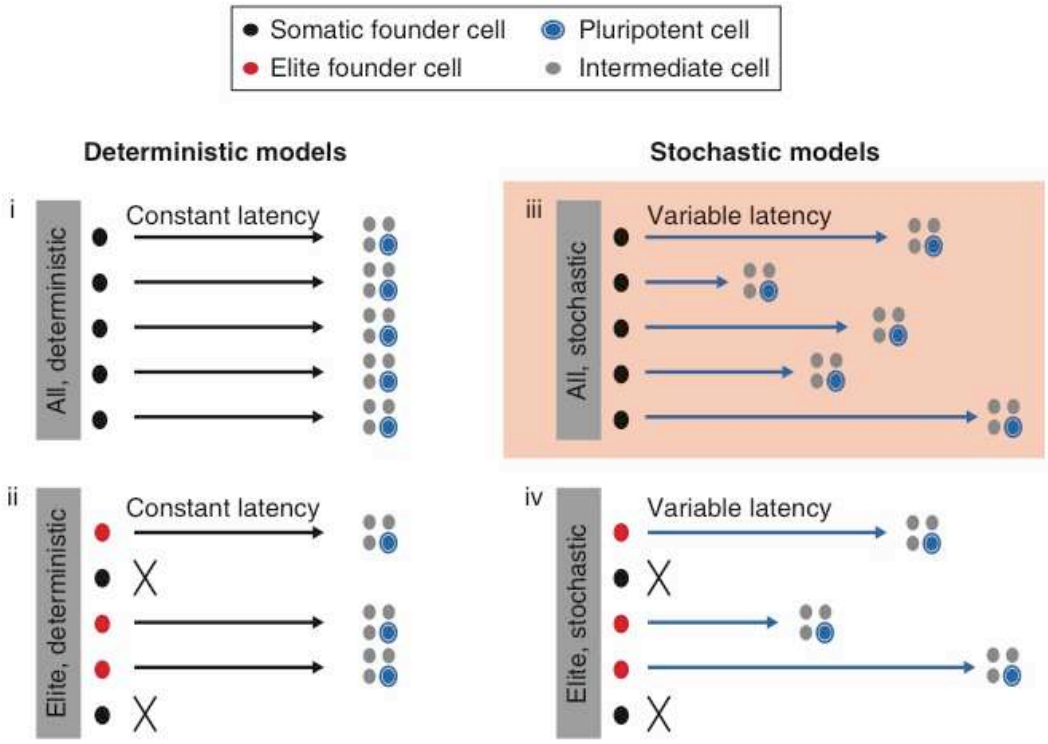


Figure 24: the different models to explain the mechanism of pluripotency induction. In pink the currently accepted one. Readapted from Hochedlinger and Jaenisch[272].

The understanding of the molecular actors involved in the reprogramming steps has been hampered by the intrinsic heterogeneity of the entire process, combined with the global low efficiency. The development of the "reprogrammable mice" or secondary reprogramming systems[275] for human cells allowed a more homogenous view of the transition and detailed mechanistic studies on the role of the OSKM factors in the early phase of the reprogramming. The more accepted view is that OCT4, SOX2 and KLF4 (OSK) act as pioneer factors, binding to very closely packed chromatin regions thanks to their ectopic high expression levels, with MYC acting as modulator (nonlinear amplifier) of this action, so easing the deployment of OSK on the chromatin[276]. Moreover, cells actively proliferating are more prone to reprogram, as the activation of the new transcription network seems to be more easily implemented by successive rounds of cell division.

Immediately after the discovery of iPSCs, one of the most urgent questions revolved about the comparison of this artificially generated cell type with their closest counterpart, the ESCs generated from the blastocyst. One of the biggest strength of the pluripotent reprogramming is its robustness: almost every cell type could be reprogrammed. Nevertheless, our understanding of the reprogramming process highlights how the cell type is affecting the efficiency and the completeness of the process. Are the newly generated iPSCs retaining any sort of "memory" of their cell of origin? While the first reports proposed a significant impact of the somatic origin on the iPS cell line epigenetic status[277], the current view, corroborated by large screening of several cell lines of different origins, describe a major impact of the genetic variability between individuals as the primary source of differences between iPSCs, with a much reduced importance of the epigenetic memory[278,279], merely reduced to non-functional differences.

Another major point of discussion related to the reprogramming is the mutagenicity associated to the process. Excluding the possible mutagenesis associated with integrative expression systems for the introduction of the OSKM factors, is the reprogramming increasing the mutational status compared to the original cell population? Like for the epigenetic memory, the first results proved that OSKM conversion cause a certain level of genomic alterations, mostly copy number variations (CNVs) arising during the early stages of cell culture and selected by the long-term culture[280]. Indeed, the continuous stress associated to the cell culture is a burden affecting the iPSCs genomic stability[281], although it is difficult to conclude if this is a specific characteristic of PSCs. But the recent development of efficient single-cell analysis techniques, combined with a more thorough approach, concluded that most of the mutations observed in newly derived iPSC clones are present in the starting cell populations, with only a reduced number of differences arising *de novo*[282]. Indeed, there is a higher than expected degree of mosaicism for certain tissues[283], thus the analysis of the iPS cell lines derived from those tissues could lead to the misleading observation that several *de novo* variants arise during the reprogramming transition. Another parameter that should be accounted for is the donor age, as there is a clear linear relationship between the age of the donor and the number of mutations, at least for the haematopoietic compartment[284]. Surprisingly enough, after a certain threshold (>90 yrs) there is a clear reduction in the number of mutations, probably as a consequence of a reduced cell pool of progenitors.

On the same issue of the ESCs-iPSCs comparison, it is of great interest to establish how the reprogramming process is inscribed in the framework of pluripotency (the naive versus primed state). To which type of pluripotency are the iPSCs associated with? Clearly, the culture conditions are dictating the pluripotency state, similarly to the ESCs; moreover, the general shut-down of the exogenous OSKM factors required for the correct reprogramming is at odds with the need for the

sustained expression of transgenic factors associated to the primed-to-naive conversion, at least for human cells. It is also necessary to remind that there are some crucial differences between human and mouse pluripotency, so is it probable that some differences could arise also during the reprogramming process. One possible tool to monitor the transition would be the X chromosome inactivation status, a classic example of epigenetic transcriptional regulation (Figure 25). In the mouse case, the late stages of reprogramming show a re-activation of the gene expression for the alleles located on the inactive X-chromosome (Xi), approximately at the same time of the activation of the endogenous *Sox2* expression. This is associated to the loss of expression of the *Xist* RNA, the major responsible of the inactivation; and it is common to the naïve pluripotency state observed in the mESCs cultivated in 2i medium. It is then clear that the reprogramming, in the mouse system, could leads to a naive pluripotent cell line. In the human case, the situation is less clear: excluded the controversy on the naive state in the hESCs, it has also been highlighted how the inactivation status seems to be dependent from the culture conditions and the time spent in culture[285]: the Eggan's lab coined the term of "X-chromosome erosion" to point out how this process is time-dependent and impacting only certain area of the chromosome, leading to the partial re-activation of only certain loci on the Xi. This erosion seems to be irreversible after differentiation or reprogramming and not associated to a different pluripotent state, as hiPSCs with an eroded X chromosome are still dependent from bFGF/Activin A, so still in a primed pluripotency. No evident genetic aberration seems to be associated to the erosion, confirming it as a pure epigenetic phenomenon[286]. Recently, the Rougeulle's lab tried to unravel the mechanisms behind the chromosome erosion, highlighting a competition between two lncRNAs, *Xact* and *Xist*: the former intervenes before the loss of *Xist* coating, mirroring the X inactivation during the early development[287,288].

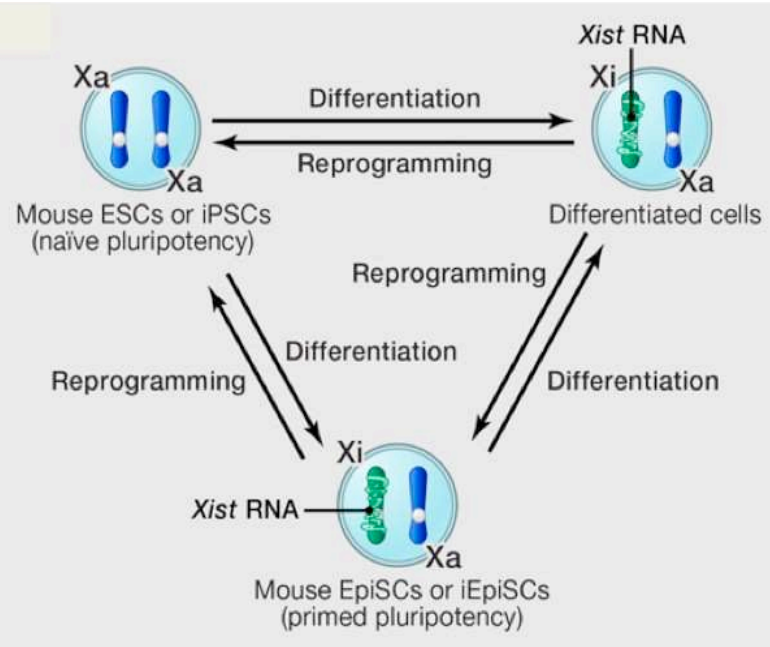


Figure 25: the different status of the X chromosome in pluripotent and differentiated cells. Readapted from Papp and Plath[276].

3.2. How pluripotent cells changed disease modelling

3.2.1. On the importance of the differentiation protocols

The impact of PSCs on haematology and disease modelling is directly dependent on the effective use of differentiation protocols that are reliable, faithful and expandable. Theoretically, a PSC should be able to give rise to every cell type specified by the three germ layers, so there is no evident reason to deny the possibility to generate every haematopoietic cell, included the stem cell pool. To support this theoretical assumption, it has been demonstrated that functional progenitors with long term-reconstitution potential are present in teratomas generated from the injection of PSCs in immune-deficient mice[289]. Nevertheless, to this day there is no formal proof of the *in vitro* generation of a functionally defined haematopoietic stem cell, intended as a cell able to reconstitute the long-term haematopoiesis in a primary recipient. This endeavor has captured the attention of many laboratories, proposing different experimental strategies, but generally drawing heavily from the knowledge of the intimate changes occurring during the embryo development. Some teams privileged a rigorous recapitulation of the *in vivo* transitions occurring during the embryonic development, preferably without the introduction of exogenous genetic elements, in a theoretical and experimental framework called direct differentiation. On the other side, there is a body of work claiming that it is not necessary to model the exact series of *in vivo* events, but that a combination of genetic and epigenetic modifications is sufficient to alter the cellular fate of a selected cell of origin, directing it toward an haematopoietic identity with an expanded set of differentiation and self-renewal properties. This approach is called cellular conversion and reminds the reprogramming to pluripotency in its core principles.

The direct differentiation approach rely on two major assumptions: the activation of different cellular pathways by extrinsic means should be sufficient to induce the necessary cellular transitions to the final product, and the *in vitro* observed phenotypes are recapitulating the *in vivo* transitions that occur during the embryonic development. This last assumption in particular require a pairwise comparison between isolated populations from both sources, and needs rigorous analysis of the cellular phenotypes and functional properties. At the moment, while there is a constant increase of new cellular markers, more or less specific, a distinct lack of effective functional assays is a real bottleneck: there is no valid *in vitro* surrogate of the haematopoietic reconstitution in a primary recipient.

For the definition of the differentiation protocols, historically the first experimental systems proposed were based on co-culture with different stromal cells, in a serum-containing medium and in presence of instructive cytokines[290,291]. This system allowed the differentiation of several

haematopoietic populations, generally re-adapting the differentiation protocols defined for adult haematopoietic progenitors[292]. Although straightforward and efficient for general purposes of disease modelling, is clearly lacking a fine-tuned control of the pathways involved in the efficient instruction of the cellular fates. The development of more advanced culture media, together with 3D and 2D differentiation protocols, allowed the switch from a serum-containing setup to a serum-free and chemically defined system, where it is possible to control the amount of each signal and greatly reduce the experimental variability.

To summarize, it is possible to describe all haematopoietic differentiation protocols in a two-steps procedure: mesoderm induction from a pluripotent source and haematopoietic specification. The study of the developmental haematopoiesis proved, in a reasonable fashion, that the haemangioblast is probably not involved in the generation of a definitive haematopoiesis. Moreover, it informed greatly about the definition of the factors necessary for the patterning of a mesoderm cell intermediate that could yield a definitive-like developmental program[293] (Figure 26). Four signalling pathways have been clearly implicated: BMP4, WNT/β-catenin, FGF and Nodal-activin[294-297]. These four pathways are crucial in the induction of the primitive streak and the subsequent specification of a migrating mesoderm (patterning). This last series of events, *in vivo*, appears to be clearly dependent from the positional gradients ensured by different cell-cell interactions and it is remarkable how an *in vitro* system (so completely deprived of any positional information) can be so closely related. Even if there is a high degree of conservation of these pathways, some species-specific differences have been noted, mostly related to the dynamics of the patterning[298], underlying once again the relevance of human models for the study of human development. The dynamic nature of development also requires a thorough assessment of the stimulation length for each factor: Sturgeon and colleagues proved quite convincingly how the primitive haematopoietic developmental program is dependent on the activation of Nodal-Activin and the inhibition of WNT/β-catenin. Vice versa, a definitive haematopoiesis requires the activation of WNT and the blockage of Nodal-Activin. They were also able to identify a combination of markers (KDR and Glycophorin A) to distinguish the primitive ($KDR^{+}GPA^{+}$) from the definitive progenitors ($KDR+GPA^{-}$)[299]. Follow up studies from different groups reported that the expression of the HOXA cluster, mediated by the members of the CDX family, is restrained to an intra-embryonic developmental program, so reasonably include the instructions for the specification of the AGM-related progenitors[300-302].

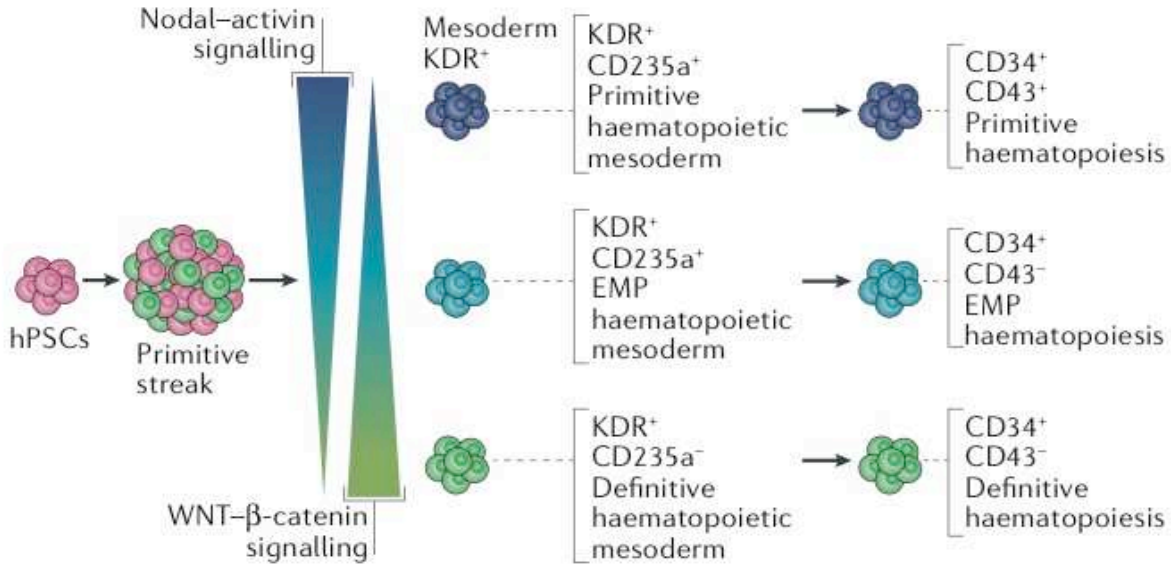


Figure 26: the gradients of Activin and WNT signalling are crucial for the derivation *in vitro* of the different type of HEs. Readapted from Ditadi, Sturgeon and Keller[40].

If the study of the mesoderm patterning allowed a first important step forward, the transition toward a haematopoietic progenitor it is still an open question, because of the convergence of several confounding factors. The first one is that *in vitro* it is not possible to discriminate spatially the yolk sac hemogenic endothelium (leading to EMP and LMPP) from the AGM hemogenic endothelium (HSC), and there is no specific marker for each one of them[303,304]. The second is that the two progenitors seems to be specified very closely in time, so it is technically challenging to discriminate them based on the emergence time. To exit this conundrum, one way is to better understand the process of endothelial to haematopoietic transition (EHT), in order to gain enough insights and highlights eventual differences between the different hemogenic endothelia (Figure 27). Two assumptions are necessary to accept this theory: first, the hemogenic endothelium is intrinsically different from the hemangioblast; second, the hemogenic endothelium is actually a transitional form of a mesodermal derivative with haematopoietic potential, that retains some phenotypic aspects of an endothelial cell, but is devoid of any differentiation potential in that sense. This theory would explain while lineage tracing experiments did not showed a major contribution to both endothelial and haematopoietic compartments, from the AGM-bound progenitors[305]. A logical consequence of this approach is that the bifurcation between haematopoietic and endothelial fate must occurs at an earlier stage, before the emergence of a functional endothelial program. To support this model, Ditadi and colleagues proved, in a culture system supporting both the haematopoietic and the endothelial differentiation, the existence of three different populations, engaged towards a haematopoietic fate ($CD34^+CD73^-CD184^-DLL4^-$) an arterial endothelium fate ($CD34^+CD73^+CD184^+$) and a venous endothelium fate ($CD34^+CD73^+CD184^-$). Effectively, the

haematopoietic-bound population proved to be endothelial-like immediately after isolation (no expression of any haematopoietic marker like CD43 or CD45), and subsequently becoming haematopoietic (more than 90% cells expressed the $CD43^+$ marker). A subset of these haematopoietic cells was assayed for their differentiation potential *in vitro*, showing a multipotent phenotype (T-cell and myeloid differentiation[306]). Nevertheless, the identification of the mesoderm common ancestor between endothelia and haematopoiesis is still missing, so the prove that the cells derived were actually LMPPs or pre-HSCs. In conclusion, to this date there are no clear evidences of a different EHT mechanism between the AGM-bound hemogenic endothelium and the yolk sac hemogenic endothelium[307].

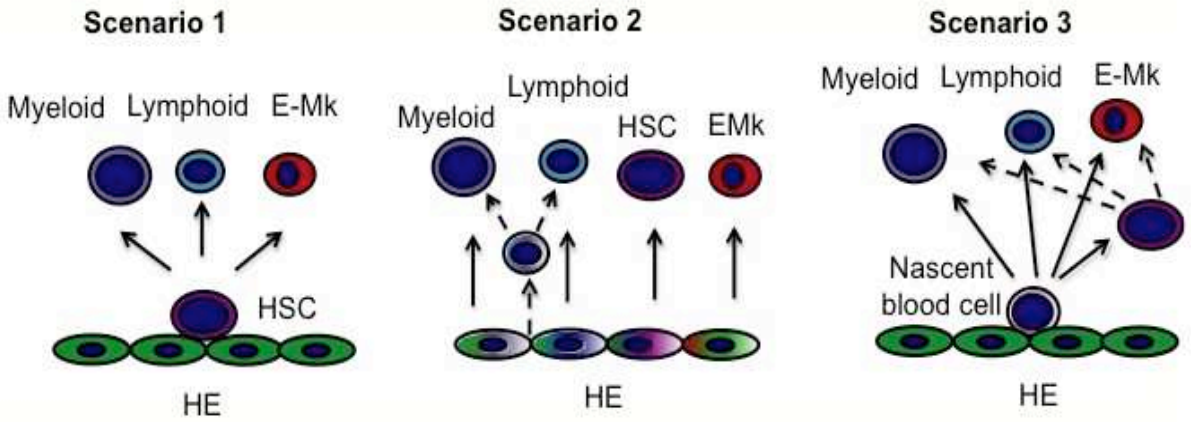


Figure 27: three different possible models for the EHT. It is worth to remind that the HE in vitro could not be able to perform a valid EHT, therefore it is crucial to compare the populations obtained in vitro with the closest counterparts in vivo. Readapted from Slukvin[307].

How we can improve these differentiation protocols? There is an urgent need for better *in vitro* readouts to assess the differentiation potential of a cell population, as the *in vivo* reconstitution assay remains an experimentally unpractical option: for example, while the globin expression pattern or the T-cell differentiation are features shared by EMPs, LMPPs and HSCs, the B-cell differentiation could be used to separate these populations, as the B-2 subtype is the product of a HSC differentiation only and is distinguishable from the B-1 subtype. This assumption was challenged recently though, drawing a common clonal origin for both subtypes[308]. As the *in vitro* maturation step (co-culture) could allow the maturation of pre-HSCs into long term engrafting progenitors[35,36], it is evident that there is a substantial lack of knowledge of the precise molecular pathways involved in the HSC maturation. Another investigation line consists in the evaluation of how different pathways, like the retinoic acid, the interferon-α and NF-kB dependent pathways, impacts the HSCs maturation. A pro-inflammatory signature has been described from HSCs of mouse and zebrafish[309], although no data is available currently about humans. Finally, it is possible

that the EHT could be mimicking a similar biological mechanism, the endothelial-to-mesenchymal transition (EMT): a recent paper from Roger Patient's group proved, in zebrafish, how similar these two phenomena are, and suggests an important role of TGFβ in the EHT[310].

The success of lineage conversion by cellular reprogramming prompted an alternative approach to the direct differentiation: instead of manipulating the signalling pathways with extracellular factors, is it possible to induce a fate conversion by the introduction of TFs or other epigenetic regulators, generating in this way an haematopoietic cell and possibly an HSC? This approach, championed by a handful of labs, has recently gained supports for its straightforward nature, as it is potentially less subject to the constraints of a step-wise differentiation approach. Furthermore, the cellular conversion is not strictly dependent from a PSC source (as it is not recapitulating the normal development), as it could be very sensitive to the epigenetic landscape of the cell of origin. In this sense, a cell ontogenically close to the target cells, like a committed myeloid precursor or an endothelial cell, is reasonably a better candidate than other cell types.
Several TFs have been described as good candidates for the role of reprogramming factor: most of them are already known to play a major role in haematopoietic derivation or specification, while some other candidates have been retrieved by comparative transcriptomic studies between the cell of origin and the cell of interest. Nevertheless, the first team to propose a cellular conversion toward a haematopoietic progenitor did not use any of these candidates, but a partial induction toward pluripotency, by the introduction of the pluripotency factor Oct4, then a haematopoietic specification by the role of instructive haematopoietic cytokines[311]. Similarly, Pulecio and colleagues proposed a different combination of TF and miRNA (Sox2 and miR-125b), but still comparable overall[312]. Both strategies were proposed in generally defined human fibroblasts, with few details about the nature of these cells of origin. These two examples of reprogramming offers some degree of engraftment and *in vitro* proofs of a haematopoietic multi-lineage potential, but no clear identification of a true HSC. A more refined approach bypassed the pluripotency induction, favouring a cocktail of TFs associated to haematopoiesis: two studies have been performed in murine fibroblasts, with two different cocktails (although they retained some common factors like Gata2). They both achieved the specification of multipotent haematopoietic progenitors, but they failed in a long-term engraftment test[313,314]. Curiously enough, these four different works reported an EHT-like conversion, with the specification of cell populations expressing both endothelial and haematopoietic markers; however, it is not known how similar this transition could be with the *in vivo* observed counterpart, or the *in vitro* EHT from direct differentiation protocols. A molecular characterization of the reprogramming events was proposed, as a three-steps process of fibroblast signature down-regulation, followed by the induction of an endothelial-like program and

haematopoietic specification. No clear parallels were drawn with the *in vivo* specification.

An interesting advancement was proposed jointly by the teams of Scandura and Rafii, using a cocktail of four TFs (*RUNX1*, *GFI1*, *SPI1* and *FOSB*), together with a specific co-culture setup of the target cells (human or mouse endothelial cells of fetal and adult origin) and a stromal layer of engineered endothelial cells, able to generate a vascular niche supplying angiocrine factors. This approach provided, for the mouse experiments, the generation of multipotent progenitors able to engraft primary and secondary recipients, capable of an immune adaptive response (so a lymphoid engraftment). The differences in engraftment properties between human and mouse cells could be attributed to species-specific features, or actually underlies a genuine different reprogramming efficiency. The clear endothelial nature of the cells of origin supports the idea that an EHT transcriptional program could be superimposed on any endothelial background[315,316]. All the presented approaches assumed that an *in vitro* setting, combined with the exogenous factors, was sufficient for the maturation of the haematopoietic progenitors. In this regard, Riddell and colleagues took a different path: starting from a haematopoietic progenitor (lymphoid or myeloid in nature) and a complex cocktail of 8 factors, they performed the maturation step in an irradiated recipient. They were able to observe a substantial engraftment, although they cannot claim convincingly any mechanistic insight[317]. Nevertheless, the *in vivo* set-up could be key to achieve a reliable maturation of any reprogrammed cell into a long-term haematopoietic stem cell.

Although the promising premises of the cell conversion approach, to this day the results are still surrounded by a certain lack of clarity and moderate skepticism. While a similar approach could reveal itself as efficiently producing functional HSCs, there are still several procedural hurdles to overcome. The more evident is linked to the selection of the reprogramming factors: while transcriptomic comparative studies could be a good starting point, the fate conversion imply a large number of modifications of the epigenetic landscape, modifications that could be possibly difficult to induce by the introduction of only a handful of TFs. In-depth analysis of the chromatin status would be desirable for future reprogramming studies. Strictly related to the reprogramming factors selection is the control of their activity: dependent from the experimental modalities chosen, a fine-tuned control of their activities is highly desirable. It is reasonable to imagine that the prolonged activity of factors linked to a state of multipotency could be deleterious for the activation of any downstream differentiation protocol. Furthermore, a recent report showed how it is possible to imagine a very dynamic activity of a master TF. Eich and colleagues described a highly unstable *Gata2* dynamics in the HSC-specification from the AGM, suggesting how complex could be the picture, with a continuous rewiring of the transcriptional landscape as a necessary condition for the

correct HSC maturation[318]. It is also very important to mention that many of the TFs proposed in the various approaches are known oncogenes or play a role in several leukemogenic programs, so an incorrect expression could have damaging consequences on the HSCs fitness.

To this day the production of HSCs, defined as cells able to fully engraft a recipient for a significant amount of time, is still out of reach. Nevertheless, significant advances have been made in the field, confirming the huge impact of those *in vitro* studies on the understanding of the human development. While the advent of single-cell resolution techniques would surely bring new insights into the cellular transitions observed in the dish, the employment of more refined cell culture methods and gene transfer systems will surely improve our capacity to manipulate the cellular machinery. The gold assay for the assessment of the stemness remains the *in vivo* transplantation though, and the reports of the successful reconstitution of the haematopoiesis in recipients injected with intermediates that are not haematopoietic, are of great relevance for the future directions. It is highly possible that the current *in vitro* settings are unable to guarantee the acquisition and/or retention of the stemness properties for human cells, as a consequence of the reduced knowledge of the *in vivo* mechanism of haematopoietic specification.
There are also other variables that should be considered. Is is well established that the onset of the blood circulation is temporally coincident with the crucial steps of the HSCs maturation and specification. Therefore, it is plausible that the circulation-associated shear stress could yield very informative insights on the EHT and lineage specification, for example by the activation of the PKA-CREB signalling, dependent on the cAMP production[319,320]. It is also important to consider that the signalling necessary for the developmental processes in analysis could be provided by other means than cell-cell interaction or paracrine stimulation. A recent paper underlies how the developing hypothalamic/pituitary/adrenal axis is actually contributing to the production of an essential signalling factor for the HSC emergence[321]. This and other important insights will require the comprehensive use of both model organisms and *in vitro* strategies.

3.2.2. Are the induced pluripotent stem cells a valid model for haematopoiesis?

The shortcomings of the *in vitro* differentiation protocols highlighted in the previous section prompt a valid objection about the importance of PSCs in the study of haematological diseases. Although there are several conditions that involve haematopoietic progenitors of embryonic and fetal origin, the large majority of the pathological haematopoiesis involves adult stem cells and progenitors with an altered fitness. If the differentiation protocols are unable to produce those types of cells, what is the impact of the model on the understanding of the pathological mechanisms? To answer to this question, I would like to debate the experimental limitations and strengths of the model, using some examples from the literature.

The first and most important limitation of the model is the variability[322]: compared to a murine model or a different organism, the iPSC model is subject to an intrinsic variability, comparable to the use of primary samples from patients. The use of iPSCs as disease models stems from the possibility to study in a dish a particular genetic background, highlighting any particular phenotype associated with a specific genotype. To account for the unavoidable phenotypic variability between individuals, it is necessary to "dilute" the genomic background effect by the inclusion of multiple genotypes in the experimental design, in order to highlight the common phenotypes and connecting them to the genetic variants. In alternative, it is possible to select a specific genotype and introduce a variant of interest by genome editing, typically using a nuclease-based strategy; the comparison would be performed between paired samples, providing an isogenic or nearly isogenic model[323].
Due to its clonal nature, the cellular reprogramming process is generating a collection of clonal cell lines from each patient, magnifying differences that could be hidden by a population-based analysis. It is acknowledged that two cell lines derived from the same individual are more similar than two cell lines derived from two individuals[324], but is still possible that some differences could be present. While the reprogramming process (in itself or the technology used to perform it) should be preferentially tailored specifically for the type of genetic composition in analysis[325], it has been proved that the observed differences between each line, issued from the same genotype, are pre-existing in the tissue of origin, and not induced by the reprogramming to pluripotency[282,326]. Nevertheless, long time culture and differentiation are major sources of mutations[281], and appropriate precautions must be considered for a high quality experimental design: no general guidelines are yet been drafted or discussed in the scientific community about the experimental aspects of iPSCs culture, a considerable issue when the reproducibility of the scientific results is took under consideration.
But the intrinsic variability of the iPSC model could also be exploited for the understanding of the

clonal evolution of a disease and the impact of different mutations in the same background. This is particularly true for the study of malignant haematopoiesis, where the iPSC reprogramming allows the generation of cell lines harbouring specific mutations associated to defined sub-clones, isolating *de facto* the sub-clone[327,328]. The isolation of several sub-clones could be an invaluable resource for the study of a specific epigenetic landscape[329], drug sensitivity[330] or disease progression[331], without the requirement of complex genomic manipulation and generation of highly complex animal model.

The second major limitation of the iPS modelling strategy is its *in vitro* nature: it is evident that the model is not suited for the analysis of the interaction between the haematopoietic cells and their microenvironment, mostly because of the lack of experimental tools able to efficiently model the haematopoietic niche in a dish. This issue, common to other experimental approaches, like the use of primary cells sampled from patients or immortalized cell lines, has been addressed by the generation of artificial humanized niches *in vivo*, both by tissue engineering[332] or genome engineering[333,334]. The latter strategy in particular proved effective for the study of specimen historically refractory to the long-term engraftment in a murine recipient[335,336]. In the context of the xenograft, only iPSC-derived haematopoietic cells harbouring potent oncogenes (MLL translocation) have been reported to engraft a recipient, and this is the direct consequence of the lack of efficient differentiation protocols for HSCs. On the other side, the pluripotency of the model could be key for the generation of a human cellular microenvironment, as it will be possible to differentiate several cell types and possibly combine them in a model of a human haematopoietic niche. The niche would be suitable for the study of the molecular interactions between haematopoietic and non haematopoietic cell types, in a highly controlled experimental setting, as it will overcome the intra-species differences of the more common co-culture systems based on murine cell lines.

Finally, an iPSC model is undeniably a time and work-consuming experimental system, as it requires special expertise for its technical handling, contrarily to most of the other available cell lines. A major effort in the development of more refined and reliable techniques for reprogramming and culture allowed the diffusion of the model. Nevertheless, the differentiation protocols are still in their technological infancy; consequently, a specific care is required for the generation of the cells of interest. In any case, the iPSC model has become one of the very few amenable experimental systems for the study of rare haematological diseases: tens of different lines have been established from patients harbouring single congenital mutations affecting one or more haematopoietic lineages[172,337,338]. These resources are a very reliable tool for the understanding of the mutations impact on the cellular biology and physiology of particular haematopoietic precursors, in a context

that is human-specific and easily scalable, two clear advantages compared to murine models and primary cells.

A good example of how an iPSC model could reliably informs on the pathological mechanisms of a rare malignant disease is the model for childhood acute B-lymphoblastic leukemia (cALL) described recently[339]. In this remarkable study the authors have been able to exploit the ability to differentiate precisely several developmentally restricted progenitors, and compare them with their *in vivo* counterparts. This insightful approach allowed the authors to identify clearly the leukemic cell of origin, a lympho-myeloid progenitor that seems to be particularly affected, in its B-lymphoid differentiation potential, by the RUNX1-ETV6 translocation, a founding mutation in cALL.
Even if there are many other examples of the successful exploitation of the PSCs as a model for haematology, the ability to produce a HSC *in vitro* is paramount for the study of the majority of haematopoietic diseases. As detailed in the previous chapter, many strategies and protocols have been proposed, but how can we improve them, for a better disease modelling? In an attempt to combine the theoretical insights of the developmental haematopoiesis with the conceptual approach of the cellular reprogramming, two USA-based teams pioneered the combination of direct differentiation with direct conversion. The team led by Igor Slukvin proved the feasibility of this approach, giving some molecular hints on the TFs involved in the EHT and how this phenomenon could be heterogeneous and responsible of different haematopoietic specification[340]. On the other side, the team led by George Daley tried to introduce, in a multipotent haematopoietic population derived with a direct differentiation protocol, a cocktail of TFs predicted to be important for HSC specification. The resulting population of progenitors was able to self-renew *in vitro* and with limited engraftment properties (mostly myeloid and erythroid[341]). One of the possible explanations for these outcomes lie in the potential of the transduced cell type: it could be possible that a haematopoietic progenitor derived from a pre-definitive developmental program is not able to undergo an HSC-like maturation. Moreover, this possibility does not exclude that other stimuli must be added to the reprogramming TFs, to achieve a correct maturation: it is worth to notice that Pereira and colleagues observed that the haematopoietic progenitor specified by their cellular conversion strategy was devoid of any clonogenic potential, and only an *in vitro* maturation step with a placental stroma was able to supply the necessary stimulation for a multipotent progenitor specification[39]. This valuable observation pre-date the updated version of the protocol described by Doulatov and colleagues: after a screening for factors enriched in fetal liver HSCs compared to HE and an analysis of other published factors, a new combination of TFs has been tested for the conferring of long term engraftment to an *in vitro*-specified human HE[342]. Although less efficient (of some orders of magnitude) than the cord blood, these cells were able to engraft primary and

secondary recipients without evident signs of leukemia, a possible concern due to the introduction of several known oncogenes. This result, even if preliminary, is a major advancement and leads the way in the process of HSCs generation *in vitro*[343].

	Szabo *et al* 2010[311]	**Pulecio *et al* 2014**[312]	**Pereira *et al* 2013**[313]	**Batta *et al* 2014**[314]	**Riddell *et al* 2014**[317]	**Doulatov *et al* 2013**[341]	**Sugimura *et al* 2017**[342]	**Sandler *et al* 2014**[315]	**Lis *et al* (2017)**[316]
Species	Human	Human	Mouse	Mouse	Mouse	Human	Human	Human	Human
Starting cell type	Fibroblast	Fibroblast	Fibroblast	Fibroblast	Haematopoietic progenitors	ES-derived haematopoietic progenitors	ES-derived mesoderm cells	Fetal Endothelial cells	Adult Endothelial cells
Transgenes	*OCT4*	*SOX2, mir125b*	*Gata2, Gfi1b, cFos, Etv6*	*Erg, Gata2, Runx1c, Scl, Lmo2*	*Runx1t1, Hlf, Lmo2, Pbx1, Zfp37, Prdm5, n-Myc, Meis1*	*ERG, HOXA9, RORA, SOX4, MYB*	*ERG, HOXA5, HOXA9, HOXA10, LCOR, RUNX1, SPI1*	*FOSB, GFI1, RUNX1, SPI1*	*FOSB, GFI1, RUNX1, SPI1*
Medium	*In vitro*	*In vitro*	*In vitro*	*In vitro*	*In vivo*	*In vitro*	*In vitro/In vivo*	*In vitro*	*In vitro*
Inducible expression	-	-	+	-	+	+	-	-	+
***In vitro* colony formation**	+	+	+	+	+	+	+	+	+
Erythroid	+	+	-	+	+	+	+	+	+
Myeloid	+	-	+	+	+	+	+	+	+
Lymphoid	-	-	-	-	+	+	+	+	+
Long Term Engraftment (including serial)	-	-	-	-	+	-	+	-	+

Table 3: the different transcription factor-based reprogramming strategies for the production of haematopoietic progenitor cells. Modified from Ebina and Rossi[344].

The value of the data generated by an iPSC model is directly dependent on the differentiation strategy used. A direct differentiation, although it is generally free from genetic manipulations that could alter the cellular physiology in a substantial way, is less prone to scalability and could yield very limited amount of cells for further analysis. In alternative, the direct conversion is a more straightforward approach and could generate larger quantities of cells of interest, but requires a step of experimental validation, as the process of cellular reprogramming could introduce bias and abnormal features that could complicate the phenotype interpretation. To give an example, as the model of a heterogeneous adult HSCs compartment is getting more experimental support, the question of the specification of this heterogeneity became more urgent than before. If it is an intrinsic property of each cell, it is tempting to speculate that it could be present already during development. If the observed heterogeneity is effectively HSC-dependent, it will be relevant to verify if the *in vitro* setup is retaining the same (possible) *in vivo* heterogeneity, and how it could be influenced by the differentiation strategy.

In conclusion, I am deeply convinced that the PSCs are a valuable experimental tool for the study of the normal and pathological haematopoiesis. The bottleneck constituted by the lack of differentiation protocol for the derivation of adult HSC *in vitro,* although a valid observation for the current modelling efforts, it is not conceptual but only technical and could be resolved with a concerted effort pointing to an increased knowledge of the human development.

3.3. Pluripotent stem cells as a therapeutic tool in haematology

The iPSC technology revolutionized the study of countless pathologies, allowing the creation of new efficient experimental models; but there is also a clear translational impact on the development of new therapeutical tools for medicine, summarized by the term "regenerative medicine". For what concerns the haematology, the field is the only one to already dispose of a widespread and validated cellular therapy: the bone marrow transplantation. This means that the combination of the available clinical expertise with the potential of a personalized therapy guaranteed by the PSCs could lead to the development of new truly curative treatments[345]. The therapeutic paradigm for iPSC-based therapeutics is a full circle, designed and conceptually tested in the early 2000s[346] (Figure 28).

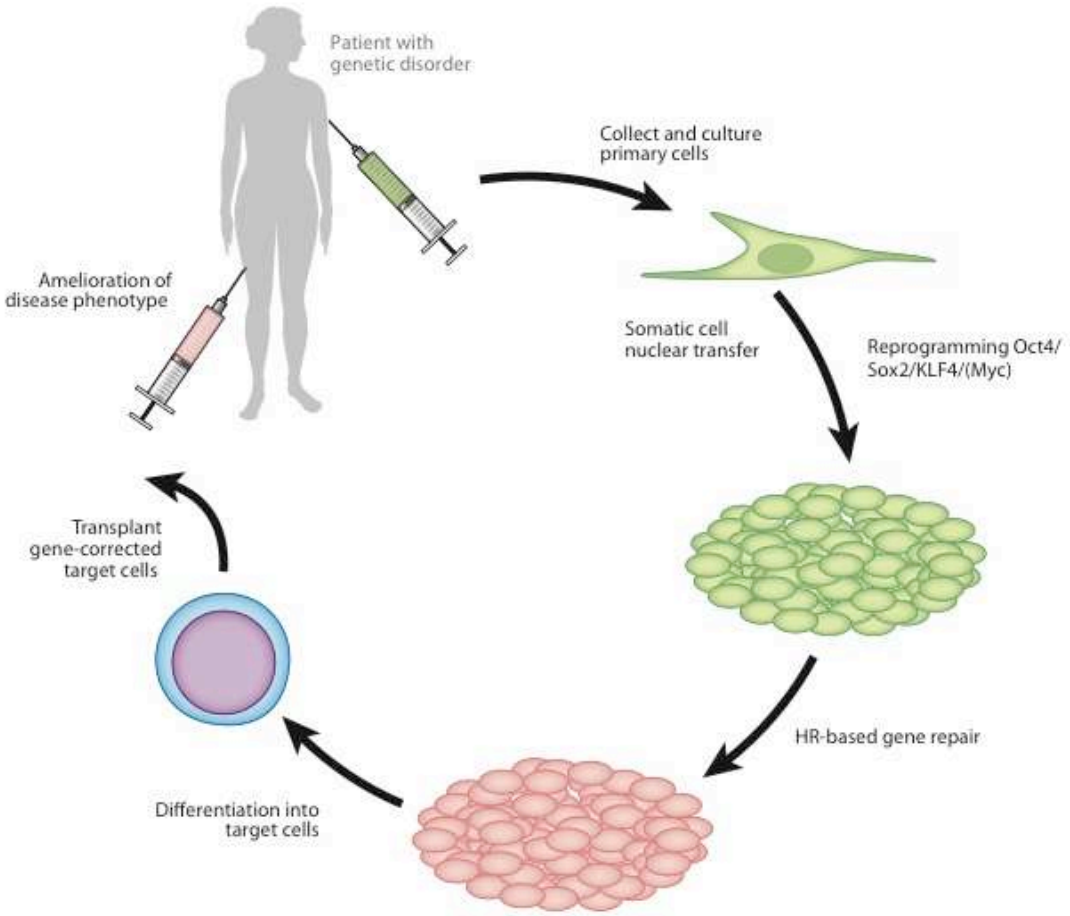


Figure 28: the therapeutic paradigm for iPS-based cell therapy. Readapted from Cherry and Daley[345].

The first step is the isolation of the cells that will be reprogrammed to the pluripotency. Originally, the current opinion was to derive cell lines specific for each genetic make-up, in order to propose autologous therapies devoid of any histocompatibility issue. Although this strategy is highly desirable, it is evident that the creation and manufacture of personalised therapies for each patient

constitutes a major economical and logistical roadblock. Currently, the therapeutic focus is switching to an allogenic approach, via the creation of cell banks of highly relevant genotypes, called "super-donors" for their high grade of compatibility with large parts of the population[347]. This approach has been championed by Yamanaka himself for regions characterized by a certain genomic homogeneity (notably the Japanese region), as an estimated few hundreds of cell lines should cover the 99% of the Japanese population. In both cases, iPS cell lines generated for translational purposes must be devoid of any tumorigenic potential, so the use of transgene-free technologies (typically mRNA reprogramming or non-integrative viral vectors) is mandatory.
The second step consists in the generation of a therapeutic population of interest. This step is disease-dependent and is the more technically difficult step of the entire circle, for the already mentioned difficulties in the production of engrafting stem cells from iPSCs. If the therapy is designed for a non-inherited condition, e.g. leukemia, it will be possible to generate an autologous cell replacement therapy devoid of adverse mutations just by selecting the non-mutated iPS cell lines. More frequently though, the iPS are proposed as a solution for rare inherited conditions, where all the cell lines derived from a specific patient will harbour the causative genotype (monogenic or multifactorial). In this case, it is necessary to eliminate the genetic cause of the disease, ideally by reverting the mutated allele to a non-pathogenic status. Currently, the first choice for this type of modification is the nuclease-mediated genome editing, with the CRISPR/Cas9 system spearheading the way. Interestingly, it could be possible to combine the iPSCs generation with the genome editing in a single step, reducing the time required for the therapy development[348]. On a different note, the genome editing could be used for the introduction of characteristics of interest and a general enhancement of the proposed therapy, e.g. a suicide gene for an improved biosafety or altering the HLA composition of the cell line for a major degree of histocompatibility. Finally, the circle terminates with the infusion of the newly generated haematopoietic cells in the patient, taking advantage of the new reduced-intensity conditioning regimes and the absence of any graft-versus-host disease, for autologous therapies.

