

Université de Montréal

**Caractérisation moléculaire d'un récent modèle
d'étude de la leucémie myéloïde aigüe à caryotype normal :
la lignée cellulaire CG-SH**

par

Géraldine Gosse

Programme de biologie moléculaire

Faculté de médecine

Mémoire présenté à la Faculté de médecine en vue de
l'obtention du grade de Maîtrise (M. SC.) en
biologie moléculaire option générale

Juillet 2013

© Géraldine Gosse, 2013

RÉSUMÉ

La leucémie myéloïde aigüe (LMA) est la forme de leucémie la plus fréquente chez l'adulte au Canada. Bien que de nombreux réarrangements chromosomiques récurrents aient été identifiés chez les patients LMA, près de la moitié des cas présentent un caryotype normal (LMA-CN). L'étude de la LMA-CN *in vitro* est rendue difficile par le fait que la survie des cellules primaires de patients est déficiente sur le long terme et que les lignées cellulaires leucémiques ont un caryotype hautement anormal. En 2009, Munker et son équipe ont établi une nouvelle lignée cellulaire, CG-SH, ayant la particularité d'avoir un caryotype normal. L'objectif principal de ce projet d'étude est de caractériser plus en détail ce nouveau modèle d'étude. Nous avons identifié l'ensemble des variants génétiques présents dans CG-SH grâce au séquençage du génome entier. Les variants susceptibles de participer à la leucémogénèse ont été isolés, tels que des insertions détectées dans *EZH2* et *GATA2*, et de nombreux variants faux-sens détectés dans des gènes pertinents pour la LMA. Nous avons montré que les cellules CG-SH sont sensibles à l'effet prolifératif d'une combinaison de cytokines, qui agissent sur le comportement des cellules en modifiant l'expression des gènes associés à la régulation de la prolifération, de l'apoptose et de la différenciation. De plus, les cytokines diminuent le taux de nécrose des cellules en culture sur le court terme. La présente étude a permis d'approfondir notre connaissance sur les caractéristiques moléculaires de la lignée cellulaire CG-SH, un nouveau modèle d'étude *in vitro* de la LMA-CN.

Mots-clés : Leucémie myéloïde aigüe, Lignée cellulaire leucémique, Séquençage nouvelle génération, Cytokines

ABSTRACT

Acute myeloid leukemia (AML) is the most frequent form of leukemia in the adult population in Canada. Although many recurrent chromosomal rearrangements have been identified in AML, almost half of all adult patients will present with a normal karyotype (NK-AML). The *in vitro* study of NK-AML is difficult because the long-term survival of primary patient samples is deficient and AML cell lines have a highly abnormal karyotype. In 2009, Munker and colleagues established a new cell line, CG-SH, with the advantage of having a normal karyotype. The main goal of this research project is to further characterize this new model system. We identified all the genetic variants present in the CG-SH cells using whole genome sequencing. We also isolated the variants that are susceptible to participate to leukemogenesis, including the insertions detected in *EZH2* and *GATA2*, and several missense mutations occurring in relevant genes for AML. We found that a combination of cytokines promotes the proliferation of CG-SH cells, and that cytokines act on the cells behavior through expression changes of the genes involved in the regulation of proliferation, apoptosis and differentiation. Moreover, cytokines trigger a decrease of the necrotic rate of CG-SH on the short-term. The current study allowed us to better appreciate molecular characteristics of the CG-SH cell line, a new model to study NK-AML *in vitro*.

Keywords: Acute myeloid leukemia, Leukemic cell line, Next generation sequencing, Cytokines

TABLE DES MATIÈRES

RÉSUMÉ	ii
ABSTRACT	iii
TABLE DES MATIÈRES.....	iv
LISTE DES TABLEAUX.....	vii
LISTE DES FIGURES.....	viii
LISTE DES ABRÉVIATIONS.....	x
REMERCIEMENTS.....	xv
CHAPITRE 1: INTRODUCTION	1
1. Le système hématopoïétique :.....	1
1.1. Définition de l'hématopoïèse :.....	1
1.2. Régulation de l'hématopoïèse :.....	4
1.2.1. Les facteurs de transcription :.....	4
1.2.2. Les mécanismes épigénétiques :.....	5
1.2.2.1. Modifications de la chromatine :.....	5
1.2.2.2. Méthylation de l'ADN :.....	6
1.2.3. Les niches hématopoïétiques et les signaux extrinsèques :.....	6
2. La leucémie myéloïde aigüe :.....	7
2.1. Incidence:.....	7
2.2. Symptômes et caractéristiques :.....	8
2.3. Diagnostic et traitement :.....	9
2.4. Classification franco-américano-britannique (FAB) des LMA :.....	10
2.5. Classification de l'Organisation Mondiale de la Santé (OMS) des LMA :.....	11
2.6. Signatures moléculaires de la LMA et facteurs pronostiques :.....	12
2.6.1. Les mutations :.....	12

2.6.1.1.	Les anomalies chromosomiques :.....	12
2.6.1.2.	Les mutations génétiques moléculaires :.....	13
2.6.2.	L'expression génique :.....	13
3.	Approches global visant à étudier la LMA:	15
3.1.	Puces à ADN :.....	15
3.2.	Séquençage à haut débit :	17
4.	Modèles d'étude de la leucémie myéloïde aigüe :	22
4.1.	Échantillons primaires de patients pour l'étude <i>in vitro</i> :.....	22
4.2.	Lignées cellulaires leucémiques pour l'étude <i>in vitro</i> :	22
4.3.	Modèles murin <i>in vivo</i> :	23
5.	Problématique :.....	24
6.	Hypothèses et objectifs :.....	25
CHAPITRE 2 : APPROCHES EXPÉRIMENTALES.....		27
<i>Culture Cellulaire</i>		27
<i>Séquençage nouvelle génération de l'ARN et de l'ADN génomique (ADNg)</i>		27
<i>Analyse des données de séquençage</i>		28
<i>Suivi de l'activité GATA2 avec un rapporteur luciférase</i>		29
<i>Suivi des changements de l'expression des gènes dans le temps</i>		29
<i>Analyse de la mort cellulaire</i>		30
CHAPITRE 3: ARTICLE		31
1.	Contributions:.....	31
2.	L'article:	33
	Abstract	34
	Introduction.....	35
	Material and methods	37
	Results	40
	Discussion	49
	Acknowledgements:	54