The entire therapeutic paradigm, although already validated for non-haematological conditions, still has to be demonstrated in haematology. But the creation of an artifical bone marrow transplant is not the only therapeutic perspective supported by the iPSC technology: with the PSCs it is feasible to produce large amounts of mature blood components, a task that has been proven elusive using different biological sources (immortalized cell lines or primary cells). Again, the major limitation is the quality of the differentiation strategy: while protocols of direct differentiation are falling short because they are unable to produce the large amounts required for the clinical practice, the direct conversion could increase the yields, but require extra work for the definition of high quality cell

products that retains an *in vivo* functionality. Nevertheless, there is a growing interest for the field, in particular thanks to the diffusion of new innovative immunotherapies[349], and it can dramatically change the current paradigm of transfusional medicine: *in vitro* produced blood cells will be homogeneous, well defined and could be easily modified, in order to acquire new features. An example would be the combination between the chimeric antigen receptor technology (CAR-T cells) and the production *in vitro* of T cells from iPSCs[350], to create an "off the shelf" therapy for solid tumours or haematological malignancies. In alternative, it is possible to generate iPSCs from a specific T cell subclone carrying a rearranged TCR, in order to generate a more efficient and antigen-specific T cells population[351]. Other notable examples are the myelo-monocytic derivatives[352] and red blood cells[353], although both cell types are confronted by technical challenges related to a correct maturation and *in vivo* efficiency proofs.

Of particular interest is the biotechnological effort dedicated to the production of platelets *in vitro*: differently from red blood cells and plasma, platelets shelf life is very short (5-7 days), primarily because the storage at 4 °C damages irreversibly the platelets. At least two teams have proposed a cell-based solution for the production of large numbers of platelets *in vitro*, and both strategies relies on the generation of large amount of MKs as a key passage for the generation of the required amounts of platelets. Nakamura and colleagues were the first to show how PSCs could give origin to an immortalized megakaryocytic cell line, by the controlled expression of *BCL-XL*, *BMI1* and *c-MYC*: while the combination of c-MYC and BMI1 induce a sustained proliferation of MK progenitors, the exogenous add of BCL-XL effectively immortalized the cells, but only if performed in a stepwise manner (c-MYC and BMI1 first, then BCL-XL). The derived immortalized progenitor could be induced to produce $CD42b^+$ particles that bear several similarities with human platelets, by switching off the three inducing factors. Considering certain variability between cell lines, each immortalized MK, in a petri dish, is able to produce between 3 and 10 platelet-like fragments in a period of 5 days[354]. On a similar note, Moreau and colleagues proposed a simpler combination of three TFs (GATA1, FLI1 and TAL1), identified by combinatorial screening, to generate a pool of highly proliferating MK progenitors, around 2×10^5 MKs/PSC, albeit not immortalised. Those cells rely on the continuous expression of the three transgenes, typically working as part of a single complex together with other partners, and are able to produce (in a stationary culture system) around three platelet-like fragments per cell[355].

Both the presented strategies for platelets production are based of the assumption that large numbers of MKs are required to produce large amounts of platelets. Although this is certainly true, it is worth to remember that *in vivo,* MKs are a very rare population of the bone marrow and the ratio of platelets produced per MK is in the order of 10^3 PLT/MK. It will be certainly interesting to focus the research efforts not only on the early phases of MK differentiation, but also on how it is possible

to efficiently reproduce *in vitro* the late stage of MKpoiesis, as a way to improve massively the efficiency of these intriguing strategies. Their combination with the development of new advanced 3D systems, for the large and cost-effective amplification of the cells, should allow the translation of this class of promising cellular drugs from the laboratorial to the industrial setting, in order to further advance towards the clinical testing.

Finally, the iPSCs could contribute to the development of new therapies indirectly, as a powerful technological platform for new, more advanced, drug discovery and repositioning efforts. As a proof of concept, Doulatov and colleagues used a Diamond-Blackfan anemia (DBA) patient-derived iPSC model to perform an unbiased chemical screen of a library of 1440 compounds, looking for any promising candidate able to rescue the erythroid maturation defect[356]. The advantage of a similar platform lies in the close relationship with the patient genetic make-up and paves the way for the personalised use of iPSCs-derived cells for assessing the *in vitro* efficacy of a drug, as a potential surrogate end-point, or to assess the *in vitro* toxicity of a drug candidate[357].

RESULTS

Increased RhoA activity due to a disrupted filamin A/$a_{IIb}b_3$ interaction induces macrothrombocytopenia *(submitted)*

Context

The term Filaminopathy A describes a large number of conditions affecting several organs. This rare syndrome affects almost only women, and encompasses a large spectrum of clinical manifestation, from extremely severe to asymptomatic. Individuals carrying mutations in the X-linked *FLNA* gene has been reported to display different bleeding episodes during their life, included thoracic aortic aneurysm[358], suggesting a compounded origin for such episodes. Furthermore, platelets studies reported the presence in the peripheral blood of two different platelets populations, one of them of a larger size and partially different in its functional properties.

An entire body of work has substantially described the role of the FLNa in mouse models, as crucial for the activity of the complex GPIb-GPIX-GPV and several other platelets integrins and receptors. The absence of FLNa leads to large fragile platelets that display abnormal cytoskeletal features and are quickly cleared from the circulation by macrophages. Nevertheless, the debate is still open on the role of FLNa in MKs, as the CD42 complex does not appear particularly modified in mouse MKs lacking FLNa, as other surface proteins like the $\alpha_{IIb}\beta_3$ integrin. No major defects in MK differentiation and maturation are reported. In human cells, there is even less clarity, as only one study has been performed on MKs from Filaminopathy A patients. As FLNA is localized on the X chromosome, any primary cell sampling will results in a mix of two populations of cells, expressing either the WT or the mutated allele. Finally, as very few individuals are currently reported to carry mutations in FLNA, it is urgent to develop a suitable, human-relevant, model for the study of the pathology.

Aim

In this work, we wanted to clarify the role of FLNA during human MKpoiesis, and understand in details the physiopathology of the macrothrombocytopenia observed in Filaminopathy A patients. The development of a suitable model could pave the way for further studies, potentially extending to other aspects of this complex syndrome.

Material and methods

To avoid the confounding effects due to the X-inactivation patterns and skewing, we harnessed the power of iPSCs, reprogramming the circulating haematopoietic progenitors from two individuals affected by Periventricular Nodular Heterotopia or PNH (CD34+ cells from peripheral blood). In

this way we hoped to generate a panel of cell lines from the same genetic background, but differing only for the expressed X-chromosome. As patients are all heterozygous, it was possibile to generate iPS cell lines expressing the X-chromosome carrying the *FLNA*^{WT} allele or the *FLNA*^{mut} allele. Those cells were screened for the allelic expression of FLNA and characterized for their pluripotency properties and genomic stability. Successively, the iPSCs were differentiated into the MK lineage via a xeno-free, serum-free protocol that was modified from a published one[359], in order to accommodate our experimental needs. To better understand the role of FLNa in MKs, we designed a series of mutant cDNA sequences, lacking different domains of the FLNa that we deemed potentially involved in the protein activity. Those mutants were introduce in an iPS cell line lacking *FLNA*, therefore free from possible confounding effects due to the endogeneous protein. We choose to insert those transgenes via the use of a robust and validated gene editing technique, already available in the lab and validated[172]. We generated a mutant lacking the C-terminal domain of interaction with the actin (ABD) called del1, a mutant lacking the interaction domain with the GPIbα protein (del2), a mutant unable to interact with the β_3 integrin (del3) and finally a mutant lacking the domain of interaction with the small GTPases (del4). This last mutant is also unable to dimerise correctly. As a positive control, we also reintroduced the full-length cDNA for *FLNA*, in the same fashion as the other mutants (mut+WT). Finally, we generated multiple iPS cell lines expressing an engineered biosensor, via lentiviral transduction. This probe is able to detect the activity of the small GTPase RhoA, via fluorescence resonance energy transfer (FRET). All those cell lines were successfully differentiated into MKs and assayed for their functional properties and for the expression of different proteins of interest.

Results and conclusion

We derived two populations of iPSCs, one expressing the full length FLNa and one expressing reduced levels of *FLNA* mRNA and no protein. We selected two clones for each population, for both patients (total of 8 cell lines), to reduce the intraclonal variability. The absence of FLNa in the FLNA^{mut} iPSCs was neither changing their pluripotency properties nor their genomic stability. No reactivation of the *FLNA*^{wt} allele, due to the Xi erosion, was detected in the FLNA^{mut} clones. In conclusion, we derived stable and validated iPSCs for further studies on the role of FLNa in MKpoiesis.

The MK differentiation was not affected by the absence of FLNa, as no statistically significant differences were detected in the surface expression of the CD41 and CD42. Nevertheless, we detected a profound reduction in the ability of FLNA^{mut} MKs to extend proplatelets in static conditions. Therefore, we conclude that the absence of FLNa is not affecting the MK differentiation, but it is severely disrupting the PPT formation. We were unable to observe any

compensation via another member of the filamin gene family, as the *FLNA* mRNA levels increase progressively during the normal differentiation, while *FLNB* mRNA is expressed at very low levels.
To understand more in details how the absence of FLNa is disrupting the PPT formation process, we exploited the collection of mutants we generated, via gene editing of one $FLNA^{mut}$ cell line. All our mutants were not disrupted in their differentiation properties, but they gave different results on the ability to extend proplatelets *in vitro*: while del1, del3 and del4 were impaired in their PPT formation similarly to the $FLNA^{mut}$ clones, the del2 and the $FLNA^{mut+WT}$ were not statistically different from the $FLNA^{wt}$ clones. This unexpected observation prompted us to verify if our FLNa mutant was effectively unable to interact with GPIbα, and we confirmed the absence of interaction between the two proteins. In conclusion we claim that the FLNa protein is affecting the PPT formation via several mechanisms, but the interaction between FLNa and CD42 is not one of them.
The study of the MKs maturation yielded an unexpected observation: while adhering on fibrinogen coated slides, the $FLNA^{mut}$ MKs were displaying the formation of large stress fibers, compared to their WT counterparts. This phenotype suggests a defect in the acto-myosin contractility and could be correlated to an abnormal activation of the RhoA pathway. Indeed, we were able to highlight the increased formation of stress fibers for all the $FLNA^{mut}$ iPS clones, for both patients. Our mutants collection was also displaying the same defects, albeit at more reduced level for del1 and del2; conversely, del3 and del4 yielded very similar results to the $FLNA^{mut}$ clones, suggesting a complete inability to restore a normal phenotype. This observation was particularly intriguing, as the del4 is lacking the interaction with the small GTPases, so it could explain the observed stress fibers formation (proxy of an overactive RhoA), but the del3 results was surprising, as the integrin β3 should not be involved in RhoA signalling. Therefore we verified the RhoA activity in presence of different substrates, notably fibrinogen, collagen I and vWF. In presence of fibrinogen, so an active signalling from the $\alpha_{IIb}\beta_3$ integrin, we were able to observe a marked increase in the RhoA activity in absence of FLNa. This activity was decreased in a dose-dependent fashion, thanks to the re-expression of the full-length FLNa. Conversely, no difference in the RhoA activity was detected in presence of the other substrates, allowing us to trace the observed RhoA increased activity directly back to the $\alpha_{IIb}\beta_3$ integrin. Finally, the RhoA signalling negative effects on stress fibers formation and PPT formation were efficiently reduced, via the chemical inhibition of one classic RhoA effector, the ROCK1/2.
In conclusion, we observed for the first time that the absence of FLNa in MKs induces a defective MK maturation because of an over-active RhoA signalling, possibly dependent from the integrin $\alpha_{IIb}\beta_3$. These observations bear a certain novelty, as the $\alpha_{IIb}\beta_3$ has never been linked to the RhoA signalling; furthermore, Glanzmann thrombasthenia and other macrothrombocytopenias are potentially sharing some similarities in their physiopathology with filaminopathy A, as the RhoA

activity is central in the correct deployment of an efficient acto-myosin cytoskeleton. Finally, the inhibition of RhoA restored an effective PPT formation, underlying a possible therapeutic avenue for those patients.

Increased RhoA activity due to a disrupted filamin A/$\alpha_{IIb}\beta_3$ interaction induces macrothrombocytopenia

Alessandro Donada[1,2,*], Nathalie Balayn[1,*], Dominika Sliwa[1], Larissa Lordier[1], Valentina Ceglia[1], Francesco Baschieri[1], Cyril Goizet[3], Rémi Favier[1,4], Lucie Tosca[5], Gérard Tachdjian[5], Cecile V. Denis[6], Isabelle Plo[1], William Vainchenker[1], Najet Debili[1], Jean-Philippe Rosa[6], Marijke Bryckaert[6] and Hana Raslova[1]

1) UMR 1170, Institut National de la Santé et de la Recherche Médicale, Univ. Paris-Sud, Université Paris-Saclay, Gustave Roussy Cancer Campus, Equipe Labellisée Ligue Nationale Contre le Cancer, 94805 Villejuif, France
2) Université Sorbonne Paris Cité, Université Paris Diderot, 75013, Paris, France
3) INSERM UMR_S 1211, Université de Bordeaux, CHU Bordeaux UNIV EA 4576, France
(4) Assistance Publique-Hôpitaux de Paris, Armand Trousseau Children Hospital, French Reference Center for Platelet Disorders, Hematological Laboratory, Paris, France.
(5) U967, Institut National de la Santé et de la Recherche Médicale, Hôpitaux Universitaires Paris-Sud, Assistance Publique-Hôpitaux de Paris Service d'Histologie Embryologie et Cytogénétique, Clamart, France;
(6) HITh, UMR_S1176, Institut National de la Santé et de la Recherche Médicale, Univ. Paris-Sud, Université Paris-Saclay, 94276, Le Kremlin-Bicêtre cedex, France.
* equal contribution
Corresponding author: Hana Raslova, INSERM UMR1170, Gustave Roussy Cancer Campus, 114 rue Edouard Vaillant, 94805, Villejuif cedex, France.
Phone number: (+33) 1 42 11 46 71
E-mail: hana.raslova@gustaveroussy.fr
Conflict of interest: The authors have declared that no conflict of interest exists.

Abstract

Filamin A (FLNa) links the cell membrane with the cytoskeleton and is central in several cellular processes. Heterozygous mutations in the X-linked *FLNA* gene are associated with a large spectrum of conditions called filaminopathies that include a macrothrombocytopenia. Here we show, using an isogenic pluripotent stem cell model derived from patients, that the absence of FLNa protein in megakaryocytes (MKs) leads to an incomplete maturation, due to the inability to produce proplatelets. The reduction in the proplatelet formation potential is associated with a defect in the acto-myosin contractility, caused surprisingly by the unexpected RhoA activation. This exacerbated RhoA activity is the result of the loss of the interaction between $\alpha_{IIb}\beta 3$ and FLNa. Finally, the pharmacological inhibition of the RhoA-associated kinase ROCK1/2 restores a normal phenotype. Overall, this work suggests a new etiology for the macrothrombocytopenia, associating an increased RhoA activity with a disrupted FLNa/$\alpha_{IIb}\beta_3$ interaction.

Introduction

The proteins of the Filamin family are actin-binding proteins, mainly involved in the organization of the three-dimensional F-actin filament network. There are 3 isoforms encoded by 3 different genes, *FLNA*, *FLNB* and *FLNC.* FLNa is the most common isoform between the three, and forms a dimer via the N-terminal domains. The interaction with the actin cytoskeleton is due principally to the actin-binding domain (ABD), at the C-terminus. The rest of the protein is composed of 24 immunoglobulin-like domains of around 100 aminoacids each, loosely organized in two rods-like super-structures. The first rod can interacts, like the ABD, with actin, while the second rod is responsible for the multiple interactions of FLNa with over 80 different partners. Amongst them are integrins, cell surface receptors essential for the regulation of cellular morphology and motility[1]. Mutations in the *FLNA* gene are associated with a wide spectrum of rare disorders, characterized by developmental anomalies affecting multiple tissues. Among these disorders called Filaminopathies A, the most frequent are the periventricular nodular heterotopia (PVNH) and the otopalatodigital spectrum[2]. These conditions are often associated with a macrothrombocytopenia and a bleeding tendency[3]. Concerning specifically the megakaryocytes (MKs), the lack of FLNa induces a macrothrombocytopenia, with some contradictory reports on the expression of the GPIb complex[4, 5, 6].

Because the gene is localized on the X-chromosome and hemizygosity is generally considered lethal, most reported patients are women carrying heterozygous mutations; in these patients, two platelets populations differing in their size are observed, and only a small fraction (10-20%) of platelets seems to lacks FLNa. Moreover, no defects in MK differentiation were observed, but only in proplatelet formation[7]. We previously showed that low platelet count was associated with a reduced FLNa content compared to control, whereas near normal FLNa levels correlated with a normal platelet count, in absence of degradation[8]. The X-inactivation skewing was excluded in the analysis of the T cells, therefore we hypothesize that there are two MK populations in the patient bone marrow: one able to correctly differentiate and give rise to normal size platelets, and one displaying a defective production of macroplatelets. This heterogeneity explains the large variability of platelet counts in filaminopathies A patients.

This hypothesis prompted us to develop a model based on the induced pluripotent stem cells; in this way we can study the megakaryopoiesis for cells that express only the wild type or the mutated gene. In fact, the cellular reprogramming allows the generation of clonal cell lines expressing only one X chromosome; this means that it is possible to compare two cell lines with the same genetic background, differing only for the expression of the mutated or the wild

type *FLNA* allele. In this way, this model prevents the variability associated to the presence of two cellular populations in patient primary hematopoietic cells (Nurden et al 2011). MK differentiation of *FLNA*^{mut}-iPSCs appeared normal, but proplatelet formation was strongly impaired. Moreover, we observed an unexpected activation of RhoA on fibrinogen matrix, but not on other extracellular matrices like collagen I and von Willebrand Factor (vWF). Deletion of the β3 and GTPases FLNa binding regions strongly affected both proplatelet formation and RhoA activation, suggesting a functional link between $\alpha_{IIb}\beta_3$ signalling and the RhoA pathway. Overall, this is the first report linking $\alpha_{IIb}\beta_3$ signalling, FLNa and RhoA, and presents such mechanism as responsible for macrothrombocytopenia.

Material and Methods

***iPSCs generation and expansion*:** $CD34^+$ cells were isolated from peripheral blood by a positive selection using immunomagnetic bead cell-sorting system (AutoMacs; Miltenyi Biotec) and cultured in serum-free medium containing EPO (1 U/mL), FLT3l (10 ng/mL), G-CSF (20 ng/mL), IL-3 (10 ng/mL), IL-6 (10 ng/mL), SCF (25 ng/mL), TPO (10 ng/mL) and GM-CSF (10 ng/mL) for 6 days. After this expansion phase, cells were transduced with the CytoTune iPS 2.0 Sendai Reprogramming Kit (Thermo Fisher) and the reprogramming was performed accordingly to the manufacturer instructions. Colonies with an ES-like morphology were manually isolated, expanded for a reduced number of passages and frozen. iPSCs were maintained on Essential 8 or Essential 8 Flex medium (Gibco), on plates coated with N-truncated Human recombinant Vitronectin (Gibco). Cell passage was routinely performed using a solution of EDTA 500 μM in PBS 1X, or TrypLE 1X (Gibco). A Mycoplasma screening was routinely performed, accordingly to the manufacturer instructions (Sigma). Cells were kept in culture for a limited amount of passages, to prevent the surge of karyotypic and genomic anomalies.

Cytogenetics analysis: karyotype analysis was performed accordingly to the standard diagnostic procedures, after colchicine treatment.

Teratoma formation assay: pluripotent cells ($1x10^6$ cells) were resuspended in Essential 8 (140 μL) and Geltrex (60 μL) to a final volume of 200 μL, and injected in NSG mice. Mice were monitored and sacrified when the size of the tumoral mass was visibly affecting the animal motility and behaviour. Teratomas were excised, fixed and embedded in paraffin, while the corresponding sections were stained with hematoxylin, eosin and saffranin.

iPSCs haematopoietic differentiation: clumps of pluripotent cells were seeded on Geltrex (Gibco)-coated plates, in presence of E8 medium, at day -1. The departing cell concentration was adjusted for each cell line and it is comprised in the 10-20% confluency range. At Day 0, cells were transferred in a xeno-free medium based on StemPro-34 SFM (Gibco), supplemented with Penicillin/Streptomycin 1% v/v (Gibco), L-Glutamine 1% v/v (Gibco), 1-Thioglycerol 0.04 mg/mL (Sigma) and ascorbic acid 50 mg/mL (Sigma). This medium was retained for the entire experiment and supplemented with different cytokines and growth factors, accordingly to the following schedule: Days 0 – 2: BMP4 (10 ng/mL), VEGF (50 ng/mL) and CHIR99021 (2 mM). Days 2 – 4: BMP4 (10 ng/mL), VEGF (50 ng/mL) and bFGF (20 ng/mL). Days 4 – 6: VEGF (15 ng/mL) and bFGF (5 ng/mL). Day 6: VEGF (50 ng/mL), bFGF (50 ng/mL), SCF (50 ng/mL) and FLT3L (5 ng/mL). Days 7-10: VEGF (50 ng/mL), bFGF (50 ng/mL), SCF (50 ng/mL), FLT3L (5 ng/mL), TPO (50 ng/mL) and IL-6 (10 ng/mL). Days 10-20: SCF (50 ng/mL), FLT3L (5 ng/mL), TPO (50 ng/mL) and IL-6 (10 ng/mL). An exhaustive list of manufacturers could be retrieved in Supplemental Table 1.

Quantitative Real Time Polymerase Chain Reaction (qRT-PCR): RNA was purified using the RNeasy Plus Micro or Mini Kit, accordingly to the cell number and the manufacturer instructions (QIAGEN). Complementary DNA synthesis was performed with the Super Script II cDNA synthesis Kit or the Super Script Vilo cDNA synthesis kit (Thermo Fisher), accordingly to the amount of extracted RNA. Real Time PCRs were performed using the Takara Bio SYBR Premix Ex Taq (Clontech) reaction mix, on the 7500 Real Time PCR system (Applied Biosystems). Gene expression was assessed by comparative CT method, using PPIA or HPRT as reference gene. A list of the primers used is available in Supplemental Table 2.

Immunoblot analysis: cells were lysed using an adequate volume of RIPA buffer or a Laemmli buffer, according to the cell number and supplemented with a cocktail of protease inhibitors. Proteins were separated by SDS-PAGE and transferred on nitrocellulose membranes. Membranes were blocked for 1 hour with a solution of BSA 5% w/v in Tween-PBS 1X. Primary antibodies were incubated overnight, secondary antibodies coupled to horseradish peroxidase for 1 hour. Both type of antibodies were diluted in a solution of BSA 5% w/v in Tween-PBS 1X. A list of the primary and secondary antibodies is available, with the respective clone names, concentrations and manufacturers (Supplemental Table 3). Detection was performed using Image Quant LAS 4000 (GE Healthcare).

***Flow Cytometry*:** single cell suspensions were stained with monoclonal antibodies directly coupled to their respective fluorochromes. A list of the used antibodies, included the isotype control, is available, with the respective clone names (Supplemental Table 3). Cells were incubated with antibodies at the concentration of 1 μL/10^6 cells, in approximately 100 μL, at 4 °C for at least 30 minutes. Cells were washed before and after incubation in PBS 1X and analyzed with a BD Canto II or BD LSRFortessa cytometers (BD Biosciences). Fluorescence Activated Cell Sorting was routinely used for MKs purification and performed on Influx, ARIA III or ARIA Fusion cell sorters (BD Biosciences).

***Immunofluorescence and confocal imaging*:** the slides were coated with fibrinogen (25 μg/mL, Sigma) for 2 hours at room temperature. Purified megakaryocytes were plated on fibrinogen for 1 or 4 hours at 37 °C. Cells were fixed with 4% paraformaldehyde, for 10 minutes. Cell membrane permeabilization was performed with a buffer containing 0.2% Triton X-100, 100 mM PIPES pH 6.9, glycerol 2 M, EGTA 1mM and $MgCl_2$ 1mM, for 5 minutes. Primary and secondary antibodies were diluted in PBS 1X containing BSA 0.1% w/v, incubation was performed for 1 hour for each antibody. A list containing concentration, clone name and manufacturer for each antibody is available (Supplemental Table 3). Slides were mounted using Vectashield with DAPI (Molecular Probes). Images were acquired under a Leica DMI 4000 SPE laser-scanning microscope, with a 63x/1.4 numeric aperture oil objective (Leica Microsystem). Slides that were analyzed for the stress fiber formation were scored by enumeration of at least 100 cells per slide. Image analysis was performed with the LASX software.

Proplatelet formation assay: megakaryocytes isolated at day 14-15 of culture were seeded at a cell density of $5x10^3$ cells/well in a 96-wells plate, in a serum-free medium containing TPO (10 ng/mL) and SCF (25 ng/mL). Proplatelet-forming cells were scored after 4-5 days by enumeration of no less than 200 cells per well using an inverted microscope (Carl Zeiss), at a 200x magnification. A proplatelet-forming megakaryocyte was considered as a cell displaying at least one cytoplasmic process with a clearly defined constriction area. Each condition was examined in triplicate.

ZFNs mediated gene editing: the gene targeting of the AAVS1 locus was performed using a donor construct (AAVS1-SA-2A-puro-pA), a kind gift of Dr. Garçon. The FLNA full cDNA and

the different mutated forms were cloned under the regulation of the CAG promoter. iPSCs were seeded at D-1 onto 6-wells Geltrex-coated plates, after PBS/EDTA-mediated dissociation, at a cellular confluence of 50%, in Essential 8 medium. Transfection of the donor construct and the two nucleases for the left and the right arm was performed at D0, using X-treme GENE 9 (Roche), in presence of the ROCK1/2 inhibitor Y-27632 10 mM (Tocris). Selection was started at D2, adding puromycin at 500 ng/mL (Sigma). Individual, puromycin-resistant clones were manually picked and genotyped after two weeks.

FRET analysis: the lentiviral RhoA FRET sensor (pPBbsr2-Raichu-2707x) was a kind gift from Dr. Matsuda. Undifferentiated iPS cells were infected and selected for the expression of the CFP and YPet by fluorescence activated cell sorting. iPS-derived MKs were isolated by fluorescence activated cell sorting and plated on fibrinogen, collagen I (Chrono-log), von Willebrand factor or polylysine for 2 hours at 37 °C. Slides were mounted with Fluoromount-G® (Southern Biotech). FRET was performed on a confocal microscope Leica SP8 with a 63X objective (1.4 NA) with the acceptor photo-bleaching method using the FRET-AB wizard of the Leica software. Briefly, the acceptor fluorophore (YPet) fluorescence was bleached at 100% laser intensity, ensuring a minimal extent of bleaching by 75% and resulting in an increased fluorescence intensity of the donor fluorophore (CFP) in the bleached region, which was subsequently measured. FRET efficiency was obtained using the formula:

$$FRET_{eff.} = \frac{CFP_{post} - CFP_{pre}}{CFP_{post}}$$

Statistics: all data are shown as mean ± SD unless differentially specified in the legend. The statistical analyses were performed using the PRISM software. Statistical significance was established using a Student's *t* test specified in legends or a One Way Analysis of Variance (ANOVA), followed by All Pairwise Multiple Comparison Procedures (Student-Newman-Keuls Method). Differences were considered significant at $P < 0.05$.

Study approval: studies were performed under the promotion of the Ethics Committee of Ile de France IV (IRB 00003835) and in accordance with the Declaration of Helsinki. All subjects provided their informed consent prior to their participation to this study.

Results

Patient iPSCs display an unstable FLNA mRNA and no FLNa

iPSC clones were derived from two female patients, characterized by a deletion of the exons 31-48 (P1) and the exons 20-48 (P2) respectively (patients P4 and P2 in our initial study[8]). The cellular reprogramming allows the generation of clonal cell lines that are stably expressing either the X chromosome carrying the *FLNA*^{WT} or the *FLNA*^{mut} allele. Therefore, a panel of isogenic lines could be issued from the same genetic background, allowing direct comparison. *FLNA*^{WT} and *FLNA*^{mut} clones were screened for each patient, using two qRT-PCRs able to discriminate between the two alleles (Figure 1A). The screening did not report any evident difference in the percentage of generated *FLNA*^{mut} clones compared to the control; nevertheless, the *FLNA* mRNA level was significantly decreased by at least 50% in all *FLNA*^{mut} clones, compared to *FLNA*^{WT} clones (Figure 1A), suggesting an early instability of the mRNA. Wild type FLNa was detected in *FLNA*^{WT} clones; however, no truncated forms of FLNa protein were detected in *FLNA*^{mut} clones (Figure 1B and Supplemental Figure 1B). This could be explained by the deletion of a region responsible for the protein stabilization, as the phosphorylation at Ser^{2152} allows protection against proteolysis[9]. The eventuality of XCI erosion[10] affecting the *FLNA* locus was excluded by immunoblot, after maintaining the iPSC clones in culture for 60 passages (Supplemental Figure 1C).

FLNa deficiency induces a deep proplatelet formation defect

Two *FLNA*^{WT} and two *FLNA*^{mut} clones for each patient were selected for further studies and validated for a stable karyotype and functional pluripotency (Supplemental Figure 2). The MK differentiation was induced using an updated version of the protocol described by Chou and colleagues[11] (Supplemental Figure 3). The absence of FLNa was confirmed by immunofluorescence in mature CD41^{+}CD42^{+} MKs (Figure 1C). Surprisingly, an abnormal distribution of F-actin was observed in *FLNA*^{mut} MKs. Indeed, *FLNA*^{mut} MKs show an unexpected presence of large stress fibers, a phenotype that suggests a GTPase RhoA overactivation. Quantification of MKs derived from *FLNA*^{WT} and *FLNA*^{mut} iPSC showed similar percentages (Figure 1Di, ii) and no difference in mean fluorescence intensity was detected for both $\alpha_{IIb}\beta_3$ (anti-CD41a) and GPIX (anti-CD42a) expression (Figure 1Diii), suggesting that both $\alpha_{IIb}\beta_3$ and GPIb-GPIX-GPV complexes were normally expressed at the cell surface of mature MKs. This is in contrast with the defective MK differentiation reported for murine MKs derived from mouse embryonic stem cells with a decreased FLNa expression[6]. Importantly, a marked defect in proplatelet formation was observed in *FLNA*^{mut} clones for both patients

(Figure 1E), consistent with the previous observation on CD34⁺-derived MKs from the patient P2[7]. The proplatelet formation defect was not related to a more rapid maturation, since the differential pattern between *FLNA*^{WT} and *FLNA*^{mut} MKs remained unchanged throughout culture. To rule out a possible partial compensation effect by the *FLNB* isoform, we measured the expression of the two genes during megakaryopoiesis (Supplemental Figure 4A). We noted a strong up-regulation of *FLNA* mRNA in the later stages of megakaryopoiesis, but most importantly no *FLNB* mRNA up-regulation was detected, in the absence of FLNa (Supplemental Figure 4B), excluding any compensatory effect between isoforms. We conclude that our model is recapitulating the main features observed in the patients, and it is suitable to address the mechanism of the defective proplatelet formation.

$\alpha_{IIb}\beta_3$ and Rho-GTPases FLNa interaction domains are essential for proplatelet formation

Taking advantage of the absence of FLNa in the *FLNA*^{mut} clones, we introduced in one P1 *FLNA*^{mut} iPSC clone several FLNa cDNAs lacking domains essential for MK maturation: i) a mutant lacking the ABD at the C-terminus (del1)[12], ii) a mutant lacking the Ig17 domain that interacts with GPIbα (del2)[13], iii) a mutant lacking the Ig21 domain the interacts with the β3 integrin (del3)[14], and iv) a mutant lacking the dimerization domain involved in the interaction with small GTPases (del4)[15]. The full-length *FLNA* cDNA (mut+WT) was also introduced, as an internal control (Figure 2A). For all these constructs, we used a zinc finger nucleases-based (ZFN) technology, targeting the AAVS1 locus as previously described[16]. Expression and relative stability of the mutant FLNa proteins were verified in all newly generated iPSC clones (Figure 2B). Concerning the MK differentiation, no difference in the percentage of CD41⁺CD42⁺ MKs was observed (Figure 2C). Surprisingly, while mutants lacking the domains interacting with actin (del1), β3 integrin (del3) and Rho-GTPAses (del4) displayed a reduced proplatelet formation potential, similar numbers of proplatelet-forming MKs were observed with the del2 mutant (unable to interact with GPIbα), the *FLNA*^{WT} and the *FLNA*$^{mut+WT}$ MKs (Figure 2D). The absence of GPIbα interaction with del2 FLNa mutant was confirmed by proximity ligation assay (Figure 2E), corroborating our results and demonstrating that the FLNa-GPIbα interaction does not play a major quantitative role in proplatelet formation.

Lack of FLNa induces an over-activation of RhoA

The adhesion of iPSC-derived *FLNA*^{mut} MKs on fibrinogen (Figure 1C) showed an unexpected increase in stress fibers formation for *FLNA*^{mut} clones. In light of this unexpected observation,

we investigated in more details the stress fiber formation on fibrinogen matrix, using our different mutants. The majority of *FLNA*^{mut} MKs display a large number of stress fibers, compared to the control and the *FLNA*$^{mut+WT}$ clone (Figure 3A), and similar percentages could be found for the mutants lacking β3 integrin-interacting Ig21 (del3), or the dimerization domain involved in the interaction with small GTPases (del4). To confirm the over-activation of RhoA pathway for the cells lacking FLNa, we directly measured RhoA activity in cells settled on fibrinogen matrix, but also on collagen I and vWF (Figure 3C), using a lentiviral FRET biosensor based on the Raichu probe[17]. On fibrinogen matrix, the absence of FLNa led to a 3-fold increase in RhoA activity, compared to the control; moreover, expression of the FLNa full-length cDNA in *FLNA*^{mut} MKs reduced the RhoA activity significantly (Figure 3Cii). Interestingly, RhoA activity appeared identical between *FLNA*^{mut} and *FLNA*^{WT} MKs adhering to collagen I (an α2β1 substrate) (Figure 3Ciii) or vWF (GPIb-GPIX-GPV substrate) (Figure 3Civ): this indicates that the up-regulation of RhoA in absence of FLNa is specific to fibrinogen and strengthens the conclusion that it is dependent on $\alpha_{IIb}\beta_3$.

To confirm the involvement of the RhoA/ROCK pathway, *FLNA*^{WT} and *FLNA*^{mut} MKs were treated with a catalytic inhibitor of ROCK1/2 (Y-27632), a major effector of RhoA-mediated signalling in MKs, and then assessed for proplatelet formation and stress fiber formation. The addition of Y-27632 led to a marked reduction in the number of stress fibers-forming MKs, restoring the normal acto-myosin contractility for all mutants (Figure 3D, Supplemental Figure 5). Importantly, proplatelet formation was partially restored as well, confirming that RhoA down-regulation is crucial for platelet formation (Figure 3E).

Discussion

FLNa is a crucial actor in linking the cytoskeleton with the different receptors, to transduce the signal after adhesion to the extracellular matrix. Among these surface proteins in megakaryocytes, the more important are the $\alpha_2\beta_1$ integrin, the complex GPIb-GPIX-GPV and the $\alpha_{IIb}\beta_3$ integrin. As the mutations in *FLNA* gene lead frequently to the macrothrombocytopenia in patients suffering from Filaminopathies A, different studies were conducted to elucidate the precise role of FLNa in megakaryopoiesis. It was reported in mouse that the homozygous *Flna* deletion is embryonic lethal, with major defects in cardiac development[2]. In females, heterozygous deletion of *Flna* leads to a mild thrombocytopenia with more than 95% of platelets of normal size expressing FLNa and only 5% of larger platelets lacking FLNa[4].

Flna homozygous deletion specific for the erythro-megakaryocytic lineage resulted in severe

macrothrombocytopenia with increased tail bleeding time. The FlnA-null platelets displayed a completely altered cytoskeleton and no linkage of GPIbα to the actin filaments. Moreover, decreased levels of GPIb-GPIX-GPV and an increase in surface expression of integrins $\alpha_{IIb}\beta3$ and GPVI were reported[4]. By using the *Flna* knock-out model specific for MKs, a premature release of large and fragile platelets and their destruction by macrophages was reported, without affecting the MK differentiation. Interestingly and contrary to platelets, FlnA-null MK displayed a normal expression of all major surface integrins and only a moderate reduction in proplatelet production[5].

Finally, a decreased surface expression of GPIbα was demonstrated in MKs derived from murine ES cells where the *FlnA* and *FlnB* expression was decreased via short-hairpin inhibition. In this last model, the MK differentiation was altered, the proplatelets exhibited a distinctly abnormal morphology with enlarged swellings and thick shafts and the platelets were two to three times larger than the FLNa positive proplatelets/platelets. The equilibrated stochiometry of FlnA and GPIbα was shown to be necessary for the optimal trafficking to the margin of the cell and to control the platelet size in an over-expression cellular model[6]. These conflicting reports, combined with the difficulties in the human modelling of this disorder, have not clearly addressed the macrothrombocytopenia etiology at the megakaryocytic level.

Using our human, patient-specific model, we demonstrated the total absence of FLNa in the iPSCs and MKs expressing the mutated *FLNA* allele. This was probably due to the deletion of the Ser^{2152}, which phosphorylation is crucial for the protein stability. One of the kinases that are able to phosphorylate FLNa is PKA, and mutations in the catalytic domain have been reported to cause a severe macrothrombocytopenia[18]. We have shown no defect in MK differentiation from $FLNA^{mut}$ iPSC, without any defect in CD42a (GPIX) and CD41a (GPIIb) surface expression. These results are consistent with the reported data[5] and confirm, in a human model, the normal expression of the complex GPIb-GPIX-GPV in MKs. Nonetheless, a major defect in proplatelet formation has been confirmed, reinforcing the observed data from Nurden and colleagues[7]. In order to understand the etiology of this defect, we tried to dissect the role of some of the more important molecular partners of FLNa in megakaryopoiesis and thrombopoiesis. While the proplatelet formation reduction in the mutant for the interaction with actin (del1) was not surprising, an almost normal formation of proplatelets for the mutant del2, unable to interact with GPIbα, was unexpected. Indeed this glycoprotein is a causal gene for macrothrombocytopenia in Bernard-Soulier syndrome[19], and the defect in its expression at surface lead to the defect in proplatelet formation[20,21]. As the GPIbα-FLNa

interaction was proved to be important for regulation of platelet activity[22], its was suggested to be involved also in proplatelet formation; however we clearly show that GPIbα-FLNa interaction is not essential for this process.

On the other hand, the mutant deleted for the domain interacting with small GTPases (del4), leads to a deep defect in proplatelet formation suggesting the importance in the interaction between RhoA and FLNa in this process. It is known that the abnormal RhoA activity inhibits the proplatelet formation, while physiologically it is down-regulated during the latest phases of megakaryopoiesis[23,24]. FLNa absence was shown to lead to an increase in RhoA activity in other cell types[25], but the relationship between these two partners have never been described in MKs. FLNa is known to interact with different GAPs, it would thus be interesting to evaluate which one is able to modulate the activity of RhoA in absence of FLNa in MKs.

Unexpectedly, we observed an increase in stress-fiber forming MK upon adhesion of MKs lacking FLNa on fibrinogen. In normal conditions, RhoA-dependent stress fiber formation is triggered by adhesion to collagen I, but not by fibrinogen, the ligand for the integrin $\alpha_{IIb}\beta_3$. The use of the mutant lacking the domain interacting with β3 integrin (del3) clearly show the crucial role of αIIbβ3-FLNa interaction in maintaining the RhoA pathway at inactive state. These results were reinforced by the fact that RhoA is over-active only on fibrinogen matrix and not on collagen I or VWF, in absence of FLNa. Finally, the abnormal activation of RhoA pathway could also explain the defective polyploidization of *Flna*$^{-/-}$ MKs[5], as this pathway plays a major role in the endomitotic process[26]. It is possible that $\alpha_{IIb}\beta_3$-dependent RhoA activation could be involved in other macrothrombocytopenias. Whether the absence of interaction between FLNa and $\alpha_{IIb}\beta_3$ is at the origin of macrothrombocytopenia in patients with mutations in *ITGA2* and *ITGB3* genes [27,28] remains to be addressed. Nevertheless, this new mechanism should be considered in the patient stratification, as we observed that the chemical inhibition of RhoA-signalling exhibits a positive effect on proplatelet formation, in absence of FLNa. Larson and Watson[29] demonstrated that the treatment of murine MKs with antagonists of $\alpha_{IIb}\beta_3$ was able to disrupt proplatelet formation, but they were unable to relate it to a signalling pathway. We solve this question showing that the interaction of FLNa with $\alpha_{IIb}\beta_3$ plays a crucial role for maintaining RhoA inactive during proplatelet formation, confirming in a significant pathological context the impact of the postulated role for FLNa as able to retain the integrin in an inactive state[14].

Overall, these results clearly demonstrate that the absence of FLNa in MKs leads to a $\alpha_{IIb}\beta_3$ dependent over-activation of the RhoA pathway, which appears as the major actor responsible for the observed deep defect in the proplatelet formation (Figure 4).

Author contributions

AD, NB, DS, VC, LL, FB, LT, GT, ND performed and analyzed experiments. AD, WV, IP, ND, JPR, MB, HR discussed results. CG and RF provided clinical and biological follow-up of patients. CD provided the biological material. HR supervised the work. AD, JPR, MB and HR wrote the article. AD and NB contributed equally to this work.

Acknowledgments

The authors thank the patients for participation in this study, O. Bawa and P. Opolon from the Preclinical research platform, Gustave Roussy, Villejuif, France, for the teratomas analysis, P. Rameau, C. Catelain and Y. Lecluse from the Imaging and cytometry platform PFIC, UMS AMMICA, Gustave Roussy Villejuif, France for the expertise in cytometry.

This work was supported by French grants from Fondation pour la Recherche Medicale (LPC20170637458) and Ligue Nationale Contre le Cancer (équipe labellisée 2016 to H.R.) and by the European grants ERA-NET (to C. Balduini, 2013) and H2020-FETOPEN-1-2016-2017-SilkFusion. AD was supported by a PhD fellowship from the Sorbonne Paris Cité and Ligue National Contre le Cancer. DS was supported by a fellowship from ANR and Ligue National Contre le Cancer.

References

1. Zhou, A. X., Hartwig, J. H. & Akyürek, L. M. Filamins in cell signaling, transcription and organ development. *Trends Cell Biol.* **20,** 113–123 (2010).

2. Feng, Y. & Walsh, C. A. The many faces of filamin: A versatile molecular scaffold for cell motility and signalling. *Nat. Cell Biol.* **6,** 1034–1038 (2004).

3. Parrini, E. *et al.* Periventricular heterotopia: Phenotypic heterogeneity and correlation with Filamin a mutations. *Brain* **129,** 1892–1906 (2006).

4. Falet, H. *et al.* A novel interaction between FlnA and Syk regulates platelet ITAM-mediated receptor signaling and function. *J. Exp. Med.* **207,** 1967–1979 (2010).

5. Jurak Begonja, A., Hoffmeister, K. M., Hartwig, J. H. & Falet, H. FlnA-null megakaryocytes prematurely release large and fragile platelets that circulate poorly. *Blood* **118,** 2285–95 (2011).

6. Kanaji, T., Ware, J., Okamura, T. & Newman, P. J. GPIb regulates platelet size by controlling the subcellular localization of filamin. *Blood* **119,** 2906–2913 (2012).

7. Nurden, P. *et al.* Thrombocytopenia resulting from mutations in filamin A can be expressed as an isolated syndrome. *Blood* **118,** 5928–5937 (2011).

8. Berrou, E. *et al.* Heterogeneity of platelet functional alterations in patients with filamin a mutations. *Arterioscler. Thromb. Vasc. Biol.* **33,** e11-8 (2013).

9. Chen, M. & Stracher, A. In situ phosphorylation of platelet actin-binding protein by cAMP-dependent protein kinase stabilizes it against proteolysis by calpain. *J. Biol. Chem.* **264,** 14282–14289 (1989).

10. Mekhoubad, S. *et al.* Erosion of dosage compensation impacts human iPSC disease modeling. *Cell Stem Cell* **10,** 595–609 (2012).

11. Chou, S. T. *et al.* Trisomy 21-associated defects in human primitive hematopoiesis revealed through induced pluripotent stem cells. *Proc. Natl. Acad. Sci.* **109,** 17573–17578 (2012).

12. Nakamura F, Hartwig JH, Stossel TP, Szymanski PT. Ca2+ and calmodulin regulate the binding of filamin A to actin filaments. *J Biol Chem*. 2005;280(37):32426-33.

13. Nakamura, F. *et al.* The structure of the GPIb-filamin A complex. *Blood* **107,** 1925–1932 (2006).

14. Liu, J. *et al.* Structural mechanism of integrin inactivation by filamin. *Nat. Struct. Mol. Biol.* **22,** 383–9 (2015).

15. Ohta, Y., Suzuki, N., Nakamura, S., Hartwig, J. H. & Stossel, T. P. The small GTPase RalA targets filamin to induce filopodia. *Proc. Natl. Acad. Sci. U. S. A.* **96,** 2122–8 (1999).

16. Antony-Debré, I. *et al.* Level of RUNX1 activity is critical for leukemic predisposition but not for thrombocytopenia. *Blood* **125,** 930–40 (2015).

17. Yoshizaki, H. *et al.* Activity of Rho-family GTPases during cell division as visualized with FRET-based probes. *J. Cell Biol.* **162,** 223–32 (2003).

18. Manchev, V. T. *et al.* A new form of macrothrombocytopenia induced by a germ-line mutation in the PRKACG gene. *Blood* **124,** 2554–2563 (2014).

19. Favier, R. & Raslova, H. Progress in understanding the diagnosis and molecular genetics of macrothrombocytopenias. *Br. J. Haematol.* **170,** 626–639 (2015).

20. Balduini, A. *et al.* Proplatelet formation in heterozygous Bernard-Soulier syndrome type Bolzano. *J. Thromb. Haemost.* **7,** 478–484 (2009).

21. Balduini, A., Malara, A., Balduini, C. L. & Noris, P. Megakaryocytes derived from patients with the classical form of Bernard-Soulier syndrome show no ability to extend proplatelets in vitro. *Platelets* **22,** 308–11 (2011).

22. Williamson, D. *et al.* Interaction between platelet glycoprotein Ibα and filamin-1 is essential for glycoprotein Ib/IX receptor anchorage at high shear. *J. Biol. Chem.* **277,** 2151–2159 (2002).

23. Sabri, S. *et al.* Differential regulation of actin stress fiber assembly and proplatelet formation by α2β1 integrin and GPVI in human megakaryocytes. *Blood* **104,** 3117–3125 (2004).

24. Chang, Y. *et al.* Proplatelet formation is regulated by the Rho/ROCK pathway. *Blood* **109,** 4229–4236 (2007).

25. Duval, D. *et al.* Valvular dystrophy associated filamin A mutations reveal a new role of its first repeats in small-GTPase regulation. *Biochim. Biophys. Acta - Mol. Cell Res.* **1843,** 234–244 (2014).

26. Mazzi, S., Lordier, L., Debili, N., Raslova, H. & Vainchenker, W. Megakaryocyte and polyploidization. *Exp. Hematol.* **57,** 1–13 (2018).

27. Ghevaert, C. *et al.* A nonsynonymous SNP in the ITGB3 gene disrupts the conserved membrane-proximal cytoplasmic salt bridge in the αIIbβ3integrin and cosegregates dominantly with abnormal proplatelet formation and macrothrombocytopenia. *Blood* **111,** 3407–3414 (2008).

28. Favier, M. *et al.* Mutations of the integrin αIIb/β3 intracytoplasmic salt bridge cause macrothrombocytopenia and enlarged platelet α-granules. *Am. J. Hematol.* **93,** 195–204 (2018).

29. Larson, M. K. & Watson, S. P. Regulation of proplatelet formation and platelet release by integrin αIIbβ3. *Blood* **108,** 1509–1514 (2006).

Figure 1

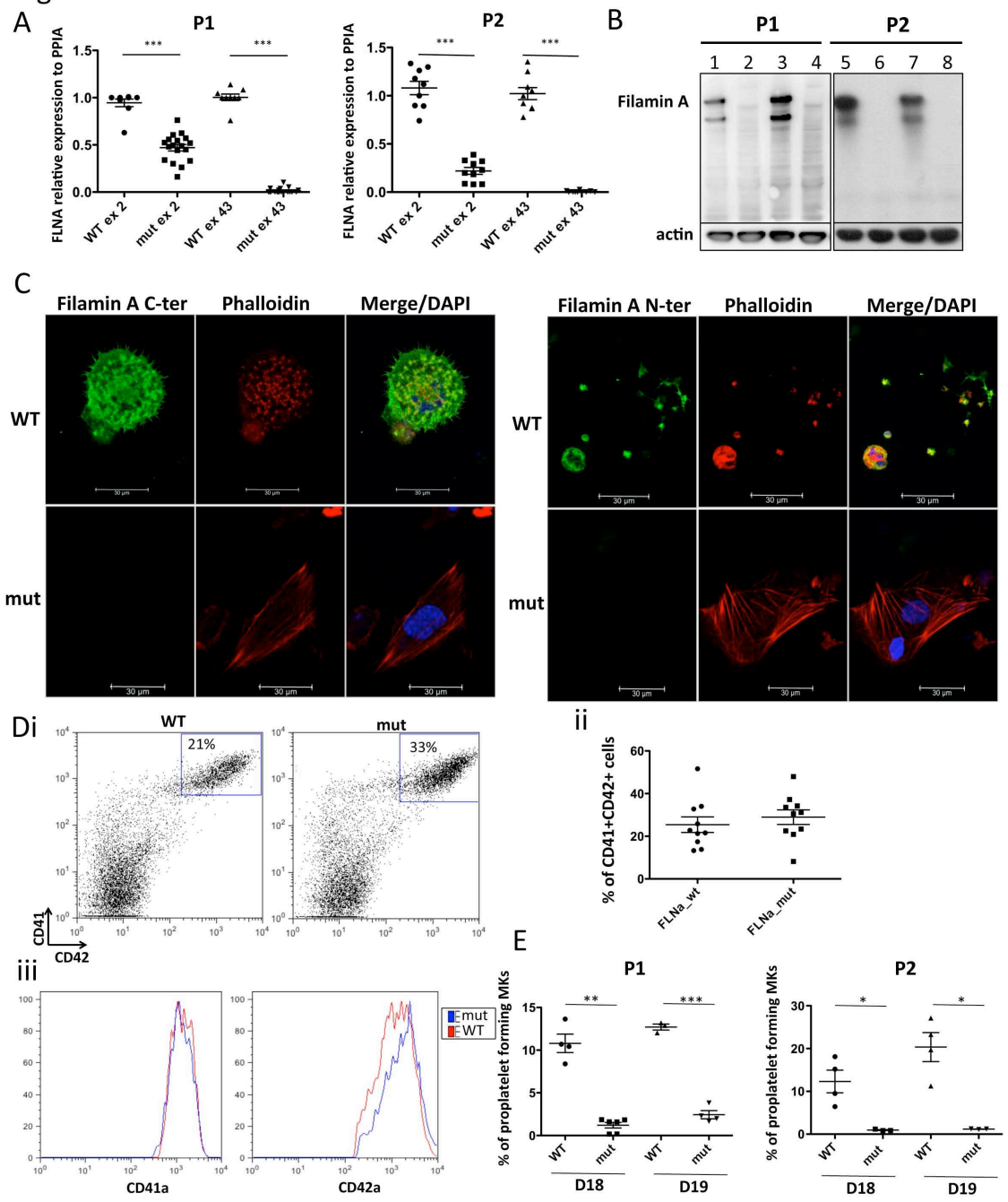


Figure 1: **Defective FLNa expression impairs normal proplatelet formation. (A)** qRT-PCR for FLNa expression on two different exons for several iPSC clones (9 FLNA^{WT} and 20 FLNA^{mut} clones for P4, 9 *FLNA^{WT}* and 10 *FLNA^{mut}* clones for P2). **(B)** Representative immunoblot for the FLNa expression on four iPSC clones for each patient, using a C-terminus specific antibody. Lines 1, 3, 5 and 7 represent *FLNA^{WT}* clones, while lines 2, 4, 6 and 8 represent FLNA^{mut} clones **(C)** Immunofluorescence staining for the expression of FLNa in iPSC-derived MKs from P1, scale bar = 30 µm. **(D)** Flow cytometry analysis of the CD41a and CD42a expression: (i) representative dot plot and (ii) relative percentages, n = 10 (iii) histogram representative of the mean fluorescence intensity, n = 5, paired Student t-test. **(E)** Proplatelet formation

potential of patient iPSC-derived clones. At least two clones were assayed for each genotype, for each patient, at two different time-points. Results are presented as mean±SEM, unpaired Student's t-test with Welch's correction was used, ***P<0.001, **P<0.01, *P<0.05, each point represent one independent experiment.

Figure 2

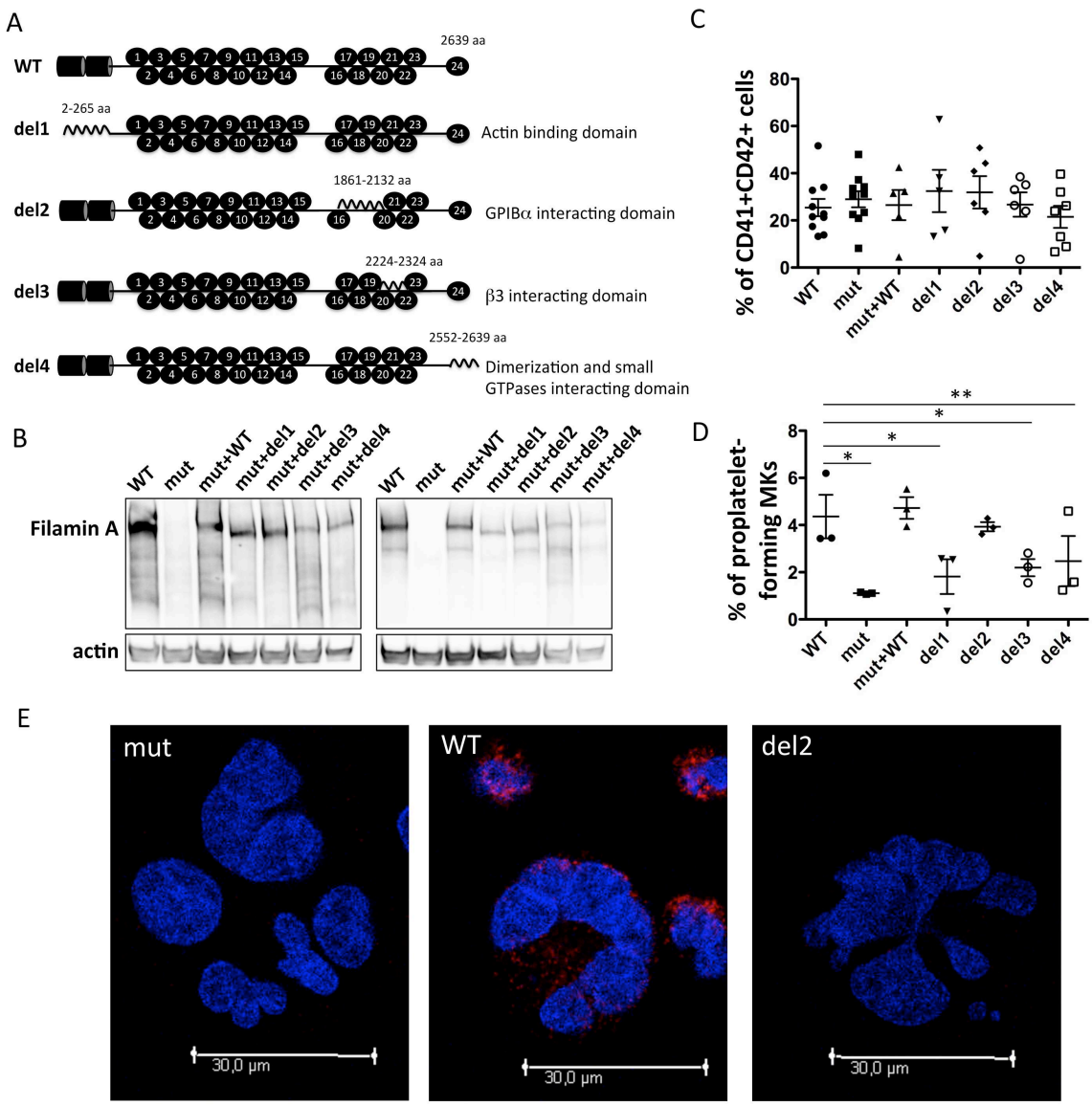


Figure 2: **Deletion of $\alpha_{IIb}\beta_3$ and Rho-GTPases FLNa interaction domains, but not of GPIbα interacting domain, deeply affect proplatelet formation. (A)** Schematic representation of the FLNa mutants introduced by ZFNs-mediated gene editing. **(B)** Representative immunoblot for the FLNa expression in the iPSC edited clones, using a N-terminus (left) and a C-terminus (right) antibody **(C)** Flow cytometry analysis of the CD41a and CD42a expression in the edited clones, each point represent one independent experiment, mean±SEM, unpaired Student's t-test with Welch's correction was used. **(D)** Proplatelet formation potential for each edited clone, n = 3, **P<0.01, *P<0.05, n=3, paired Student's t-test. **(E)** Representative pictures of proximity ligation assay for the GPIbα–FLNa interaction. The red staining represents the interaction between GPIbα and FLNa.

Figure 3

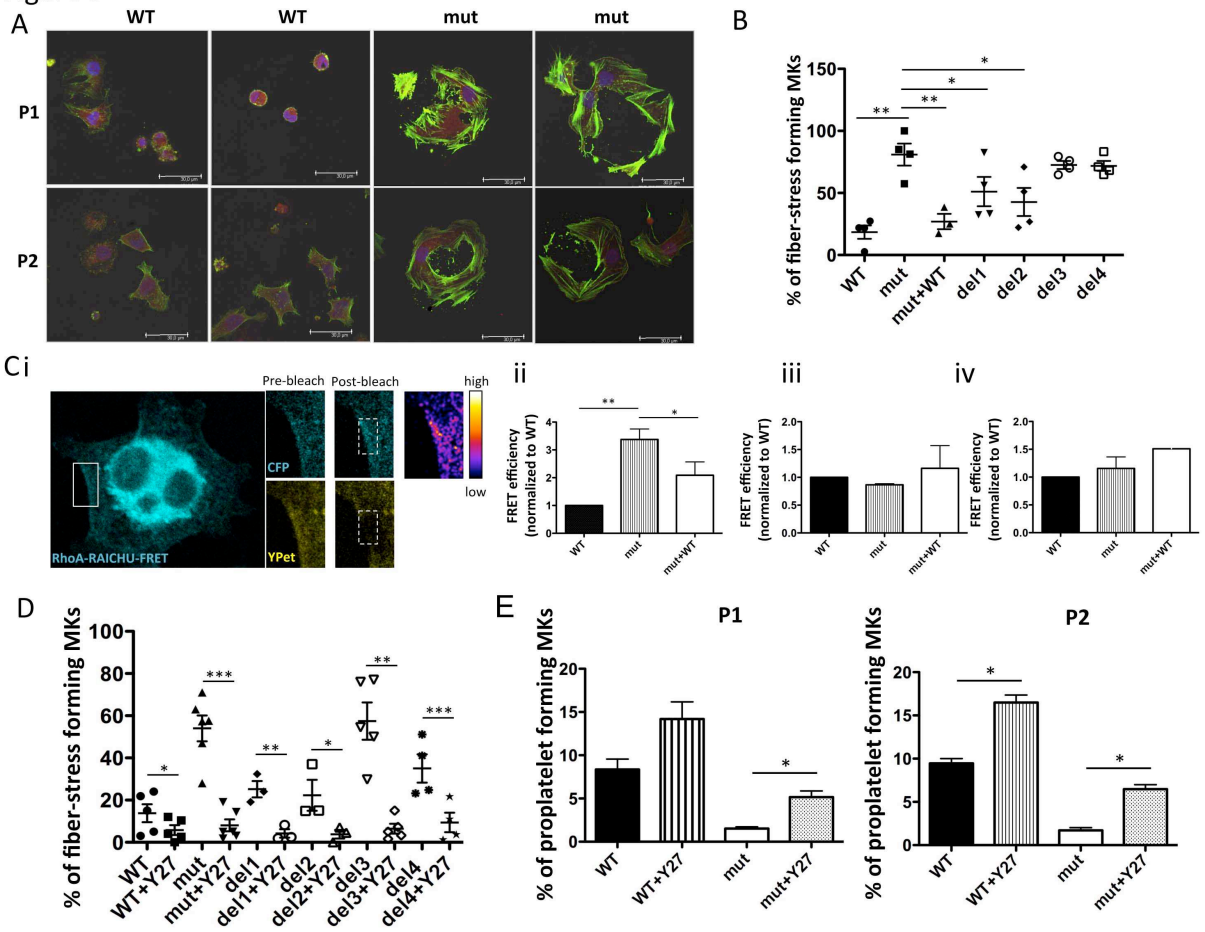


Figure 3: **RhoA is over-activated in FLNa-deficient MKs. (A)** Immunofluorescence staining of G-actin (in red) and F-actin (in green) after adhesion on fibrinogen. The nucleus is stained with DAPI (in blue), scale bar = 30 μm. **(B)** Quantification of the percentage of stress fibers-forming MKs for each edited clone, each point represent one independent experiment, mean±SEM, unpaired Student's t-test with Welch's correction was used. **P<0.01, *P<0.05. **(C)** FRET analysis for the RhoA activation on different substrates: (i) representative image, (ii) adhesion on fibrinogen (n = 4), (iii) on collagen I (n = 3) and (iv) vWF (n = 3). At least 15 cells per condition were measured. Data were analyzed performing One Way Analysis of Variance (ANOVA), followed by All Pairwise Multiple Comparison Procedures (Student-Newman-Keuls Method), ***P<0.001, **P<0.01, *P<0.05. **(D)** Stress fibers formation in presence and absence of ROCK1/2 inhibitor Y-27632 for the edited clones, n=5 for WT and del 3, n=6 for mut, n=3 for del1 and del2, n=4 for del4, paired Student's t-test, ***P<0.001, **P<0.01, *P<0.05 **(E)** Proplatelet formation potential in presence and absence of ROCK1/2 inhibitor Y-27632 for both patient-iPSC clones, n = 3, ***P<0.001, **P<0.01, *P<0.05, paired Student's t-test.

Figure 4

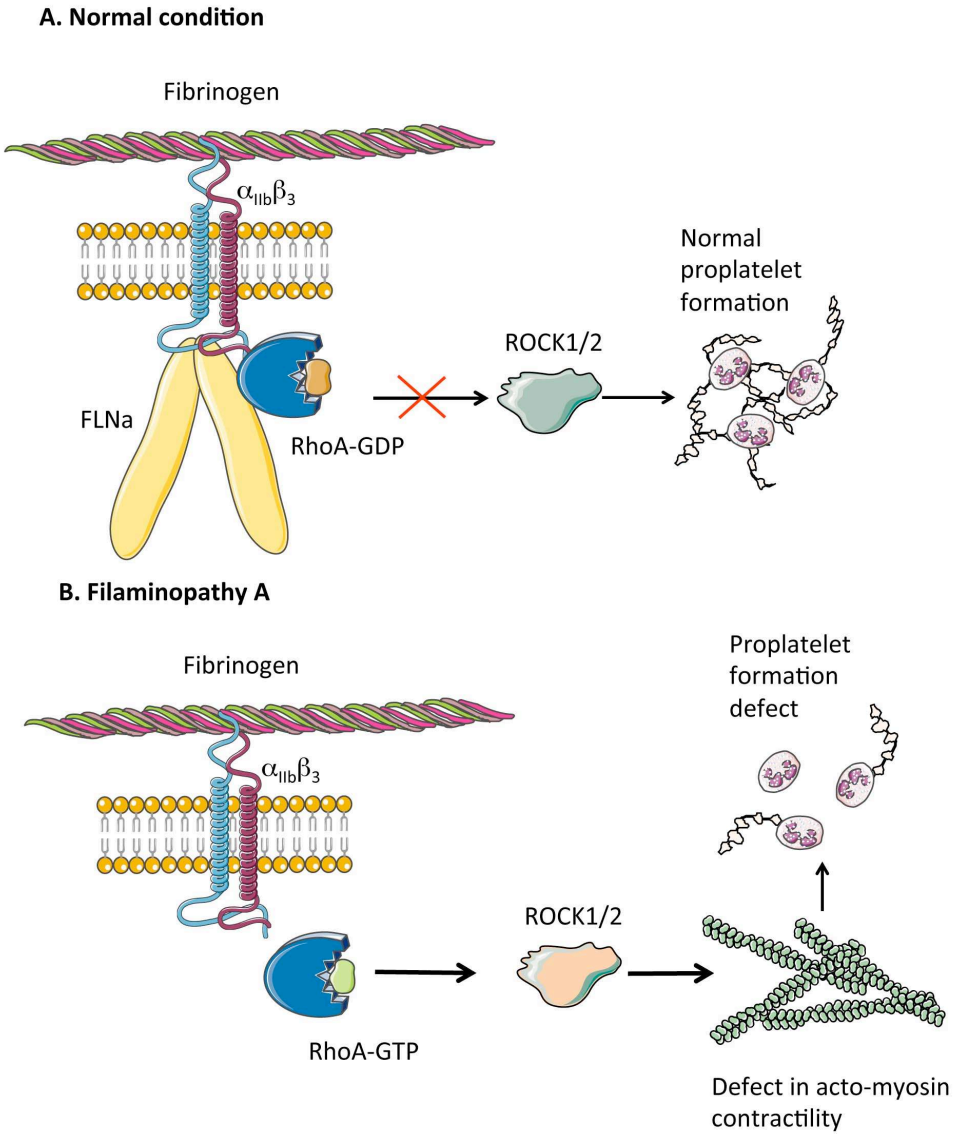


Figure 4. **Pathological mechanism for FLNa-deficient MKs.** (A) In presence of FLNa, the interaction between fibrinogen and its receptor $\alpha_{IIb}\beta_3$ does not trigger the RhoA pathway activation. No anomalies in the proplatelet formation could be observed. (B) In absence of FLNa, the interaction between fibrinogen and $\alpha_{IIb}\beta_3$ leads to an increase in RhoA activity. Consequently, the normal acto-myosin contractility is disrupted, via ROCK1/2 activity, and this leads to a deeply flawed proplatelet formation. This increased RhoA activity in absence of FLNa is specifically dependent on fibrinogen, and absent in presence of other extracellular matrices like collagen I or vWF.

Supplementary Data

Reagent	Source	Identifier
hBMP4	Peprotech	#AF-120-05ET
hVEGF	Peprotech	#100-20
hFGF-basic	Peprotech	#100-18B
hIL-6	Peprotech	#200-06
hEPO	Peprotech	#100-64
hG-CSF	Peprotech	#300-23
hGM-CSF	Peprotech	#300-03
hIL-3	Peprotech	#200-03
hTPO	Kirin Brewery, Tokyo, Japan	/
hFLT3l	Celldex Therapeutics, Inc., Needham, USA	/
hSCF	Biovitrum AB, Stockolm, Sweden	/
CHiR 99021 trihydrochloride	TOCRIS	#4953
Y-27632 dihydrochloride	TOCRIS	#1254
VTN-N	Gibco/Thermo Fisher Scientific	#A14700
Geltrex	Gibco/Thermo Fisher Scientific	#A1413202
StemPro®-34 SFM	Gibco/Thermo Fisher Scientific	#10639011
E8	Gibco/Thermo Fisher Scientific	#A1517001
E8 Flex	Gibco/Thermo Fisher Scientific	#A2858501
CytoTune™-iPS 2.0 Sendai Reprogramming Kit	Invitrogen/Thermo Fisher Scientific	#A16517
1-thioglycerol	Sigma	#M6145
Fibrinogen	Sigma	#F8630
Puromycin dihydrochloride	Sigma	#P8833
X-tremeGENE™ 9	Sigma/Roche	#000000006365779001
cOmplete™ inhibitor cocktail	Sigma/Roche	#4693159001
Collagen I	Chrono-log	#P/N385
von Willebrand Factor (WILFACTIN)	LFB	/

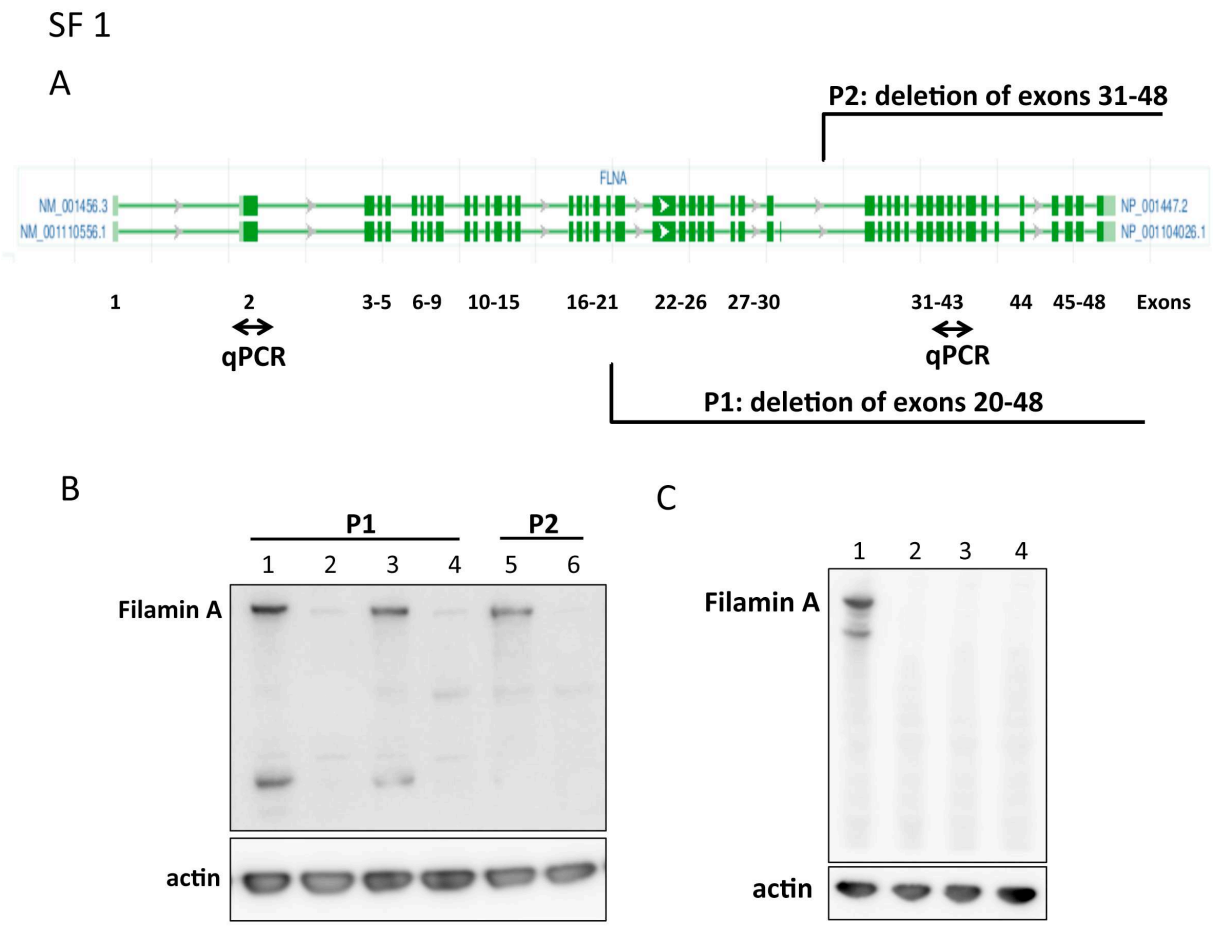
Supplementary table 1: **list of reagents used, including source and identifier**

Oligonucleotide	Source	Sequence
FLNA_Ex2_F	Eurogentec	ATGCAGCTTGAGAACGTGTC
FLNA_Ex2_R	Eurogentec	GGATCAGGGTCCAGATGAGG
FLNA_Ex43_F	Eurogentec	AGGAGCCAGGTGACTACGAA
FLNA_Ex43_R	Eurogentec	GCCCTGACTCCTGAAGGCTA
PPIA_F	Eurogentec	GTCAACCCCACCGTGTTCTT
PPIA_R	Eurogentec	CTGCTGTCTTTGGGACCTTGT
FLNB_F	Eurogentec	CATCAAGCTCGTGTCCATCG
FLNB_R	Eurogentec	TGGCATCATCATCCCCTTCA

Supplementary table 2: **list of oligonucleotides used, including source and sequence**

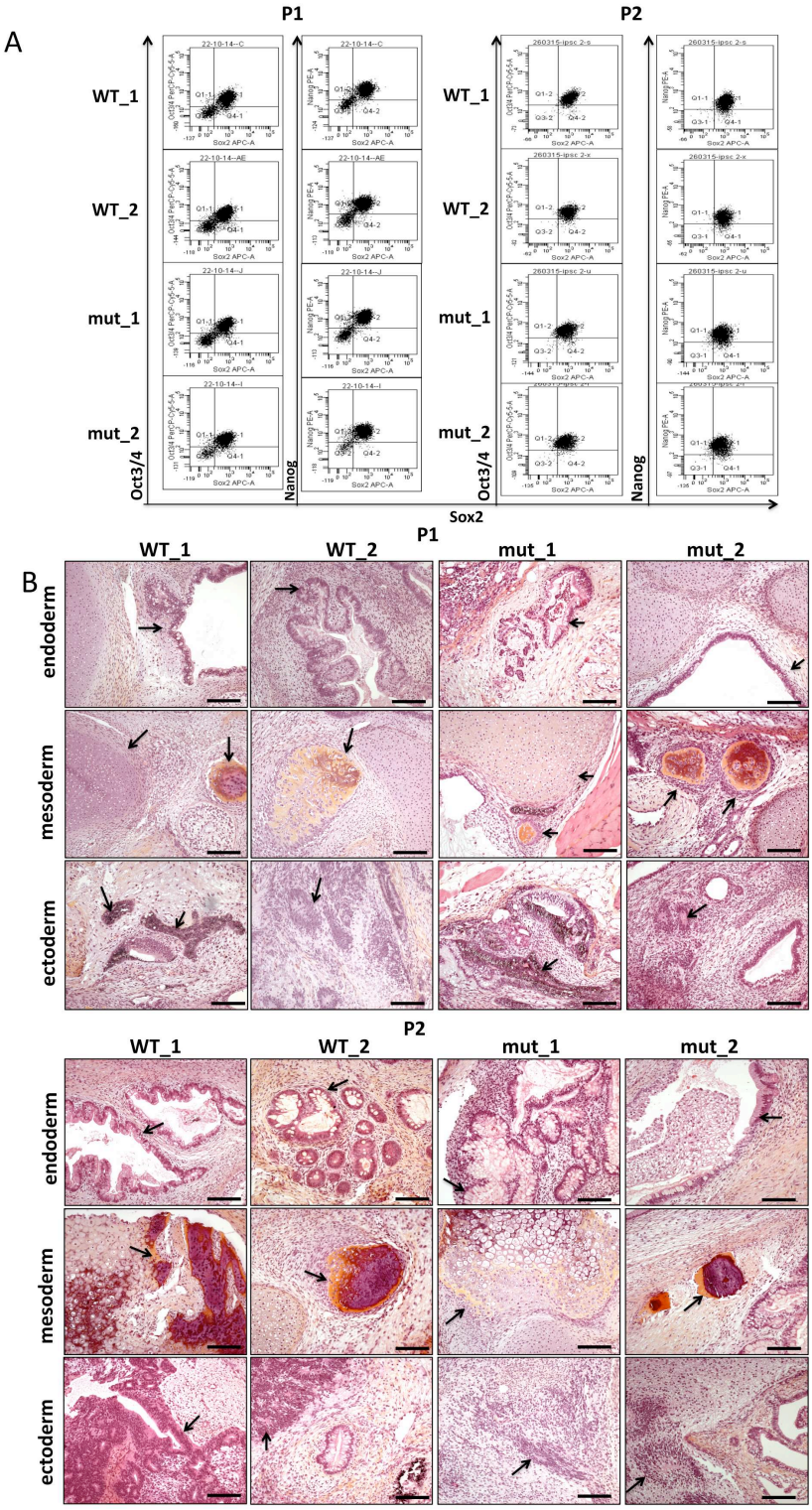
Antibody or other reagent	Source	Identifier	Working Dilution
CD41a - APC	BD Biosciences	#559777 / HIP8	1:100 (flow cytometry)
CD42a - PE	BD Biosciences	#558819 / ALMA.16	1:100 (flow cytometry)
Oct3/4 - PerCP-Cy™5.5	BD Biosciences	#560794 / 40/Oct-3	1:100 (flow cytometry)
Sox2 - Alexa Fluor® 647	BD Biosciences	#560302 / 245610	1:100 (flow cytometry)
Nanog - PE	BD Biosciences	#560483 / N31-355	1:100 (flow cytometry)
IgM - APC	BD Biosciences	#555585 / G155 - 228	1:100 (flow cytometry)
IgG1 - PE	BD Biosciences	#555749 / MOPC-21	1:100 (flow cytometry)
IgG1 - PerCP-Cy™5.5	BD Biosciences	#552834 / MOPC-21	1:100 (flow cytometry)
IgG2a - Alexa Fluor® 647	BD Biosciences	#565365 / MOPC-21	1:100 (flow cytometry)
FLNB	Abcam	#ab97457	1:1000 (WB)
FLNA (C-TERM)	Abcam	#ab51217	1:200 (IF) – 1:1000 (WB)
FLNA (N-TERM)	Millipore	#MAB1678 / PM6/317	1:200 (IF) – 1:1000 (WB)
Phalloidin	Sigma	#P5282	1:400 (IF)
a-Tubulin	Sigma	#T5168 / B-5-1-2	1:200 (IF)
b-Tubulin	Sigma	#T8328 / AA2	1:200 (IF)
Actin	Sigma	#A5441 / AC-15	1:1000 (WB)
IgG anti-rabbit H+L Alexa Fluor 633	Thermo Fisher Scientific	#A-21071	1:400 (IF)
IgG anti-mouse H+L Alexa Fluor 546	Thermo Fisher Scientific	#A-11003	1:400 (IF)
IgG anti-rabbit HRP-linked	Cell Signaling Technology	#7074	1:2000 (WB)
IgG anti-mouse HRP-linked	Cell Signaling Technology	#7076	1:2000 (WB)
GP1bα	Thermo Fisher Scientific	#MA5-16565/AK2	1:100 (PLA)
Vectashield with DAPI	Vector laboratories	#H-1200	/
Fluoromount-G®	SouthernBiotech	#0100-01	/

Supplementary table 3: **list of antibodies used, including source, identifier and working dilution**

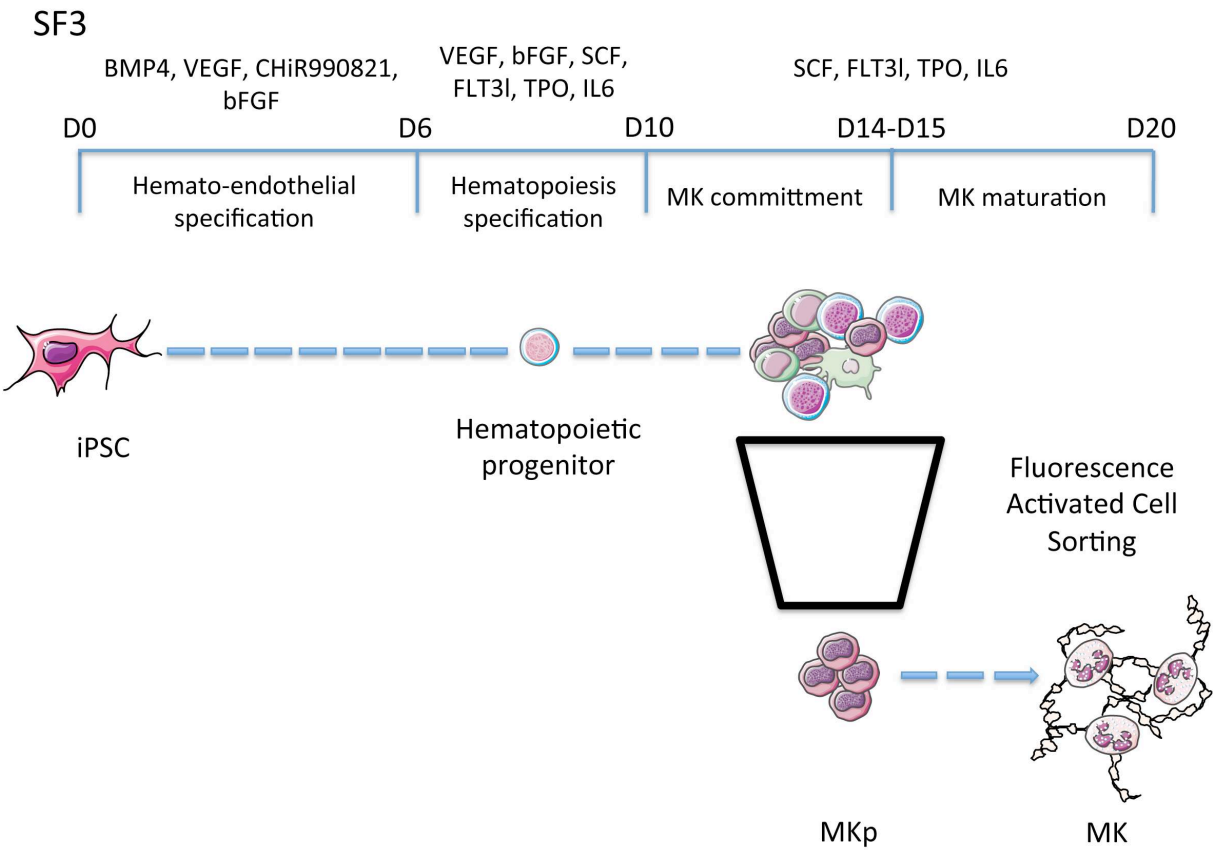


Supplementary figure 1: **FLNA absence in iPSC clones expressing the mutated X chromosome. (A)** qRT-PCR primer design for the identification of the expressed X chromosome. Amplification of exon 2 allows to detect both WT and mutated mRNA, that of exon 43 only the WT mRNA. **(B)** Immunoblot for the FLNA expression in 2 $FLNA^{WT}$ and 2 $FLNA^{mut}$ iPSC clones for P4 and 1 $FLNA^{WT}$ and 1 $FLNA^{mut}$ iPSC clone for P2 using a N-terminal antibody. **(C)** Immunoblot for the FLNA expression on iPSC cultivated for 60 passages using C-terminal antibody.

SF2

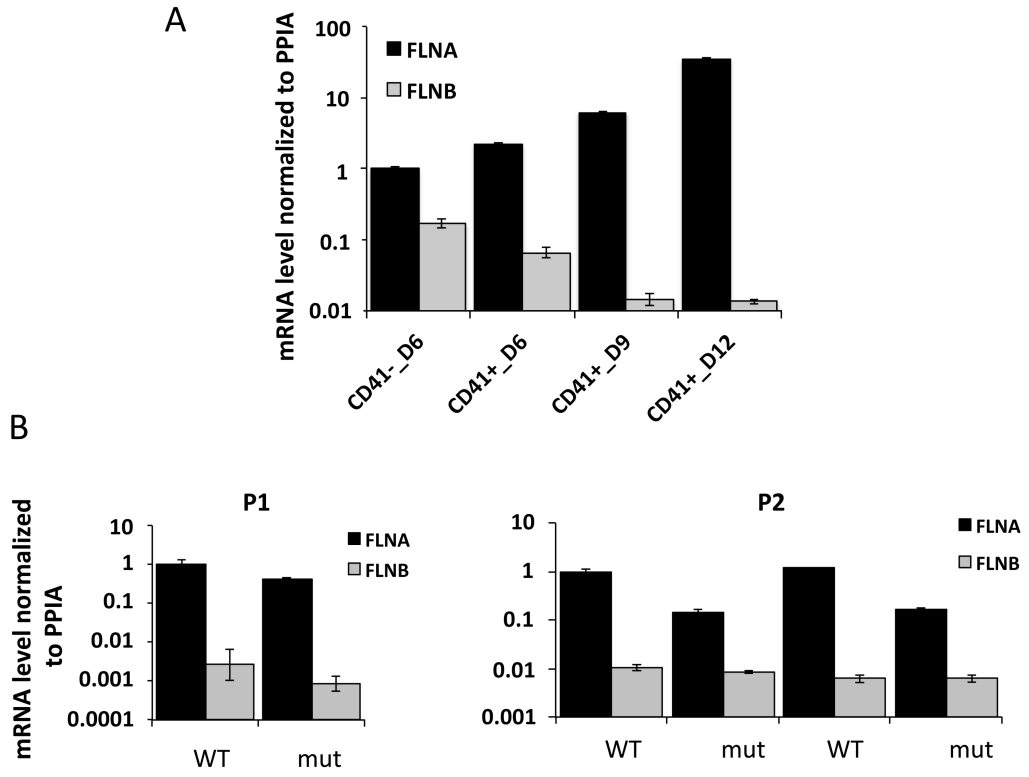


Supplementary figure 2: **iPSCs characterization and validation. (A)** Expression of the transcription factors Oct3/4, Sox2 and Nanog by flow cytometry. Four clones (two *FLNA*^{WT} and two *FLNA*^{mut}) for each patient have been assessed. **(B)** Teratoma formation assay. Four clones (two *FLNA*^{WT} and two *FLNA*^{mut}) for each patient have been assessed Scale bar = 100 μm.

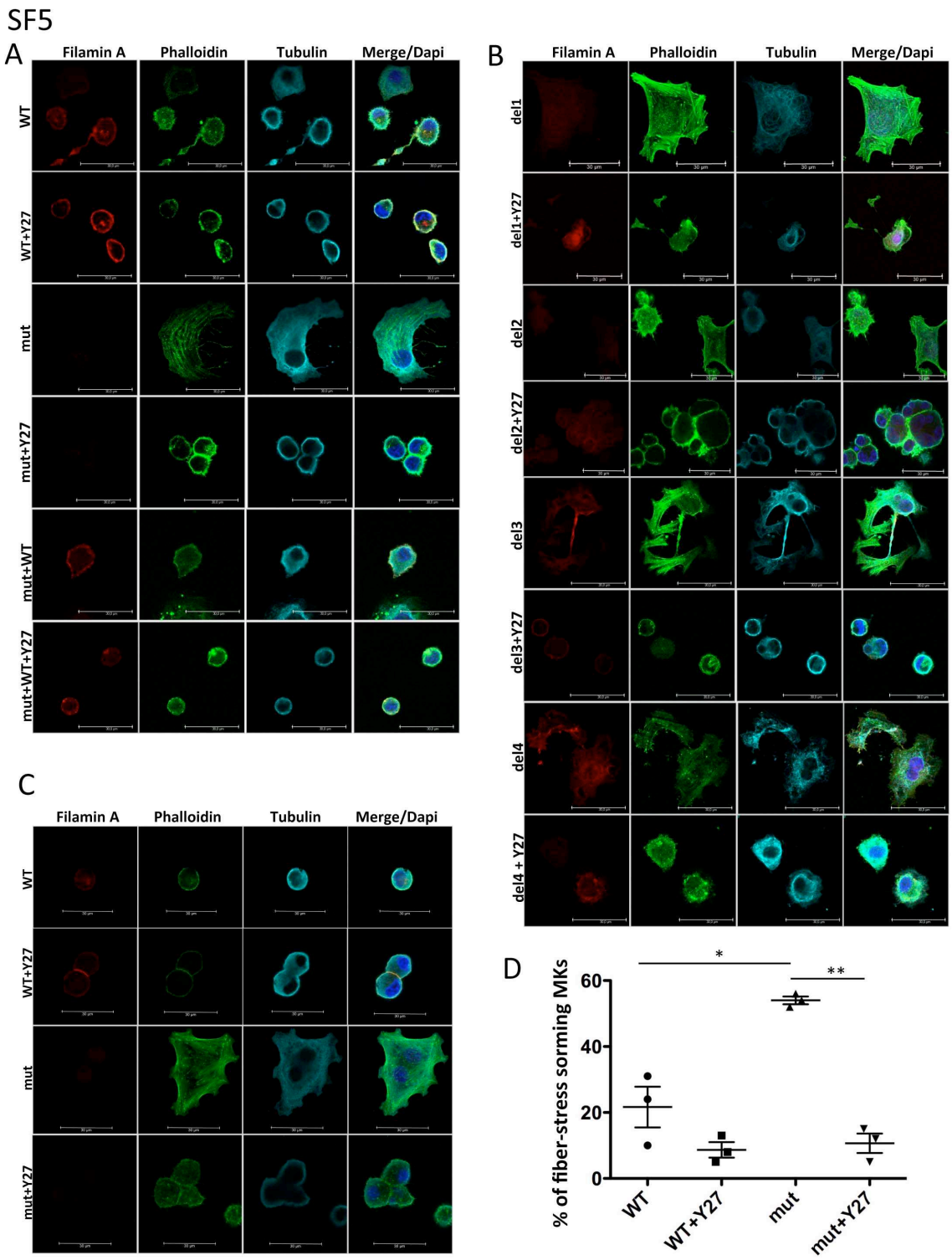


Supplementary figure 3: **Schematic representation of the hematopoietic differentiation protocol.** The described protocol could be subdivided in four different stages: in the first phase the iPSCs are differentiated toward the hemato-endothelial progenitor fate, by a mesodermal induction operated in the first two days of differentiation. In the second stage, the culture conditions are adjusted for the support of the endothelial-to-hematopoietic transition of the first hematopoietic cells. The latter two stages recapitulate the early MK differentiation and the MK maturation, so the culture conditions are adapted for the hematopoietic cells.