Authorship contributions:.....	54
Disclosure of conflicts of interest:	54
Figure legends:.....	55
Figures	58
Tables	66
Supplementary figures.....	69
Supplementary tables	71
CHAPITRE 4: DISCUSSION	77
CHAPITRE 5 : PERSPECTIVES	82
CHAPITRE 6 : CONCLUSION.....	83
CHAPITRE 7: BIBLIOGRAPHIE	84

LISTE DES TABLEAUX

CHAPITRE 1 : INTRODUCTION

Tableau 1.1: Les types de cellules sanguines et leurs fonctions	3
Tableau 1.2: Facteurs pronostiques défavorables pour les patients avec la LMA.	10
Tableau 1.3: Effets de la surexpression des gènes sur le risque clinique.	14

CHAPITRE 2 : ARTICLE

Table 2.1: Genes recurrently mutated in AML carrying damaging annotated SNPs in the CG-SH cells	66
Table 2.2: Genes with novel damaging variants in the CG-SH cells reported to be implicated in AML and cancer in the literature	67
Table 2.3: Mutated genes in the CG-SH cells are associated with biological functions involved in leukemia development	68
Supplementary table 2.1: Identification of novel variants in the CG-SH cells in genes unknown to be frequently mutated in NK-AML	71
Supplementary table 2.2: GO-terms involving myeloid differentiation in the functional analysis of up-regulated genes with cytokines	74
Supplementary table 2.3 GO-terms involving cell proliferation in the functional analysis of up-regulated genes with cytokines	75
Supplementary table 2.4 GO-terms involving myeloid differentiation in the functional analysis of down-regulated genes with cytokines	76

LISTE DES FIGURES

CHAPITRE 1 : INTRODUCTION

Figure1.1: L'hématopoïèse et les facteurs de transcription impliqués	2
Figure1.2: Auto-renouvellement et survie des CSH.	4
Figure1.3: Comparaison entre les puces à oligonucléotides et les puces à polymorphisme nucléotidique	17
Figure1.4: Étapes successives du séquençage traditionnel versus le séquençage de seconde-génération	21

CHAPITRE 2 : ARTICLE

Figure2.1: Single base pair insertion in <i>EZH2</i> and loss of the SET domain	58
Figure2.2: Duplication/insertion in <i>GATA2</i> and loss of the protein end	59
Figure2.3: The activity of <i>GATA2</i> is altered in the CG-SH cells	60
Figure2.4: Analysis of expression levels in the CG-SH cells of the genes recurrently mutated in AML	61
Figure2.5: CG-SH cells grow faster in the presence of cytokines	62
Figure2.6: Differential expression in the CG-SH cells cultured with and without cytokines	63
Figure2.7: Global expression profiles of the CG-SH cells versus other AML cell lines and patient samples	64
Figure2.8: Flow cytometry analysis of cytokines effects on the CG-SH cells death.	65
Supplementary figure2.1: Time course analysis of differentially expressed genes in the presence of cytokines.....	69

Supplementary figure2.2: Time course analysis of up-regulated genes involved in
negative regulation of apoptosis.....70

LISTE DES ABRÉVIATIONS

A = Adénine

ADN = Acide désoxyribonucléique

ADNc = Acide désoxyribonucléique complémentaire

ADNg = Acide désoxyribonucléique génomique

AF10 = ALL1-fused gene from chromosome 10 protein

AF6 = ALL1-fused gene from chromosome 6 protein

AF9 = ALL1-fused gene from chromosome 9 protein

AGM = Aire aorte-gonade-mésonephros

AML = Acute myeloid leukemia

ARN = Acide ribonucléique

ASE = Allele specific expression

ASXL1 = Additional sex combs like 1 (Drosophila)

BMI1 = B lymphoma Mo-MLV insertion region 1 homolog

C = Cytosine

CAN = Calcineurin inhibitor (Cain)

CBF = Core binding factor

CBF β = Core binding factor, beta subunit

CBP = C-AMP Response Element-binding protein (CREB)-binding protein

cDNA = Complementary deoxyribonucleic acid

CEBPA = CCAAT enhancer-binding protein- α

ChIP-seq = Chromatine immunoprecipitation-sequencing

CLP = common lymphoid progenitor

CMP = common myeloid progenitor

COSMIC = Catalogue Of Somatic Mutations In Cancer

CSE = Cellule souche embryonnaire

CSH = Cellule souche hématopoïétique

CXCL12 = Chemokine (C-X-C motif) ligand 12

ddNTP = Didésoxynucléotides

DEK = DEK oncogene

DNA = Deoxyribonucleic acid

DNMT = Deoxyribonucleic acid methyltransferase

DNMT1 = Deoxyribonucleic acid (cytosine-5-)-methyltransferase 1

DNMT3a = Deoxyribonucleic acid (cytosine-5-)-methyltransferase 3 alpha

ELL = Elongation factor RNA polymerase II

ESC = embryonic stem cells

ETV6 = Ets variant 6

EZH2 = enhancer of zeste homolog 2

FAB = Classification franco-américano-britannique

FACS = Fluorescence-activated cell sorting

FBS = Fetal bovin serum

FISH = Fluorescence in situ hybridization

FITC = fluorescein isothiocyanate

FL = Fms-related tyrosine kinase 3 ligand

FLT3 = Fms-related tyrosine kinase 3

G = Guanine

GATA1 = GATA binding protein 1, ou globin transcription factor 1

GATA2 = GATA binding protein 2

G-CSF = Granulocyte-colony stimulating factor

gDNA = Genomic Deoxyribonucleic acid

GMP = granulocyte/macrophage progenitor

GO = Gene Ontology

HKMT = Histones lysine méthyl transférases

HLA = Human leukocyte antigen

HOXA9 = Homeobox A9

HOXB4 = Homeobox B4

HSC = hematopoietic stem cells

IDH1 = Isocitrate dehydrogenase 1 (NADP+), soluble

IDH2 = Isocitrate dehydrogenase 2 (NADP+), mitochondrial

IL3 = Interleukin-3

IL6 = Interleukin-6

IMDM = Iscove's Modified Dulbecco's Medium

ITD = Internal tandem duplication

JAK2 = Janus kinase 2

LMA = Leucémie myéloïde aigüe

LMA-CN = Leucémie myéloïde aigüe à caryotype normal

LMO2 = LIM domain only 2 (rhombotin-like 1)

LT-CSH = Cellules souches hématopoïétiques capables de reconstituer l'hématopoïèse au long terme