SF4



Supplementary figure 4: ***FLNB* does not compensate the lack of *FLNA*. (A)** qRT-PCR of *FLNA* and *FLNB* transcript levels along megakaryopoiesis in cord blood-derived MK progenitors and mature MKs; the image is representative of three independent experiments. **(B)** qRT-PCR of FLNA and FLNB transcript level 1 $FLNA^{WT}$ and 1 $FLNA^{mut}$ clone for P1 and in 2 $FLNA^{WT}$ and 2 $FLNA^{mut}$ clones for P2.



Supplementary figure 5: **ROCK1/2 inhibition reduces stress fibers formation. (A)** Representative images from *FLNA*^{WT}, *FLNA*^{mut} and *FLNA*$^{mut+WT}$ clones in presence or absence of Y-27632, scale bar = 30 μm. **(B)** Representative images from the different mutants del1, del2, del3 and del4 in presence or absence of Y-27632, scale bar = 30 μm. **(C)** Representative images of *FLNA*^{WT}, *FLNA*^{mut} clones for P2 in presence or absence of Y-27632, scale bar = 30 μm. **(D)** Quantification of the percentage of stress fibers-forming MKs for P2 clones, n = 3, **P<0.01, *P<0.05, paired Student's t-test.

Abnormal ANKRD26 expression leads to the hyperactivation of the G-CSFR signalling pathway in ANKRD26-related thrombocytopenia patients *(in preparation)*

Context

ANKRD26-related thrombocytopenia or Thrombocytopenia 2 (THC2) is one of the most commonly diagnosed inherited thrombocytopenia. The clinical picture includes autosomal dominant inheritance pattern, dysmegakaryopoiesis, mild to severe thrombocytopenia and an increased risk to develop myeloid malignancies. This last feature in particular requires attention, as THC2 patients could benefit from a close surveillance and management of their haematological parameters. THC2 is one of the three ITs currently described that predispose to the development of haematological malignancies. While the other two disorders (FPD/AML and ETV6-RT) are associated to mutations in transcription factors, THC2 is caused by mutations in the 5' UTR of *ANKRD26*, a gene relatively unknown and not associated to oncogenesis and leukemic transformation. Those mutations are mostly localized in a short stretch of nucleotides that includes a binding motif for the transcriptional complex formed by RUNX1 and FLI1. Mutations in the 5' UTR are effectively reducing the DNA affinity of these two transcription factors, and in this way the gene is not downregulated anymore along the megakaryopoiesis, in patients cells.

Described as the ancestor of an entire, primate-specific, gene family called POTE, the only reports on the functional role of this gene describe it as involved in adipogenesis and predict a functional impact on the modulation of several signalling pathways. A first report on the role of ANKRD26, in the context of the thrombocytopenia, describes an abnormal activity of the signalling pathways dependent from the axis TPO/MPL. In particular, the MAPK pathway seems positively affected by ANKRD26, as patients MKs displayed an increased phosphorylation of ERK1/2, compared to the control MKs. This is sufficient to induce a defective proplatelet formation, therefore a possible explanation for the inherited thrombocytopenia.

Nevertheless, it is currently unknown if these mutations have a broader impact than the one described for the MK lineage. Early clinical data report in some individual an increase in the leukocytes count, but no evident anomalies in the bone marrow other than the dysmegakaryopoiesis have been described[359]. Moreover, no hypothesis have been formulated that could explain the observed predisposition to leukemia.

Aim

In this project our major focus has been the elucidation of the mechanisms underlying the leukemic predisposition. In order to do it efficiently, we developed an iPSC-based model and we combined it with primary cell culture from patient haematopoietic progenitors and an engineered cell line.

Material and Methods

We generated several iPS clones from three patients, in order to increase the robustness of our observations. Moreover, we compared directly the results obtained using the iPSCs with some primary samples obtained from patients carrying the same mutations, in order to confirm the solidity of our observations. The patients cell lines have been compared to three cell lines generated from three different genomic backgrounds, further reducing the non-specificity of our results. The differentiation protocols used for this project are largely based on the one described for the FLNa project: we optimized the culture conditions to support the proliferation and differentiation of the granulocytes, both from iPSCs and primary sources. We used a battery of tests to characterize the granulopoiesis of mutated cells, and those preliminary observations must be further validated, in order to increase our confidence in the presented results.

We also managed to enginered a cell line that is cytokine-dependent, therefore constitutes a valid tool for evaluating any changes in the signalling pathway dependent from the cytokine, in our case G-CSF. The advantage of such model is the versatility and robustness of a cancer cell line, but the translability of these results could be limited, therefore it is paramount to confirm those results in a more pathologically relevant context.

Results and conclusion

We investigated the *ANKRD26* expression in different populations, and we observed an increased expression of this gene in granulocytes and cells already committed towards the granulo-monocytic fate. Conversely, immature cells do not display any significant difference in their *ANKRD26* expression levels, an observation that prompted us to investigate the granulocytic lineage. We observed an increased proliferation, coupled with a delayed maturation, two features that are common for several haematological disorders affecting the signalling mediated by the G-CSF and its cognate receptor. These results are generated from different patients genomic backgrounds, but they need to be systematically confirmed for all the generated cell lines, in particular the data referring to the maturation, as such phenotype is highly dependent from the experimental conditions.

Based on the preliminary data generated on THC2-derived cells granulopoiesis, we decided to investigate the receptor activity and role, in a context where we could modulate the ANKRD26

expression. Exploiting two different strategies, we have been able to generate two models that allowed us to evaluate how the ANKRD26 dosage could affect the G-CSFR activity. Indeed, with both models we observed different levels of G-CSFR surface expression, corroborating an active role of ANKRD26 in cytokine-mediated signalling modulation. Moreover, the analysis of the receptor activation dynamics revealed an increased signalling output for STAT3 and MAPK, both in a time and dose-dependent fashion. This latter feature in particular captured our attention, as cytokine hypersensitivity is a feature of certain congenital disorders that can progress to overt leukemia, like severe congenital neutropenia (SCN).

We finally started to confirm the observations obtained in our cancer cell model in our iPS model. Indeed we have been able to highlight anomalies in both MAPK and STAT3 signalling output, albeit in a preliminary way. All the iPS cell lines will be thoroughly examined for those signalling defects. We also began to evaluate the impact of the deregulated signalling on the transcriptomic landscape. Further analysis is required to proper exploit the data generated by RNA-sequencing, but we have identified a promising candidate, a cyclin called CCNI2, which gene expression is increased in patient cells, compared with the control cells. We plan to investigate the impact of this gene on the cell cycle regulation.

In conclusion, we propose a novel model of preleukemic predisposition, dependent on an increased sensitivity of the granulocytic lineage to the G-CSF. This feature underlies the existence of a “fertile substrate” for the acquisition of other genomic hits and the emergence of mutated clones, therefore the described progression toward a full-blown leukemia.

Abnormal ANKRD26 expression leads to the hyperactivation of the G-CSFR signalling pathway in ANKRD26-related thrombocytopenia patients.

Alessandro Donada[1,2], Vladimir T. Manchev[1], Nathalie Balayn[1], Francesca Basso-Valentina[1], Manuel Lisetto[1], Remi Favier[1,3], Lucie Tosca[4], Gérard Tachdjian[4], Isabelle Plo[1], William Vainchenker[1] and Hana Raslova[1*].

1) UMR 1170, Institut National de la Santé et de la Recherche Médicale, Univ. Paris-Sud, Université Paris-Saclay, Gustave Roussy Cancer Campus, Equipe Labellisée Ligue Nationale Contre le Cancer, 94805 Villejuif, France

2) Université Sorbonne Paris Cité, Université Paris Diderot, 75013, Paris, France

3) Assistance Publique-Hôpitaux de Paris, Armand Trousseau Children Hospital, French Reference Center for Platelet Disorders, Hematological Laboratory, Paris, France.

4) U967, Institut National de la Santé et de la Recherche Médicale, Hôpitaux Universitaires Paris-Sud, Assistance Publique-Hôpitaux de Paris Service d'Histologie Embryologie et Cytogénétique, Clamart, France;

* **Corresponding author**: Hana Raslova, INSERM UMR1170, Gustave Roussy Cancer Campus, 114 rue Edouard Vaillant, 94805, Villejuif cedex, France.
Phone number: (+33) 1 42 11 46 71
E-mail: hana.raslova@gustaveroussy.fr
Competing interest: The authors declare no competing interests.

Abstract

Thrombocytopenia 2 (THC2 or ANKRD26-RT) is an inherited thrombocytopenia associated to an increased risk of developing myeloid malignancies. It is caused by mutations in the 5' UTR of the gene *ANKRD26*, and these anomalies induce a deregulation of the gene expression and a persistent expression during megakaryopoiesis. Moreover, megakaryocytes expressing higher levels of ANKRD26 display an increased activation of the signalling pathways downstream of the TPO receptor. We have investigated the possible role of this type of mutations in a different hematopoietic cell type and their impact on the observed predisposition. Using a combination of primary cells, patient-derived induced pluripotent stem cells and leukemic cell lines, we demonstrate that THC2 patient granulocytes show an increased proliferation and a delayed maturation *in vitro.* Those cells express increased levels of *ANKRD26* compared to the healthy control, albeit the difference is restricted to the more committed stages of the granulopoiesis. Cells expressing higher levels of *ANKRD26* display an increased surface expression of the G-CSF receptor and a hypersensitivity to the G-CSF, leading to an increased proliferation due to an abnormal activation of the signalling pathways downstream of the receptor, notably STAT3 and MAPK. The inhibition of the JAK2 kinase activity restores a normal proliferation in patient cells, suggesting that the deregulation of the G-CSF-dependent signalling is the causative mechanism for the observed anomalies in the granulopoiesis. We identified a member of the cyclin family, (CCNI2), as a promising candidate to be involved in the observed phenotype. In conclusion we propose a new model for the leukemic predisposition in THC2 patients, and we support a broader role for ANKRD26 as an important modulator of the cytokine-mediated cell signalling in the hematopoiesis.

Introduction

Inherited thrombocytopenias (ITs) with predisposition to leukemia are congenital blood disorders characterized by autosomal dominant transmission, a mild to moderate reduction in platelet count, normal mean platelet volume, bone marrow dysmegakaryopoiesis and an increased risk of hematological malignancies. To this date, mutations in three genes have been described as responsible of such conditions: *RUNX1* alterations, inducing familial platelet disorder with predisposition myelogenous leukemia (FPD/AML)[1], mutations in the 5' UTR region of *ANKRD26* gene in thrombocytopenia 2 (THC2)[2] and *ETV6* mutations in thrombocytopenia 5 (THC5)[3]. While *RUNX1* and *ANKRD26* mutations predispose to myeloid malignancies, the *ETV6* mutations are more often linked to lymphoid malignancies. The penetrance of the leukemic predisposition is very variable between the different ITs. In THC2 around 10% of patients develop a myeloid malignancy (AML, MDS, CML)[2] while in THC5 and FPD/AML it could reach 50%. ETV6 and RUNX1 are two well-known oncogenes, described in multiple acquired disorders and even associated in a prototypical gene fusion responsible for childhood B cell leukemia (cALL)[4]. In the context of germline mutation, the ETV6 mechanism leading to the development of leukemia is currently unknown. On the other hand, RUNX1 dominant negative-like mutations of RUNX1 lead to a loss of function, inducing an exacerbated amplification of myeloid progenitors displaying an increased genomic instability that could predispose to leukemia[5].

ANKRD26 is the ancestor of a family of primate-specific genes termed POTE (Prostate-, Ovary-, Testis-, and placenta-Expressed genes), that are expressed only in few normal tissues and a larger number of pathological ones. This gene encodes for a protein of 192 kDa, containing both spectrin helices and ankyrin repeats, protein domains known to interact with cytoskeletal and signalling proteins[6,7]. In a murine KO model, it was shown that Ankrd26 controls the response of white adipose tissue to insulin. Its mutations leads to hyperphagia and diabetes, in an obesity-dependent manner[8,9]. A marked adipogenesis in a *Ankrd26* KO mice was shown to be mediated by an increased activation of the ERK and the mTOR pathways[10]; furthermore, ANKRD26 was also described to interact with different proteins, including hyaluronan-mediated motility receptor (HMMR)[11]. This body of work suggests a role in the modulation of signalling pathways.

In THC2 patients, the described mutations are not affecting the coding sequence, but they all localize in a patch of few nucleotides in the 5' UTR. Those mutations affect the binding of the complex RUNX1/FLI1, responsible of the regulation of several genes implicated in the final stages of the megakaryopoiesis[12,13]. Patient cells exhibiting those mutations are unable to

down-regulate ANKRD26, therefore the gene is constantly expressed across the entire differentiation and its mRNA could be detected in patient platelets. Moreover, megakaryocytes harboring the 5' UTR mutations display an over-activation of the TPO/MPL-associated signalling pathway, notably the MAPK pathway. The sustained activation of this pathway is negatively affecting the proplatelet formation process, and it has been described as the major responsible of the observed thrombocytopenia[14].

In an attempt to elucidate the molecular details that could explain the observed leukemic predisposition, we generated several iPSC clones from three patients carrying different THC2 mutations. Combining this experimental tool with blood samples from the same patients and an engineered cellular model, we have been able to observe a patient-specific deregulation of ANKRD26 in the granulo-monocytic lineage. This deregulation leads to an increased proliferation and delayed maturation of the granulocytic lineage. These data, combined with the high degree of functional similarities between MPL and G-CSFR, prompted us to test the hypothesis of a hyperactivation of the G-CSF-dependent signalling pathways. Indeed, we have been able to highlight an increased activation of different downstream effectors of G-CSFR, in particular the MAPK and the STAT3 pathways. The inhibition of the G-CSFR signalling was able to restore a normal proliferation. In conclusion, we propose a model of leukemic predisposition based on the abnormal amplification of the granulocytic compartment, via increased G-CSF-mediated signalling. This anomaly creates the conditions for the acquisition of a secondary hit and the emergence of a truly malignant clone, and underlies a more general involvement of ANKRD26 in the modulation of the cytokine-mediated signalling.

Material and Methods

Primary cells culture: $CD34^+$ cells were isolated from peripheral blood by a positive selection using immunomagnetic bead cell-sorting system (AutoMacs; Miltenyi Biotec) and cultured in serum-free medium containing G-CSF (20 ng/mL), IL-3 (10 ng/mL) and SCF (25 ng/mL).

UT-7 cell line culture and generation: the original UT-7 G-CSFR-HA cell line was a kind gift of Dr. Nathalie Droin. This cell line and all the descendent ones were cultured in MEMa medium, in presence of 10% v/v of heat inactivated fetal bovine serum (FBS) and GM-CSF (10 ng/mL). UT-7 cells were transduced with lentiviral virus encoding for a short hairpin RNA against ANKRD26 (shANK) as previously described[14], and isolated via fluorescence-activated cell sorting (FACS) 48 hours later, based on the reporter gene expression (mCherry). In a similar way, UT-7 cells were transduced with a lentiviral vector encoding for the full-length ANKRD26 cDNA (PRRL-PGK-mCherry), and isolated via FACS 48 hours later, based on the reporter gene expression.

***iPSCs generation and expansion*:** Patient $CD34^+$ cells were isolated as previously described and cultured in serum-free medium containing EPO (1 U/mL), FLT3l (10 ng/mL), G-CSF (20 ng/mL), IL-3 (10 ng/mL), IL-6 (10 ng/mL), SCF (25 ng/mL), TPO (10 ng/mL) and GM-CSF (10 ng/mL) for 6 days. After this expansion phase, cells were transduced with the CytoTune iPS 2.0 Sendai Reprogramming Kit (Thermo Fisher) and the reprogramming was performed accordingly to the manufacturer instructions. In alternative, cells were transduced with VSV-G pseudotyped retroviruses encoding for the OSKM combination (OCT4, SOX2, KLF4 and c-MYC). Colonies with an ES-like morphology were manually isolated, expanded for a reduced number of passages and frozen. iPSCs were maintained on Essential 8 or Essential 8 Flex medium (Gibco), on plates coated with N-truncated Human recombinant vitronectin (Gibco). Cell passage was routinely performed using a solution of EDTA 500 μM in PBS 1X, or TrypLE 1X (Gibco). A Mycoplasma screening was routinely performed, accordingly to the manufacturer instructions (Sigma). Cells were kept in culture for a limited amount of passages, to prevent the surge of karyotypic and genomic anomalies. The iPSC control cell lines were already established and previously characterized[5,15].

Cytogenetics analysis: karyotype analysis was performed accordingly to the standard diagnostic procedures, after colchicine treatment.

Teratoma formation assay: pluripotent cells ($1x10^6$ cells) were resuspended in Essential 8 (140 μL) and Geltrex (60 μL) to a final volume of 200 μL, and injected intramuscularly in NSG mice. Mice were monitored and sacrificed when the size of the tumoral mass was visibly affecting the animal motility and behavior. Teratomas were excised, fixed and embedded in paraffin, while the corresponding sections were stained with hematoxylin, eosin and saffranin.

iPSCs haematopoietic differentiation: clumps of pluripotent cells were seeded on Geltrex-coated plates (Gibco), in presence of E8 medium, at day -1. The departing cell concentration was adjusted for each cell line and was comprised in the 10-20% confluence range. At Day 0, cells were transferred in a xeno-free medium based on StemPro-34 SFM (Gibco), supplemented with Penicillin/Streptomycin 1% v/v (Gibco), L-Glutamine 1% v/v (Gibco), 1-Thioglycerol 0.04 μg/mL (Sigma) and ascorbic acid 50 μg/mL (Sigma). This medium was retained for the entire experiment and supplemented with different cytokines and growth factors, accordingly to the following schedule: Days 0 – 2: BMP4 (10 ng/mL), VEGF (50 ng/mL) and CHIR99021 (2 mM). Days 2 – 4: BMP4 (10 ng/mL), VEGF (50 ng/mL) and FGF2 (20 ng/mL). Days 4 – 6: VEGF (15 ng/mL) and FGF2 (5 ng/mL). Day 6: VEGF (50 ng/mL), FGF2 (50 ng/mL), SCF (50 ng/mL) and FLT3L (5 ng/mL). Days 7-10: VEGF (50 ng/mL), FGF2 (50 ng/mL), SCF (50 ng/mL), FLT3L (5 ng/mL), TPO (50 ng/mL) and IL-6 (10 ng/mL). Days 10-20: SCF (50 ng/mL), G-CSF (25 ng/mL), and IL-3 (10 ng/mL). An exhaustive list of manufacturers could be retrieved in Supplemental Table 1.

Methylcellulose culture assay: hematopoietic progenitors were plated in triplicate at different cell densities (1000 cells/plate for primary cells, 3000 cells/plate for iPSC-derived cells) and scored for the presence of colonies 14 days after. Cells were plated in human methylcellulose medium H4434 (Stem Cell Technologies), containing recombinant human cytokines.

Proliferation assay in liquid medium: UT-7 G-CSFR-HA cells, expressing different levels of ANKRD26, were seeded at a cell density of $5x10^4$ cells/well in a 12-wells plate, in presence of GM-CSF (5 ng/mL) or G-CSF at the indicated concentrations. Cell numbers were quantified on a daily basis in triplicate, using a hemocytometer. Primary cells and iPS-derived cells were seeded at a cell density of $5x10^4$ cells/well in a 12-wells plate, in serum-free medium containing G-CSF (25 ng/mL), SCF (20 ng/mL) and IL-3 (10 ng/mL). Cell numbers were

quantified at the indicated days.

Cytospin analysis and staining: cells were centrifuged on slides for 5 minutes at 500 rpm, on a Cytospin 2 centrifuge (Shandon). Slides were stained with May-Grünwald Giemsa (MGG) staining; at least 100 cells per slide were scored. Images were obtained using a Leica DMRB microscope, using a 63x magnification objective.

RNA sequencing: cells were cell sorted and flash-frozen at -80 °C. RNA extraction and purification were performed with the RNeasy Mini Kit (Qiagen). The quality of the RNA was assessed via Bioanalyser and samples with a RIN>7 were retained for downstream cDNA synthesis, fragmentation and sequencing.

Quantitative Real Time Polymerase Chain Reaction (qRT-PCR): RNA was purified using the RNeasy Plus Micro or Mini Kit, accordingly to the cell number and the manufacturer instructions (Qiagen). Complementary DNA synthesis was performed with the Super Script II cDNA synthesis Kit or the Super Script Vilo cDNA synthesis kit (Thermo Fisher), accordingly to the amount of extracted RNA. Real Time PCRs were performed using the Takara Bio SYBR Premix Ex Taq (Clontech) reaction mix, on the 7500 Real Time PCR system (Applied Biosystems). Gene expression was assessed by comparative CT method, using PPIA or HPRT as reference genes. A list of the primers used is available in Supplemental Table 2.

Flow Cytometry**:** single cell suspensions were stained with monoclonal antibodies directly coupled to their respective fluorochromes. A list of the used antibodies, included the isotype control, is available, with the respective clone names (Supplemental Table 3). Cells were incubated with antibodies at the concentration of 1 μL/10^6 cells, in approximately 100 μL, at 4 °C for at least 30 minutes. Cells were washed before and after incubation in PBS 1X and analyzed with a BD Canto II or BD LSRFortessa cytometers (BD Biosciences). Fluorescence Activated Cell Sorting was routinely used for haematopoietic cells purification and performed on Influx, ARIA III or ARIA Fusion cell sorters (BD Biosciences).

Immunoblot analysis: cells were lysed using an adequate volume of RIPA buffer or a Laemmli buffer, according to the cell number and supplemented with a cocktail of protease inhibitors (Roche). For primary cells and iPS-derived cells, the lysis buffer was supplemented with phosphatases. Proteins were separated by SDS-PAGE and transferred on nitrocellulose

membranes. Membranes were blocked for 1 hour with a solution of BSA 5% w/v in 0.1% Tween-PBS 1X. Primary antibodies were incubated overnight, secondary antibodies coupled to horseradish peroxidase for 1 hour. Both type of antibodies were diluted in a solution of BSA 5% w/v in 0.1% Tween-PBS 1X. A list of the primary and secondary antibodies is available, with the respective clone names, concentrations and manufacturers (Supplemental Table 3). Detection was performed using Image Quant LAS 4000 (GE Healthcare).

G-CSF stimulation analysis in UT-7: cells were washed extensively to remove the GM-CSF and starved overnight in presence of 5% FBS. At least $4x10^5$ cells per condition were plated. The successive day, cells were stimulated with different concentration of G-CSF for 10 minutes or at different time points at the indicated concentration. After stimulation, cells were harvested and lysed accordingly to the Immunoblot analysis section.

G-CSF stimulation in primary cells and iPS-derived cells: cells were washed extensively to remove all present cytokines and starved overnight in serum-free medium deprived of insulin-transferrin-selenium. At least $4x10^5$ cells per condition were plated. The successive day, cells were stimulated with 25 ng/mL of G-CSF at different time points. After stimulation, cells were harvested in ice cold PBS, extensively washed to remove the bovine serum albumin presents in the serum-free medium and immediately lysed in ice-cold Laemmli buffer containing DTT 10 mM, protease inhibitor cOmplete (5X), phosphatase inhibitor cocktail 3 (25X), sodium orthovanadate 50 μM, sodium fluoride 500 μM and PMSF 1 mM. Samples were boiled for approximately 10 minutes and sonicate (20" pulse, 10" rest) for 5 minutes. They were then treated accordingly to the immunoblot analysis section.

Statistics: all data are shown as mean ± SD unless differentially specified in the legend. The statistical analyses were performed using the PRISM software. Statistical significance was established using a Student's *t* test specified in legends. Differences were considered significant at $P < 0.05$.

Study approval: studies were performed in accordance with the Declaration of Helsinki. All subjects provided their informed consent prior to their participation to this study.

Results

Defect in *in vitro* granulopoiesis derived from primary patient cells.

Previously we have shown that the mutations in 5' UTR region of *ANKRD26* gene lead to the overexpression of ANKRD26 in megakaryocytes and platelets of patients and that this exacerbated ANKRD26 expression overactivates TPO/MPL signalling and induces a proplatelet formation defect, therefore the observed thrombocytopenia[14]. Due to the functional homologies and the clinical data, describing a certain leukocytosis for ANKRD26 patients, we focused on the granulo-monocytic lineage. As shown on Figure 1A, the expression of ANKRD26 is not modified in early progenitors ($CD34^+$ cells) from patient peripheral blood, but there is a significant increase in mature granulocytes ($CD15^+$ cells). When we assessed the colony forming potential of $CD34^+$ cells, we observed a significant, more than two-fold, increase in the number of CFU-GM colonies in patients compared with the healthy controls (Figure 1B). In order to restrict our attention on the granulocytic lineage, we evaluated the *in vitro* proliferation of patient $CD34^+$ cells in a liquid medium containing only SCF, IL3 and G-CSF, so a cytokine cocktail able to induce the granulocytic differentiation[16]. We focused on the granulocytic lineage because the G-CSF receptor is part of the same cytokine receptor family of MPL. As shown on Figure 1C, an increase in proliferation was observed for 2 patients carrying different mutations, suggesting that ANKRD26 overexpression alters the G-CSFR/G-CSF signalling. To assess whether this increased proliferation is linked to a defect in maturation, the May-Grünwald and Giemsa staining was performed and cells were scored accordingly to their maturation. In both patients, the frequency of immature cells (myeloblasts, promyelocytes and myelocytes) was increased compared to controls, while the percentages of more mature metamyelocytes and polynuclear neutrophils were clearly decreased (Figure 1D). Nevertheless, the percentage of mature granulocytes for patient cells increases over time (Figure 1E), suggesting only a delay in the maturation, instead of an arrest. Overall these results show that the ANKRD26 overexpression leads to defects in the granulopoiesis, with an increased proliferation and a delay in maturation.

Defect in *in vitro* granulopoiesis derived from patient iPS cells.

In order to overcome the limitations associated to the rarity of THC2 patients, we generated an efficient model to study the physiopathological mechanisms underlying the leukemic predisposition. We derived the iPSC lines from three patients harbouring different mutations: c.-118A>C (ANKR1, P1_I), c.-127delAT (ANKR2, P2_II) and c.-127A>T (ANKR3, P3_III) mutations (Table 1). The iPSC were generated from $CD34^+$ cells isolated from peripheral

blood, both via integrative (ANKR1 and ANKR2) and non-integrative (ANKR3) reprogramming. As a control, we used three iPS cell lines derived from healthy individual and already described[5,15]. All new iPSC lines were characterized for their phenotypic (data not shown) and functional pluripotency, via teratoma formation assay (Supplemental Figure 1). We verified also the genomic stability by karyotypic analysis and comparative genomic hybridization, and selected two clones for each patient for further studies (total of 6 cell lines).

We used a serum-free, xeno-free differentiation protocol, adapted from the one described by Chou and colleagues[17]. A detailed description of the protocol can be found in Supplemental Figure 2. We isolated the hematopoietic progenitors after 14 days of culture ($CD34^+CD43^+$), as this population has been described to have a bias towards the granulo-monocytic differentiation output[5]. First, we checked the expression of *ANKRD26* in $CD34^+CD43^+$ cells, observing a three-fold increase in cells from all the three genetic backgrounds, compared to the healthy controls (Figure 2A). Compared to the CD34+ cells isolated from peripheral blood, the CD34+/43+ are embryonic in origin, therefore restricted in their self-renewal and differentiation properties. Nevertheless, they constitute a valid proxy for the adult CD34+ cells, as they are able to differentiate towards the grnaulo-monocytic lineage. The investigation of the colony formation potential of these cells revealed a significant three-fold increase in the number of granulo-monocytic progenitors for all patient cell lines (Figure 2B). Similarly to primary cells, an increased proliferative rate was also detected for one patient cell lines (ANK1) (Figure 2C). In order to verify the existence of any differentiation abnormalities, we tested the differentiation kinetics of iPS-derived cells cultivated in presence of G-CSF, IL3 and SCF, by flow cytometry (Figure 2D). The expression of CD11b alone within the hematopoietic $CD43^+$ cells ($CD43^+$, $CD11b^+$, $CD14^-$, $CD15^-$) is characteristic for immature GM progenitors, while the mature monocytes express CD14 and the mature granulocytes are positive for CD15. Clearly, a delay in granulocytic differentiation is detected, as a larger population of patient cells is not even expressing CD11b (83% for patient cells, compared to 42% for control cells) and importantly, patient cells are blocked at the $CD11b^+CD15^-$ stage and do not acquire the granulocytic marker CD15 (Figure 2E). These preliminary results confirm the observations obtained for primary patient cells and clearly show a similar defect in granulopoiesis, combined with an increased proliferation rate.

ANKRD26 overexpression induces a hypersensitivity of G-CSF receptor to G-CSF.

To further investigate the effect of ANKRD26 overexpression on G-CSFR/G-CSF signalling, we generated several cell models based on the UT-7 cell line, modified to express the G-CSF receptor tagged with an HA tag. One model is based on the overexpression of the cDNA of ANKRD26, tagged with a HaloTag; a second model is based on a shRNA against *ANKRD26*, as the UT-7 cell line already express basal levels of *ANKRD26*. In both cases, we are able to modulate the levels of ANKRD26, therefore study the dose-dependent effects of this gene on the G-CSF-dependent signalling. First, the increase and the decrease of ANKRD26 in both models were confirmed by qRT-PCR (Figure 3 A,C). As shown on Figure 3B, an increase in ANKRD26 expression level leads to an increase in G-CSFR levels at the cell surface, while a decrease in ANKRD26 expression level reduces the G-CSFR expression (Figure 3D). These results demonstrate that ANKRD26 plays a direct or indirect role in retaining the G-CSF receptor at the cell surface. To know whether this increased level of G-CSFR could lead to the deregulation in the downstream signalling effectors, we analyzed the phosphorylation status of four classic effectors of the G-CSF receptor. While the phospho-STAT5 level was not affected (data not shown), phospho-STAT3 was increased in presence of higher levels of ANKRD26. The same result was obtained for phosphoERK1/2, but not for AKT (Figure 3E). We also verified if the increased levels of receptor at the surface could spur an increased sensitivity to lower levels of G-CSF. We tested growing concentrations of the cytokine and indeed we observed that even at the lowest concentration it was already evident a certain activation of STAT3 and MAPK (Figure 3F). To test the functional consequences of the hypersensitivity to G-CSF, we plated UT-7 cells in presence of different concentration of G-CSF and assessed their proliferation at different days of culture. Indeed, a dose of 0.1 ng/mL of G-CSF was able to induce the proliferation of cells with higher level of ANKRD26, while the cells transduced with shANK did not proliferate at the same cytokine concentration. Using more elevated doses of G-SCF (more than 1 ng/mL), no statistically significant differences in proliferation rate were detected anymore (Figure 3G), suggesting a saturation of the cell surface pool of available G-CSF receptor. These experiments demonstrate that the ANKRD26 overexpression leads to the increased expression of G-CSFR at cell surface and to the cell hypersensitivity to G-CSF. Further studies would be looking at the precise mechanism linking the increase ANKRD26 levels with the receptor dynamics.

Hyperactivation of JAK2/STAT pathway leads to the increased proliferation of the granulocytic lineage in THC2 patients.

To confirm these results in patient cells, we checked the G-CSFR/G-CSF signalling activation during the granulocytic differentiation. We observed an increased STAT3 signal (Fig 4Ai) in patient-derived iPSC, and an increased MAPK signalling (Fig 4Aii) in patient $CD34^+$ cells. Both results must be confirmed in other iPSC lines and possibly in primary cells. To confirm that the deregulation of the signalling pathways is directly impacting the observed increase in the proliferation rate, we treated patient $CD34^+$ cells with a JAK1/2 inhibitor (Ruxolitinib) and an inhibitor of MAPK activity (PD080259). The concentrations for each inhibitor were derived from previous results[14,18]. In two independent experiments (Figure 4B), the JAK2 inhibitor led to a marked decrease in proliferation, while MAPK has no or only a mild effect. Moreover, the proliferation decrease induced by the inhibition of JAK2 signalling restored a proliferation rate that is comparable to what is observed for control cells (Fig 4Bii). These results corroborated the hypothesis that the JAK2/STAT3 signalling is enhanced in presence of ANKRD26, however they should be confirmed by using a specific approach targeting STAT3.

Finally, in order to understand the difference induced by the abnormal signalling, we investigated the transcriptomic profile of the patient granulocytic progenitors: iPSC-derived granulocytic progenitors ($CD43^+CD11b^+CD14^-$) were generated and sorted for all three iPSC lines (Fig 4Ci), and the RNA-seq analysis was performed, comparing their transcriptomic profile with granulocytes derived from three control iPS cell lines. Interestingly, only 24 genes were significantly up-regulated and 7 down-regulated ($P<1x10^5$) in patient cells compared to controls (Figure 4Cii). Within the up-regulated genes, the *CCNI2* gene (Cyclin I Family member 2), a cyclin responsible of the cyclin-dependent kinase 5 (CDK5) activity, caught our attention, as CDK5 was previously described to phosphorylate NOXA in presence of glucose. This phosphorylation promotes the cytosolic sequestration of NOXA and suppresses its apoptotic function in leukemic cells[19]. The CCNI2 implication in increased granulocytic proliferation in THC2 patients is under investigation.

Discussion

The germline mutations in 5' UTR region of gene encoding for ANKRD26 induces the thrombocytopenia and also predispose to the development of myeloid malignancies. Previously we have reported a hyperactivation of MAPK pathway in patient megakaryocytes as causal of dysmegakaryopoiesis and disrupted proplatelet formation, while in the present

work we focused on the granulo-monocytic lineage, to describe a potential mechanism leading to the preleukemic state.

The hypothesis of a deregulation in the signalling mediated by the G-CSF receptor stems from two main observations: the increased presence of granulo-monocytic colonies observed in the methylcellulose assay, and the detection of the increased *ANKRD26* levels in the $CD15^+$ cells. Both observations, together with the increased leukocytes levels found in several patients[2], supported the idea that this gene could have a broader effect, impacting other hematopoietic lineages and not only the megakaryocytic one. The G-CSF receptor is part of the same type I cytokine receptor family that includes MPL and the EPOR, and all three share a considerable degree of structural and functional homology. For this reason, we made the hypothesis that ANKRD26 could be involved in the signalling modulation of these three receptors. To investigate the potential role of ANKRD26 deregulation in establishment of pre-leukemic state in THC2 patients, we focused here on G-CSFR.

Defects in the G-CSF-dependent signalling have been described to severely impact the granulocytic lineage. Mutations in the G-CSF receptor have been associated both to neutropenia and myeloproliferative neoplasm[20]. In the first case, it is generally the case of truncating mutations, disrupting the binding of the cytokine to the extracellular portion of the receptor and making it unable to activate the downstream effectors. In the second case, mutations are affecting the transmembrane and intracellular parts of the protein, and induce a constitutive activation of the receptor. The classic example of this type of mutations is the T618I substitution, that have been described in more than 80% of the chronical neutrophilic leukemia (CNL) cases. CNL is a rare myeloproliferative neoplasm affecting the granulocytic lineage, and the T618I mutation induces an excessive proliferation of neutrophils[21]. The signalling pathways activated by the G-CSF are primarily involved in the proliferation and differentiation of hematopoietic progenitors into mature granulocytes, and play an essential role in modulating the neutrophils production under basal and emergency conditions[22]. Our results present a more nuanced picture than the one due to the mutations in the G-CSF receptor. We have not observed a constitutive activation in absence of the G-CSF, but a hypersensitivity of the receptor to G-CSF. This milder phenotype could explain the lower predisposition penetrance, compared with the mutations affecting directly the receptor physiology. Indeed, we have not observed a block of the differentiation, as typically observed in MDS or congenital neutropenia, but only a delayed maturation, a feature that could be dependent from the increased proliferation. The defective granulopoiesis observed *in vitro* constitutes a novel observation for the THC2 patients and the ITs in general. The bone

marrow examinations do not suggest any major anomaly for the granulocytic lineage, albeit an unexplained leukocytosis was described for several patients, as an increase in hemoglobin levels and possibly erythrocytosis[2,23]. This picture fits well with the absence of a marked alteration of the G-CSFR signalling.

Anomalies in the proliferation and/or maturation of the granulocytic compartment could lead to the generation of a "fertile substrate" for the acquisition of a secondary hit that will support the emergence of a malignant clone, as it is the case for the severe congenital neutropenia (SCN). This hereditary condition could be broadly described as an arrest of the granulocytic maturation and a reduced number of circulating neutrophils. Although mutations in more than 15 different genes have been described as causal for SCN, the underlying physiopathology point out towards a fundamental disruption of the G-CSF dependent signalling[24]. In this context, some patients could acquire additional mutations, notably in the G-CSF receptor, or RUNX1 and further progress towards the evolution of a true MDS or AML. The G-CSFR mutations are typically truncating mutations of the cytoplasmic domain that induces higher levels of receptor at the surface and cytokine hypersensitivity. The mutated receptor induces a disruption of the balance phosphoSTAT5/phosphoSTAT3 and the abnormal activation of the downstream effectors PI3K-AKT and MAPK, leading to the increased production of ROS.

In normal conditions, the G-CSFR, like MPL and EPOR, undergoes conformational changes in response to cytokine binding, and this will leads to the activation of the associated JAK kinase and downstream signalling cascades. At the same time, different mechanisms that down-regulate these receptors activity, including internalization, endocytosis and subsequent degradation, are rapidly turned on to weaken the receptor signalling. The role of the ubiquitin/proteasome system was well demonstrated in G-CSFR internalization[25]. This receptor was shown to be ubiquitinated via both ligand-dependent and independent mechanisms, and its endocytosis was inhibited after disruption of the ubiquitin machinery. The generation of a murine mutant (K762R/G-CSFR), that abrogates the ubiquitinylation of the lysine residue at position 762 of the receptor, showed an impaired mono/polyubiquitinylation in response to G-CSF. Consequently, an enhanced sensitivity to G-CSF and increased cell proliferation and cell survival after cytokine depletion were observed. Moreover, the region containing this residue is frequently deleted in SCN patients that progress to AML (SCN/AML), underscoring a pathological relevance of this type of mechanism in the leukemic progression. These phenotypes were linked to the prolonged activation of STAT5 and AKT. On a similar note, the mutation of five lysine residues to arginine in the same

motif were also reported to result in hyperproliferative responses to G-CSF and with prolonged activation of Stat5 and Stat3[26]. Overall, all these observations are comparable with our results on THC2 patient granulocytes, and thus a defect in receptor degradation, due to alterations in the ubiquitinylation, is an attractive hypothesis to explain the preleukemic model in THC2 patients. The ubiquitinylation of the EGF receptor was also described as a major driver of its endocytosis[27]. The internalization of this ubiquitinylated receptor was described as being regulated by ANKRD13, a protein containing several ubiquitin interacting motifs (UIM). ANKRD26 does not contain an UIM domain, however it contains a CCDC144C protein coiled-coiled region. The members of coiled-coil domain-containing protein (CCDC) family can potentially interact with ubiquitin specific peptidases 25 and 32 (USP25 and USP32), mediating the ubiquitin release from the receptor. It would be therefore of interest to further explore the potential role of ANKRD26 in the ubiquitin-dependent internalization of G-SCFR

To decipher the downstream mechanism leading to the deregulation of granulocytic lineage, we performed a RNA-sequencing analysis of the immature progenitors derived from patient iPS cells. Among the deregulated genes, CCNI2 is of particular interest, because it is a cyclin-dependent kinase-5 activator (CDK5). CDK5 is a kinase regulating the pro-apoptotic function of NOXA, as it phosphorylates a serine residue (S^{13}) in presence of glucose, and this phosphorylation promotes the cytosolic sequestration of NOXA and the suppression of its apoptotic function in leukemic cells[19]. The role of increased CCNI2 level in granulocyte progenitors should be further investigated.

Overall, our study show that an increased ANKRD26 expression in granulo-monocytic progenitors induces a defect in G-CSFR internalization and increased cells hypersensitivity to the cytokine, leading then to the overactivation of the G-SCFR/G-SCF signalling cascades. This sequence of events triggers the hyperproliferation of the granulocytic progenitors and a delayed maturation, creating a cellular compartment more sensitive to the leukemic transformation.

Acknowledgments

The authors thank the patients for participation in this study, O. Bawa and P. Opolon from the Preclinical research platform, Gustave Roussy, Villejuif, France, for the teratomas analysis, P. Rameau, C. Catelain and Y. Lecluse from the Imaging and cytometry platform PFIC, UMS AMMICA, Gustave Roussy Villejuif, France for the expertise in cytometry.

This work was supported by French grant from Ligue Nationale Contre le Cancer (équipe labellisée 2016 to H.R.), by the European grants ERA-NET and H2020-FETOPEN-1-2016-2017- SilkFusion. AD was supported by a PhD fellowship from the Sorbonne Paris Cité and Ligue National Contre le Cancer.

Author contributions

AD, VTM, NB, FBB, ML, LT, GT performed and analyzed experiments. AD, WV, IP, HR discussed results. RF provided clinical and biological follow-up of patients. HR supervised the work. AD and HR wrote the article.

Disclosure of Conflicts of Interest

The authors have no competing financial interests to declare.

References

1. Song, W. J. *et al.* Haploinsufficiency of CBFA2 causes familial thrombocytopenia with propensity to develop acute myelogenous leukaemia. *Nat. Genet.* **23,** 166–175 (1999).

2. Noris, P. *et al.* Mutations in ANKRD26 are responsible for a frequent form of inherited thrombocytopenia: analysis of 78 patients from 21 families. *Blood* **117,** 6673–80 (2011).

3. Zhang, M. Y. *et al.* Germline ETV6 mutations in familial thrombocytopenia and hematologic malignancy. *Nat. Genet.* **47,** 180–185 (2015).

4. Greaves, M. A causal mechanism for childhood acute lymphoblastic leukaemia. *Nat. Rev. Cancer* 1–14 (2018).

5. Antony-Debré, I. *et al.* Level of RUNX1 activity is critical for leukemic predisposition but not for thrombocytopenia. *Blood* **125,** 930–40 (2015).

6. Djinovic-Carugo, K., Gautel, M., Ylänne, J. & Young, P. The spectrin repeat: A structural platform for cytoskeletal protein assemblies. *FEBS Letters* **513,** 119–123 (2002).

7. Mosavi, L. K., Cammett, T. J., Desrosiers, D. C. & Peng, Z. The ankyrin repeat as molecular architecture for protein recognition. *Protein Sci.* **13,** 1435–1448 (2004).

8. Bera, T. K. *et al.* A model for obesity and gigantism due to disruption of the Ankrd26 gene. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 270–5 (2008).

9. Raciti, G. A., Bera, T. K., Gavrilova, O. & Pastan, I. Partial inactivation of Ankrd26 causes diabetes with enhanced insulin responsiveness of adipose tissue in mice. *Diabetologia* **54,** 2911–2922 (2011).

10. Fei, Z., Bera, T. K., Liu, X., Xiang, L. & Pastan, I. Ankrd26 gene disruption enhances adipogenesis of mouse embryonic fibroblasts. *J. Biol. Chem.* **286,** 27761–8 (2011).

11. Liu, X.-F. *et al.* ANKRD26 and its interacting partners TRIO, GPS2, HMMR and DIPA regulate adipogenesis in 3T3-L1 cells. *PLoS One* **7,** e38130 (2012).

12. Bluteau, D. *et al.* Dysmegakaryopoiesis of FPD/AML pedigrees with constitutional RUNX1 mutations is linked to myosin II deregulated expression. *Blood* **120,** 2708–18 (2012).

13. Lordier, L. *et al.* RUNX1-induced silencing of non-muscle myosin heavy chain IIB contributes to megakaryocyte polyploidization. *Nat. Commun.* **3,** 717 (2012).

14. Bluteau, D. *et al.* Thrombocytopenia-associated mutations in the ANKRD26 regulatory region induce MAPK hyperactivation. *J. Clin. Invest.* **124,** 580–91 (2014).

15. Saliba, J. *et al.* Heterozygous and Homozygous JAK2(V617F) States Modeled by Induced Pluripotent Stem Cells from Myeloproliferative Neoplasm Patients. *PLoS One* **8,** e74257 (2013).

16. Lachmann, N. *et al.* Large-scale hematopoietic differentiation of human induced pluripotent stem cells provides granulocytes or macrophages for cell replacement therapies. *Stem Cell Reports* **4,** 282–296 (2015).

17. Chou, S. T. *et al.* Trisomy 21-associated defects in human primitive hematopoiesis revealed through induced pluripotent stem cells. *Proc. Natl. Acad. Sci.* **109,** 17573–17578 (2012).

18. Chachoua, I. *et al.* Thrombopoietin receptor activation by myeloproliferative neoplasm associated calreticulin mutants. *Blood* **127,** 1325–1335 (2016).

19. Lowman, X. H. *et al.* The Proapoptotic Function of Noxa in Human Leukemia Cells Is Regulated by the Kinase Cdk5 and by Glucose. *Mol. Cell* **40,** 823–833 (2010).

20. Dwivedi, P. & Greis, K. D. Granulocyte colony-stimulating factor receptor signaling in severe congenital neutropenia, chronic neutrophilic leukemia, and related

malignancies. *Exp. Hematol.* **46,** 9–20 (2017)

21. Maxson, J. E. & Tyner, J. W. Genomics of chronic neutrophilic leukemia. *Blood* **129,** 715–722 (2017).

22. Touw, I. P., Palande, K. & Beekman, R. Granulocyte Colony-Stimulating Factor Receptor Signaling. *Hematol. Oncol. Clin. North Am.* **27,** 61–73 (2013).

23. Perez Botero, J. *et al.* Clinical spectrum and clonal evolution in germline syndromes with predisposition to myeloid neoplasms. *Br. J. Haematol.* **182,** 141–145 (2018).

24. Skokowa, J., Dale, D. C., Touw, I. P., Zeidler, C. & Welte, K. Severe congenital neutropenias. *Nat. Rev. Dis. Prim.* **3,** 17032 (2017).

25. Ai, J., Druhan, L. J., Loveland, M. J. & Avalos, B. R. G-CSFR ubiquitination critically regulates myeloid cell survival and proliferation. *PLoS One* **3,** e3422 (2008).

26. Irandoust, M. I. *et al.* Suppressor of cytokine signaling 3 controls lysosomal routing of G-CSF receptor. *EMBO J.* **26,** 1782–93 (2007).

27. Tanno, H., Yamaguchi, T., Goto, E., Ishido, S. & Komada, M. The Ankrd 13 family of UIM-bearing proteins regulates EGF receptor endocytosis from the plasma membrane. *Mol. Biol. Cell* **23,** 1343–1353 (2012).

Table 1

Patient ID	Mutation	Sex	Family history of malignancy	Platelet count (x10^9/L)	Bone marrow morphology	Reprogramming Method
P1_I (P1)	c. -118A>C	M	no	45	Dysmegakaryopoiesis	Integrative (Retrovirus)
P2_II (P2)	c. -127delAT	F	no	26	Dysmegakaryopoiesis	Integrative (Retrovirus)
P3_III (P3)	c. -127A>T	M	yes	40	Dysmegakaryopoiesis	Non Integrative (Sendai Virus)
P4_II (P4)	c. -127delAT	M	no	55	ND	
P5_II (P5)	c. -127delAT	M	no	56	ND	

Table 1. **Summary for patient mutations and main clinical features.** The reprogrammed patients and the used strategy are also highlighted.

Figure 1

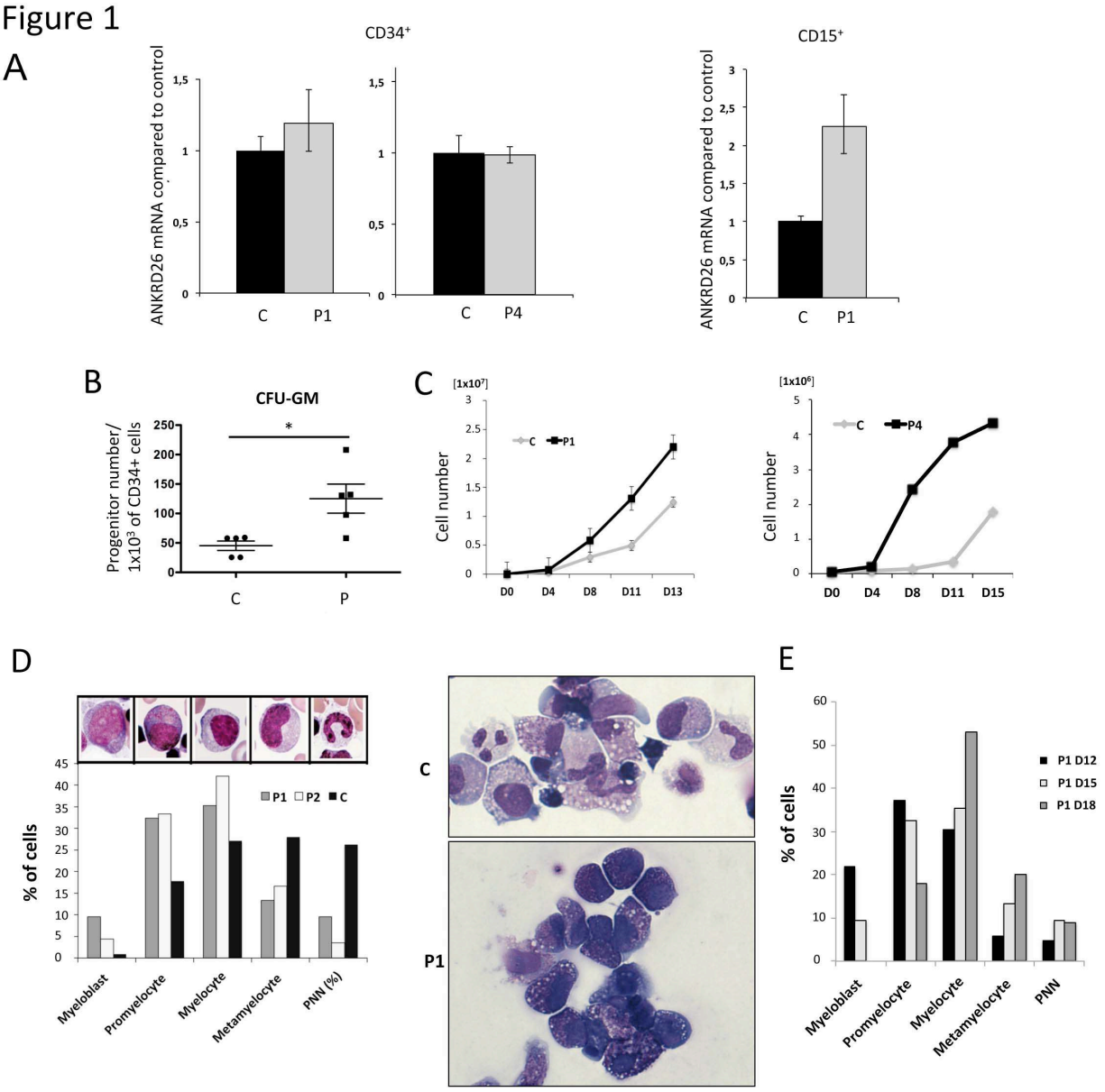


Figure 1. **A defective in *in vitro* granulopoiesis is observed in THC2 patient primary cells.** **(A)** RT-qPCR analysis of *ANKRD26* transcript level normalized to *PPIA* in CD34⁺ and D15⁺ cells from THC2 patients relative to healthy subjects (n = 2 for CD34⁺ cells and n = 1 for CD15⁺ cells). **(B)** Methylcellulose culture assay for CD34⁺ cells isolated from patients peripheral blood. The histograms represent the number of colonies CFU- GM evaluated at day 14 of culture (n = 5 for patients, n = 5 for healthy controls, **P*<.05, paired t-test). **(C)** Proliferation assay for CD34⁺ cells, isolated from the peripheral blood of two patients (P1 and P4) and two independent controls. Cells were plated in triplicate and counted at the indicated time points (D4-D15). **(D-E)** Percentage of cells at different stages of granulopoiesis, from *in vitro* culture of CD34+ of patients P1 and P2 and a healthy control. Cells were kept in culture for 18 days and evaluated via May Grünwald Giemsa staining (n = 1).

Figure 2

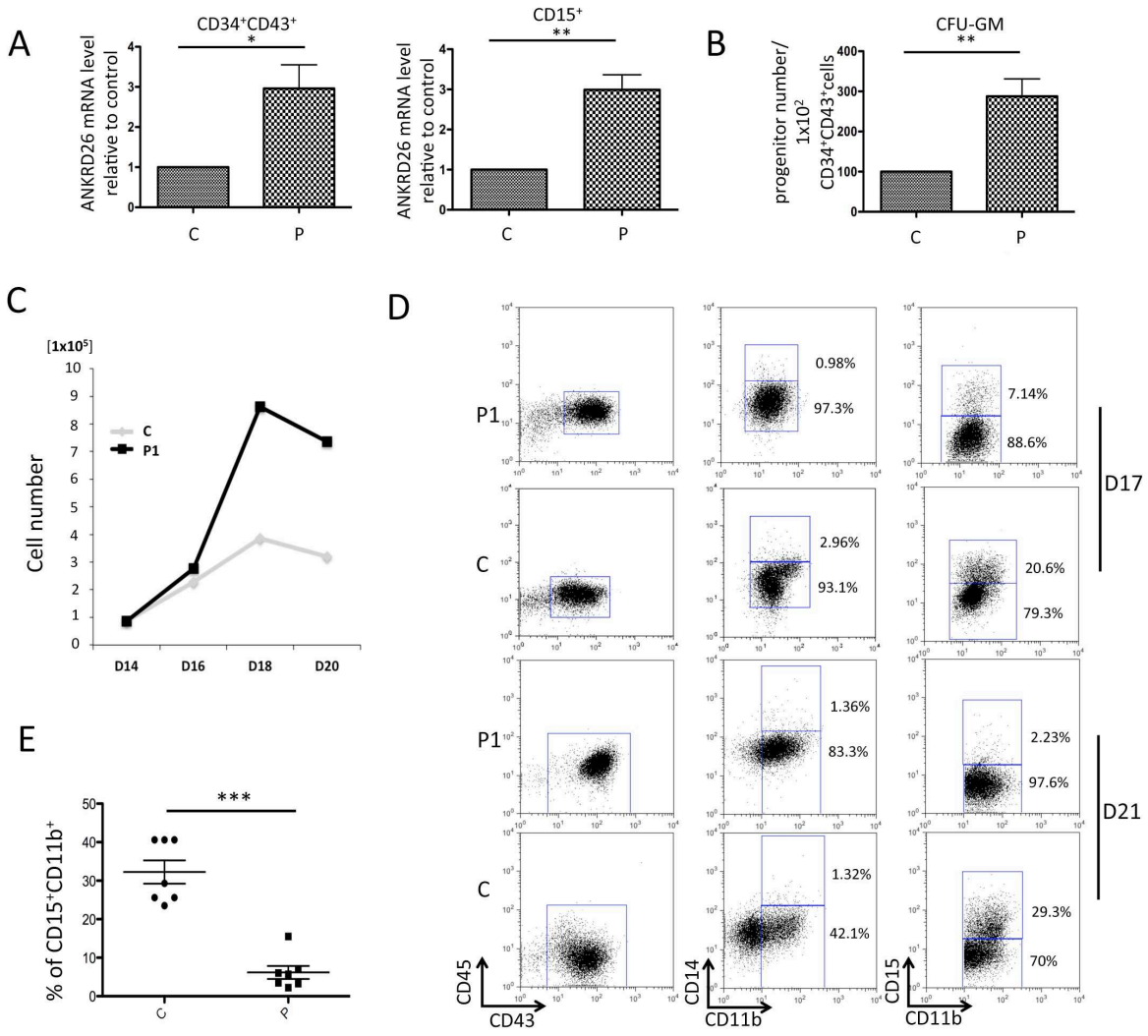


Figure 2. **Patient iPSC-derived granulocytes recapitulates the same defects observed in primary cells. (A)** RT-qPCR analysis of *ANKRD26* transcript level normalized to *PPIA* in CD34⁺CD43⁺ and CD15⁺ cells from patient-derived iPSC relative to control iPSC (n=4 for CD34⁺CD43⁺ cells and n = 6 for CD15⁺ cells, **P*<.05, ***P*<.001, paired t-test). **(B)** The CD34⁺CD43⁺ were sorted at day 14 and plated in methycellulose in triplicate. The histograms present the number of myeloid progenitors (CFU-GM) evaluated at day 14 of culture (n = 7, ***P*<.01, paired t-test). **(C)** Proliferation assay for iPSC-derived CD34⁺CD43⁺ progenitors, sorted at day 14. Cells were plated and counted at the indicated days (D16-D20) **(D)** Representative data of flow cytometry analysis at day 17 and 21 showing a delay in granulocytic maturation. **(E)** Percentage of CD15⁺CD11b⁺ cells at day 21-22 of culture (n = 7 for patient iPSC lines, n = 4 for control iPSC cell lines, ****P*<.001, paired t-test).

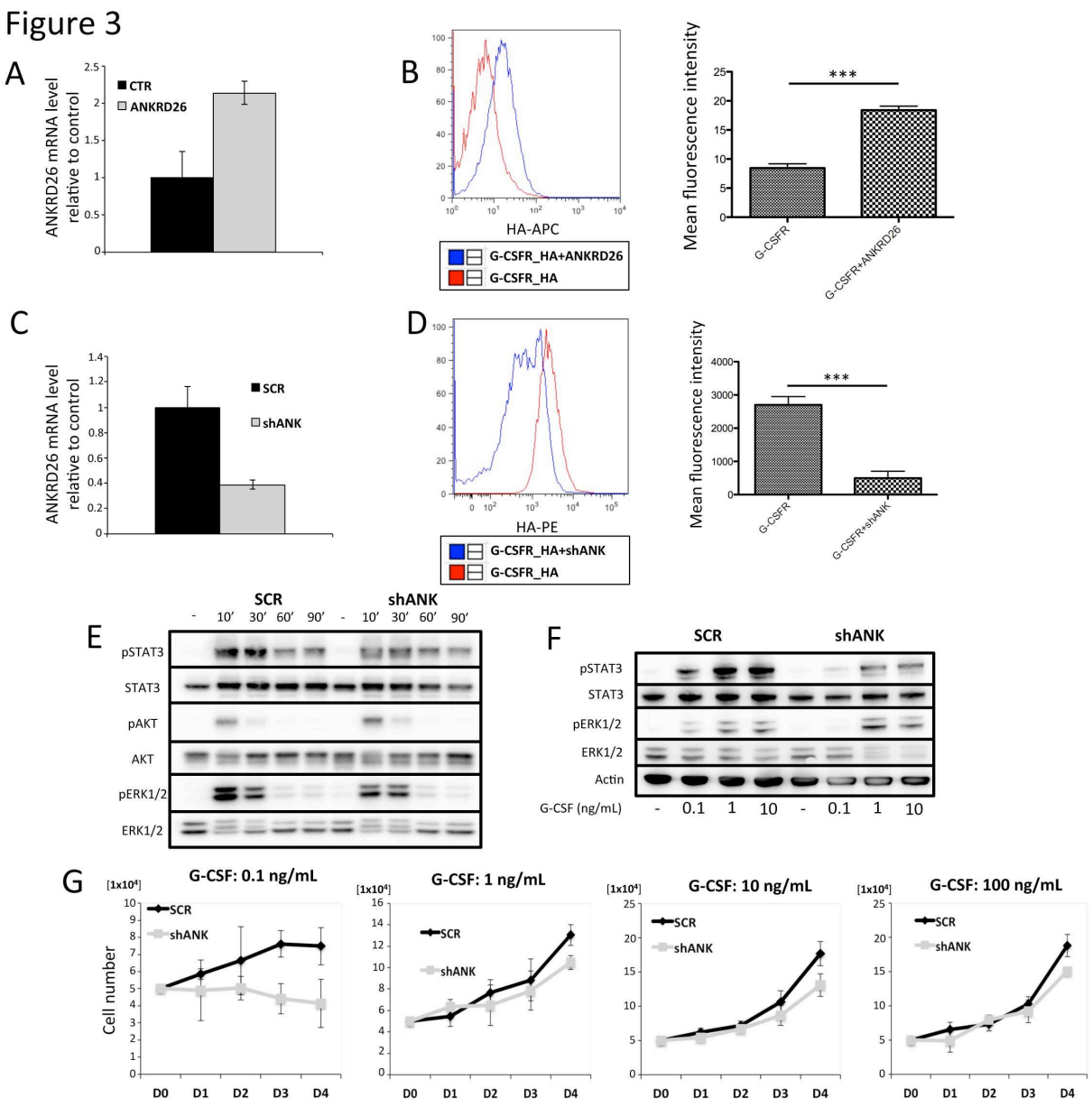
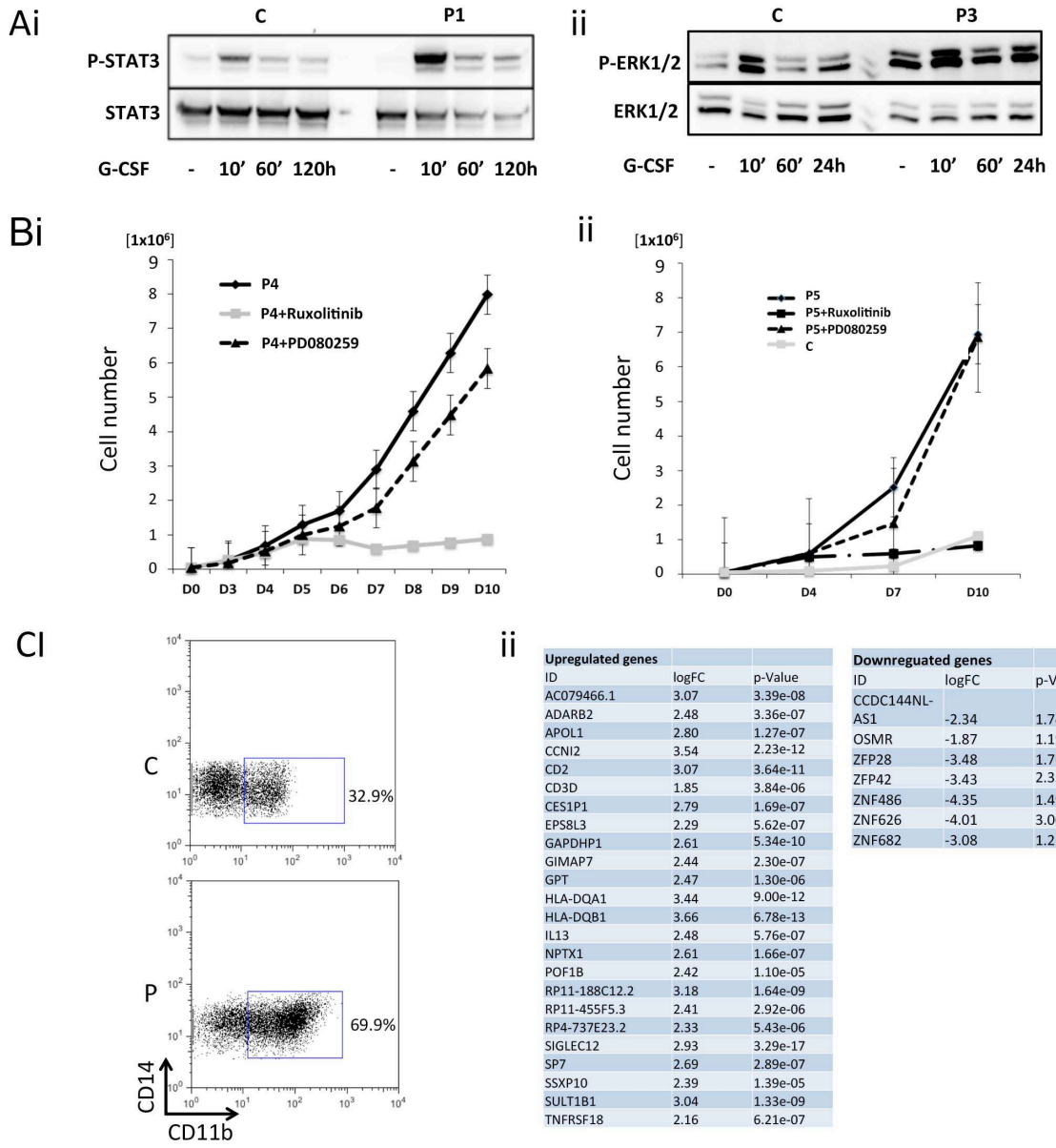


Figure 3. **ANKRD26 increased levels induce a hypersensitivity to G-CSF. (A)** Expression levels of *ANKRD26* mRNA in the UT-7 G-CSFR-HA cell line over-expressing the full-length cDNA, compared with the control cell line, normalized on PPIA. **(B)** Surface levels of G-CSFR-HA, compared to the control cell line. The histogram represents the mean fluorescence intensity (MFI) on the APC channel, n = 3, ***P<.001. **(C)** Expression levels of *ANKRD26* mRNA in UT-7 G-CSFR-HA cells transduced with the short-hairpin RNA against ANKRD26 (ShANK), compared with a control short-hairpin (SCR). The *ANKRD26* expression is normalized on PPIA. **(D)** Surface levels of G-CSFR-HA. The histogram represents the mean fluorescence intensity (MFI) on the APC channel, n = 3, ***P<.001. **(E,F)** Phosphorylation of STAT3, AKT and MAPK, in UT-7 G-CSFR SCR and ShANK. **(E)** Cells were starved overnight and stimulated at the indicated time points with 25 ng/mL of G-CSF. **(F)** Cells were starved overnight and stimulated for 10 minutes at the indicated concentration of G-CSF. **(G)** Proliferation assay for

UT-7 G-CSFR SCR and ShANK, cultured in presence of the indicated concentration of G-CSF. The error bars represent ± SD of triplicate. One of three experiments with similar results is shown.

Figure 4



ii

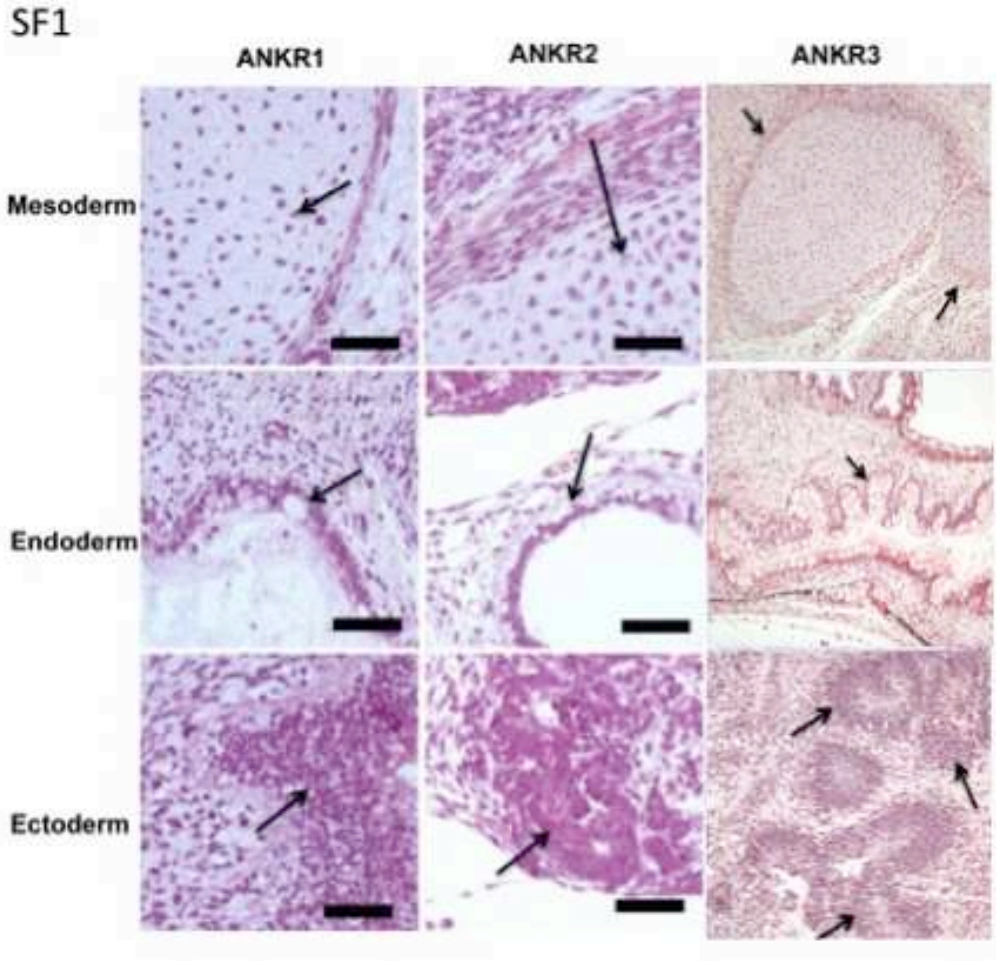
Upregulated genes		
ID	logFC	p-Value
AC079466.1	3.07	3.39e-08
ADARB2	2.48	3.36e-07
APOL1	2.80	1.27e-07
CCNI2	3.54	2.23e-12
CD2	3.07	3.64e-11
CD3D	1.85	3.84e-06
CES1P1	2.79	1.69e-07
EPS8L3	2.29	5.62e-07
GAPDHP1	2.61	5.34e-10
GIMAP7	2.44	2.30e-07
GPT	2.47	1.30e-06
HLA-DQA1	3.44	9.00e-12
HLA-DQB1	3.66	6.78e-13
IL13	2.48	5.76e-07
NPTX1	2.61	1.66e-07
POF1B	2.42	1.10e-05
RP11-188C12.2	3.18	1.64e-09
RP11-455F5.3	2.41	2.92e-06
RP4-737E23.2	2.33	5.43e-06
SIGLEC12	2.93	3.29e-17
SP7	2.69	2.89e-07
SSXP10	2.39	1.39e-05
SULT1B1	3.04	1.33e-09
TNFRSF18	2.16	6.21e-07

Downreguated genes		
ID	logFC	p-Value
CCDC144NL-AS1	-2.34	1.74e-07
OSMR	-1.87	1.19e-05
ZFP28	-3.48	1.76e-14
ZFP42	-3.43	2.33e-10
ZNF486	-4.35	1.49e-16
ZNF626	-4.01	3.00e-14
ZNF682	-3.08	1.22e-08

Figure 4. **Hyperactivation of JAK2/STAT pathway leads to the increased proliferation of the granulocytic lineage in THC2 patients. (A)** Western blot analysis of STAT3 and MAPK pathways in granulocytic lineage. **(i)** $CD34^{+}CD43^{+}$ cells derived from patient's iPSC were sorted at day 14, cultured in presence of G-SCF, SCF and IL3 for six days, starved over night and stimulated with G-CSF 20 ng/mL at the indicated time points. **(ii)** Patient granulocytes derived from CD34+ *in vitro* were starved over night and stimulated with G-CSF 20 ng/mL at the indicated time points. **(B)** Proliferation assay of $CD34^{+}$ cells from patient peripheral blood. Ruxolitinib and PD080259 were added at D0 to the culture. The cells were counted at the indicated days (D4-D10). The error bars represent ± SD of triplicate. Two independent experiments were performed for P4 and P5 (i and ii). **(Ci)** Granulocytic progenitors from iPSCs ($CD43^{+}CD14^{-}CD11b^{+}$) were sorted at day 16 of culture and analyzed via RNA

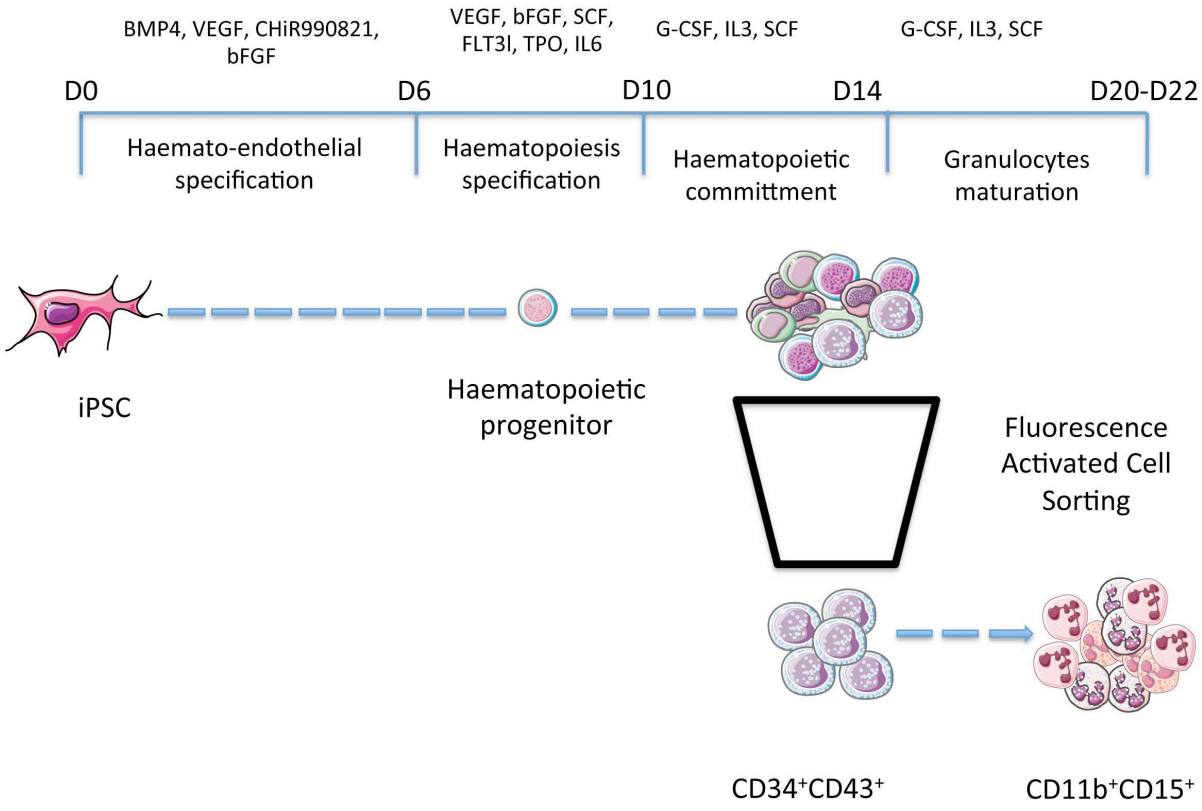
sequencing. Two clones per each patient iPS cell line (n = 6) and three control cell lines (n = 3) were analyzed. **(ii)** The tables show lists of up- and down-regulated genes with $P<1x10^{5}$.

Supplementary Data



Supplemental figure 1. **Characterization of the functional pluripotency for the THC2 patient-iPSCs.** Teratoma formation assay, images are representative of one clone per each different genomic background. Scale bar = 25 μm for ANKR1 and ANKR2, 100 μm for ANKR3.

SF2



Supplemental figure 2. **Schematic representation of the hematopoietic differentiation protocol.** The described protocol could be subdivided in four different stages: in the first phase the iPSCs are differentiated toward the hemato-endothelial progenitor fate, by a mesodermal induction operated in the first two days of differentiation. In the second stage, the culture conditions are adjusted for the support of the endothelial-to-hematopoietic transition of the first hematopoietic cells. The latter two stages recapitulate the early and late granulopoiesis, so the culture conditions are adapted for the hematopoietic cells.

Reagent	Source	Identifier
hBMP4	Peprotech	#AF-120-05ET
hVEGF	Peprotech	#100-20
hFGF-basic	Peprotech	#100-18B
hIL-6	Peprotech	#200-06
hEPO	Peprotech	#100-64
hG-CSF	Peprotech	#300-23
hGM-CSF	Peprotech	#300-03
hIL-3	Peprotech	#200-03
hTPO	Kyowa Kirin, Tokyo, Japan	/
hFLT3l	Celldex Therapeutics, Inc., Needham, USA	/
hSCF	Biovitrum AB, Stockolm, Sweden	/
CHiR 99021 trihydrochloride	TOCRIS	#4953
Y-27632 dihydrochloride	TOCRIS	#1254
VTN-N	Gibco/Thermo Fisher Scientific	#A14700
Geltrex	Gibco/Thermo Fisher Scientific	#A1413202
StemPro®-34 SFM	Gibco/Thermo Fisher Scientific	#10639011
E8	Gibco/Thermo Fisher Scientific	#A1517001
E8 Flex	Gibco/Thermo Fisher Scientific	#A2858501
CytoTune™-iPS 2.0 Sendai Reprogramming Kit	Invitrogen/Thermo Fisher Scientific	#A16517
Minimum Essential Medium a	Gibco/Thermo Fisher Scientific	#22561021
Fetal Bovine Serum	Hyclone	#SV30160.03
1-thioglycerol	Sigma	#M6145
MethoCult H4434 Classic	Stem Cell Technologies	#04434
cOmplete™ inhibitor cocktail	Sigma/Roche	#4693159001
Phosphatase Inhibitor Cocktail 3	Sigma	#P0044
PMSF	Sigma	#P7626
Sodium Orthovanadate	Sigma	#S6508
Sodium Fluoride	Sigma	#201154

Supplementary table 1: **list of reagents, with source and reference code.**

Oligonucleotide	Source	Sequence
ANKRD26_F	Eurogentec	CTATGTCAGAGGCTTCACTGGAG
ANKRD26_R	Eurogentec	CTCAGCACATCTGACAGCTTCTG
PPIA_F	Eurogentec	GTCAACCCCACCGTGTTCTT
PPIA_R	Eurogentec	CTGCTGTCTTTGGGACCTTGT

Supplementary table 2: **list of oligonucleotides used, with source and sequence**

Antibody or other reagent	Source	Identifier	Working Dilution
CD34 PE	BD Biosciences	#581	1/100 (flow cytometry)
CD34 PE-Cy7	Biolegend	#581	1/100 (flow cytometry)
CD43 APC	BD Biosciences	#1G10	1/100 (flow cytometry)
CD11b - PE	BD Biosciences	#ICRF44	1/100 (flow cytometry)
CD14 APC-H7	BD Biosciences	#M5E2	1/100 (flow cytometry)
CD15 V450	BD Biosciences	#HI98	1/100 (flow cytometry)
CD45	BD Biosciences	#HI30	1/100 (flow cytometry)
Oct3/4 - PerCP-Cy™5.5	BD Biosciences	#560794 / 40/Oct-3	1/100 (flow cytometry)
Sox2 - Alexa Fluor® 647	BD Biosciences	#560302 / 245610	1/100 (flow cytometry)
Nanog - PE	BD Biosciences	#560483 / N31-355	1/100 (flow cytometry)
IgM - APC	BD Biosciences	#555585 / G155 - 228	1/100 (flow cytometry)
IgG1 - PE	BD Biosciences	#555749 / MOPC-21	1/100 (flow cytometry)
IgG1 - PerCP-Cy™5.5	BD Biosciences	#552834 / MOPC-21	1/100 (flow cytometry)
IgG2a - Alexa Fluor® 647	BD Biosciences	#565365 / MOPC-21	1/100 (flow cytometry)
STAT3	Cell Signaling Technology	#D1A5	1/1000 (WB)
STAT5	Cell Signaling Technology	#D3N2B	1/1000 (WB)
AKT (pan)	Cell Signaling Technology	#C67E7	1/1000 (WB)
p42/44 MAPK (Erk1/2)	Cell Signaling Technology	#9102	1/1000 (WB)
Phospho STAT3 (Y705)	Cell Signaling Technology	#D3A7	1/1000 (WB)
Phospho STAT5 (Y694)	Cell Signaling Technology	#D47E7	1/1000 (WB)
Phospho AKT (S473)	Cell Signaling Technology	#D9E	1/1000 (WB)
Phospho p42/44 MAPK (Erk1/2) (T202/Y204)	Cell Signaling Technology	#9101	1/1000 (WB)
Actin	Sigma	#A5441 / AC-15	1/1000 (WB)
IgG anti-rabbit H+L Alexa Fluor 633	Invitrogen/Thermo Fisher Scientific	#A-21071	1/400 (IF)
IgG anti-mouse H+L Alexa Fluor 546	Invitrogen/Thermo Fisher Scientific	#A-11003	1/400 (IF)
IgG anti-rabbit HRP-linked	Cell Signaling Technology	#7074	1/2000 (WB)
IgG anti-mouse HRP-linked	Cell Signaling Technology	#7076	1/2000 (WB)

Supplemental table 3: **list of antibodies, with respective source, clone identifier and working dilution**

DISCUSSION
AND
PERSPECTIVES

Disease modelling via induced pluripotent stem cells

The study of rare diseases, including congenital platelet disorders, has typically been hampered by the urgent needs of valid and efficient tools and models. The reduced availability of cellular models reproducing efficiently the pathological features observed in patients, combined with the limitations of animal models that are not fully recapitulating the disorder, are the biggest obstacles. The accessibility to patient samples, while relatively straightforward for the study of malignant haematopoiesis, it has been historically challenging, due to the small number of individuals correctly diagnosed as affected by IT. Even now, with a certain awareness of the existence of this pathological entity and the diagnostic efforts championed by the BRIDGE consortium in the UK, we are still unable to offer a diagnosis for almost half of the individuals that are not affected by ITP and display some degree of thrombocytopenia[144]. Nevertheless, the entire field is progressively increasing the number of genes involved in ITs, and we can estimate that the creation of other multi-center joint efforts would efficiently reduce the number of thrombocytopenic individual without a genetic diagnosis.

If the major efforts have been dedicated to the identification of the genetic causes of congenital platelet disorders, we just begun as a field to understand the molecular mechanisms responsibles of the defects observed in ITs patients. The implications of this type of work are clear, as it is reasonable to expect that ITs physiopathology describes crucial aspects of the MKpoiesis. This knowledge could be exploited for biotechnological and medical purposes, for example expanding this type of research to diseases with a somatic origin. Our work could be related to this ongoing effort, as we tried to unravel some of the mechanisms underlying the correct maturation of MKs, one of the most fascinating processes in haematology. We had chose the Filaminopathy A as a candidate, because of its unusual clinical manifestation and for the well-known role of FLNa into several cellular processes.

At the same time and surprisingly, ITs are not affecting exclusively the MK lineage. As presented in Chapter II, several forms of ITs modify the haematopoiesis in a more disrupting way, affecting the entire haematopoietic output like in CAMT, or a restricted number of lineages, like in the case of GATA1-related thrombocytopenia. In this category we can surely include the ITs predisposing to haematological malignancies. Our lab has a long-standing interest in the events that arise prior to a full-blown malignancy, and in this context the THC2 is a valuable example of an original predisposition mechanism, albeit not strictly related to the megakaryopoiesis. There is a growing interest in understanding the cellular events that leads to the development of leukemia, and our work is pointing out towards an original mechanism, that could be more potentially extended to other lineage than the granulocytic one.

The choice to generate iPSC clones from different IT patients is not a novel approach, but there is a growing interest into alternative experimental models that could yield novel and insightful results on human haematopoiesis[360]. Our lab has already experience in this type of experimental approach, both for congenital and acquired disorders[172,327,361]. Therefore, the technical expertise required for such challenging projects has been already available in the lab, and has been very helpful for implementing novel and improved techniques. This transition towards new methodological tools allowed us to change our experimental strategies, notably increasing the amount of cell lines in culture and differentiated at the same time. In regards to the haematological differentiation, the described protocols are improved versions of the one described by Chou and colleagues[359], modified only in the cytokine requirements for the haematopoietic specification and maturation (MKs versus granulocytes). The original version was embryoid bodies (EB)-based, and in our protocol we have been able to adapt the same culture conditions to a 2D system, yielding highly comparable results but removing the technical variables associate to the EBs, notably the increased cell death in the process of EB formation. The protocol is robust enough to allow parameters optimization and troubleshooting in a relatively quick fashion, as it is not limited by the constraints of a serum-full medium (notably the batch variability) or a co-culture approach on an instructive stroma. Nevertheless, the protocol is highly sensitive to many experimental parameters: confluency and pluripotency of the original PSCs culture plays the major role, as the starting cells dictate the final results. Another parameter is the cell density at the very beginning of the differentiation: this variable is highly dependent from the used cell line, so it takes a certain effort to calibrate efficiently experiments containing multiple lines. In the FLNa case, the task was easier, as isogenic cell lines do not differ sensibly in this regard. In the ANKRD26 project, to reduce the cell-line variability[324], we opted for the use of at least two clones for each patient, for a total of 6 cell lines actively differentiated. We compared those results to 3 cell lines from 3 healthy individuals. Accordingly to the published results, we support the idea that the protocol used is prevalently recapitulating the early primitive haematopoiesis or the transient-definitive wave. The main argument is based on the ability to differentiate haematopoietic cells that are morphologically, phenotypically and functionally acting like a granulocytic progenitor. This cell type has not been described in any study of the primitive haematopoiesis, and it is associated to a transient definitive haematopoiesis[27]. Nonetheless, we cannot discuss the cell origin of the other cell types, notably MKs, as the ontogenic differences between primitive- and transient definitive-derived MKs are unknown and probably very subtle[362]. A plausible possibility is that both types of haematopoietic differentiation are actually occurring, and it would be of great interest to manipulate the WNT/β-catenin pathway in the early days of the differentiation, in order to increase the output of one of the two ontogenic programs[299].

As already mentioned in the Chapter III, we are not able to differentiate HSCs starting from pluripotent stem cells. While this model limitation is not greatly affecting the results presented in this work, as our main focus has been the megakaryopoiesis and the granulopoiesis, it is nonetheless limiting the type of conclusions we can draw. It would be of great interest to extend the conclusion to an experimental setting closer to *in vivo*. A possible solution would be the use of xenograft murine models for short-term engraftment of MK precursors. While devoid of interest for the study of platelet production in the long term (only stem cells could colonize the bone marrow efficiently and in a stable fashion), in this way it would be possible to recapitulate some of the variables that affect the MK physiology *in vivo*, notably in the context of proplatelet formation and release into the circulation. This type of approach requires a certain degree of technical expertise, notably in the engraftment procedures as MKs could be ill equipped for a proper homing to the murine bone marrow niche. In this regard, Lefrançais and colleagues proposed recently that the lung could be an important site for platelet biosynthesis in mouse[363], therefore there is the risk that the engrafted human MKs can find themselves trapped in the recipients lungs. It would be suitable then to proceed with an injection directly in the femoral bones. In alternative, it could be interesting to combine our iPS-derived haematopoietic populations with the new humanized mouse model described by the Majeti's lab[332]: in this clever model, recipients are injected with human mesenchymal stem cells subcutaneously, in order to create a humanized, bone marrow-like, niche that could be successively injected with human haematopoietic cells. The major intest of this approach is the study of haematopoietic cells historically difficult to engraft, like MPN- or MDS-derived cells. In our case, it would be possible to create those artificial niches directly from the same source of the haematopoietic cells, as pluripotent stem cells could be successfully differentiated into mesenchymal cells[364]. Both approaches are, nonetheless, very challenging from the technical point of view; a more reductionist solution could be the use of better refined artificial niches, like the one described by Balduini and colleagues: a silk-based system mimicking the bone marrow microenvironment, as it can be easily integrated with other cell types (endothelial cells) and component of the extracellular matrix[365]. The combination of these advanced experimental systems with the versatility of the iPS model would greatly extend the scientific interest in our model.