LT-HSC = long-term hematopoietic stem cell

MEIS1 = Meis homeobox 1

MEP = megakaryocyte/erythroid progenitor

MLL = Mixed lineage-leukemia

MLLT1 = Mixed-lineage leukemia; translocated to, 1

MOZ = Monocytic leukemia zinc-finger protein

MPO = Myeloperoxidase

MYH11 = Myosin, heavy chain 11, smooth muscle

NK-AML = Normal karyotype acute myeloid leukemia

NPM1 = Nucleophosmin 1

NRAS = Neuroblastoma RAS viral (v-ras) oncogene homolog

OMS = Organisation mondiale de la santé

PCR = Polymerase chain reaction

PEG = Profil d'expression génique

PI = Propidium iodide

PML = Promyelocytic leukemia

PRC1 = Polycomb-group repressive complex 1

PRC2 = Polycomb-group repressive complex 2

PTEN = phosphatase and tensin homolog

RARA = Retinoic acid receptor, alpha

RBCs = red blood cells

RLU = relative light units

RNA = Ribonucleic acid

RNA-seq = Ribonucleic acid-sequencing

RPKM = Reads per kilo base per million

RPM = rotation per minute

RT-PCR = Reverse transcription polymerase chain reaction

RUNX1 = Runt-related transcription factor 1

RUNX1T1 = Runt-related transcription factor 1

SCF = Stem cell factor

SCF = Stem cell factor

SET = Su(var)3-9, 'Enhancer of zeste', Trithorax

SNG = Séquençage nouvelle génération

SNP = Single-nucléotide polymorphism

SNV = Single nucleotide polymorphism

ST-HSC = short term hematopoietic stem cell

T = Thymine

tal1 = T-cell acute lymphocytic leukemia 1

TET2 = Tet methylcytosine dioxygenase 2

TP53 = tumor protein p53

USP42 = Ubiquitin specific peptidase 42

VCAM1 = Vascular cell adhesion molecule 1

VEP = Variant Effect Predictor

REMERCIEMENTS

J'aimerais remercier les personnes de l'Institut de Recherche en Immunologie et Cancérologie qui m'ont accompagnée tout au long de ma formation, et qui ont fait de ma maîtrise une expérience enrichissante autant d'un point de vu professionnel que personnel.

Je tiens particulièrement à remercier mon directeur de recherche, Dr Brian Wilhelm, sans qui rien de tout cela n'aurait été possible. Merci de m'avoir donné l'opportunité de rejoindre votre laboratoire, j'ai beaucoup apprécié travailler avec vous et c'est grâce à vous que j'ai pu découvrir la recherche. Merci également de votre patience et de votre disponibilité, j'ai toujours trouvé réponse à mes questions qu'elles aient été pertinentes ou non.

J'aimerais remercier tous les membres de mon laboratoire. Je remercie Audrey Forest qui était présente dans mes débuts et m'a beaucoup appris sur les techniques de notre laboratoire. Je remercie également Magalie Celton pour l'assistance qu'elle m'a apportée dans l'analyse de mes résultats. Sans leur expertise il m'aurait été plus difficile de mener à bien mon projet de maîtrise.

Je tiens également à remercier le laboratoire de Guy Sauvageau pour tout le matériel prêté et l'aide technique apportée par ses membres. Je remercie particulièrement Jalila Chagraoui pour sa gentillesse et pour tous ses bons conseils, ainsi que Caroline Pabst pour son aide précieuse.

J'aimerais notamment remercier les membres de la plateforme de cytométrie en flux d'avoir pris le temps de m'expliquer le fonctionnement complexe de leurs machines et pour leur patience.

Enfin, je remercie ma famille pour tous ses encouragements, même à distance. Merci à mes amis d'avoir été à mes côtés. Ainsi qu'un grand merci à Alexandre Jaubert pour son attention et son précieux soutien durant ces deux années.

Merci.

CHAPITRE 1: INTRODUCTION

1. Le système hématopoïétique :

1.1. Définition de l'hématopoïèse :

Le sang est un tissu hautement régénératif en raison de la courte vie des cellules sanguines ¹, puisqu'il génère un billion de cellules par jour dans la moelle osseuse d'un adulte ². L'hématopoïèse est le processus par lequel les cellules souches hématopoïétiques (CSH), qui sont au sommet de la hiérarchie de la lignée hématopoïétique, produisent différentes cellules progénitrices qui se différencient progressivement pour donner à leur tour les cellules du sang (**Figure 1.1**) ^{1,3}. Les cellules sanguines matures incluent les hématocytes (ou érythrocytes et globules rouges), les plaquettes générées à partir de mégacaryocytes et les leucocytes (ou globules blancs, qui comportent les granulocytes, les monocytes et les lymphocytes), ayant des fonctions diverses dans l'organisme (**Tableau 1.1**) ³.

Pour pouvoir régénérer le système hématopoïétique au complet, les CSH doivent à la fois s'auto-renouveler pour maintenir le pool de CSH et se différencier ⁴. Les cellules souches embryonnaires (CSE) sont dérivées du mésoderme embryonnaire ¹ et migrent d'un site hématopoïétique à un autre pendant le développement ⁵⁻⁷. Les sites de l'hématopoïèse durant le développement sont séquentiellement le sac vitellin, l'aire aortogonade-mésonephros (AGM), plus récemment découvert : le placenta ⁸, qui forment le système primitif; le foie fœtal, le thymus, la rate et la moelle osseuse, qui forment le système définitif ¹. Il existe des différences phénotypiques entre les CSH adultes et embryonnaires ⁹. Les CSE sont pluripotentes et donc capables de donner naissance à toutes les cellules d'un organisme, alors que les cellules adultes sont multipotentes et ne génèrent que les cellules du sang. Les CSE sont capables d'expansion via des divisions

symétriques donnant deux cellules souches identiques, alors que les CSH adultes effectuent préférentiellement des divisions asymétriques (désignées par le terme « d'auto-renouvellement ») donnant une cellule souche et une cellule engagée dans la différenciation. De plus, les phénomènes de quiescence et de sénescence ne sont observés que chez les CSH adultes (**Figure 1.2**). Les CSH adultes sont principalement quiescentes (en phase G0 du cycle cellulaire), mais elles peuvent aussi reprendre le cycle en se renouvelant ou rentrer en différenciation, en sénescence ou en apoptose. On estime qu'une CSH se divise en moyenne 1 fois tous 145 jours.

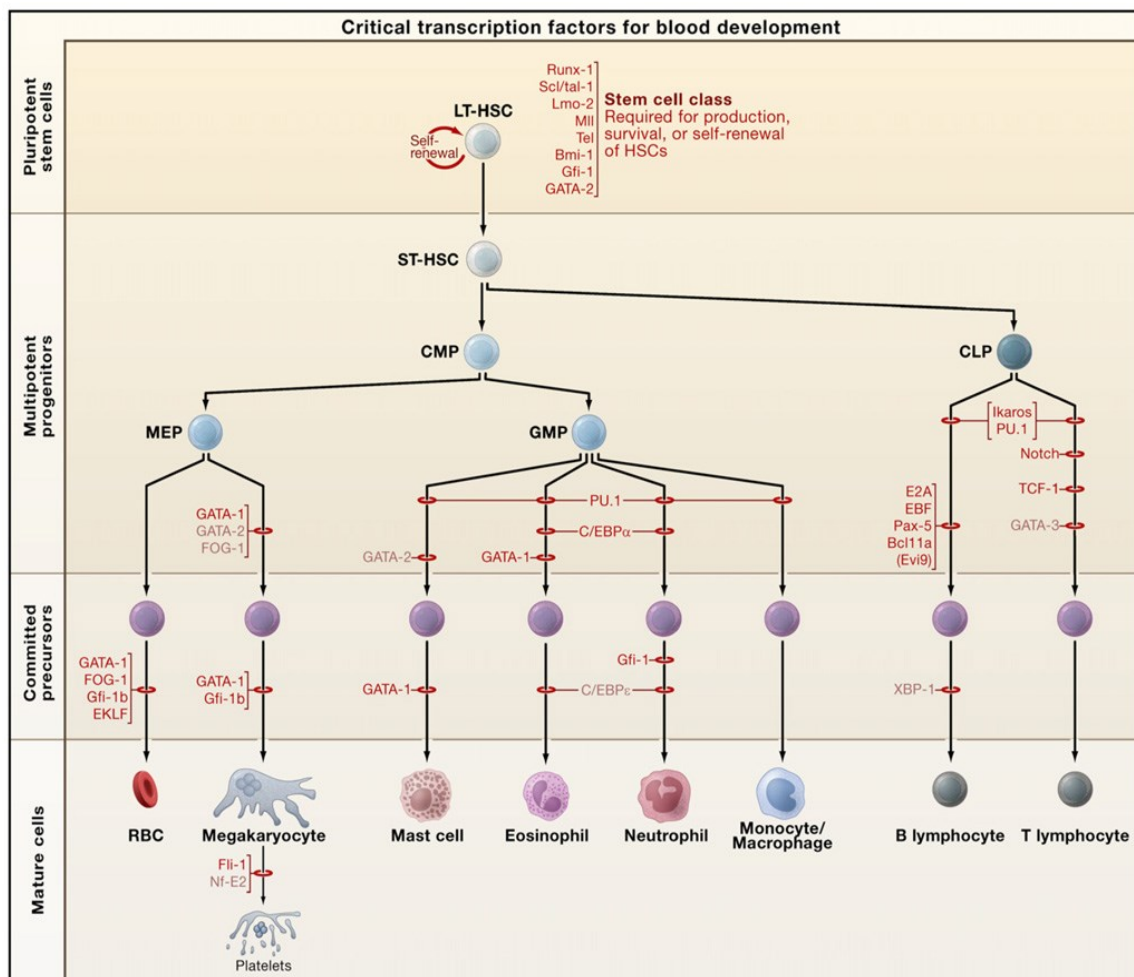


Figure 1.1 : L'hématopoïèse et les facteurs de transcription impliqués.

Réédité à partir de *Cell* ¹ © 2008, avec la permission d' Elsevier.

Abréviations : LT-HSC, cellule souche hématopoïétique de long terme; ST-HSC, cellule souche hématopoïétique de court terme; CMP, progéniteur myéloïde commun; CLP, progéniteur lymphoïde commun; MEP, progéniteur mégacaryocyte/érythroïde; GMP, progéniteur granulocyte/macrophage; RBCs, globules rouges.

Tableau 1.1: Les types de cellules sanguines et leurs fonctions.

Spécialisation	Cellules matures du sang	Fonction des cellules	
lignée lymphoïde	Lymphocyte T auxiliaire (Th1 et Th2)	Cellule CD4+ qui reconnaît les antigènes présentés par le complexe majeur d'histocompatibilité de type II. Les Th1 participent aux réponses retardées de type hypersensibilité et protègent des pathogènes et virus. Les Th2 aident les lymphocytes B et aide à exterminer et les parasites.	
	Lymphocyte T régulateur	Participe à l'homéostasie et au maintien de la tolérance immunitaire, en inhibant la prolifération des autres lymphocytes T.	
	Lymphocyte T cytotoxique	Lie les antigènes du complexe majeur d'histocompatibilité de type I et libère alors de la perforine qui lyse des cellules cibles.	
	Lymphocyte B	Cellule lymphoïde qui produit les anticorps.	
lignée myéloïde	Érythrocyte	Globule rouge qui contient l'hémoglobine pour oxygéner les tissus.	
	Polynucléaires	Éosinophile	Leucocyte associé aux infections par les vers parasites et à l'hypersensibilité.
		Granulocyte (neutrophile)	Leucocyte destiné à l'ingestion et la destruction de microorganismes (tels que les bactéries).
		Basophile	Leucocyte rare qui intervient dans les réactions allergiques par l'intermédiaire de l'histamine et facilitent la diapédèse.
	Monocyte/Macrophage	Le monocyte est le précurseur mononucléaire du macrophage. Les macrophages sont des phagocytes qui libèrent les autres cellules des corps étrangers. Ils sont aussi impliqués dans la présentation d'antigènes.	
Mégacaryocyte	Cellule polyploïde géante qui libère des centaines de plaquettes ayant un rôle dans la coagulation.		