FLNa impact on megakaryopoiesis

The case of Filaminopathy A is a good example of the advantages of an iPSC-based model: the main clinical manifestation of Filaminopathy A is the PNH, a condition associated to a defective migration of the neural progenitors during the embryo development. Individuals affected by PNH carry heterozygous mutations in *FLNA*. The genetic knock out of *FlnA* is not reproducing any defect in the nervous system development, suggesting a substantial difference between humans and mice and ruling out this model as a valid tool for the study of the PNH. Rats have been more successful, as they display a similar phenotype to human, albeit using an intensive and time-consuming approach[364]. In this regard, our iPS-based, isogenic model could be of interest for the study of the physiopathological mechanisms behind the multiple clinical manifestations of Filaminopathy A. Our work constitutes a demonstration of the flexibility of this type of models, as they can be modified in a relatively straightforward fashion via gene editing in a locus that is easily expressed (a "safe harbour"): introducing the *FLNA* cDNA mutants in a $FLNA^{mut}$ clone, so in absence of any residual copy of FLNa, avoids possible confounding effects in the phenotype interpretation. This experimental setup is more efficient than a simple transgenic over-expression, as it is finely tuned in presence of an appropriate promoter. Nevertheless it is not perfect, as the expression of the constructs does not depends on the many epigenetic constraints associated to the original *FLNA* locus. In this case, the most appropriate solution would be the engineering of the original locus via CRISPR/Cas9, an approach extremely laborious and time-consuming. Another evident advantage, already mentioned in the Chapter III, is that a pluripotency-based model is a powerful tool for the study of complex syndromes affecting multiple organs and cell types. In this regard, our model could be nicely combined with the new advanced 3D differentiation protocols, notably the generation of cerebral organoids, as a possible model for the defective neural migration observed in PNH patients[365]. Finally, the X-linked nature of the disease, combined with the reprogramming approach, offers an advantage that is technically very interesting: as the human iPS cell lines are clonal by definition and they retain one X chromosome inactivated, it is possible to derive isogenic cell lines from the same genetic background, as they will differ only in the expressed X chromosome. In this way, the confounding effects due to the genomic differences between individuals are removed, undoubtedly an advantage in practical terms, as less cell lines are required for each experiment.

In our work, we have been able to assess the damaging impact of two different intragenic deletions. The choice of these two mutations stems from the previous work from Berrou and colleagues[247]: in their work, they were not able to detect any truncated form of mutated FLNa in patient platelets, but they did not analyse the MKs. They propose that maybe the absence of truncated proteins in platelets is the consequence of the platelets half-life, as it accelerates the decay of the proteins

produced in MKs. In alternative, no truncated protein is produced whatsoever, because of the non-sense mediated decay of the *FLNA* mRNA. Our results support this latter hypothesis, as we were able to observe a reduced expression of the *FLNA* mRNA in *FLNA*^{mut} iPSCs, leading to the complete absence of the protein, both in PSCs and MKs. To be more rigorous, we used two antibodies recognizing two different epitopes at the N- and C-terminal of FLNa, and obtained similar results. No Xi erosion was observed for the FLNA locus after a high number of passages (more than 60), supporting the stability of our cell lines. In any case, it is good practice to never keep cells in culture for such a long period, in order to avoid the inevitable introduction of genomic mutations that could alter the stability of our cell lines.

The analysis of the MK lineage confirmed the early observation of Nurden and colleagues: MKs from FLNA^{mut} genomes are not altered in their early differentiation properties, as they express normal surface levels of the two major protein complexes CD41 and CD42. It is well known that primitive MKs displays very low degree of modal ploidy, even if they are correctly able to maturate and release proplatelet extensions[75]. This is a regretful limitation of the PSC model and other fetal-based systems, but luckily it is also the only one, as the rest of MK differentiation and maturation is not altered between adult and fetal/embryonic haematopoiesis. It would be interesting to understand more on this ontogenic difference, and how to manipulate the underlying mechanisms, in order to generate better disease models that could recapitulate also the endomitotic process[366]. The deep defect in PPT formation was a further validation of the value of our model for physiopathological studies. As for the ploidization, there are some ontogenic differences between the embryonic and the adult MKs, notably the proplatelets are smaller in size and less branched than their adult counterparts. This is the reflection of the smaller MK size, therefore a smaller DMS. The percentage of PPT-forming MKs is slightly reduced compared to the adult cells, but it is nonetheless robust enough for a rigoruous evaluation of the PPT formation potential *in vitro*.

The central role of FLNa in the cytoskeleton dynamics is well known and supported by a large body of work[241]. Essentially, FLNa and the other members of the Filamin family are responsible of the structural stability of the actin cytoskeleton. To accomplish this role, the FLNa forms a homodimer thanks to the N-terminal domain and organized in a V-shape. The protein central body contains a rod-like portion, accomodating up to 24 repeats of approximately 100 aminoacids each, each one folded as an immunoglobulin. For this reasons, those domains are called Ig-like domains. The big rod-like structure is actually separated in two smaller rods, by two hinge domains localised between the Ig-like domain 15 and 16 (hinge 1 or H1) and between Ig-like 23 and 24 (hinge 2 or H2). This domain organization allows a certain degree of flexibility in the FLNa protein conformation, at the point that the 3D interaction between actin and FLNa relies on two points of contact: the actin

binding domain at the C-terminal and part of the Rod 1, from the Ig-like 9 to Ig-like 15.

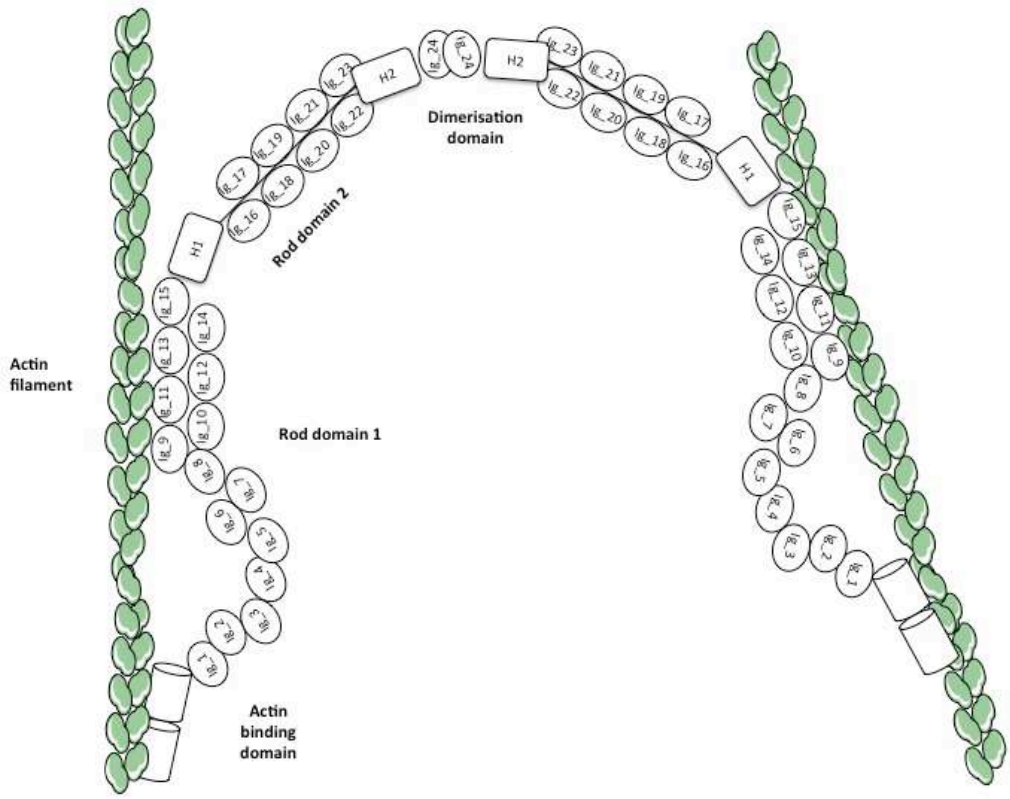


Figure 29: the actin-FLNa interaction, mediated by the ABD and part of the Rod 1.

The ABD sequence is relatively common between different actin binding proteins, therefore it is expected a certain degree of competitiveness between FLNa and other structural protein, like α-actinin, for actin binding. Nevertheless, the second interaction with the Rod 1 confers a higher affinity and allows the FLNa to be a structural support to any mechanical stress affecting the cytoskeleton[367]. The ABD binding is necessary to initiate the interaction and the rod 1 domain increases the binding affinity for the filament. In this way, each FLNa monomer is interacting very strongly with actin, but only one filament at the time; therefore, as FLNa is only active under the dimer form, the final result is that two different filaments are effectively connected. This is the reason why the FLNa dimer supplies a certain structural integrity to the entire cytoskeleton, acting as a prop and supporting the structure in presence of mechanical stress (Figure 29). Our collection of mutants could be a very helpful experimental tools for the fine understanding of the FLNa-actin dynamics: the mutant named del1, as lacking the ABD, is theoretically unable to bind the actin filaments, but still retains the ability to form a dimer; conversely, the mutant named del4 is lacking the dimerization domain but not the ABD. The comparison between these two mutants could shed light on the interaction dynamics between actin and FLNa.

Although our results were not particularly focused on these aspects of the FLNa activity, we have been able to show that the inability of FLNa to bind actin is affecting the late MKpoiesis: the cells are unable to form and extend proplatelets in the same fashion as the WT counterparts and the acto-myosin contractility is partially de-regulated, albeit not in a substantial way like in the complete absence of FLNa. It would be of great interest to observe more in details the ability of del1 mutant cells to react to mechanical stress, therefore evaluating the contribution of the FLNa protein to the MK cytoskeleton stability under stress.

Excluded the rod 1 and the ABD, the remaining parts of the protein are free to interact with all the others numerous partners, therefore complying with the protein role of molecular relay between the cell membrane and the cytoskeleton. This second role implies a close association to surface receptors, membrane integrins and other proteins involved in signalling pathways. In this regard, FLNa has several interactors that are crucial in the different steps of MKpoiesis. One of the first to be described as important for signalling is between FlnA and the tyrosine kinase Syk in mouse platelets, as the absence of the former induces an abnormal spatial distribution of the latter, therefore influencing the SYK docking to GPVI, the collagen receptor[243]. But the two FLNa most important interactors are the complex GPIb-GPIX-GPV and the $\alpha_{IIb}\beta_3$ integrin.
Most of the data discussing these interactions has been generated in murine platelets, and the generalization of those conclusions to MK could be misleading. A good example of this discrepancy between the two contexts is the results obtained by Jurak Begonja and colleagues in 2011: while platelets lacking FlnA were expressing reduced levels of all subunits of the CD42 receptor, the originating MKs were showing no statistically significant differences in the surface expression. The authors explain the discrepancy in terms of increased protein degradation mediated by the metalloproteases MMP9 and ADAM17, therefore exonerating the FlnA from a possible role in defining the correct localization of the GPIb-GPIX-GPV complex. Furthermore, no differences were observed in the expression of CD61 (β_3), either in platelets or MKs lacking FlnA[244]. In this regard, our results confirm this body of work, as we were not able to highlight any difference in the surface expression of CD41 and CD42, in absence of FLNa. As expected, also our mutants were not displaying any alteration, therefore we prove convincingly that the FLNa is not involved in the correct localization at the cell surface of both complexes. We can also argue that no evident compensation, mediated by *FLNB*, could explain the observed phenotypes, as this *FLNA* homolog is expressed at very low levels during the late stages of MKpoiesis. As our system is not suitable for platelet analysis, we do not have the data to fully compare our interpretation with the one from Jurak Begonja and colleagues[244]. Nevertheless, the platelets isolated from the blood of the Filaminopathy A patients that we reprogrammed are not displaying any decrease in CD42 surface

expression, supporting our hypothesis that in human cells, FLNa is not necessary for the correct localization and stability of both complexes at the membrane. Nevertheless, we support the interpretation referring to FLNa as a crucial linker of the CD42 receptor with the cytoskeleton, via GPIbα[368]. FLNa acts as a direct mediator of the CD42-dependent signalling, positively modulating the receptor function. The interaction FLNa-GPIbα is fundamental for the platelet adhesion to vWF[369] and has been implicated in the transendothelial migration of MKs for proplatelet extension in the sinusoid vessels, via Cdc42[81]. Our results prove that the lack of interaction between FLNa and GPIbα is not affecting the PPT formation, but we do not exclude other functional consequences of this disruption. Almost all mutations in GPIbα describe in BSS patients are affecting the extracellular portion of the protein[370]. The data supporting an impact of the interaction GPIbα-FLNa has been generated in cell line models over-expressing the two components[245] as transgenes, so they cannot be fully recapitulating the phenotype observed in human MKs.

Our observation of the abnormal spreading of $FLNA^{mut}$ MKs on fibrinogen, with the observed large stress fibers that were absent in the control cells, was extremely surprising. The robustness of this observation was confirmed in both patients, for several clones, therefore we excluded a single cell line-specific defect. Stress fibers formation is one of the hallmarks of the small GTPase RhoA activity[371] and underlies an actively contracting cytoskeleton. MKs displaying stress fibers are less prone to form proplatelets, as the cytoskeleton is not supporting the microtubule elongation necessary for proplatelet extension. This could be dependent from the MK localization in the bone marrow, as the different component of the extracellular matrix (ECM) impact directly the MKs activity and PPT formation[372,373]. The BM niche is virtually separated into two compartments, the osteoblastic niche found near the endosteum and the second compartment is the vascular niche near the sinusoids. These two niches consist of different cell types and different ECM components. In the osteoblastic niche, collagen I is the most abundant component, and binding of MKs to collagen I inhibits proplatelet formation. The platelet release is inhibited through adhesion collagen I via activation of the Rho/ROCK1/2 signalling cascade. On the other hand, in the vascular niche, MKs interact with collagen IV at the microenvironment which allow MKs to mature and form proplatelet.

In our dataset, there is a partial dissociation between stress fibers presence and PPT formation: the inhibition of ROCK1/2, a classic effector of RhoA and a well-known actor in the process of PPT formation, albeit able to increase the number of PPT-forming MKs in the $FLNA^{mut}$ cells, is not restoring the potential to the same levels of the $FLNA^{WT}$ MKs treated with the same inhibitor. At the same time, the ROCK1/2 inhibition was sufficient to abrogate completely the stress fibers formation, proving that stress fibers formation is not the only cellular mechanism having a negative impact on PPT formation. A similar observation was done before, as MKs adhering on collagen and

treated with a ROCK1/2 inhibitor were displaying no stress fibers, but they were still unable to give PPTs[372]. The authors suggested that collagen I was not activating only RhoA, but also other cytoskeleton and microtubules regulator that have a negative impact on PPT formation. Furthermore, ROCK1/2 is not the exclusive effector of RhoA, as DIAPH1 has been described to be dependent from RhoA and negatively affecting the PPT formation[235] (Figure 30). Therefore it would be very interesting to verify if the double inhibition of ROCK1/2 and DIAPH1 is sufficient to restore completely the phenotype observed in FLNA^{WT} MKs.

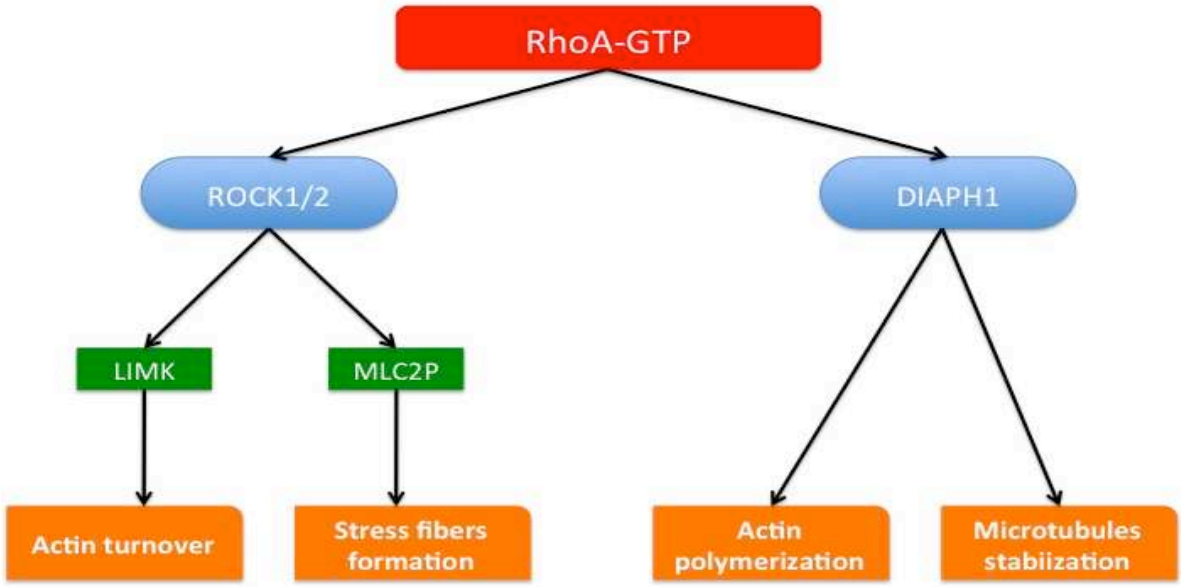


Figure 30: the different effector of the active RhoA small GTPase.

Nevertheless, we do not exclude that the lack of FLNa is potentially affecting the proplatelet formation via other mechanisms that are not RhoA-dependent. As an example, the del1 mutant displays an intermediary phenotype when analysed for the presence of stress fibers and PPT formation, therefore it could hide an unknown mechanism related to the actin cytoskeleton dynamics.

The claim that the lack of FLNa is inducing a defective cytoskeleton contractility because of an over-active RhoA prompted us to verify the specificity of this model: as already mentioned, FLNa is a molecular hub that links cell surface receptors and integrins with the cytoskeleton, and two notable FLNa interactors are CD42 (via GPIbα) and CD41 (via β_3). If GPIbα has been proven to not be involved in the PPT formation, this is not the case for the interaction with β_3. The FLNa mutant lacking the domain interacting with β_3 (del3 mutant) displays the same phenotype as the FLNA^{mut} for both stress fibers and PPT formation, underlying a crucial role for this interaction in the normal MK maturation. This mutant is lacking the Ig21 domain that interacts also with other β_3-

containing integrins, supporting the notion that Ig21 is a relatively promiscuous binding site for β integrins. This lower degree of specificity could be a specific feature of the interaction, as it has been proved that there is a competition between FLNa and other proteins like talin and kindlin, modulators of integrin activity[374]. The proposed mechanism, in resting platelets, is that FLNa is interacting with a bent $\alpha_{IIb}\beta_3$ complex, and the displacement via talin binding is deemed necessary for the activation of the inside out signalling of the integrin complex[242]. It is therefore plausible to imagine how the lack of FLNa is not retaining anymore the integrin complex in an inactive state, so the $\alpha_{IIb}\beta_3$ could be potentially more prone to be active in MKs and platelets. It would be interesting to evaluate more in details the activation dynamics of $\alpha_{IIb}\beta_3$ in absence of FLNa. Moreover, the del4, like the del3, is showing a phenotype very similar to the $FLNA^{mut}$, corroborating our model of a predominant role of the RhoA increased activity in the physiopathology of Filaminopathy A-linked macrothrombocytopenia. FLNa is able to interact directly with several small GTPases, but also with some of their effectors; we can suppose that FLNa probably acts as a signalling scaffold that is bringing together the main components of the different pathways[375,376]. FLNa could bind ROCK, a downstrem effector, but also GEFs and GTPase-activating proteins (GAPs), therefore upstream effectors. Individuating those crucial activators of RhoA would be a massive task, as there is little information about the expression of GAPs/GEFs in MKpoiesis, but nevertheless it would shed light on the regulation of RhoA activation. The portion of FLNa that interacts with small GTPases and their effectors is also very close to the domain necessary for dimerisation, the Ig24, therefore it could be possible that the dimerisation process is improving the recruitment of the GTPases. The del4 is lacking a large protein portion that encompasses the Ig24 and the small GTPases interaction domain, so we cannot exclude one effect from the other. It is also possible that the absence of FLNa dimerisation is completely abrogating the functional activity of the protein, and not just its simple structural role in actin cytoskeleton.

While the RhoA activity has been the major focus of our efforts, we cannot exclude that the absence of FLNa can have an impact on the signalling mediated by CDC42. The role of this small GTPase has been recently linked to the transendothelial elongation of proplatelets fragments and associated to the GPIbα[81]. Moreover, CDC42 and one of its GEF is directly interacting with FLNa, suggesting the presence of a circuit involving the CD42 receptor, FLNa and CDC42. It would be extremely interesting to verify the existence of such functional pathway in the context of Filaminopathy A, as it could add a further layer on the understanding of the disease physiopathology. In our lab we have at our disposal a biosensor for the measurement of the activity of CDC42, based on the same principle of the one used for the measurement of the RhoA activity, thus it would be feasible to assess the activation state of this crucial small GTPase in absence of FLNa.

We support the specificity of the axis $\alpha_{IIb}\beta_3$–FLNa-RhoA thanks to our experiments on the RhoA activity in presence of different extracellular matrices: only the fibrinogen was able to induce an overactive RhoA in FLNA^{mut} MKs, a phenotype that depends, in a dose-dependent fashion, from the FLNa expression levels. This final observation allows us to draw parallels with other macrothrombocytopenias and could define a new physiopathological mechanism for other ITs. The first example is the PRKACG-mutated thrombocytopenia: the biallelic mutations are impacting a key component of the PKA, and the PKA is responsible of the protein stability of FLNa (via phosphorylation of the Ser^{2152}). Consequently, patient MKs display greatly reduced levels of FLNa and the defect in PPT formation, characteristics that are very similar to the observed one in FLNA^{mut} cells. Evaluating the RhoA activity of these cells could be a simple way to confirm the similar cellular mechanism. Another very reasonable candidate that could share some aspects of our proposed mechanism is GT: mutations in the $\alpha_{IIb}\beta_3$ complex are responsible for this condition, so it would be of great interest to test our proposed model in this context. Indeed several mutations affecting the intracellular tail of the β_3 subunit are causing GT variants with macrothrombocytopenia and reduced PPT formation. Mechanistically, some mutations have been associated to a defective inside out signalling, due to a disrupted interaction talin/α-kindlin with the β_3 subunit tail[377]. This type of mutations could be defined as loss-of-function, as the MK loses the ability to activate correctly the $\alpha_{IIb}\beta_3$-dependent signalling. But there are also several GT mutations that induce an increased activation of this integrin-dependent signalling, via mutations that are affecting the cytoplasmic tail of β_3. The first report of such type of mutations date to 2008 and describe a pedigree characterized by autosomal dominant transmission of a GT with macrothrombocytopenia. A single base substitution (2245G>C) was described as responsible of the phenotype and leads to a non-synonymous substitution of a highly conserved residue (D723H), involved in the establishment of a salt-bridge between the two integrin subunits. The functional consequence is a constitutive activation of the integrin and an abnormal production of PPT, with macrothrombocytopenia[210]. It could be reasonable to expect that the disrupted salt-bridge in the cytoplasm is affecting the ability of the integrin to interact with the other partners, therefore changing the equilibrium between FLNa-bound and talin/α-kindlin-bound integrin, therefore leading to a modified signalling dynamic. Moreover, other families with similar mutations have been described recently, underlying a more common mechanism than expected for such GT subtypes to other mutations affecting the interaction between the integrin subunit β_3 and its signalling regulators[378]. A third example of macrothrombocytopenia that could share the physiopathology with Filaminopathy A is the α-actinin (ACTN1)-related IT. This protein has been described as able to bind β_3 before the triggering of the inside-out signalling cascade, and this

complex dissociates once the integrin is activated. Moreover, down-regulation of α-actinin introduce a higher degree of $\alpha_{IIb}\beta_3$ activity in platelets, supporting a model where the binding of α-actinin is keeping the integrin inactive, and its displacement is necessary for the pathway activation[236]. It would be therefore interesting to observe if this activation also involves the RhoA pathway, or is mediated by alternative signalling routes.

Most of our attention has been dedicated to the final stages of MKpoiesis, but we cannot exclude that the FLNA dependent over-activation of RhoA is not playing a role in the polyploidization. The already published works were not focused on this aspect of MKpoiesis, but Falet and colleagues reported a marked decrease in the higher ploidy class (N>16) and, conversely, an increase in the total number of MKs, possibly suggesting a defective endomitosis[244]. Unfortunately, our experimental system is not particularly suitable to study this part of MKpoiesis, but it would be valuable to verify if the $\alpha_{IIb}\beta_3$–FLNa-RhoA axis is also important for the process of endomitosis.

Finally, the FLNa protein has been recently described as a factor able to promote the transcriptional activity of the TFs SRF and MKL1, via a mechanism involving the actin[379]. The authors describe a transcriptional decrease of gene targets of MKL1 like *MYH9*, when FLNa is absent, and link this difference to two mechanisms: FLNa interferes with MKL1 phosphorylation and avoid the repressive effect of the actin monomer, as the binding of MKL1 with globular-actin impede the nuclear translocation of the TF. It is therefore possible that the absence of FLNa in Filaminopathy A patient underscores another pathological mechanism that it is unrelated to the most canonical roles of the protein. It would be interesting to verify the transcriptomic profiles of our iPSC-derived MKs, and using our mutants to confirm the specificity of the eventual differences, as the del3 mutant lacks the domain of interaction with MKL1.

In conclusion, we propose an original model for the explanation of the observed macrothrombocytopenia in Filaminopathy A patients, that links a major membrane integrin ($\alpha_{IIb}\beta_3$) with a classic signalling pathway in MKpoiesis (RhoA). In this context, the FLNa acts as a negative regulator of the integrin activity, and the absence of FLNa is therefore triggering a cascade that is disrupting the process of PPT formation. In this context, acting on the RhoA downstream effectors via a pharmacological strategy could restore an acceptable platelet level and be an efficiency therapy for the most severe cases of macrothrombocytopenia.

ANKRD26-RT, a novel example of leukemic predisposition

THC2 is one of the most diffused types of ITs, with hundreds of pedigrees described worldwide. This number is progressively increasing, because of the increased awareness towards conditions like inherited platelet disorders. Due to the clinical manifestation that underlies no evident defects in platelet function and no alterations in the platelet size, THC2 is frequently misdiagnosed as immune thrombocytopenia or even MDS, with potentially disastrous consequences in the patient care[380]. Frequently, ITs predisposing to leukemia are associated to other form of hereditary myeloid malignancy syndromes (HMMS), and the clinical practice suggests to consider also the possibility of a congenital origin for the diagnose of myeloid malignancies[164]. Therefore it would be extremely helpful for the future to 1) develop rapid screening methods for the identification of this type of conditions, and 2) elucidate the mechanism behind the predisposition, in order to define a possible therapeutic approach and organize an efficient clinical surveillance and management for these patients[381]. For the first point, our lab developed a rapid diagnostic test to diagnose ITs, using MYH10 expression in platelets as a biomarker[382]. This type of test could be also very helpful to rapidly distinguish THC2 from the other two predisposing ITs, directing the diagnostic investigation towards the presence of the causing mutations in the 5' UTR. We delved into the second point after the first study on the physiopathology mechanism behind the thrombocytopenia[127]. Many questions were raised by that study, mostly related to the functional link between the increased levels of ANKRD26 found in patients MKs and the deregulated signalling output in response to TPO. The few studies previously performed were describing a membrane associate gene that is negatively regulating the IGF-dependent signalling. Consequently, mouse knock out model for *Ankrd26* display obesity and diabetes, before of the enhanced adipogenesis[383,384]. Few other studies have conclusively investigated the ANKRD26 function, and they were mostly focused on the adipogenesis.

We started to investigate the role of ANKRD26 outside of MKpoiesis, in an attempt to highlight any anomalies affecting other haematopoietic lineages. The first large clinical study described a statistically significant leukocytosis and increased haemoglobin levels for several THC2 patients, so we speculated that other myeloid lineages were affected by the ANKRD26 mutations. The semisolid culture assays we performed were confirming this suspect, as we were able to detect an increase in the number of CFU-GM and erythroid colonies. This observation, summed up with the increase in the ANKRD26 levels in CD15+ cells, prompted us to focus our attention on the granulocytic compartment, as a possible candidate for the preluekemic predisposition. Nevertheless, we plan to perform a thorough analysis of the other major myeloid lineages, notably the erythroid lineage. ANKRD26 appear to be expressed in highly immature progenitors and HSCs, therefore it is reasonable to imagine that the gene is not strictly specific for MKs, but could be more broadly

expressed. Consequently, it is of great interest to understand the impact of the 5' UTR alterations observed in THC2 patients for other haematopoietic lineages. In MKs, the mutations disrupt the binding of the repressor complex RUNX1/FLI1, a central component of the transcriptional network of those cells. While RUNX1 is expressed and active in granulocytes and monocytes, and directly impact their physiology, FLI1 is much less implicated. Nevertheless, other TFs of the ETS family could possibly be involved, and play an active role in the regulation of *ANKRD26* expression. One hypothetical candidate could be PU.1, as it is part of the same family of FLI1, is a master regulator of the myeloid differentiation[385] and is able to interact with RUNX1[386].

The analysis of the *in vitro* granulopoiesis was revealing an increased proliferation of the granulocytic lineage, combined with a delayed maturation. This data is highly valuable, as it has been generated from patient samples, and indeed proved that there is a certain degree of anomalies in this cell lineage, implicitly broadening the impact of THC2 anomalies to other haematopoietic cells. Nevertheless, we needed to establish a higher degree of confidence in these preliminary observations, as very few patients have been tested. Moreover, a more robust characterization of the observed phenotype was required for excluding the impact of unpredictable variables, like the presence of undetected genetic lesions in the analysed samples. In regard to these limitations, the development of several iPS cell lines was an attempt to overcome the limits associated to the reduced number of patient samples available. The establishment of the cell lines and their characterization did not reveal any major anomalies, so the routine maintenance of these cell lines in culture. Compared to the Filaminopathy A project, we needed some control lines from healthy individuals: in a more appropriate way, we used iPS cell lines derived from individuals not affected by any haematological disorder. Due to the experimental design, including at least six different genomic backgrounds, it has been considerably more difficult to perform the differentiation experiments: in particular, the optimization of the experimental parameters necessary for a good quality differentiation experiment has been particularly challenging.

Nevertheless, we have been able to generate haematopoietic progenitors and granulocytes from every single patient cell line we generated, in a demonstration of the robustness of our differentiation protocol. With our iPS cell lines we confirmed the preliminary observations from patient samples, effectively validating our hypothesis of a defective granulopoiesis for THC2 patients. We plan to rigorously validate this claim with additional experiments, concerning both the proliferation and the maturation aspects highlighted in patient cells. This point could be particularly crucial, as we cannot exclude some differences between patients, due to the nature of the mutation: in the previous work, all patients have been pooled together, as they all show a clinical thrombocytopenia. On the contrary, the predisposition penetrance is not complete, therefore it could be possible that the mutation is directly affecting the pre-leukemic phenotype, in a similar way to

what is observed in FPD/AML, where mutations dominant negative-like are proner to progress to leukemia, compared to the haploinsufficiency-like[163].

The anomalies observed in the granuocytic lineage are reminiscent of different conditions that are clearly leukemic or predisposing to leukemia. The increased proliferation of the granulocytic lineage is the main feature of a rare myeloproliferative neoplasm called chronic neutrophilic leukemia (CNL). This disease is mainly caused by alterations of the signalling dependent from the G-CSF receptor, and the most common (almost 80% of lesions) is the T618I substitution, which induce a constitutive activation of the receptor[387]. Albeit not classified as a MPN, Plo and colleagues described a family affected by hereditary chronic neutrophilia with high penetrance (75% of the family members). These individuals displayed, among the many symptoms, a moderate block in the granulocytic differentiation, with 20% immature granulocytes in the peripheral blood and no increased in the number of blasts in the bone marrow. Genetic analysis revealed the presence of a mutation affecting the transmembrane domain of the G-CSF receptor, namely T617N (also called T640N), and this anomaly induces a constitutive activity of the receptor. In vitro analysis of the cell proliferation did not reveal a major increased in the proliferation, but an increased sensitivity to low doses of G-CSF. It is worth noticing that at least one member of the family developed MDS[388]. Another disorder affecting the granulocytic lineage and intertwined with the progression to AML is the severe congenital neutropenia (SCN). In this hereditary condition, patients display a life-threating neutropenia and are susceptible of developing AML. Several genes have been described as mutated in such inherited disorders, but most of them have an impact on the G-CSF signalling. This cytokine has a role that is highly comparable with the TPO role for megakaryopoiesis, therefore is not surprising that anomalies that are abrogating or reducing the signalling output are severely affecting granulopoiesis. The classic leukemic progression, for SCN patients, includes the acquisition of an oncogenic hit on the G-CSF receptor. Typically the mutated receptor lost parts of its cytoplasmic domain, altering the delicate balance between proliferation and differentiation and promoting the clonal expansion of the mutated cells. The acquisition of a secondary hit, notably in RUNX1, is the other event that is sufficient to induce an overt MDS or AML[389].

All these observations, combined with the well-known impact of ANKRD26 on TPO/MPL signalling, prompted us to investigate the dynamics of the G-CSF-dependent signalling in presence of different levels of ANKRD26. We decided to turn, for the mechanistic details, to a robust model like the cytokine-dependent cell line UT-7[390]: this cancer cell line could be engineered to be strictly dependent from a specific cytokine for its proliferation. We have been able to obtain a UT-7 cell line that was overexpressing a tagged version of the G-CSFR (a kind gift from Dr. Nathalie Droin). Moreover, we already knew that this cell line was expressing *ANKRD26*, so we designed an

experimental setup where we reduced the gene expression levels via shRNA. At the same time, we planned to increase further the ANKRD26 levels, this time via transgenic overexpression of the full-length cDNA. Those two experimental systems proved that the receptor expression at the cell surface correlate directly with the ANKRD26 expression levels: more ANKRD26 leads to more receptor at the surface. This observation must be validated with an in-depth characterization of the underlying mechanisms. An important body of work describes the dynamics of the receptor and its regulation, and many possible mechanisms could explain the increased surface expression, ranging from a defective internalization, an increased re-routing and a defective degradation and/or an increased stability of the receptor[391]. This observation naturally led to the analysis of the activation kinetics of the signalling output of the G-CSF receptor: we analysed the classic downstream effectors, and we highlighted a stronger activation of the STAT3 and MAPK pathways. More in details, those two pathways appear to be activated more efficiently at lower doses of G-CSF and more intensively, in presence of higher levels of ANKRD26. Nevertheless, we did not observe any phosphorylation in absence of G-CSF, excluding a constitutive activation of the receptor; we also did not observe a prolonged signal, albeit we did not perform any wash-up of the cytokine[392]. To confirm the impact of this deregulated signalling, we used the proliferation of UT-7 cells at different concentration of G-CSF as an output. Indeed, at lower concentrations of G-CSF, cells expressing more ANKRD26 proliferate more, and this difference is progressively less evident at higher concentrations. In conclusion, all these observations point towards a stronger activation of STAT3 and MAPK, as a probable consequence of the increased surface levels of the G-CSF receptor. This phenotype is highly reminiscent of the observations obtained for MPL: unpublished data obtained with the same type of model suggest a widespread mechanism, valid also for the TPO and EPO receptors. We plan to investigate the dynamics of these receptors, in order to highlight any common features between them and in relationship with ANKRD26. We did not observe any alteration for the activation of STAT5, but only for STAT3; we suspect that the alteration of the balance between these two effectors of the G-CSF is at the core of the defective granulopoiesis, and it will surely be investigated in the future, in particular using specific inhibition strategies for the two different TFs.

Albeit very efficient for this sort of studies, the UT-7 model lacks the physiological resemblance with the patient context: the ANKRD26 expression level are not necessarily comparable, so the levels of receptor at the cell surface. To corroborate our observation we checked the G-CSF signalling in iPS-derived and patient-derived granulocytes, and we were able to observe similar deregulation for the two downstream pathways. It is paramount to confirm those preliminary observations for the other mutations. We finally started to look in more details into the observed increase in cell proliferation. Via transcriptomic analysis of granulocytic progenitors coming for

patient iPSCs and control iPSCs, we discovered an upregulation of a relatively unknown cyclin, *CCNI2*. This gene, not expressed in normal cells, is one of the activating cyclin for CDK5, a cyclin dependent kinase that has been associated to leukemic progression and glucose metabolism[393]. We are also planning to perform more thorough analysis on the data generated via RNA sequencing, in particularly searching for any enrichment in gene ontology pathways that could describe in better details the anomalies between normal and patient granulocytes. To improve further the experimental setup for this type of analys, it would be particularly interesting to introduce or correct the mutation in a wild-type or mutated cell line, as an isogenic comparison would be extremely beneficial: as the signalling defect is milder, compared to a direct anomaly of the cytokine receptor, the genomic variability associated to the background could dilute the impact of the mutation.

In conclusion, we propose a novel and original mechanism for the predisposition to leukemia in THC2 patients (Figure 31). The pathogenic mutations affects the regulation of ANKRD26, prompting an abnormal increase in the gene expression along granulopoiesis, a defect that mimic the phenotype observed for the megakaryocytic lineage. The granulopoiesis for cells derived from a THC2 genomic background is abnormal, with an increased proliferation and a delayed maturation. Those defects are probably related to the abnormal activation of STAT3 and MAPK in response to G-CSF. This response is due to an increased receptor expression at the cell surface, a mechanism that should underlie a defect in the receptor trafficking. Consequently, the granulocytic compartment could be proner to the acquisition of genomic lesions that could lead to the development of an over myeloid leukemia.

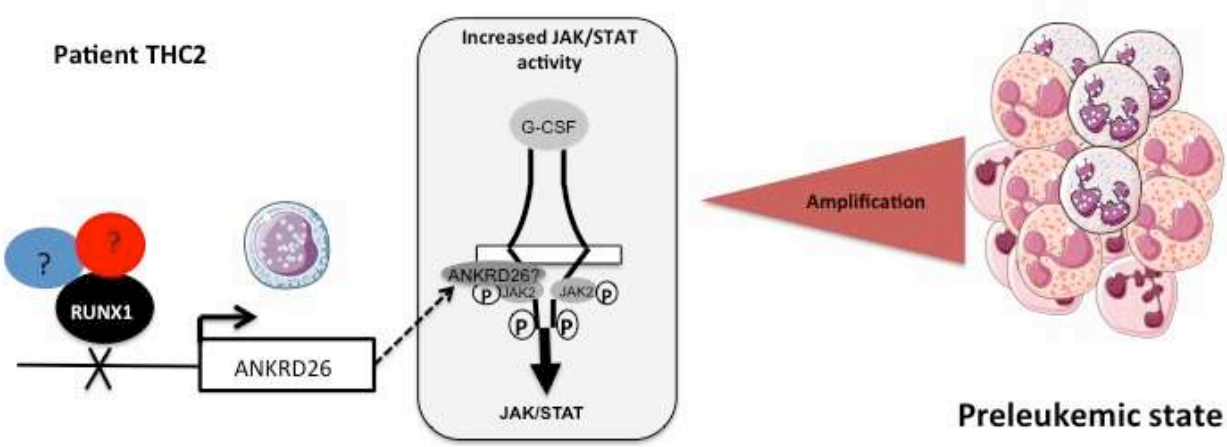


Figure 31: the proposed model for the leukemic predisposition in ANKRD26-RT patients.