Adapté avec la permission de Macmillan Publishers Ltd: *Nat Rev Genet* ³, © 2000

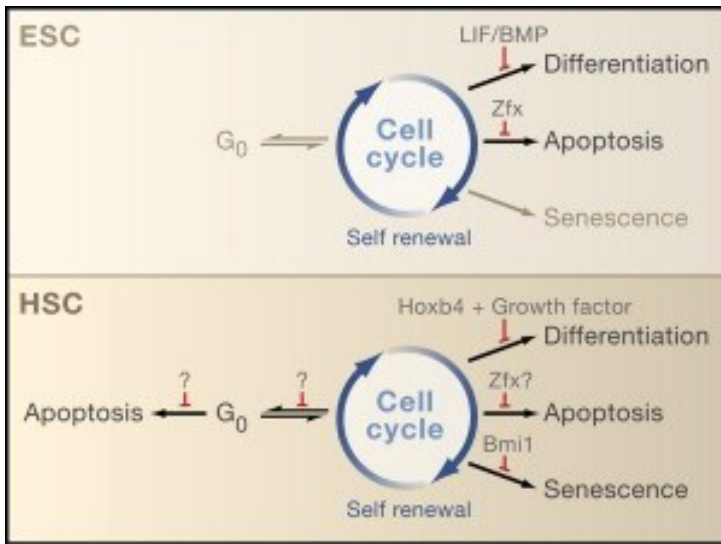


Figure 1.2 : Auto-renouvellement et survie des CSH.

Réédité à partir de *Cell*⁹ © 2007, avec la permission d' Elsevier.

Abbréviations : ESC, cellules souches embryonnaires; HSC, cellules souches hématopoïétiques.

1.2. Régulation de l'hématopoïèse :

1.2.1. Les facteurs de transcription :

L'hématopoïèse est un processus hautement régulé faisant intervenir des changements d'expression génique. Parmi les facteurs de transcription qui régulent le développement du système hématopoïétique, il y a ceux qui sont impliqués dans la formation et la fonction des CSH, tels que MLL, RUNX1, ETV6 , tal1 et LMO2; et ceux qui ont un rôle dans la spécialisation cellulaire pendant la différenciation, incluant PU.1, Gfi-1 et CEBPA¹. Il en est de même pour l'hématopoïèse adulte, bien que les facteurs de transcription puissent être différents ou avoir d'autres fonctions. En effet, RUNX1 qui est un facteur de transcription essentiel au développement du système hématopoïétique embryonnaire, n'est plus impliqué dans le maintien des CSH chez la souris adulte et devient nécessaire à la différenciation mégacaryocytaire et lymphocytaire¹⁰. Certains

autres facteurs régulent la prolifération des CSH, tels que HOXB4^{11,12}, et d'autres participent à chaque étape de la spécialisation des progéniteurs (**Figure 1.1**), tels que GATA1^{1,3}. La régulation de la différenciation s'effectue grâce à une combinaison de facteurs de transcription, associés à des destins cellulaires différents, qui agissent au sein d'un complexe protéique et ont un effet antagoniste les uns sur les autres^{1,3}. Ainsi, la cellule doit activer le facteur de transcription qui va permettre l'engagement dans son propre destin cellulaire, et en même temps inhiber les autres facteurs associés à des spécialisations alternatives. Par exemple, le destin cellulaire diffère selon si GATA-1 (érythroïde/mégacaryocyte/éosinophile) ou PU.1 (myéloïde) est le plus fortement exprimé dans la cellule¹³. Les facteurs de transcription impliqués dans l'hématopoïèse peuvent fonctionner en coopération avec des cofacteurs (co-activateurs ou co-répresseurs) qui régulent leur effet sur la transcription, en recrutant la machinerie de remodelage de la chromatine, ou en modifiant de façon covalente les facteurs de transcription¹⁴.

1.2.2. Les mécanismes épigénétiques :

Le terme « épigénétique » fait référence à des changements stables et héréditaires de l'expression des gènes, sans que la séquence ADN ne soit modifiée. Les mécanismes épigénétiques fonctionnent de façon concertée afin d'activer ou de réprimer l'expression des gènes régulateurs de l'hématopoïèse¹⁴.

1.2.2.1. Modifications de la chromatine :

Les modifications post-traductionnelles des queues d'histones modulent l'accessibilité de l'ADN aux protéines de la machinerie d'expression des gènes telles que les facteurs de transcription (comme énoncé en section 1.2.1)¹⁴. Parmi ces modifications, l'acétylation des histones (catalysée par les histones acétyl transférase) participe à

l'activation de certains gènes impliqués dans l'hématopoïèse. Des histones acétyl transférasés telles que CBP recrutée par GATA1, et MOZ sont respectivement nécessaires à la production de β -globine ¹⁵, et de cellules B ¹⁶. Une autre modification des histones est la méthylation. Alors que la méthylation des arginines est associée à l'activation des gènes, l'effet de la méthylation des lysines dépend du résidu sur lequel elle a lieu ¹⁷. Parmi les histones lysine méthyl transférasés (HKMT), deux types de familles protéiques majeures appelées Polycomb et Trithorax, répriment et activent respectivement les gènes du développement embryonnaire, tels que les gènes *HOX* ^{2,14,18}. Certains composants de ces complexes participent à l'hématopoïèse. Par exemple BMI1 (appartenant au complexe Polycomb PRC1) est essentiel à l'auto-renouvellement des CSH ^{18,19} et EZH2 (appartenant au complexe Polycomb PRC2) est impliqué dans la multipotence des CSH et dans la leucémogénèse ^{20,21}.

1.2.2.2. Méthylation de l'ADN :

La méthylation de l'ADN consiste en l'ajout d'un groupement méthyl en C5 d'une cytidine et a lieu au niveau des îlots CpG de l'ADN (lorsque les cytosines sont liées par une liaison phosphodiester aux guanines) et est associée à des régions de la chromatine transcriptionnellement inactives ¹⁴. Cette réaction est catalysée par une famille de méthyl-tranférase appelées DNMT, incluant DNMT1 qui est essentielle à l'hématopoïèse. En effet, les CSH adultes murines DNMT1-déficientes présentent des défauts d'auto-renouvellement, de différenciation et de rétention des CSH dans les niches (voir la définition des niches à la section 1.2.3) ²².

1.2.3. Les niches hématopoïétiques et les signaux extrinsèques :

Les CSH sont régulées par leur microenvironnement dans la moelle osseuse (située dans la partie centrale des os longs), appelé « niche » ¹⁴. Dans la niche endostéale, les

ostéoblastes modulent la prolifération des CSH et leurs fonctions cellulaires en interagissant avec les CSH via des jonctions adhérentes, ou en sécrétant des molécules signalétiques telles que l'angiopoïétine ²³⁻²⁵. L'importance des macrophages spécialisés de l'os endostéal (appelés « *ostéomacs* ») dans le maintien et la rétention des CSH a été reportée ^{26,27}. Par la suite, des CSH capables de reconstituer l'hématopoïèse à long terme (LT-CSH) ont été détectées au niveau de l'endothélium des vaisseaux sinusoides, suggérant l'existence d'une niche vasculaire ²⁸. Les cellules réticulaires de cette niche expriment le ligand chimiotactique CXCL12 induisant la migration des CSH ^{29,30}. D'autres cellules localisées autour des sinusoides, des cellules souches mésenchymales (nestine+), sécrètent des facteurs hématopoïétiques importants, tels que SCF, l'interleukine-7, ou la molécule d'adhésion des cellules vasculaires 1 (VCAM1) ³¹.

2. La leucémie myéloïde aigüe :

La leucémie est le terme utilisé pour désigner l'ensemble des cancers qui touchent les cellules souches de la moelle osseuse. La LMA, débute à partir de cellules progénitrices de la lignée myéloïde, et est à opposer aux leucémies lymphoblastiques qui atteignent les précurseurs des lymphocytes. C'est une forme aigüe de leucémie, c'est-à-dire à évolution rapide, pouvant prendre quelques jours ou quelques semaines pour se développer. Elle est à opposer à la forme chronique de leucémie, dont l'évolution peut prendre quelques mois à quelques années, et dont les symptômes sont souvent inexistantes au début ³².

2.1. Incidence:

D'après les chiffres de la Société Canadienne du cancer la LMA est la forme de leucémie la plus fréquente chez l'adulte au Canada, avec 1114 nouveaux cas de LMA reportés en 2007, le nombre de décès par LMA en 2008 était de 815, et 1300 cas de LMA étaient attendus en 2010 ³³. Aux États-Unis, les statistiques estimées pour 2013 sont 14 590

nouveaux cas de LMA et 10 370 décès dus à la LMA, majoritairement des adultes ³⁴. Malgré de nombreux cas infantiles, cette maladie est liée à la vieillesse, la moyenne d'âge des patients LMA est environ de 70 ans. La prévalence de la maladie est de 3,8 cas sur 100 000, allant jusqu'à 17,9 cas sur 100 000 pour les adultes de 65 ans et plus ³⁵. Chez les patients âgés, les LMA dites « secondaires », qui surviennent suite au traitement par chimiothérapie pour un autre cancer ou suite à un désordre hématologique antécédent, représentent 30 à 40% des cas. De plus, l'incidence de la LMA *de novo*, émergeant spontanément augmente avec l'âge ³⁶.

2.2. Symptômes et caractéristiques :

La LMA se caractérise par une accumulation de cellules myéloïdes immatures dans la moelle osseuse qui sont incapables de se différencier en cellules matures du sang et qui prolifèrent anormalement jusqu'à ce que leur nombre dépasse celui des cellules normales ^{36,37}. C'est alors que les symptômes de la maladie apparaissent, car il y a perte des cellules hématopoïétiques (insuffisance médullaire), telles que les granulocytes (granulocytopenie), les thrombocytes (thrombocytopenie), les globules rouges (résultant en une anémie), et parfois les leucocytes (leucopenie). Les symptômes de la LMA, tels que fatigue, hémorragie, infection, fièvre et dyspnée expiratoire (difficulté respiratoire lors de l'expiration), sont variés et ne sont pas spécifiques de la maladie ³⁷. L'anémie est le signe le plus commun de l'insuffisance médullaire et les patients sont alors suivis de près ³⁶. Les blastes leucémiques peuvent s'infiltrer dans d'autres tissus, tels que le cerveau, les poumons, le foie, la peau, les nœuds lymphatiques, en passant par le sang et provoquer d'autres symptômes ou la mort ^{36,37}. Une autre cause de décès est l'infection liée à la perte des neutrophiles ³⁶.

2.3. Diagnostic et traitement :

Le diagnostic de la LMA repose sur le fait que plus de 20% (critère de l'Organisation Mondiale de la Santé) ou 30% (critère du comité Français-américain-britannique) des cellules de la moelle osseuse ou du sang sont des blastes ³⁶⁻³⁸. Les patients avec les anomalies chromosomiques suivantes : t(8;21)(q22;q22), inv(16)(p13q22), t(16;16)(p13;q22), ou t(15;17)(q22;q12) sont diagnostiqués pour la LMA indépendamment du pourcentage en cellules hématopoïétiques (voir la section 2.5) ³⁸.

Le traitement de la LMA est divisé en deux phases : l'induction, une chimiothérapie qui vise une rémission complète, et la post-induction, qui vise à empêcher la rechute ^{36,37,39}. La décision concernant les premiers traitements dépend de facteurs pronostiques variés, qui permettent en général de prédire la réponse à la chimiothérapie d'induction, ainsi que le risque de rechute après le traitement (**Tableau 1.2**) ³⁷. L'induction consiste généralement en un traitement de chimiothérapie avec des composés cytotoxiques, alternant 7 jours d'injection de cytarabine et 3 jours d'anthracycline, ou thérapie « 3+7 » ^{39,40}. Il existe trois thérapies post-induction possibles pour les jeunes patients : la chimiothérapie, la greffe de moelle osseuse provenant d'un donneur HLA-compatible ou non (transplantation allogénique ou allogreffe) et la greffe de moelle provenant des propres cellules du patient (transplantation autologue ou autogreffe) ³⁷. Il existe également des thérapies ciblées dirigées spécifiquement contre des marqueurs de la LMA, telles que l'azacitidine, un agent déméthylant ^{41,42}, les inhibiteurs FLT3 tyrosine kinase ⁴³ et le Mylotarg combinant un anticorps anti-CD33 avec un antibiotique antitumoral ⁴⁴.

Tableau 1.2 : Facteurs pronostiques défavorables pour les patients avec la LMA.

Facteurs utilisés pour prédire la réponse à la chimiothérapie d'induction	Facteurs utilisés pour prédire la rechute
Caryotype défavorable	Caryotype défavorable
Âge > 60 ans	Âge > 60 ans
LMA secondaire	Réponse retardée à la chimiothérapie d'induction
Faible score de performance	Caractéristiques de résistance à de multiples substances
Caractéristiques de résistance à de multiples substances	Nombre de leucocytes > 20 000/mm ³
Nombre de leucocytes > 20 000/mm ³	Sexe féminin
Immunophénotype défavorable	Niveau de lactate déshydrogénase élevé
Cellules CD33 positives	Prolifération autonome des cellules leucémiques

Reproduit et traduit avec la permission de *N Engl J Med* ³⁷, Copyright Massachusetts Medical Society

2.4. Classification franco-américano-britannique (FAB) des LMA :

La classification FAB des leucémies aiguës est fondée sur les critères morphologiques des cellules leucémiques, qui permettent de déterminer leur degré de maturité et leur spécialisation. Elle inclut les leucémies aiguës de type lymphoblastiques et myéloïdes. Cette classification proposée en 1976 par des hématologues français, américains et britanniques après l'étude de 200 cas de leucémies repose sur des techniques morphologiques et cytologiques basiques. Ainsi, on distingue 7 groupes de LMA : M1 (leucémie myéloblastique sans maturation), M2 (leucémie myéloblastique avec maturation), M3 (leucémie promyélocytaire hypergranulaire), M4 (leucémie myélomonocytaire), M5 (leucémie monocytique), M6 (érythroleucémie) et M7 (leucémie mégacaryocytaire) ^{45,46}.