REFERENCES

1. Sabin F "Preliminary note on the differentiation of angioblasts and the method by which they produce blood-vessels, blood-plasma and red blood-cells as seen in the living chick (1917)." *J Hematother Stem Cell Res*. **11,** 5–7 (2002); DOI: 10.1089/152581602753448496
2. Murray P. D. F. "The development in vitro of the blood of the early chick embryo." *Proc. R. Soc. Lond. B* **111,** 497-521 (1932); DOI: 10.1098/rspb.1932.0070
3. Huber, T. L. *et al.* "Haemangioblast commitment is initiated in the primitive streak of the mouse embryo." *Nature* **432,** 625–30 (2004); DOI: 10.1038/nature03122
4. Choi, K. *et al.* "A common precursor for hematopoietic and endothelial cells." *Development* **125,** 725–32 (1998); DOI:
5. Padrón-Barthe, L. *et al.* "Clonal analysis identifies hemogenic endothelium as the source of the blood-endothelial common lineage in the mouse embryo." *Blood* **124,** 2523–32 (2014); DOI: 10.1182/blood-2013-12-545939
6. De Bruijn, M. "The hemangioblast revisited." *Blood* **124,** 2472–2473 (2014); DOI: 10.1182/blood-2014-09-597674
7. Dieterlen-Lievre, F. "On the origin of haemopoietic stem cells in the avian embryo: an experimental approach." *J Embryol Exp Morphol* **33,** 607–619 (1975); DOI:
8. Godin, I. E. *et al.* "Para-aortic splanchnopleura from early mouse embryos contains B1a cell progenitors." *Nature* **364,** 67–70 (1993); DOI: 10.1038/364067a0
9. Medvinsky, A. L. *et al.* "An early pre-liver intraembryonic source of CFU-S in the developing mouse." *Nature* **364,** 64–67 (1993); DOI: 10.1038/364064a0
10. Tavian, M. *et al.* "Aorta-associated CD34+ hematopoietic cells in the early human embryo." *Blood* **87,** 67–72 (1996); DOI:
11. Palis, J. "Primitive and definitive erythropoiesis in mammals." *Front. Physiol.* **5,** 1–9 (2014); DOI: 10.3389/fphys.2014.00003
12. Hegyi, E. *et al.* "Developmental changes in human megakaryocyte ploidy." *Exp. Hematol.* **19**, 87-94 (1991); DOI:
13. Naito, M. *et al* "Development, differentiation, and maturation of fetal mouse yolk sac macrophages in cultures." *J Leukoc Biol.* 46, 1-10 (1989); DOI:
14. Silver, L. & Palis, J. "Initiation of murine embryonic erythropoiesis: a spatial analysis." *Blood* **89,** 1154–64 (1997); DOI:
15. Tavian, M., Hallais, M. F. & Péault, B. "Emergence of intraembryonic hematopoietic precursors in the pre-liver human embryo." *Development* **126,** 793–803 (1999); DOI:

16. Medvinsky, A. & Dzierzak, E. "Definitive hematopoiesis is autonomously initiated by the AGM region." *Cell* **86,** 897–906 (1996); DOI: 10.1016/S0092-8674(00)80165-8

17. De Bruijn, M. *et al* "Definitive hematopoietic stem cells first develop within the major arterial regions of the mouse embryo." *EMBO J.* **19,** 2465–2474 (2000); DOI: 10.1093/emboj/19.11.2465

18. Jaffredo, T., *et al* "Tracing the progeny of the aortic hemangioblast in the avian embryo." *Dev. Biol.* **224,** 204–214 (2000); DOI: 10.1006/dbio.2000.9799

19. Ivanovs, A. *et al.* "Highly potent human hematopoietic stem cells first emerge in the intraembryonic aorta-gonad-mesonephros region." *J. Exp. Med.* **208,** 2417–2427 (2011); DOI: 10.1084/jem.20111688

20. Gekas, C., *et al* "The placenta is a niche for hematopoietic stem cells." *Dev. Cell* **8,** 365–375 (2005); DOI: 10.1016/j.devcel.2004.12.016

21. Li, Z. *et al.* "Mouse embryonic head as a site for hematopoietic stem cell development." *Cell Stem Cell* **11,** 663–675 (2012); DOI: 10.1016/j.stem.2012.07.004

22. Yokomizo, T. & Dzierzak, E. "Three-dimensional cartography of hematopoietic clusters in the vasculature of whole mouse embryos." *Development* **137,** 3651–3661 (2010); DOI: 10.1242/dev.051094

23. Li, W. *et al.* "Endothelial Cells in the Early Murine Yolk Sac Give Rise to CD41-expressing Hematopoietic Cells." *Stem Cells Dev.* **14,** 44–54 (2005); DOI: 10.1089/scd.2005.14.44

24. Lancrin, C. *et al.* "The haemangioblast generates haematopoietic cells through a haemogenic endothelium stage." *Nature* **457,** 892–895 (2009); DOI: 10.1038/nature07679

25. Eilken, H. M., Nishikawa, S. I. & Schroeder, T. "Continuous single-cell imaging of blood generation from haemogenic endothelium." *Nature* **457,** 896–900 (2009); DOI: 10.1038/nature07760

26. Migliaccio, G. *et al.* "Human embryonic hemopoiesis. Kinetics of progenitors and precursors underlying the yolk sac-liver transition." *J. Clin. Invest.* **78,** 51–60 (1986); DOI: 10.1172/JCI112572

27. Palis, J. *et al.* "Development of erythroid and myeloid progenitors in the yolk sac and embryo proper of the mouse." *Development* **126,** 5073–84 (1999); DOI:

28. Böiers, C. *et al.* "Lymphomyeloid Contribution of an Immune-Restricted Progenitor Emerging Prior to Definitive Hematopoietic Stem Cells." *Cell Stem Cell* **13,** 535–548 (2013); DOI: 10.1016/j.stem.2013.08.012

29. Sturgeon, C. M. et al. “Defining the path to hematopoietic stem cells.” *Nat. Biotechnol.* **31,** 416–418 (2013); DOI: 10.1038/nbt.2571

30. Squarzoni, P. *et al.* "Microglia Modulate Wiring of the Embryonic Forebrain." *Cell Rep.* **8,** 1271–1279 (2014); DOI: 10.1016/j.celrep.2014.07.042

31. Nielsen, M. M., Witherden, D. A. & Havran, W. L. "γδ T cells in homeostasis and host defence of epithelial barrier tissues." *Nat. Rev. Immunol.* **17,** 733–745 (2017); DOI: 10.1038/nri.2017.101

32. Montecino-Rodriguez, E. & Dorshkind, K. "B-1 B Cell Development in the Fetus and Adult." *Immunity* **36,** 13–23 (2012); DOI: 10.1016/j.immuni.2011.11.017

33. Yoder, M. C. *et al.* "Characterization of definitive lymphohematopoietic stem cells in the day 9 murine yolk sac." *Immunity* **7,** 335–344 (1997); DOI: 10.1016/S1074-7613(00)80355-6

34. Taoudi, S. *et al.* "Extensive hematopoietic stem cell generation in the AGM region via maturation of VE-Cadherin+CD45+pre-definitive HSCs." *Cell Stem Cell* **3,** 99–108 (2008); DOI: 10.1016/j.stem.2008.06.004

35. Rybtsov, S. *et al.* "Hierarchical organization and early hematopoietic specification of the developing HSC lineage in the AGM region." *J. Exp. Med.* **208,** 1305–1315 (2011); DOI: 10.1084/jem.20102419

36. Rybtsov, S. *et al.* "Tracing the origin of the HSC hierarchy reveals an SCF-dependent, IL-3 independent CD43- embryonic precursor." *Stem Cell Reports* **3,** 489–501 (2014); DOI: 10.1016/j.stemcr.2014.07.009

37. Hadland, B. K. *et al.* "Endothelium and NOTCH specify and amplify aorta- gonad-mesonephros–derived hematopoietic stem cells." *J. Clin. Invest.* **125,** 1–14 (2015); DOI: 10.1172/JCI80137DS1

38. Tober, J. *et al.* Maturation of hematopoietic stem cells from pre-hematopoietic stem cells is accompanied by up-regulation of PD-L1. *J. Exp. Med.* **215,** 645–659 (2017); DOI: 10.1084/jem.20161594

39. Pereira, C. F. *et al.* "Hematopoietic Reprogramming In Vitro Informs In Vivo Identification of Hemogenic Precursors to Definitive Hematopoietic Stem Cells." *Dev. Cell* **36,** 525–539 (2016); DOI: 10.1016/j.devcel.2016.02.011

40. Ditadi, A., Sturgeon, C. M. & Keller, G. "A view of human haematopoietic development from the Petri dish." *Nat. Rev. Mol. Cell Biol.* **18,** 56–67 (2016); DOI: 10.1038/nrm.2016.127

41. Ema, H. & Nakauchi, H. "Expansion of hematopoietic stem cells in the developing liver of a mouse embryo." *Blood* **95,** 2284–8 (2000); DOI:

42. Traver, D. *et al.* "Fetal liver myelopoiesis occurs through distinct, prospectively isolatable progenitor subsets." *Blood* **98,** 627–635 (2001); DOI: 10.1182/blood.V98.3.627

43. Draper, J.E. *et al.* "A novel prospective isolation of murine fetal liver progenitors to study in utero hematopoietic defects." *PLoS Genet.* 14, e1007127 (2018); DOI: 10.1371/journal.pgen.1007127

44. Ivanovs, A. *et al.* "Human haematopoietic stem cell development: from the embryo to the dish." *Development* **144,** 2323–2337 (2017); DOI: 10.1242/dev.134866

45. Laurenti, E. & Göttgens, B. "From haematopoietic stem cells to complex differentiation landscapes." *Nature* **553,** 418–426 (2018); DOI: 10.1038/nature25022

46. Becker, A. J., McCulloch, E. A. & Till, J. E. "Cytological demonstration of the clonal nature of spleen colonies derived from transplanted mouse marrow cells." *Nature* **197,** 452–4 (1963); DOI: 10.1088/1751-8113/44/8/085201

47. Spangrude, G., Heimfeld, S. & Weissman, I. "Purification and characterization of mouse hematopoietic stem cells." *Science* **241,** 58–62 (1988); DOI: 10.1126/science.2898810

48. Akashi, K. *et al.* "A clonogenic common myeloid progenitor that gives rise to all myeloid lineages." *Nature* **404,** 193–197 (2000); DOI: 10.1038/35004599

49. Kiel, M. J. *et al.* "SLAM Family Receptors Distinguish Hematopoietic Stem and Progenitor Cells and Reveal Endothelial Niches for Stem Cells." *Cell* **121,** 1109–1121 (2005); DOI: 10.1016/j.cell.2005.05.026

50. Wilson, A. *et al.* "Hematopoietic Stem Cells Reversibly Switch from Dormancy to Self-Renewal during Homeostasis and Repair." *Cell* **135,** 1118–1129 (2008); DOI: 10.1016/j.cell.2008.10.048

51. Bernitz, J. M. *et al.* "Hematopoietic Stem Cells Count and Remember Self-Renewal Divisions." *Cell* **167,** 1296–1309.e10 (2016); DOI: 10.1016/j.cell.2016.10.022

52. Muller-Sieburg, C. E. *et al.* "Myeloid-biased hematopoietic stem cells have extensive self-renewal capacity but generate diminished lymphoid progeny with impaired IL-7 responsiveness." *Blood* **103,** 4111–8 (2004); DOI: 10.1182/blood-2003-10-3448

53. Dykstra, B. *et al.* "Long-Term Propagation of Distinct Hematopoietic Differentiation Programs In Vivo." *Cell Stem Cell* **1,** 218–229 (2007); DOI: 10.1016/j.stem.2007.05.015

54. Adolfsson, J. *et al.* "Identification of Flt3+ lympho-myeloid stem cells lacking erythro-megakaryocytic potential: A revised road map for adult blood lineage commitment." *Cell* **121,** 295–306 (2005); DOI: 10.1016/j.cell.2005.02.013

55. Doulatov, S. *et al.* "Revised map of the human progenitor hierarchy shows the origin of macrophages and dendritic cells in early lymphoid development." *Nat. Immunol.* **11,** 585–593 (2010); DOI: 10.1038/ni.1889

56. Sanjuan-Pla, A. *et al.* "Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy." *Nature* **502,** 232–236 (2013); DOI: 10.1038/nature12495

57. Haas, S. *et al.* “Inflammation-Induced Emergency Megakaryopoiesis Driven by Hematopoietic Stem Cell-like Megakaryocyte Progenitors.” *Cell Stem Cell* **17,** 422–434 (2015); DOI: 10.1016/j.stem.2015.07.007

58. Notta, F. *et al.* “Distinct routes of lineage development reshape the human blood hierarchy across ontogeny.” *Science* **351,** aab2116 (2016); DOI: 10.1126/science.aab2116

59. Carrelha, J. *et al.* “Hierarchically related lineage-restricted fates of multipotent haematopoietic stem cells.” *Nature* **554,** 106–111 (2018); DOI: 10.1038/nature25455

60. Gekas, C. & Graf, T. “CD41 expression marks myeloid-biased adult hematopoietic stem cells and increases with age.” *Blood* **121,** 4463–72 (2013); DOI: 10.1182/blood-2012-09-457929

61. Cabezas-Wallscheid, N. *et al.* “Identification of regulatory networks in HSCs and their immediate progeny via integrated proteome, transcriptome, and DNA methylome analysis.” *Cell Stem Cell* **15,** 507–522 (2014); DOI: 10.1016/j.stem.2014.07.005

62. Haas, S., Trumpp, A. & Milsom, M. D. “Causes and Consequences of Hematopoietic Stem Cell Heterogeneity.” *Cell Stem Cell* **22,** 627–638 (2018); DOI: 10.1016/j.stem.2018.04.003

63. Yu, V. W. C. *et al.* “Epigenetic Memory Underlies Cell-Autonomous Heterogeneous Behavior of Hematopoietic Stem Cells.” *Cell* **167,** 1310–1322.e17 (2016); DOI: 10.1016/j.cell.2016.10.045

64. Rodriguez-Fraticelli, A. E. *et al.* “Clonal analysis of lineage fate in native haematopoiesis.” *Nature* **553,** 212–216 (2018); DOI: 10.1038/nature25168

65. Perié, L. *et al.* “The Branching Point in Erythro-Myeloid Differentiation.” *Cell* **163,** 1655–1662 (2015); DOI: 10.1016/j.cell.2015.11.059

66. Karamitros, D. *et al.* “Single-cell analysis reveals the continuum of human lympho-myeloid progenitor cells.” *Nat. Immunol.* **19,** 85–97 (2018); DOI: 10.1038/s41590-017-0001-2

67. Velten, L. *et al.* "Human haematopoietic stem cell lineage commitment is a continuous process." *Nat. Cell Biol.* **19,** (2017); DOI: 10.1038/ncb3493

68. Biasco, L. *et al.* “In Vivo Tracking of Human Hematopoiesis Reveals Patterns of Clonal Dynamics during Early and Steady-State Reconstitution Phases.” *Cell Stem Cell* **19,** 107–119 (2015); DOI: 10.1016/j.stem.2016.04.016

69. Ganuza, M. *et al.* “Lifelong haematopoiesis is established by hundreds of precursors throughout mammalian ontogeny.” *Nat. Cell Biol.* **19,** 1153–1163 (2017); DOI: 10.1038/ncb3607

70. Bowman, R. L., Busque, L. & Levine, R. L. “Clonal Hematopoiesis and Evolution to Hematopoietic Malignancies.” *Cell Stem Cell* **22,** 157–170 (2018); DOI: 10.1016/j.stem.2018.01.011

71. Raslova, H. *et al.* “Megakaryocyte polyploidization is associated with a functional gene amplification.” *Blood* **101,** 541–544 (2003); DOI: 10.1182/blood-2002-05-1553

72. Gao Y. *et al.* "Role of RhoA-Specific Guanine Exchange Factors in Regulation of Endomitosis in Megakaryocytes." *Dev. Cell* **22,** 573–584 (2012); DOI: 10.1016/j.devcel.2011.12.019

73. Lordier, L. *et al.* "RUNX1-induced silencing of non-muscle myosin heavy chain IIB contributes to megakaryocyte polyploidization." *Nat. Commun.* **3,** 717 (2012); DOI: 10.1038/ncomms1704

74. Leysi-Derilou, Y. *et al.* "Polyploid megakaryocytes can complete cytokinesis." *Cell Cycle* **9,** 2589–2599 (2010); DOI: 10.4161/cc.9.13.12078

75. Mazzi, S. et al. "Megakaryocyte and polyploidization." *Exp. Hematol.* **57,** 1–13 (2018); DOI: 10.1016/j.exphem.2017.10.001

76. Trakala, M. *et al.* "Functional Reprogramming of Polyploidization in Megakaryocytes." *Dev. Cell* **32,** 155–167 (2015); DOI: 10.1016/j.devcel.2014.12.015

77. Chang, Y. *et al.* "Proplatelet formation is regulated by the Rho/ROCK pathway." *Blood* **109,** 4229–4236 (2007); DOI: 10.1182/blood-2006-04-020024

78. Kovar, D. R. "Molecular details of formin-mediated actin assembly." *Curr. Opin. Cell Biol.* **18,** 11–17 (2006); DOI: 10.1016/j.ceb.2005.12.011

79. Pan, J. *et al.* "The formin DIAPH1 (mDia1) regulates megakaryocyte proplatelet formation by remodeling the actin and microtubule cytoskeletons." *Blood* **124,** 3967–3977 (2014); DOI: 10.1182/blood-2013-12-544924

80. Kauskot, A. *et al.* "LIM kinase/cofilin dysregulation promotes macrothrombocytopenia in severe von Willebrand disease-type 2B." *JCI Insight* **1,** 1–17 (2016); DOI: 10.1172/jci.insight.88643

81. Dütting, S. *et al.* "A Cdc42/RhoA regulatory circuit downstream of glycoprotein Ib guides transendothelial platelet biogenesis." *Nat. Commun.* **8,** 15838 (2017); DOI: 10.1038/ncomms15838

82. Nishimura, S. *et al.* "IL-1α induces thrombopoiesis through megakaryocyte rupture in response to acute platelet needs." *J. Cell Biol.* **209,** 453–66 (2015); DOI: 10.1083/jcb.201410052

83. Orkin, S. H. & Zon, L. I. Hematopoiesis: an evolving paradigm for stem cell biology. *Cell* **132,** 631–44 (2008); DOI: 10.1016/j.cell.2008.01.025

84. Tijssen, M. R. *et al.* "Genome-wide Analysis of Simultaneous GATA1/2, RUNX1, FLI1, and SCL Binding in Megakaryocytes Identifies Hematopoietic Regulators." *Dev. Cell* **20,** 597–609 (2011); DOI: 10.1016/j.devcel.2011.04.008

85. Crispino, J. D. & Horwitz, M. S. "GATA factor mutations in hematologic disease." *Blood* **129,** 2103–2110 (2017); DOI: 10.1182/blood-2016-09-687889

86. Pimkin, M. *et al.* “Divergent functions of hematopoietic transcription factors in lineage priming and differentiation during erythro-megakaryopoiesis.” *Genome Res.* **24,** 1932–1944 (2014); DOI: 10.1101/gr.164178.113

87. Pozner, A. *et al.* “Developmentally regulated promoter-switch transcriptionally controls Runx1 function during embryonic hematopoiesis.” *BMC Dev. Biol.* **7,** 84 (2007); DOI: 10.1186/1471-213X-7-84

88. Roberts, I. *et al.* “GATA1-mutant clones are frequent and often unsuspected in babies with Down syndrome: identification of a population at risk of leukemia.” *Blood* **122,** 3908–17 (2013); DOI: 10.1182/blood-2013-07-515148

89. Wilting, R. H. *et al.* “Overlapping functions of Hdac1 and Hdac2 in cell cycle regulation and haematopoiesis.” *EMBO J.* **29,** 2586–97 (2010); DOI: 10.1038/emboj.2010.136

90. Messaoudi, K. *et al.* “Critical role of the HDAC6–cortactin axis in human megakaryocyte maturation leading to a proplatelet-formation defect.” *Nat. Commun.* **8,** 1786 (2017); DOI: 10.1038/s41467-017-01690-2

91. Yu, M. *et al.* “Direct recruitment of polycomb repressive complex 1 to chromatin by core binding transcription factors.” *Mol. Cell* **45,** 330–43 (2012); DOI: 10.1016/j.molcel.2011.11.032

92. Cooper, S. *et al.* “Jarid2 binds mono-ubiquitylated H2A lysine 119 to mediate crosstalk between Polycomb complexes PRC1 and PRC2.” *Nat. Commun.* **7,** 13661 (2016); DOI: 10.1038/ncomms13661

93. Sashida, G. & Iwama, A. “Multifaceted role of the polycomb-group gene EZH2 in hematological malignancies.” *Int. J. Hematol.* **105,** 23–30 (2017); DOI: 10.1007/s12185-016-2124-x

94. Quivoron, C. *et al.* “TET2 Inactivation Results in Pleiotropic Hematopoietic Abnormalities in Mouse and Is a Recurrent Event during Human Lymphomagenesis.” *Cancer Cell* **20,** 25–38 (2011); DOI: 10.1016/j.ccr.2011.06.003

95. Ge, L. *et al.* “TET2 Plays an Essential Role in Erythropoiesis by Regulating Lineage-Specific Genes via DNA Oxidative Demethylation in a Zebrafish Model.” *Mol. Cell. Biol.* **34,** 989–1002 (2014); DOI: 10.1128/MCB.01061-13

96. Xu, W. *et al.* “Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of α-ketoglutarate-dependent dioxygenases.” *Cancer Cell* **19,** 17–30 (2011); DOI: 10.1016/j.ccr.2010.12.014

97. Evans, B. *et al.* “Registered report: Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of α-ketoglutarate-dependent dioxygenases.” *Elife* **4,** e07420 (2015); DOI: 10.7554/eLife.07420

98. Raffel, S. *et al.* “BCAT1 restricts αKG levels in AML stem cells leading to IDHmut-like DNA hypermethylation.” *Nature* **551,** 384–388 (2017); DOI: 10.1038/nature24294

99. Mazumdar, C. & Majeti, R. “The role of mutations in the cohesin complex in acute myeloid leukemia.” *International Journal of Hematology* **105,** 31–36 (2017); DOI: 10.1007/s12185016-2119-7

100. Deininger, M. W. N., Tyner, J. W. & Solary, E. “Turning the tide in myelodysplastic/myeloproliferative neoplasms.” *Nat. Rev. Cancer* **17,** 425–440 (2017); DOI: 10.1038/nrc.2017.40

101. Vainchenker, W. & Kralovics, R. “Genetic basis and molecular pathophysiology of classical myeloproliferative neoplasms.” *Blood* **129,** 667–679 (2017); DOI: 10.1182/blood-2016-10-695940

102. Varghese, L. N. *et al.* “Functional characterization of c-Mpl ectodomain mutations that underlie congenital amegakaryocytic thrombocytopenia.” *Growth Factors* **32,** 18–26 (2014); DOI: 10.3109/08977194.2013.874347

103. Fox, N. E. *et al.* “F104S c-Mpl responds to a transmembrane domain-binding thrombopoietin receptor agonist: proof of concept that selected receptor mutations in congenital amegakaryocytic thrombocytopenia can be stimulated with alternative thrombopoietic agents.” *Exp. Hematol.* **38,** 384–91 (2010); DOI: 10.1016/j.exphem.2010.02.007

104. Sabath, D. F., Kaushansky, K. & Broudy, V. C. “Deletion of the extracellular membrane-distal cytokine receptor homology module of Mpl results in constitutive cell growth and loss of thrombopoietin binding.” *Blood* **94,** 365–7 (1999); DOI:

105. Millot, G. A *et al.* “Differential signalling of NH2-terminal flag-labelled thrombopoietin receptor activated by TPO or anti-FLAG antibodies.” *Cell. Signal.* **16,** 355–363 (2004); DOI: 10.1016/j.cellsig.2003.08.010

106. Sattler, M. *et al.* “The thrombopoietin receptor c-MPL activates JAK2 and TYK2 tyrosine kinases.” *Exp. Hematol.* **23,** 1040–8 (1995); DOI:

107. Tanner, J. W. *et al.* “The conserved box 1 motif of cytokine receptors is required for association with JAK kinases.” *J. Biol. Chem.* **270,** 6523–30 (1995); DOI: 10.1074/jbc.270.12.6523

108. Ungureanu, D. *et al.* “The pseudokinase domain of JAK2 is a dual-specificity protein kinase that negatively regulates cytokine signaling.” *Nat. Struct. Mol. Biol.* **18,** 971–976 (2011); DOI: 10.1038/nsmb.2099

109. James, C. *et al.* “A unique clonal JAK2 mutation leading to constitutive signalling causes polycythaemia vera.” *Nature* **434,** 1144–1148 (2005); DOI: 10.1038/nature03546

110. Kralovics, R. *et al.* “A Gain-of-Function Mutation of JAK2 in Myeloproliferative Disorders.” *N. Engl. J. Med.* **352,** 1779–1790 (2005); DOI: 10.1056/NEJMoa051113

111. Levine, R. L. *et al.* “Activating mutation in the tyrosine kinase JAK2 in polycythemia vera, essential thrombocythemia, and myeloid metaplasia with myelofibrosis.” *Cancer Cell* **7,** 387–397 (2005); DOI: 10.1016/j.ccr.2005.03.023

112. Baxter, E. J. *et al.* “Acquired mutation of the tyrosine kinase JAK2 in human myeloproliferative disorders.” *Lancet* **365,** 1054–1061 (2005); DOI: 10.1016/S0140-6736(05)71142-9

113. Malinge, S. *et al.* “Novel activating JAK2 mutation in a patient with Down syndrome and B-cell precursor acute lymphoblastic leukemia.” *Blood* **109,** 2202–4 (2007); DOI: 10.1182/blood-2006-09-045963

114. Leroy, E. *et al.* “His499 Regulates Dimerization and Prevents Oncogenic Activation by Asparagine Mutations of the Human Thrombopoietin Receptor.” *J. Biol. Chem.* **291,** 2974–87 (2016); DOI: 10.1074/jbc.M115.696534

115. Defour, J.P. *et al.* “Tryptophan at the transmembrane-cytosolic junction modulates thrombopoietin receptor dimerization and activation.” *Proc. Natl. Acad. Sci. U. S. A.* **110,** 2540–5 (2013); DOI: 10.1073/pnas.1211560110

116. Pikman, Y. *et al.* “MPLW515L is a novel somatic activating mutation in myelofibrosis with myeloid metaplasia.” *PLoS Med.* **3,** 1140–1151 (2006); DOI: 10.1371/journal.pmed.0030270

117. Staerk, J. *et al.* “An amphipathic motif at the transmembrane-cytoplasmic junction prevents autonomous activation of the thrombopoietin receptor.” *Blood* **107,** 1864–1871 (2006); DOI: 10.1182/blood-2005-06-2600

118. Klampfl, T. *et al.* “Somatic mutations of calreticulin in myeloproliferative neoplasms.” *N. Engl. J. Med.* **369,** 2379–90 (2013); DOI: 10.1056/NEJMoa1311347

119. Chachoua, I. *et al.* “Thrombopoietin receptor activation by myeloproliferative neoplasm associated calreticulin mutants.” *Blood* **127,** 1325–1335 (2016); DOI: 10.1182/blood-2015-11-681932

120. Araki, M. *et al.* “Activation of the thrombopoietin receptor by mutant calreticulin in CALR-mutant myeloproliferative neoplasms.” *Blood* **127,** 1307–1316 (2016); DOI: 10.1182/blood-2015-09-671172

121. Vainchenker, W. & Constantinescu, S. N. “JAK/STAT signaling in hematological malignancies.” *Oncogene* **32,** 2601–13 (2013); DOI: 10.1038/onc.2012.347

122. Walz, C. *et al.* “Essential role for Stat5a/b in myeloproliferative neoplasms induced by BCR-ABL1 and JAK2 V617F in mice.” *Blood* **119,** 3550–3560 (2012); DOI: 10.1182/blood-2011-12-397554

123. Yan, D., Hutchison, R. E. & Mohi, G. “Critical requirement for Stat5 in a mouse model of polycythemia vera.” *Blood* **119,** 3539–3549 (2012); DOI: 10.1182/blood-2011-03-345215

124. Reich, N. C. “STATs get their move on.” *JAK-STAT* **2,** e27080 (2013); DOI: 10.4161/jkst.27080

125. Park, H. J. *et al.* “Cytokine-induced megakaryocytic differentiation is regulated by genome-wide loss of a uSTAT transcriptional program.” *EMBO J.* **35,** 580–594 (2016); DOI: 10.15252/embj.201592383

126. Séverin, S., Ghevaert, C. & Mazharian, A. “The mitogen-activated protein kinase signaling pathways: role in megakaryocyte differentiation.” *J. Thromb. Haemost.* **8,** 17–26 (2010); DOI: 10.1111/j.1538-7836.2009.03658.x

127. Bluteau, D. *et al.* “Thrombocytopenia-associated mutations in the ANKRD26 regulatory region induce MAPK hyperactivation.” *J. Clin. Invest.* **124,** 580–91 (2014); DOI: 10.1172/JCI71861

128. Besancenot, R. *et al.* “A senescence-like cell-cycle arrest occurs during megakaryocytic maturation: Implications for physiological and pathological megakaryocytic proliferation.” *PLoS Biol.* **8,** e1000476 (2010); DOI: 10.1371/journal.pbio.1000476

129. De Graaf, C. A. & Metcalf, D. “Thrombopoietin and hematopoietic stem cells.” *Cell Cycle* **10,** 1582–1589 (2011); DOI: 10.4161/cc.10.10.15619

130. Kuter, D. J. & Rosenberg, R. D. “The reciprocal relationship of thrombopoietin (c-Mpl ligand) to changes in the platelet mass during busulfan-induced thrombocytopenia in the rabbit.” *Blood* **85,** 2720–30 (1995)

131. Fielder, P. J. *et al.* “Human platelets as a model for the binding and degradation of thrombopoietin.” *Blood* **89,** 2782–8 (1997)

132. Favale, F. *et al.* “An incomplete trafficking defect to the cell-surface leads to paradoxical thrombocytosis for human and murine MPL P106L.” *Blood* **128,** 3146–3158 (2016); DOI: 10.1182/blood-2016-06-722058

133. Lannutti, B. J. *et al.* “Incomplete restoration of Mpl expression in the mpl-/- mouse produces partial correction of the stem cell-repopulating defect and paradoxical thrombocytosis.” *Blood* **113,** 1778–1785 (2009); DOI: 10.1182/blood-2007-11-124859

134. Stoffel, R., Wiestner, A. & Skoda, R. “Thrombopoietin in thrombocytopenic mice: evidence against regulation at the mRNA level and for a direct regulatory role of platelets.” *Blood* **87,** (1996); DOI:

135. Hitchcock, I. S. *et al* "YRRL motifs in the cytoplasmic domain of the thrombopoietin receptor regulate receptor internalization and degradation." *Blood* **112,** 2222–31 (2008); DOI: 10.1182/blood-2008-01-134049

136. Dahlen, D. D., Broudy, V. C. & Drachman, J. G. "Internalization of the thrombopoietin receptor is regulated by 2 cytoplasmic motifs." *Blood* **102,** 102–108 (2003); DOI: 10.1182/blood-2002-11-3468

137. Saur, S. J. *et al.* "Ubiquitination and degradation of the thrombopoietin receptor c-Mpl." *Blood* **115,** 1254–63 (2010); DOI: 10.1182/blood-2009-06-227033

138. Hookham, M. B. *et al.* "The myeloproliferative disorder-associated JAK2 V617F mutant escapes negative regulation by suppressor of cytokine signaling 3." *Blood* **109,** 4924–4929 (2007); DOI: 10.1182/blood-2006-08-039735

139. McMullin, M. F. & Cario, H. "LNK mutations and myeloproliferative disorders." *Am. J. Hematol.* **91,** 248–51 (2016); DOI: 10.1002/ajh.24259

140. Currao, M., Balduini, C. L. & Balduini, "A High Doses of Romiplostim Induce Proliferation and Reduce Proplatelet Formation by Human Megakaryocytes." *PLoS One* **8,** e54723 (2013); DOI: 10.1371/journal.pone.0054723

141. Di Buduo, C. A. *et al.* "Revealing eltrombopag's promotion of human megakaryopoiesis through AKT/ERK-dependent pathway activation." *Haematologica* **101,** 1479–1488 (2016); DOI: 10.3324/haematol.2016.146746

142. Wang, X. *et al.* "A thrombopoietin receptor antagonist is capable of depleting myelofibrosis hematopoietic stem and progenitor cells." *Blood* **127,** 3398–3409 (2016); DOI: 10.1182/blood-2015-10-674465

143. Nurden, A. T. & Caen, J. P. "An Abnormal Platelet Glycoprotein Pattern in Three Cases of Glanzmann's Thrombasthenia." *Br. J. Haematol.* **28,** 253–260 (1974); DOI: 10.1111/j.1365-2141.1974.tb06660.x

144. Lentaigne, C. *et al.* "Inherited platelet disorders: toward DNA-based diagnosis." *Blood* **127,** 2814–23 (2016); DOI: 10.1182/blood-2016-03-378588

145. Simeoni, I. *et al.* "A high-throughput sequencing test for diagnosing inherited bleeding, thrombotic, and platelet disorders." *Blood* **127,** 2791–2803 (2016); DOI: 10.1182/blood-2015-12-688267

146. Favier, R. & Raslova, H. "Progress in understanding the diagnosis and molecular genetics of macrothrombocytopenias." *Br. J. Haematol.* **170,** 626–639 (2015); DOI: 10.1111/bjh.13478

147. Eto, K. & Kunishima, S. "Linkage between the mechanisms of thrombocytopenia and thrombopoiesis." *Blood* **127,** 1234–41 (2016); DOI: 10.1182/blood-2015-07-607903

148. Ballmaier, M. & Germeshausen, M. “Congenital amegakaryocytic thrombocytopenia: Clinical presentation, diagnosis, and treatment.” *Semin. Thromb. Hemost.* **37,** 673–681 (2011); DOI: 10.1055/s-0031-1291377

149. Alexander, W. S. *et al.* “Deficiencies in progenitor cells of multiple hematopoietic lineages and defective megakaryocytopoiesis in mice lacking the thrombopoietic receptor c-Mpl.” *Blood* **87,** 2162–2170 (1996).

150. Hirata, S. *et al.* “Congenital amegakaryocytic thrombocytopenia iPS cells exhibit defective MPL-mediated signaling.” *J. Clin. Invest.* **123,** 3802–14 (2013); DOI: 10.1172/JCI64721

151. Seo, A. *et al.* “Bone marrow failure unresponsive to bone marrow transplant is caused by mutations in thrombopoietin.” *Blood* **130,** 875–880 (2017); DOI: 10.1182/blood-2017-02-768036

152. Pecci, A. *et al.* “Thrombopoietin mutation in congenital amegakaryocytic thrombocytopenia treatable with romiplostim.” *EMBO Mol. Med.* **10,** 63–75 (2018); DOI: 10.15252/emmm.201708168

153. Kim, A. R. & Sankaran, V. G. “Thrombopoietin: tickling the HSC’s fancy.” *EMBO Mol. Med.* **10,** 10–12 (2018); DOI: 10.15252/emmm.201708450

154. Albers, C. a *et al.* “Compound inheritance of a low-frequency regulatory SNP and a rare null mutation in exon-junction complex subunit RBM8A causes TAR syndrome.” *Nat. Genet.* **44,** 435–9, S1-2 (2012); DOI: 10.1038/ng.1083

155. Jacobsen, P. *et al.”* An (11; 21) translocation in four generations with chromosome 11 abnormalities in the offspring: A clinical, cytogenetical, and gene marker study.” *Hum. Hered.* **23,** 568–585 (1973); DOI: 10.1159/000152624

156. Favier, R. *et al.* “A novel genetic thrombocytopenia (Paris-Trousseau) associated with platelet inclusions, dysmegakaryopoiesis and chromosome deletion AT 11q23.” *C.R.Acad. Sci.* **316,** 698–701 (1993).

157. Krishnamurti, L. *et al.* “Paris-Trousseau syndrome platelets in a child with Jacobsen’s syndrome.” *Am. J. Hematol.* **66,** 295–299 (2001); DOI: 10.1002/ajh.1061

158. Hart, A. *et al.* “Fli-1 is required for murine vascular and megakaryocytic development and is hemizygously deleted in patients with thrombocytopenia.” *Immunity* **13,** 167–177 (2000); DOI: 10.1016/S1074-7613(00)00017-0

159. Stockley, J. *et al.* “Enrichment of FLI1 and RUNX1 mutations in families with excessive bleeding and platelet dense granule secretion defects.” *Blood* **122,** 4090–4093 (2013); DOI: 10.1182/blood-2013-06-506873

160. Stevenson, W. S. *et al.* “Paris-Trousseau thrombocytopenia is phenocopied by the autosomal recessive inheritance of a DNA-binding domain mutation in FLI1.” *Blood* **126,** 2027–2030 (2015); DOI: 10.1182/blood-2015-06-650887

161. Raslova, H. *et al.* “FLI1 monoallelic expression combined with its hemizygous loss underlies Paris-Trousseau/Jacobsen thrombopenia.” *J. Clin. Invest.* **114,** 77–84 (2004); DOI: 10.1172/JCI200421197

162. Song, W. J. *et al.* “Haploinsufficiency of CBFA2 causes familial thrombocytopenia with propensity to develop acute myelogenous leukaemia.” *Nat. Genet.* **23,** 166–175 (1999); DOI: 10.1038/13793

163. Matheny, C. J. *et al.* “Disease mutations in RUNX1 and RUNX2 create nonfunctional, dominant-negative, or hypomorphic alleles.” *EMBO J.* **26,** 1163–1175 (2007); DOI: 10.1038/sj.emboj.7601568

164. Churpek, J. E. & Godley, L. A. “How I diagnose and manage individuals at risk for inherited myeloid malignancies.” *Blood* **128,** 1800–1813 (2016); DOI: 10.1182/blood-2016-05-670240

165. Antony-Debré, I. *et al.* “Somatic mutations associated with leukemic progression of familial platelet disorder with predisposition to acute myeloid leukemia.” *Leukemia* **30,** 999–1002 (2016); DOI: 10.1038/leu.2015.236

166. North, T. *et al.* “Cbfa2 is required for the formation of intra-aortic hematopoietic clusters.” *Development* **126,** 2563–75 (1999)

167. Ichikawa, M. *et al.* “AML-1 is required for megakaryocytic maturation and lymphocytic differentiation, but not for maintenance of hematopoietic stem cells in adult hematopoiesis.” *Nat. Med.* **10,** 299–304 (2004); DOI: 10.1038/nm997

168. Growney, J. D. *et al.* “Loss of Runx1 perturbs adult hematopoiesis and is associated with a myeloproliferative phenotype.” *Blood* **106,** 494–504 (2005); DOI: 10.1182/blood-2004-08-3280

169. Bluteau, D. *et al.* “Down-regulation of the RUNX1-target gene NR4A3 contributes to hematopoiesis deregulation in familial platelet disorder/acute myelogenous leukemia.” *Blood* **118,** 6310–6320 (2011); DOI: 10.1182/blood-2010-12-325555

170. Sakurai, M. *et al.* “Impaired hematopoietic differentiation of RUNX1-mutated induced pluripotent stem cells derived from FPD/AML patients.” *Leukemia* **28,** 2344–2354 (2014); DOI: 10.1038/leu.2014.136

171. Connelly, J. P. *et al.* “Targeted correction of RUNX1 mutation in FPD patient-specific induced pluripotent stem cells rescues megakaryopoietic defects.” *Blood* **124,** 1926–1930 (2014); DOI: 10.1182/blood-2014-01-550525

172. Antony-Debré, I. *et al.* "Level of RUNX1 activity is critical for leukemic predisposition but not for thrombocytopenia." *Blood* **125,** 930–40 (2015); DOI: 10.1182/blood-2014-06-585513

173. Bluteau, D. *et al.* "Dysmegakaryopoiesis of FPD/AML pedigrees with constitutional RUNX1 mutations is linked to myosin II deregulated expression." *Blood* **120,** 2708–18 (2012); DOI: 10.1182/blood-2012-04-422337

174. Zhang, M. Y. *et al.* "Germline ETV6 mutations in familial thrombocytopenia and hematologic malignancy." *Nat. Genet.* **47,** 180–185 (2015); DOI: 10.1038/ng.3177

175. Greaves, M. "A causal mechanism for childhood acute lymphoblastic leukaemia." *Nat. Rev. Cancer* 1–14 (2018); DOI: 10.1038/s41568-018-0015-6

176. Kwiatkowskit, B. A. *et al.* "The ets family member tel binds to the Fli-1 oncoprotein and inhibits its transcriptional activity." *J. Biol. Chem.* **273,** 17525–17530 (1998); DOI: 10.1074/jbc.273.28.17525

177. Feurstein, S. & Godley, L. A. "Germline ETV6 mutations and predisposition to hematological malignancies." *Int. J. Hematol.* **106,** 189–195 (2017); DOI: 10.1007/s12185-017-2259-4

178. Poggi, M. *et al.* "Germline variants in ETV6 underlie reduced platelet formation, platelet dysfunction and increased levels of circulating CD34+progenitors." *Haematologica* **102,** 282–294 (2017); DOI: 10.3324/haematol.2016.147694

179. Turro, E. *et al.* "A dominant gain-of-function mutation in universal tyrosine kinase SRC causes thrombocytopenia, myelofibrosis, bleeding, and bone pathologies." *Sci. Transl. Med.* **8,** 1–16 (2016); DOI: 10.1126/scitranslmed.aad7666

180. Ferreira, R. et al. "GATA1 function, a paradigm for transcription factors in hematopoiesis." *Mol. Cell. Biol.* **25,** 1215–27 (2005); DOI: 10.1128/MCB.25.4.1215-1227.2005

181. Nichols, K. E. *et al.* "Familial dyserythropoietic anaemia and thrombocytopenia due to an inherited mutation in GATA 1." *Nat. Genet.* **24,** 266–270 (2000); DOI: 10.1038/73480

182. Freson, K., Wijgaerts, A. & Van Geet, C. "GATA1 gene variants associated with thrombocytopenia and anemia." *Platelets* **28,** 731–734 (2017); DOI: 10.1080/09537104.2017.1361525

183. Thambyrajah, R. *et al.* "GFI1 proteins orchestrate the emergence of haematopoietic stem cells through recruitment of LSD1." *Nat. Cell Biol.* **18,** 21–32 (2016); DOI: 10.1038/ncb3276

184. Rabbolini, D. J. et al. "GFI1B variants associated with thrombocytopenia." *Platelets* **28,** 525–527 (2017); DOI: 10.1080/09537104.2017.1317734

185. Monteferrario, D. *et al.* "A Dominant-Negative GFI1B Mutation in the Gray Platelet Syndrome." *N. Engl. J. Med.* **370,** 245–253 (2014); DOI: 10.1056/NEJMoa1308130

186. Saleque, S., Cameron, S. & Orkin, S. H. “The zinc-finger proto-oncogene Gfi-1b is essential for development of the erythroid and megakaryocytic lineages.” *Genes Dev.* **16,** 301–306 (2002); DOI: 10.1101/gad.959102

187. Raccuglia, G. “Gray platelet syndrome. A variety of qualitative platelet disorder.” *Am. J. Med.* **51,** 818–828 (1971); DOI: 10.1016/0002-9343(71)90311-1

188. Gunay-Aygun, M. *et al.* “Gray platelet syndrome: natural history of a large patient cohort and locus assignment to chromosome 3p.” *Blood* **116,** 4990–5001 (2010); DOI: 10.1182/blood-2010-05-286534

189. Albers, C. A. *et al.* “Exome sequencing identifies NBEAL2 as the causative gene for gray platelet syndrome.” *Nat. Genet.* **43,** 735–737 (2011); DOI: 10.1038/ng.885

190. Kahr, W. H. *et al.* “Mutations in NBEAL2, encoding a BEACH protein, cause gray platelet syndrome.” *Nat. Genet.* **43,** 738–740 (2011); DOI: 10.1038/ng.884

191. Gunay-Aygun, M. *et al.* “NBEAL2 is mutated in gray platelet syndrome and is required for biogenesis of platelet γ-granules.” *Nat. Genet.* **43,** 732–734 (2011); DOI: 10.1038/ng.883

192. Kahr, W. H. A. *et al.* “Abnormal megakaryocyte development and platelet function in Nbeal2-/- mice.” *Blood* **122,** 3349–3358 (2013); DOI: 10.1182/blood-2013-04-499491

193. Di Buduo, C. A. *et al.* “Abnormal proplatelet formation and emperipolesis in cultured human megakaryocytes from gray platelet syndrome patients.” *Sci. Rep.* **6,** 23213 (2016); DOI: 10.1038/srep23213

194. Deppermann, C. *et al.* “Gray platelet syndrome and defective thrombo-inflammation in Nbeal2-deficient mice.” *J. Clin. Invest.* **123,** 3331–3342 (2013); DOI: 10.1172/JCI69210

195. Guerrero, J. A. *et al.* “Gray platelet syndrome: Proinflammatory megakaryocytes and α-granule loss cause myelofibrosis and confer metastasis resistance in mice.” *Blood* **124,** 3624–3635 (2014); DOI: 10.1182/blood-2014-04-566760

196. Sowerby, J. M. *et al.* “NBEAL2 is required for neutrophil and NK cell function and pathogen defense.” *J. Clin. Invest.* **127,** 3521–3526 (2017); DOI: 10.1172/JCI91684

197. Mayer, L. *et al.* “Nbeal2 interacts with Dock7, Sec16a, and Vac14.” *Blood* **131,** 1000–1011 (2018); DOI: 10.1182/blood-2017-08-800359

198. Wijgaerts, A. *et al.* “The transcription factor GATA1 regulates NBEAL2 expression through a long-distance enhancer.” *Haematologica* **102,** 695–706 (2017); DOI: 10.3324/haematol.2016.152777

199. Massaad, M. J., Ramesh, N. & Geha, R. S. “Wiskott-Aldrich syndrome: A comprehensive review.” *Ann. N. Y. Acad. Sci.* **1285,** 26–43 (2013); DOI: 10.1111/nyas.12049

200. Haddad, E. *et al.* “The thrombocytopenia of Wiskott Aldrich syndrome is not related to a defect in proplatelet formation.” *Blood* **94,** 509–18 (1999).

201. Aiuti, A. *et al.* “Lentiviral hematopoietic stem cell gene therapy in patients with Wiskott-Aldrich syndrome.” *Science* **341,** 1233151 (2013); DOI: 10.1126/science.1233151

202. Newman, P. J. *et al.* “The molecular genetic basis of Glanzmann thrombasthenia in the Iraqi-Jewish and Arab populations in Israel.” *Proc. Natl. Acad. Sci. U. S. A.* **88,** 3160–3164 (1991); DOI: 10.1073/pnas.88.8.3160

203. George, J. N., Caen, J. P. & Nurden, A. T. “Glanzmann’s thrombasthenia: the spectrum of clinical disease.” *Blood* **75,** 1383–95 (1990).