2.5. Classification de l'Organisation Mondiale de la Santé (OMS) des LMA :

La classification FAB des tumeurs des tissus hématopoïétiques et lymphoïdes par l'OMS vise à différencier les maladies qui sont cliniquement significatives, qui sont mutuellement exclusives et dont le diagnostic est possible en utilisant les techniques aujourd'hui disponibles ⁴⁷. Ce classement repose sur des critères cliniques, morphologiques, immunophénotypiques, génétiques et biologiques ^{47,48}. La première classification de 2001 a été révisée en 2008 pour y ajouter de nouvelles catégories basées sur des découvertes plus récentes. Cependant, des « entités provisoires » subsistent dans ce nouveau classement, car les récentes informations méritent d'être approfondies ⁴⁸. Par exemple, dans les LMA à caryotype normal, les cas présentant des petites mutations dans les gènes *NPM1* et *CEBPA* pourraient constituer des entités à part entière. Certaines anomalies cytogénétiques fréquemment retrouvées dans les cas de LMA, telles que les translocations $t(8;21)(q22;q22)$ et $t(16;16)(p13.1;q22)$, ou l'inversion $inv(16)(p13.1q22)$, suffisent à différencier quelques groupes, sinon la classification des types de LMA dépend de critères morphologiques, cliniques et phénotypiques ⁴⁸.

2.6. Signatures moléculaires de la LMA et facteurs pronostiques :

2.6.1. Les mutations :

2.6.1.1. Les anomalies chromosomiques :

La plupart des LMA présentent des anomalies chromosomiques récurrentes, telles que des translocations et des inversions, impliquant notamment des oncogènes de fusion⁴⁹. Les mutations cytogénétiques sont des facteurs pronostiques qui permettent de prédire la sensibilité et la résistance des cellules leucémiques aux traitements conventionnels³⁶. Le caryotype est le facteur de prédiction du pronostic le plus puissant pour la LMA, et permet de classer les jeunes patients adultes dans des catégories de risque favorable, intermédiaire et défavorable⁵⁰. Les catégories de risque en fonction du caryotype ont été révisées par le Conseil de Recherche Médicale (*Medical Research Council*) après l'étude de 5876 patients (entre 16 et 59 ans) qui étaient classés dans 54 groupes cytogénétiques différents, et incluent des anomalies cytogénétiques rares⁵¹.

Parmi les anomalies cytogénétiques de risque favorable, on compte la translocation (15;17)(q22;q21) qui conduit à la fusion *PML-RARA*⁵², et les réarrangements impliquant le complexe hétérodimérique de transcription « *core binding factor* » (ou CBF) qui sont présentes dans 10 à 20% des cas de LMA^{36,50}. Ceux-ci incluent t(8;21)(q22;q22) entraînant la fusion *RUNX1-RUNX1T1*, inv(16)(p13.1q22) impliquant *CBFβ*, et plus rarement t(16;16)(p13.1;q22) conduisant à la fusion *CBFβ-MYH11*⁵³⁻⁵⁵.

De plus, certaines anomalies chromosomiques prédisent un pronostic défavorable. C'est le cas de celles qui impliquent le gène *MLL*, un facteur de transcription majeur de l'hématopoïèse (voir section 1.2.1) qui compte plus de 60 partenaires de fusion^{56,57}. La fusion *MLL-AF9* (t(9;11)(p22;q23)) est la plus fréquente⁵⁸. D'autres translocations impliquent *ELL*, *MLLT1*, *AF6* ou encore *AF10*^{57,59}. La translocation rare t(6;9)(p23;q34) entraîne la fusion *DEK-CAN* et est également associée à un pronostic défavorable^{60,61}.

2.6.1.2. Les mutations génétiques moléculaires :

Cependant, un peu moins de la moitié (40 à 45%) des cas de LMA adultes ne présentent pas de réarrangements chromosomiques ayant une valeur pronostic, ceux sont des LMA à caryotype normal (LMA-CN) ⁵⁰. Le séquençage de génomes de LMA-CN a permis la détection de mutations moléculaires génétiques récurrentes plus petites, telles que des insertions/délétions, et des mutations ponctuelles ^{62,63}. Ces mutations génétiques moléculaires seules, ensembles, ou en combinaison avec les mutations cytogénétiques constituent également des facteurs pronostiques ⁵⁶.

Les mutations des gènes *NPM1* (petites insertions hétérozygotes) et *FLT3* (surtout une duplication interne en tandem (ITD)) sont les plus fréquentes pour les LMA-CN, et représentent respectivement environ 50% et 30-50% des cas ^{48,56,64}. De plus, les mutations de *NPM1* prédisent un pronostic favorable ⁶⁵, surtout en l'absence de mutation dans *FLT3* ⁶⁶⁻⁶⁸, et la mutation *FLT3*-ITD confère un pronostic défavorable ⁶⁹⁻⁷¹ et est associée à la fusion *PML-RARA* ^{72,73}. Le gène *CEBPA* est muté dans 5 à 10% des cas de LMA, plus fréquemment dans les LMA-CN et dans quelques cas de LMA familiales ^{50,56,74}. Les mutations de ce gène confèrent un pronostic favorable aux patients LMA-CN ou de risque intermédiaire ^{75,76}. D'autres mutations dans des gènes codant des régulateurs de l'hématopoïèse tels que *TET2* ⁷⁷⁻⁷⁹, *IDH1/2* ⁸⁰⁻⁸³, *ASXL1* ⁸⁴⁻⁸⁶ et *DNMT3A* ^{87,88}, ont été récemment découvertes.