204. Nurden, A. T. & Pillois, X. “ITGA2B and ITGB3 gene mutations associated with Glanzmann thrombasthenia.” *Platelets* **29,** 98–101 (2018); DOI: 10.1080/09537104.2017.1371291

205. Mitchell, W. B. et al. “αIIbβ3 biogenesis is controlled by engagement of αIIb in the calnexin cycle via the N15-linked glycan.” *Blood* **107,** 2713–2719 (2006); DOI: 10.1182/blood-2005-07-2990

206. Zhu, J. *et al.* “The Structure of a Receptor with Two Associating Transmembrane Domains on the Cell Surface: Integrin αIIbβ3.” *Mol. Cell* **34,** 234–249 (2009); DOI: 10.1016/j.molcel.2009.02.022

207. Watson, S. P. *et al.* “GPVI and integrin alphaIIb beta3 signaling in platelets.” *J. Thromb. Haemost.* **3,** 1752–62 (2005); DOI: 10.1111/j.1538-7836.2005.01429.x

208. Nurden, A. T. *et al.* “Glanzmann thrombasthenia: A review of ITGA2B and ITGB3 defects with emphasis on variants, phenotypic variability, and mouse models.” *Blood* **118,** 5996–6005 (2011); DOI: 10.1182/blood-2011-07-365635

209. Ley, K. et al. “Integrin-based therapeutics: biological basis, clinical use and new drugs.” *Nat. Rev. Drug Discov.* **15,** 173–183 (2016); DOI: 10.1038/nrd.2015.10

210. Ghevaert, C. *et al.* “A nonsynonymous SNP in the ITGB3 gene disrupts the conserved membrane-proximal cytoplasmic salt bridge in the αIIbβ3integrin and cosegregates dominantly with abnormal proplatelet formation and macrothrombocytopenia.” *Blood* **111,** 3407–3414 (2008); DOI: 10.1182/blood-2007-09-112615

211. Schaffner-Reckinger, E. *et al.* “Overexpression of the partially activated αIIb β3D723H integrin salt bridge mutant downregulates RhoA activity and induces microtubule-dependent proplatelet-like extensions in Chinese hamster ovary cells.” *J. Thromb. Haemost.* **7,** 1207–1217 (2009); DOI: 10.1111/j.1538-7836.2009.03494.x

212. Bury, L. et al. “Outside-In Signalling Generated by a Constitutively Activated Integrin αIIbβ3 Impairs Proplatelet Formation in Human Megakaryocytes.” *PLoS One* **7,** e34449 (2012); DOI: 10.1371/journal.pone.0034449

213. Reynolds, A. R. *et al.* "Elevated Flk1 (vascular endothelial growth factor receptor 2) signaling mediates enhanced angiogenesis in β3-integrin-deficient mice." *Cancer Res.* **64,** 8643–8650 (2004); DOI: 10.1158/0008-5472.CAN-04-2760

214. Bernard, J. & Soulier, J.P. "Sur une nouvelle variete de dystrophic thrombocytaire-hemorragipare congenitale." *Sem Hop Paris*, 24, 3217– 3223 (1948)

215. Nurden, A. T. et al. "Analysis of the glycoprotein and protein composition of Bernard-Soulier platelets by single and two-dimensional sodium dodecyl sulfate-polyacrylamide gel electrophoresis." *J. Clin. Invest.* **67,** 1431–40 (1981)

216. Lu, Q. et al. "Snake venom C-type lectins interacting with platelet receptors. Structure–function relationships and effects on haemostasis." *Toxicon* **45,** 1089–1098 (2005); DOI: 10.1016/j.toxicon.2005.02.022

217. Luo, S. Z. *et al.* "Glycoprotein Ibα forms disulfide bonds with 2 glycoprotein Ibβ subunits in the resting platelet." *Blood* **109,** 603–609 (2007); DOI: 10.1182/blood-2006-05-024091

218. Mo, X. *et al.* "The transmembrane domain of glycoprotein Ib alpha is critical to efficient expression of glycoprotein Ib-IX Complex in the plasma membrane." *J. Biol. Chem.* **281,** 23050–23059 (2006); DOI: 10.1074/jbc.M600924200

219. Luo, S. Z. *et al.* "Role of the transmembrane domain of glycoprotein IX in assembly of the glycoprotein Ib-IX complex." *J. Thromb. Haemost.* **5,** 2494–2502 (2007); DOI: 10.1111/j.1538-7836.2007.02785.x

220. Mo, X. *et al.* "Binding of platelet glycoprotein Ibβ through the convex surface of leucine-rich repeats domain of glycoprotein IX." *J. Thromb. Haemost.* **7,** 1533–1540 (2009); DOI: 10.1111/j.1538-7836.2009.03536.x

221. Li, R. & Emsley, J. "The organizing principle of the platelet glycoprotein Ib-IX-V complex." *J. Thromb. Haemost.* **11,** 605–14 (2013); DOI: 10.1111/jth.12144

222. Kahn, M. L. *et al.* "Glycoprotein V-deficient platelets have undiminished thrombin responsiveness and do not exhibit a Bernard-Soulier phenotype." *Blood* **94,** 4112–21 (1999).

223. Finch, C. N. et al. "Evidence that an abnormality in the glycoprotein Ib alpha gene is not the cause of abnormal platelet function in a family with classic Bernard-Soulier disease." *Blood* **75,** 2357–2362 (1990).

224. Wright, S. D. et al. "Double heterozygosity for mutations in the platelet glycoprotein IX gene in three siblings with Bernard-Soulier syndrome." *Blood* **81,** 2339–47 (1993).

225. Kunishima, S. *et al.* "Missense mutations of the glycoprotein (GP) Ib beta gene impairing the GPIb alpha/beta disulfide linkage in a family with giant platelet disorder." *Blood* **89,** 2404–2412 (1997).

226. Savoia, A. *et al.* “Autosomal dominant macrothrombocytopenia in Italy is most frequently a type of heterozygous Bernard-Soulier syndrome.” *Blood* **97,** 1330–1335 (2001); DOI: 10.1182/blood.V97.5.1330

227. Balduini, A. *et al.* “Proplatelet formation in heterozygous Bernard-Soulier syndrome type Bolzano.” *J. Thromb. Haemost.* **7,** 478–484 (2009); DOI: 10.1111/j.1538-7836.2008.03255.x

228. Vettore, S. *et al.* “A A386g biallelic GPIbα gene mutation with anomalous behavior: A new mechanism suggested for Bernard-Soulier syndrome pathogenesis.” *Haematologica* **96,** 1878–1882 (2011); DOI: 10.3324/haematol.2010.039008

229. Balduini, A., et al. “Megakaryocytes derived from patients with the classical form of Bernard-Soulier syndrome show no ability to extend proplatelets in vitro.” *Platelets* **22,** 308–11 (2011); DOI: 10.3109/09537104.2010.547960

230. Kanaji, T., Russell, S. & Ware, J. “Amelioration of the macrothrombocytopenia associated with the murine Bernard-Soulier syndrome.” *Blood* **100,** 2102–2107 (2002); DOI: 10.1182/blood-2002-03-0997

231. Strassel, C. *et al.* “Intrinsic impaired proplatelet formation and microtubule coil assembly of megakaryocytes in a mouse model of Bernard-Soulier syndrome.” *Haematologica* **94,** 800–810 (2009); DOI: 10.3324/haematol.2008.001032

232. Othman, M. “Platelet-type Von Willebrand disease: Three decades in the life of a rare bleeding disorder.” *Blood Rev.* **25,** 147–153 (2011); DOI: 10.1016/j.blre.2011.03.003

233. Rosa, R. F. M. *et al.* “Hematological abnormalities and 22q11.2 deletion syndrome.” *Rev. Bras. Hematol. Hemoter.* **33,** 151–154 (2011); DOI: 10.5581/1516-8484.20110037

234. Stritt, S. *et al.* “A gain-of-function variant in DIAPH1 causes dominant macrothrombocytopenia and hearing loss.” *Blood* **127,** 2903–2914 (2016); DOI: 10.1182/blood-2015-10-675629

235. Pan, J. *et al.* “The formin DIAPH1 (mDia1) regulates megakaryocyte proplatelet formation by remodeling the actin and microtubule cytoskeletons.” *Blood* **124,** 3967–3977 (2014); DOI: 10.1182/blood-2013-12-544924

236. Tadokoro, S. *et al.* “A potential role for α-actinin in inside-out αIIbβ3 signaling.” *Blood* **117,** 250–8 (2011); DOI: 10.1182/blood-2009-10-246751

237. Westbury, S. K., Shoemark, D. K. & Mumford, A. D. “ACTN1 variants associated with thrombocytopenia.” *Platelets* **28,** 625–627 (2017); DOI: 10.1080/09537104.2017.1356455

238. Wang, K., Ash, J. F. & Singer, S. J. “Filamin, a new high-molecular-weight protein found in smooth muscle and non-muscle cells.” *Proc. Natl. Acad. Sci. U. S. A.* **72,** 4483–6 (1975).

239. Parrini, E. *et al.* “Periventricular heterotopia: Phenotypic heterogeneity and correlation with Filamin a mutations.” *Brain* **129,** 1892–1906 (2006); DOI: 10.1093/brain/awl125

240. Feng, Y. & Walsh, C. A. “The many faces of filamin: A versatile molecular scaffold for cell motility and signalling.” *Nat. Cell Biol.* **6,** 1034–1038 (2004); DOI: 10.1038/ncb1104-1034

241. Zhou, A. X., Hartwig, J. H. & Akyürek, L. M. “Filamins in cell signaling, transcription and organ development.” *Trends Cell Biol.* **20,** 113–123 (2010); DOI: 10.1016/j.tcb.2009.12.001

242. Falet, H. “New insights into the versatile roles of platelet FlnA.” *Platelets* **24,** 1–5 (2013); DOI: 10.3109/09537104.2011.654004

243. Falet, H. *et al.* “A novel interaction between FlnA and Syk regulates platelet ITAM-mediated receptor signaling and function.” *J. Exp. Med.* **207,** 1967–1979 (2010); DOI: 10.1084/jem.20100222

244. Jurak Begonja, A. *et al.* “FlnA-null megakaryocytes prematurely release large and fragile platelets that circulate poorly.” *Blood* **118,** 2285–95 (2011); DOI: 10.1182/blood-2011-04-348482

245. Kanaji, T. *et al.* “GPIb regulates platelet size by controlling the subcellular localization of filamin.” *Blood* **119,** 2906–2913 (2012); DOI: 10.1182/blood-2011-08-376566

246. Nurden, P. *et al.* “Thrombocytopenia resulting from mutations in filamin A can be expressed as an isolated syndrome.” *Blood* **118,** 5928–5937 (2011); DOI: 10.1182/blood-2011-07-365601

247. Berrou, E. *et al.* “Heterogeneity of platelet functional alterations in patients with filamin a mutations.” *Arterioscler. Thromb. Vasc. Biol.* **33,** e11-8 (2013); DOI: 10.1161/ATVBAHA.112.300603

248. Chen, M. & Stracher, A. “In situ phosphorylation of platelet actin-binding protein by cAMP dependent protein kinase stabilizes it against proteolysis by calpain.” *J. Biol. Chem.* **264,** 14282–14289 (1989).

249. Manchev, V. T. *et al.* “A new form of macrothrombocytopenia induced by a germ-line mutation in the PRKACG gene.” *Blood* **124,** 2554–2563 (2014); DOI: 10.1182/blood-2014-01-551820

250. Kunishima, S. *et al.* “Mutation of the β1-tubulin gene associated with congenital macrothrombocytopenia affecting microtubule assembly.” *Blood* **113,** 458–461 (2009); DOI: 10.1182/blood-2008-06-162610

251. Davis, B. *et al.* “Mutation in β1-tubulin correlates with macrothrombocytopenia in Cavalier King Charles Spaniels.” *J. Vet. Intern. Med.* **22,** 540–545 (2008); DOI: 10.1111/j.1939-1676.2008.0085.x

252. Pleines, I. *et al.* “Mutations in tropomyosin 4 underlie a rare form of human macrothrombocytopenia.” *J. Clin. Invest.* **127,** 814–829 (2017); DOI: 10.1172/JCI86154

253. Pecci, A. & Balduini, C. L. “Lessons in platelet production from inherited thrombocytopenias.” *Br. J. Haematol.* **165,** 179–192 (2014); DOI: 10.1111/bjh.12752

254. De Los Angeles, A. *et al.* "Hallmarks of pluripotency." *Nature* **525,** 469–478 (2015); DOI: 10.1038/nature15515

255. Evans, M. J. & Kaufman, M. H. "Establishment in culture of pluripotential cells from mouse embryos." *Nature* **292,** 154–156 (1981); DOI: 10.1038/292154a0

256. Ying, Q. L. & Smith, A. "The Art of Capturing Pluripotency: Creating the Right Culture." *Stem Cell Reports* **8,** 1457–1464 (2017); DOI: 10.1016/j.stemcr.2017.05.020

257. Hackett, J. A. & Surani, M. A. "Regulatory principles of pluripotency: from the ground state up." *Cell Stem Cell* **15,** 416–430 (2014); DOI: 10.1016/j.stem.2014.09.015

258. Brons, I. G. M. *et al.* "Derivation of pluripotent epiblast stem cells from mammalian embryos." *Nature* **448,** 191–195 (2007); DOI: 10.1038/nature05950

259. Tesar, P. J. *et al.* "New cell lines from mouse epiblast share defining features with human embryonic stem cells." *Nature* **448,** 196–199 (2007); DOI: 10.1038/nature05972

260. Guo, G. *et al.* “Klf4 reverts developmentally programmed restriction of ground state pluripotency.” *Development* **136,** 1063–1069 (2009); DOI: 10.1242/dev.030957

261. Mathieu, J. & Ruohola-Baker, H. "Metabolic remodeling during the loss and acquisition of pluripotency." *Development* **144,** 541–551 (2017); DOI: 10.1242/dev.128389

262. Smith, A. "Formative pluripotency: the executive phase in a developmental continuum." *Development* **144,** 365–373 (2017); DOI: 10.1242/dev.142679

263. Thomson, J. A. *et al.* "Embryonic stem cell lines derived from human blastocysts." *Science* **282,** 1145–7 (1998)

264. Theunissen, T. W. & Jaenisch, R. “Mechanisms of gene regulation in human embryos and pluripotent stem cells.” *Development* **144,** 4496–4509 (2017); DOI: 10.1242/dev.157404

265. Dodsworth, B. T., Flynn, R. & Cowley, S. A. "The Current State of Naive Human Pluripotency." *Stem Cells* **33,** 3181–3186 (2015); DOI: 10.1002/stem.2085

266. Cowan, C. A. *et al.* “Nuclear reprogramming of somatic cells after fusion with human embryonic stem cells.” *Science* **309,** 1369–73 (2005); DOI: 10.1126/science.1116447

267. Gurdon, J. B., Elsdale, T. R. & Fischberg, M. “Sexually mature individuals of Xenopus laevis from the transplantation of single somatic nuclei” *Nature* **182**, 64-65 (1958); DOI: 10.1038/182064a0

268. Wilmut, I. *et al.* “Viable offspring derived from fetal and adult mammalian cells.” *Nature* **385,** 810–813 (1997); DOI: 10.1038/385810a0

269. Tachibana, M. *et al.* “Human embryonic stem cells derived by somatic cell nuclear transfer.” *Cell* **153,** 1228–1238 (2013); DOI: 10.1016/j.cell.2013.05.006

270. Graf, T. “Historical origins of transdifferentiation and reprogramming.” *Cell Stem Cell* **9,** 504–516 (2011); DOI: 10.1016/j.stem.2011.11.012

271. Takahashi, K. & Yamanaka, S. "Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors." *Cell* **126,** 663–76 (2006); DOI: 10.1016/j.cell.2006.07.024

272. Hochedlinger, K. & Jaenisch, R. “Induced Pluripotency and Epigenetic Reprogramming.” *Cold Spring Harb. Perspect. Biol.* **7,** a019448 (2015); DOI: 10.1101/cshperspect.a019448

273. Hou, P. *et al.* “Pluripotent Stem Cells Induced from Mouse Somatic Cells by Small-Molecule Compounds.” *Science* **341,** 651–654 (2013); DOI: 10.1126/science.1239278

274. Blanchard, J. W. *et al.* “Replacing reprogramming factors with antibodies selected from combinatorial antibody libraries.” *Nat. Biotechnol.* **35,** 960–968 (2017); DOI: 10.1038/nbt.3963

275. Hanna, J., Carey, B. W. & Jaenisch, R. “Reprogramming of somatic cell identity.” *Cold Spring Harb. Symp. Quant. Biol.* **73,** 147–55 (2008); DOI: 10.1101/sqb.2008.73.025

276. Papp, B. & Plath, K. “Epigenetics of reprogramming to induced pluripotency.” *Cell* **152,** 1324–1343 (2013); DOI: 10.1016/j.cell.2013.02.043

277. Lister, R. *et al.* “Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells.” *Nature* **471,** 68–73 (2011); DOI: 10.1038/nature09798

278. Burrows, C. K. *et al.* “Genetic Variation, Not Cell Type of Origin, Underlies the Majority of Identifiable Regulatory Differences in iPSCs.” *PLoS Genet.* **12,** e1005793 (2016); DOI: 10.1371/journal.pgen.1005793

279. Kyttälä, A. *et al.* “Genetic Variability Overrides the Impact of Parental Cell Type and Determines iPSC Differentiation Potential.” *Stem Cell Reports* **6,** 200–212 (2016); DOI: 10.1016/j.stemcr.2015.12.009

280. Hussein, S. M. *et al.* “Copy number variation and selection during reprogramming to pluripotency.” *Nature* **471,** 58–62 (2011); DOI: 10.1038/nature09871

281. Merkle, F. T. *et al.* "Human pluripotent stem cells recurrently acquire and expand dominant negative P53 mutations." *Nature* **545,** 229–233 (2017); DOI: 10.1038/nature22312

282. Kwon, E. M. *et al.* "iPSCs and fibroblast subclones from the same fibroblast population contain comparable levels of sequence variations." *Proc. Natl. Acad. Sci.* **114,** 1964–1969 (2017); DOI: 10.1073/pnas.1616035114

283. Vijg, J. “Somatic mutations, genome mosaicism, cancer and aging.” *Curr. Opin. Genet. Dev.* **26,** 141–149 (2014); DOI: 10.1016/j.gde.2014.04.002

284. Lo Sardo, V. *et al.* “Influence of donor age on induced pluripotent stem cells.” *Nat. Biotechnol.* **35,** 69–74 (2017); DOI: 10.1038/nbt.3749

285. Mekhoubad, S. *et al.* “Erosion of dosage compensation impacts human iPSC disease modeling.” *Cell Stem Cell* **10,** 595–609 (2012); DOI: 10.1016/j.stem.2012.02.014

286. Anguera, M. C. *et al.* “Molecular signatures of human induced pluripotent stem cells highlight sex differences and cancer genes.” *Cell Stem Cell* **11,** 75–90 (2012); DOI: 10.1016/j.stem.2012.03.008

287. Vallot, C. *et al.* “Erosion of X chromosome inactivation in human pluripotent cells initiates with XACT coating and depends on a specific heterochromatin landscape.” *Cell Stem Cell* **16,** 533–546 (2015); DOI: 10.1016/j.stem.2015.03.016

288. Vallot, C. *et al.* “XACT Noncoding RNA Competes with XIST in the Control of X Chromosome Activity during Human Early Development.” *Cell Stem Cell* **20,** 102–111 (2017); DOI: 10.1016/j.stem.2016.10.014

289. Amabile, G. *et al.* "In vivo generation of transplantable human hematopoietic cells from induced pluripotent stem cells." *Blood* **121,** 1255–64 (2013); DOI: 10.1182/blood-2012-06-434407

290. Kaufman, D. S. *et al.* "Hematopoietic colony-forming cells derived from human embryonic stem cells." *Proc. Natl. Acad. Sci. U. S. A.* **98,** 10716–21 (2001); DOI: 10.1073/pnas.191362598

291. Vodyanik, M. A. *et al.* "Human embryonic stem cell-derived CD34+ cells: efficient production in the coculture with OP9 stromal cells and analysis of lymphohematopoietic potential." *Blood* **105,** 617–26 (2005); DOI: 10.1182/blood-2004-04-1649

292. Ackermann, M. *et al.* "Lost in translation: pluripotent stem cell-derived hematopoiesis." *EMBO Mol. Med.* **7,** 1388–402 (2015); DOI: 10.15252/emmm.201505301

293. Dzierzak, E. & Speck, N. A. "Of lineage and legacy: The development of mammalian hematopoietic stem cells." *Nat. Immunol.* **9,** 129–136 (2008); DOI: 10.1038/ni1560

294. Chadwick, K. *et al.* "Cytokines and BMP-4 promote hematopoietic differentiation of human embryonic stem cells." *Blood* **102,** 906–15 (2003); DOI: 10.1182/blood-2003-03-0832

295. Kennedy, M. *et al.* "Development of the hemangioblast defines the onset of hematopoiesis in human ES cell differentiation cultures." *Blood* **109,** 2679–87 (2007); DOI: 10.1182/blood-2006-09

296. Pick, M., *et al.* "Differentiation of human embryonic stem cells in serum-free medium reveals distinct roles for bone morphogenetic protein 4, vascular endothelial growth factor, stem cell factor, and fibroblast growth factor 2 in hematopoiesis." *Stem Cells* **25,** 2206–14 (2007); DOI: 10.1634/stemcells.2006-0713

297. Kennedy, M. *et al.* "T lymphocyte potential marks the emergence of definitive hematopoietic progenitors in human pluripotent stem cell differentiation cultures." *Cell Rep.* **2,** 1722–35 (2012); DOI: 10.1016/j.celrep.2012.11.003

298. Kinder, S. J. *et al.* "The orderly allocation of mesodermal cells to the extraembryonic structures and the anteroposterior axis during gastrulation of the mouse embryo." *Development* **126,** 4691–4701 (1999); DOI: 10.1002/aja.1002030109

299. Sturgeon, C. M. *et al.* "Wnt signaling controls the specification of definitive and primitive hematopoiesis from human pluripotent stem cells." *Nat. Biotechnol.* **32,** 554–561 (2014); DOI: 10.1038/nbt.2915

300. Dou, D. R. *et al.* "Medial HOXA genes demarcate haematopoietic stem cell fate during human development." *Nat. Cell Biol.* **18,** 595–606 (2016); DOI: 10.1038/ncb3354

301. Ng, E. S. *et al.* "Differentiation of human embryonic stem cells to HOXA + hemogenic vasculature that resembles the aorta-gonad-mesonephros." *Nat. Biotechnol.* **34,** 1168–1179 (2016); DOI: 10.1038/nbt.3702

302. Creamer, J. P. *et al.* "Human definitive hematopoietic specification from pluripotent stem cells is regulated by mesodermal expression of CDX4." *Blood* **129,** 2988–2992 (2017); DOI: 10.1182/blood-2016-11-749382

303. Yoshimoto, M. *et al.* "Embryonic day 9 yolk sac and intra-embryonic hemogenic endothelium independently generate a B-1 and marginal zone progenitor lacking B-2 potential." *Proc. Natl. Acad. Sci. U. S. A.* **108,** 1468–1473 (2011); DOI: 10.1073/pnas.1015841108

304. McGrath, K. E. *et al.* "Distinct Sources of Hematopoietic Progenitors Emerge before HSCs and Provide Functional Blood Cells in the Mammalian Embryo." *Cell Rep.* **11,** 1892–1904 (2015); DOI: 10.1016/j.celrep.2015.05.036

305. Swiers, G. *et al.* "Early dynamic fate changes in haemogenic endothelium characterized at the single-cell level." *Nat. Commun.* **4,** 2924 (2013); DOI: 10.1038/ncomms3924

306. Ditadi, A. *et al.* "Human definitive haemogenic endothelium and arterial vascular endothelium represent distinct lineages." *Nat. Cell Biol.* **17,** 580–91 (2015); DOI: 10.1038/ncb3161

307. Slukvin, I. I. "Generating human hematopoietic stem cells in vitro–exploring endothelial to hematopoietic transition as a portal for stemness acquisition." *FEBS Lett.* **590,** 4126–4143 (2016); DOI: 10.1002/1873-3468.12283

308. Hadland, B. K. *et al.* "A Common Origin for B-1a and B-2 Lymphocytes in Clonal Pre-Hematopoietic Stem Cells." *Stem Cell Reports* **8,** 1563–1572 (2017); DOI: 10.1016/j.stemcr.2017.04.007

309. Espín-Palazón, R. *et al.* "Proinflammatory Signaling Regulates Hematopoietic Stem Cell Emergence." *Cell* **159,** 1070–1085 (2014); DOI: 10.1016/j.cell.2014.10.031

310. Monteiro, R. *et al.* "Transforming Growth Factor β Drives Hemogenic Endothelium Programming and the Transition to Hematopoietic Stem Cells." *Dev. Cell* **38,** 358–370 (2016); DOI: 10.1016/j.devcel.2016.06.024

311. Szabo, E. *et al.* "Direct conversion of human fibroblasts to multilineage blood progenitors." *Nature* **468,** 521–526 (2010); DOI: 10.1038/nature09591

312. Pulecio, J. *et al.* "Conversion of Human Fibroblasts Into Monocyte-Like Progenitor Cells." *Stem Cells* **32,** 2923–2938 (2014); DOI: 10.1002/stem.1800

313. Pereira, C. F. *et al.* "Induction of a hemogenic program in mouse fibroblasts." *Cell Stem Cell* **13,** 205–218 (2013); DOI: 10.1016/j.stem.2013.05.024

314. Batta, K. *et al.* "Direct Reprogramming of Murine Fibroblasts to Hematopoietic Progenitor Cells." *Cell Rep.* **9,** 1871–1885 (2014); DOI: 10.1016/j.celrep.2014.11.002

315. Sandler, V. M. *et al.* "Reprogramming human endothelial cells to haematopoietic cells requires vascular induction." *Nature* **511,** 312–8 (2014); 10.1038/nature13547

316. Lis, R. *et al.* "Conversion of adult endothelium to immunocompetent haematopoietic stem cells." *Nature* **545,** 439–445 (2017); DOI: 10.1038/nature22326

317. Riddell, J. *et al.* "Reprogramming committed murine blood cells to induced hematopoietic stem cells with defined factors." *Cell* **157,** 549–564 (2014); DOI: 10.1016/j.cell.2014.04.006

318. Eich, C. *et al.* "In vivo single cell analysis reveals Gata2 dynamics in cells transitioning to hematopoietic fate." *J. Exp. Med.* **215,** 233–248 (2018); DOI: 10.1084/jem.20170807

319. Adamo, L. *et al.* "Biomechanical forces promote embryonic haematopoiesis." *Nature* **459,** 1131–1135 (2009); DOI: 10.1038/nature08073

320. North, T. E. *et al.* "Hematopoietic Stem Cell Development Is Dependent on Blood Flow." *Cell* **137,** 736–748 (2009); DOI: 10.1016/j.cell.2009.04.023

321. Kwan, W. *et al.* "The Central Nervous System Regulates Embryonic HSPC Production via Stress- Responsive Glucocorticoid Receptor Signaling" *Cell Stem Cell* **19,** 370–82 (2016); DOI: 10.1016/j.stem.2016.06.004

322. Kilpinen, H. *et al.* "Common genetic variation drives molecular heterogeneity in human iPSCs." *Nature* **546,** 370–375 (2017); DOI: 10.1038/nature22403

323. Hendriks, W. T., Warren, C. R. & Cowan, C. A. "Genome Editing in Human Pluripotent Stem Cells: Approaches, Pitfalls, and Solutions." *Cell Stem Cell* **18,** 53–65 (2016); DOI: 10.1016/j.stem.2015.12.002

324. Mills, J. a *et al.* "Clonal genetic and hematopoietic heterogeneity among human-induced pluripotent stem cell lines." *Blood* **122,** 2047–51 (2013); DOI: 10.1182/blood-2013-02-484444

325. Müller, L.U. *et al.* "Overcoming reprogramming resistance of Fanconi anemia cells." *Blood* **119**, 5449–5457 (2012); DOI: 10.1182/blood-2012-02-408674

326. Young, M. A. *et al.* "Background mutations in parental cells account for most of the genetic heterogeneity of induced pluripotent stem cells." *Cell Stem Cell* **10,** 570–582 (2012); DOI: 10.1016/j.stem.2012.03.002

327. Saliba, J. *et al.* "Germline duplication of ATG2B and GSKIP predisposes to familial myeloid malignancies." *Nat. Genet.* **47,** 1131–1140 (2015); DOI: 10.1038/ng.3380

328. Kotini, A. G. *et al.* "Functional analysis of a chromosomal deletion associated with myelodysplastic syndromes using isogenic human induced pluripotent stem cells." *Nat. Biotechnol.* **33,** 646–655 (2015); DOI: 10.1038/nbt.3178

329. Chao, M. P. *et al.* Human AML-iPSCs Reacquire Leukemic Properties after Differentiation and Model Clonal Variation of Disease. *Cell Stem Cell* **20,** 329–344 (2017); DOI: 10.1016/j.stem.2016.11.018

330. Kotini, A. G. *et al.* "Stage-Specific Human Induced Pluripotent Stem Cells Map the Progression of Myeloid Transformation to Transplantable Leukemia." *Cell Stem Cell* **20,** 315–328 (2017); DOI: 10.1016/j.stem.2017.01.009

331. Papapetrou, E. P. "Patient-derived induced pluripotent stem cells in cancer research and precision oncology." *Nat. Med.* **22,** 1392–1401 (2016); DOI: 10.1038/nm.4238

332. Reinisch, A. *et al.* "A humanized bone marrow ossicle xenotransplantation model enables improved engraftment of healthy and leukemic human hematopoietic cells. "*Nat. Med.* **22,** 812–821 (2016); DOI: 10.1038/nm.4103

333. Rongvaux, A. *et al.* "Development and function of human innate immune cells in a humanized mouse model." *Nat. Biotechnol.* **32,** 364–72 (2014); DOI: 10.1038/nbt.2858

334. Yurino, A. *et al.* "Enhanced Reconstitution of Human Erythropoiesis and Thrombopoiesis in an Immunodeficient Mouse Model with KitWvMutations." *Stem Cell Reports* **7,** 425–438 (2016); DOI: 10.1016/j.stemcr.2016.07.002

335. Yoshimi, A. *et al.* "Robust patient-derived xenografts of MDS/MPN overlap syndromes capture the unique characteristics of CMML and JMML." *Blood* **130,** 397–407 (2017); DOI: 10.1182/blood-2017-01-763219

336. Zhang, Y. *et al.* "Engraftment of chronic myelomonocytic leukemia cells in immunocompromised mice supports disease dependency on cytokines." *Blood Adv.* **1,** 972–979 (2017); DOI: 10.1182/bloodadvances.2017004903

337. Garçon, L. *et al.* "Ribosomal and hematopoietic defects in induced pluripotent stem cells derived from Diamond Blackfan anemia patients." *Blood* **122,** 912–21 (2013); DOI: 10.1182/blood-2013-01-478321

338. Nayak, R. C. *et al.* "Pathogenesis of ELANE-mutant severe neutropenia revealed by induced pluripotent stem cells." *J. Clin. Invest.* **125,** 3103–3116 (2015); DOI: 10.1172/JCI80924

339. Böiers, C. *et al.* "A Human IPS Model Implicates Embryonic B-Myeloid Fate Restriction as Developmental Susceptibility to B Acute Lymphoblastic Leukemia-Associated ETV6-RUNX1." *Dev. Cell* **44,** 362–377.e7 (2018); DOI: 10.1016/j.devcel.2017.12.005

340. Elcheva, I. *et al.* "Direct induction of haematoendothelial programs in human pluripotent stem cells by transcriptional regulators." *Nat. Commun.* **5,** 1–11 (2014); DOI: 10.1038/ncomms5372

341. Doulatov, S. *et al.* Induction of Multipotential Hematopoietic Progenitors from Human Pluripotent Stem Cells via Respecification of Lineage-Restricted Precursors. *Cell Stem Cell* **13,** 459–470 (2013); DOI: 10.1016/j.stem.2013.09.002

342. Sugimura, R. *et al.* "Haematopoietic stem and progenitor cells from human pluripotent stem cells." *Nature* **545,** 432–438 (2017); DOI: 10.1038/nature22370

343. Hamey, F. K. & Göttgens, B. "Demystifying blood stem cell fates." *Nat. Cell Biol.* **19,** 261–263 (2017); DOI: 10.1038/ncb3494

344. Ebina, W. & Rossi, D. J. "Transcription factor-mediated reprogramming toward hematopoietic stem cells." *EMBO J.* **34,** 694–709 (2015); DOI: 10.15252/embj.201490804

345. Cherry, A. B. C. & Daley, G. Q. "Reprogrammed Cells for Disease Modeling and Regenerative Medicine." *Annu. Rev. Med.* **64,** 277–290 (2013); DOI: 10.1146/annurev-med-050311-163324

346. Rideout, W. M. *et al.* "Correction of a Genetic Defect by Nuclear Transplantation and Combined Cell and Gene Therapy. *Cell* **109,** 17–27 (2002); DOI: 10.1016/S0092-8674(02)00681-5

347. Shi, Y. *et al.* "Induced pluripotent stem cell technology: A decade of progress." *Nat. Rev. Drug Discov.* **16,** 115–130 (2017); DOI: 10.1038/nrd.2016.245

348. Tidball, A. M. *et al.* "Rapid Generation of Human Genetic Loss-of-Function iPSC Lines by Simultaneous Reprogramming and Gene Editing." *Stem Cell Reports* **9,** 725–731 (2017); DOI: 10.1016/j.stemcr.2017.07.003

349. Zhu, H. *et al.* "Human Pluripotent Stem Cells to Produce Cell-Based Cancer Immunotherapy. *Stem Cells* (2018); DOI:10.1002/stem.2754

350. Themeli, M. *et al.* "Generation of tumor-targeted human T lymphocytes from induced pluripotent stem cells for cancer therapy." *Nat. Biotechnol.* **31,** 928–933 (2013); DOI: 10.1038/nbt.2678

351. Vizcardo, R. *et al.* "Regeneration of human tumor antigen-specific T cells from iPSCs derived from mature CD8+ T cells." *Cell Stem Cell* **12,** 31–36 (2013); DOI: 10.1016/j.stem.2012.12.006

352. Lachmann, N. *et al.* "Large-scale hematopoietic differentiation of human induced pluripotent stem cells provides granulocytes or macrophages for cell replacement therapies." *Stem cell reports* **4,** 282–96 (2015); DOI: 10.1016/j.stemcr.2015.01.005

353. Rousseau, G. F., Giarratana, M. C. & Douay, L. "Large-scale production of red blood cells from stem cells: What are the technical challenges ahead?" *Biotechnol. J.* **9,** 28–38 (2014); DOI: 10.1002/biot.201200368

354. Nakamura, S. *et al.* "Expandable megakaryocyte cell lines enable clinically applicable generation of platelets from human induced pluripotent stem cells." *Cell Stem Cell* **14,** 535–48 (2014); DOI: 10.1016/j.stem.2014.01.011

355. Moreau, T. *et al.* "Large-scale production of megakaryocytes from human pluripotent stem cells by chemically defined forward programming." *Nat. Commun.* **7,** 11208 (2016); DOI: 10.1038/ncomms11208

356. Doulatov, S. *et al.* "Drug discovery for Diamond-Blackfan anemia using reprogrammed hematopoietic progenitors." *Sci. Transl. Med.* **9,** eaah5645 (2017); DOI: 10.1126/scitranslmed.aah5645

357. Burridge, P. W. *et al.* "Human induced pluripotent stem cell-derived cardiomyocytes recapitulate the predilection of breast cancer patients to doxorubicin-induced cardiotoxicity." *Nat. Med.* **22,** 547–556 (2016); DOI: 10.1038/nm.4087

358. Chen, M. H. *et al.* "Thoracic aortic aneurysm in patients with loss of function Filamin A mutations: Clinical characterization, genetics, and recommendations." *Am. J. Med. Genet. Part A* **176,** 337–350 (2018); DOI: 10.1002/ajmg.a.38580

359. Chou, S. T. *et al.* "Trisomy 21-associated defects in human primitive hematopoiesis revealed through induced pluripotent stem cells." *Proc. Natl. Acad. Sci.* **109,** 17573–17578 (2012); DOI: 10.1073/pnas.1211175109

360. Engert, A. *et al.* "The european hematology association roadmap for european hematology research: A consensus document." *Haematologica* **101,** 115–208 (2016); DOI: 10.3324/haematol.2015.136739

361. Saliba, J. *et al.* "Heterozygous and Homozygous JAK2(V617F) States Modeled by Induced Pluripotent Stem Cells from Myeloproliferative Neoplasm Patients." *PLoS One* **8,** e74257 (2013); DOI: 10.1371/journal.pone.0074257

362. Bluteau, O. *et al.* "Developmental changes in human megakaryopoiesis." *J. Thromb. Haemost.* **11,** 1730–1741 (2013); DOI: 10.1111/jth.12326

363. Lefrançais, E. *et al.* "The lung is a site of platelet biogenesis and a reservoir for haematopoietic progenitors." *Nature* **544,** 105–109 (2017); DOI: 10.1038/nature21706

364. Barberi, T. *et al.* "Derivation of multipotent mesenchymal precursors from human embryonic stem cells." *PLoS Med.* **2,** 0554–0560 (2005); DOI: 10.1371/journal.pmed.0020161

365. Di Buduo, C. A. *et al.* "Programmable 3D silk bone marrow niche for platelet generation ex vivo and modeling of megakaryopoiesis pathologies." *Blood* **125,** 2254–2264 (2015); DOI: 10.1182/blood-2014-08-595561

366. Carabalona, A. *et al.* "A glial origin for periventricular nodular heterotopia caused by impaired expression of Filamin-A." *Hum. Mol. Genet.* **21,** 1004–1017 (2012); DOI: 10.1093/hmg/ddr531

367. Lancaster, M. A. *et al.* "Cerebral organoids model human brain development and microcephaly." *Nature* **501,** 373–379 (2013); DOI: 10.1038/nature12517

368. Elagib, K. E. *et al.* "Neonatal expression of RNA-binding protein IGF2BP3 regulates the human fetal-adult megakaryocyte transition." *J. Clin. Invest.* **127,** 2365–2377 (2017); DOI: 10.1172/JCI88936

369. Nakamura, F. *et al.* "Structural basis of filamin A functions." *J. Cell Biol.* **179,** 1011–1025 (2007); DOI: 10.1083/jcb.200707073

370. Nakamura, F. *et al.* "The structure of the GPIb-filamin A complex." *Blood* **107,** 1925–1932 (2006); DOI: 10.1182/blood-2005-10-3964

371. Williamson, D. *et al.* "Interaction between platelet glycoprotein Ibα and filamin-1 is essential for glycoprotein Ib/IX receptor anchorage at high shear." *J. Biol. Chem.* **277,** 2151–2159 (2002); DOI: 10.1074/jbc.M109384200

372. Savoia, A. *et al.* "Spectrum of the mutations in bernard-soulier syndrome." *Hum. Mutat.* **35,** 1033–1045 (2014); DOI: 10.1002/humu.22607

373. Ridley, A. J. & Hall, A. "The small GTP-binding protein rho regulates the assembly of focal adhesions and actin stress fibers in response to growth factors." *Cell* **70,** 389–399 (1992); DOI: 10.1016/0092-8674(92)90163-7

374. Sabri, S. *et al.* "Differential regulation of actin stress fiber assembly and proplatelet formation by α2β1 integrin and GPVI in human megakaryocytes." *Blood* **104,** 3117–3125 (2004); DOI: 10.1182/blood-2003-12-4398

375. Larson, M. K. & Watson, S. P. "Regulation of proplatelet formation and platelet release by integrin αIIbβ3." *Blood* **108,** 1509–1514 (2006); DOI: 10.1182/blood-2005-11-011957

376. Kiema, T. *et al.* "The molecular basis of filamin binding to integrins and competition with talin." *Mol. Cell* **21,** 337–347 (2006); DOI: 10.1016/j.molcel.2006.01.011

377. Ohta, Y. *et al.* “The small GTPase RalA targets filamin to induce filopodia.” *Proc. Natl. Acad. Sci. U. S. A.* **96,** 2122–8 (1999); DOI: 10.1073/PNAS.96.5.2122

378. Ohta, Y., Hartwig, J. H. & Stossel, T. P. “FilGAP, a Rho- and ROCK-regulated GAP for Rac binds filamin A to control actin remodelling.” *Nat. Cell Biol.* **8,** 803–814 (2006); DOI: 10.1038/ncb1437

379. Moser, M. et al. “Kindlin-3 is essential for integrin activation and platelet aggregation.” *Nat. Med.* **14,** 325–330 (2008); DOI: 10.1038/nm1722

380. Favier, M. *et al.* “Mutations of the integrin αIIb/β3 intracytoplasmic salt bridge cause macrothrombocytopenia and enlarged platelet α-granules.” *Am. J. Hematol.* **93,** 195–204 (2018); DOI: 10.1002/ajh.24958

381. Kircher, P. *et al.* “Filamin A interacts with the coactivator MKL1 to promote the activity of the transcription factor SRF and cell migration.” *Sci. Signal.* **8,** ra112 (2015); DOI: 10.1126/scisignal.aad2959

382. Zaninetti, C. *et al.* “Inherited thrombocytopenia caused by ANKRD26 mutations misdiagnosed and treated as myelodysplastic syndrome: report on two cases.” *J. Thromb. Haemost.* **15,** 2388–2392 (2017); DOI: 10.1111/jth.13855

383. Godley, L. A. & Shimamura, A. "Genetic predisposition to hematologic malignancies: management and surveillance." *Blood* **130,** 424–432 (2017); DOI: 10.1182/blood-2017-02-735290

384. Antony-Debré, I. *et al.* “MYH10 protein expression in platelets as a biomarker of RUNX1 and FLI1 alterations.” *Blood* **120,** 2719–22 (2012); DOI: 10.1182/blood-2012-04-422352

385. Bera, T. K. *et al.* “A model for obesity and gigantism due to disruption of the Ankrd26 gene.” *Proc. Natl. Acad. Sci. U. S. A.* **105,** 270–5 (2008); DOI: 10.1073/pnas.0710978105

386. Fei, Z. *et al.* “Ankrd26 gene disruption enhances adipogenesis of mouse embryonic fibroblasts.” *J. Biol. Chem.* **286,** 27761–8 (2011); DOI: 10.1074/jbc.M111.248435

387. Scott, E. W. *et al.* “Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages.” *Science* **265,** 1573–1577 (1994).

388. Hosokawa, H. *et al.* “Transcription Factor PU.1 Represses and Activates Gene Expression in Early T Cells by Redirecting Partner Transcription Factor Binding.” *Immunity* **48,** 1119–1134.e7 (2018); DOI: 10.1016/j.immuni.2018.04.024

389. Maxson, J. E. & Tyner, J. W. “Genomics of chronic neutrophilic leukemia.” *Blood* **129,** 715–722 (2017); DOI: 10.1182/blood-2016-10-695981

390. Plo, I. *et al.* “An activating mutation in the CSF3R gene induces a hereditary chronic neutrophilia.” *J. Exp. Med.* **206,** 1701–7 (2009); DOI: 10.1084/jem.20090693

391. Skokowa, J. *et al.* “Severe congenital neutropenias.” *Nat. Rev. Dis. Prim.* **3,** 17032 (2017); DOI: 10.1038/nrdp.2017.32

392. Komatsu, N. *et al.* “Establishment and characterization of a human leukemic cell line with megakaryocytic features: dependency on granulocyte-macrophage colony-stimulating factor, interleukin 3, or erythropoietin for growth and survival.” *Cancer Res.* **51,** 341–348 (1991)

393. Touw, I. P., Palande, K. & Beekman, R. “Granulocyte Colony-Stimulating Factor Receptor Signaling.” *Hematol. Oncol. Clin. North Am.* **27,** 61–73 (2013); DOI: 10.1016/j.hoc.2012.10.002

394. Wölfler, A. *et al.* Site-specific ubiquitination determines lysosomal sorting and signal attenuation of the granulocyte colony-stimulating factor receptor. *Traffic* **10,** 1168–1179 (2009); DOI: 10.1111/j.1600-0854.2009.00928.x

395. Lowman, X. H. *et al.* “The Proapoptotic Function of Noxa in Human Leukemia Cells Is Regulated by the Kinase Cdk5 and by Glucose.” *Mol. Cell* **40,** 823–833 (2010); DOI: 10.1016/j.molcel.2010.11.035

ANNEX

Balduini, A., Raslova, H., Di Buduo, C.A., **Donada A.**, Ballmaier M., Germeshausen M., Balduini C.L. "Clinic, pathogenic mechanisms and drug testing of two inherited thrombocytopenias, ANKRD26 -related Thrombocytopenia and MYH9 -related diseases." *Eur. J. Med. Genet.* (2018). doi:10.1016/j.ejmg.2018.01.014