2.6.2. L'expression génique :

Comme montré en section 2.6.1, de nombreux régulateurs de l'expression génique sont affectés dans la LMA. Grâce au développement des techniques de puces à ADN et séquençage, des profils d'expression génique (PEG) représentant des changements d'expression des gènes caractéristiques de la LMA, ont pu être identifiés ⁸⁹. Dans de

nombreux cas de LMA-CN des gènes surexprimés ayant une valeur pronostique ont pu être détectés (**Tableau 1.3**). De plus, l'étude de 242 patients LMA-CN, a montré que la sous-expression de *HOXA4* et la surexpression de *HOXA9*, *MEIS1* et *DNMT3A*, sont associées à des résultats cliniques peu favorables ⁹⁰.

Outre l'expression différentielle de gènes codant des protéines, certaines aberrations dans l'expression des microARN constituent des marqueurs pour certains sous-types de LMA ⁹¹. Par exemple, l'analyse de l'expression de 260 microARN chez des patients avec des LMA génétiquement définies, montre une corrélation entre l'expression de microARN et des sous-types de LMA, tels que les LMA avec les mutations t(8;21), t(15;17), inv(16), *NPM1*, et *CEBPA* ⁹². De plus, une valeur pronostique a pu être attribuée à de nombreux microARN différentiellement exprimés ⁹¹, tels que miR-191 et miR-199a dont la surexpression est associée à un pronostic défavorable ⁹³.

Tableau 1.3 : Effets de la surexpression des gènes sur le risque clinique.

Gène surexprimé	Pourcentage de cas	Pronostic associé
<i>BAALC</i> (<i>Brian and acute leukaemia cytoplasmic gene</i>)	~ 50	Défavorable
<i>ERG</i> (<i>Ets-related gene</i>)	~ 25	Défavorable
<i>MN1</i> (<i>Meningioma-1 gene</i>)	~ 25-50	Défavorable
<i>EVI1</i> (<i>Ecotropic viral integration-1 gene</i>)	6-11	Défavorable
<i>CXCR4</i> (<i>Chemokine (C-X-C motif) receptor 4</i>)	~ 33	Inconnu
<i>VEGFC</i> (<i>Vascular endothelial growth factor C</i>)	~ 50	Défavorable
<i>CDKN2A, p16INK4A</i> (<i>Cyclin-dependent kinase inhibitor 2A</i>)	~ 75	Défavorable

Adapté et traduit de la Table1 de la revue de littérature de Jonge H.J.M *et al* ⁸⁹ © 2011

3. Approches global visant à étudier la LMA:

3.1. Puces à ADN :

La technique d'hybridation sur puce à ADN consiste en l'immobilisation de molécules d'acides nucléiques sur une surface solide, qui servent de sondes à l'hybridation de fragments d'acide nucléiques provenant d'un échantillon ⁹⁴. Il existe des puces d'hybridation génomique comparatives, appelées également puces à ADNc, et des puces à ADN à polymorphisme nucléotidique, autrement appelées puces à oligonucléotides (**Figure 1.3**) ^{94,95}.

Pour les puces à ADNc, des sondes d'ADNc (résultant de l'amplification des ADNc de gènes d'intérêts insérés dans des clones, créant une bibliothèque d'ADNc) sont déposées régulièrement sur des filtres de nitrocellulose ou des lames de verre, de façon à créer des tâches pour chaque ADNc (**Figure 1.3**) ⁹⁴. Les sondes d'ADNc mesurent généralement entre 0,6 à 2,4 kb et représentent des gènes spécifiques. Une fois hybridées à la puces, elles sont détectées par radioactivité, par fluorescence simple ou par fluorescence double ⁹⁶. Cette dernière technique repose sur la compétition entre l'ADN test (de tumeur par exemple) et l'ADN contrôle, marqués par un fluorochrome différent. Un changement de la fluorescence témoigne d'une variation du nombre de copies d'un gène, pouvant aller de quelques paires de bases (pb) à la taille d'un chromosome ^{95,97}. Cette approche a permis de détecter de nouvelles anomalies chromosomiques dans 8 cas de LMA, telles que les amplifications 1p36, 10q26, 11p15, 20q13 et 22q23 ⁹⁸. La comparaison du signal des deux échantillons permet d'éliminer les artéfacts et d'avoir un haut niveau de précision, un avantage par rapport aux puces à oligonucléotides ⁹⁵.

Pour les puces à oligonucléotides, un pattern de sondes oligonucléotidiques correspondant à différents gènes est déposé ou synthétisé directement sur une plaque de silicone. Les fragments d'ADN test marqués par fluorescence sont hybridés aux sondes oligonucléotidiques (**Figure 1.3**) ⁹⁴. La taille des oligonucléotides peut atteindre 60

nucléotides, ce qui permet de limiter les hybridations non spécifiques. Cette méthode permet d'uniformiser la taille des oligonucléotides et de détecter l'épissage alternatif, mais elle est limitée par la disponibilité des séquences du génome humain, qui rend la construction de sondes oligonucléotidiques plus difficile ⁹⁴. L'application la plus fréquente de cette technique est la détection de polymorphismes d'un seul nucléotide (SNP) en utilisant des oligonucléotides représentant une version allélique d'un gène. L'intensité du signal permet également de détecter les variations du nombre de copies de gènes ⁹⁵. Kolquist et son équipe ont montré l'application de la puce à oligonucléotides à la détection de nouvelles variations du nombre de copies géniques impliquées dans la LMA sur l'ensemble du génome ⁹⁹.

Ces deux techniques sont également applicables à l'étude des changements de l'expression des gènes impliqués dans le cancer ^{94,96}. L'étude par puce à ADNc de l'expression différentielle entre des échantillons de patients LMA et des cellules hématopoïétiques normales, a permis la découverte de PEG spécifiques de la LMA ¹⁰⁰. L'analyse informatique des résultats d'expression génique peut être faite avec un apprentissage non-supervisé, ou par un apprentissage supervisé. Grâce à la méthode non-supervisée, l'analyse d'une hybridation sur puce à oligonucléotides sur 285 patients LMA a permis de détecter des PEG et d'isoler 16 groupes de patients avec des signatures moléculaires et des classes pronostiques distinctes ¹⁰¹.

Les technologies d'hybridation sur puces à ADN confèrent une variabilité importante entre les résultats, notamment due à la façon dont l'ADN est extrait, au marquage des sondes, à l'hybridation non-spécifique, au type de plateforme de puce utilisée, au type d'échantillons, à la méthode utilisée pour analyser les résultats ¹⁰².

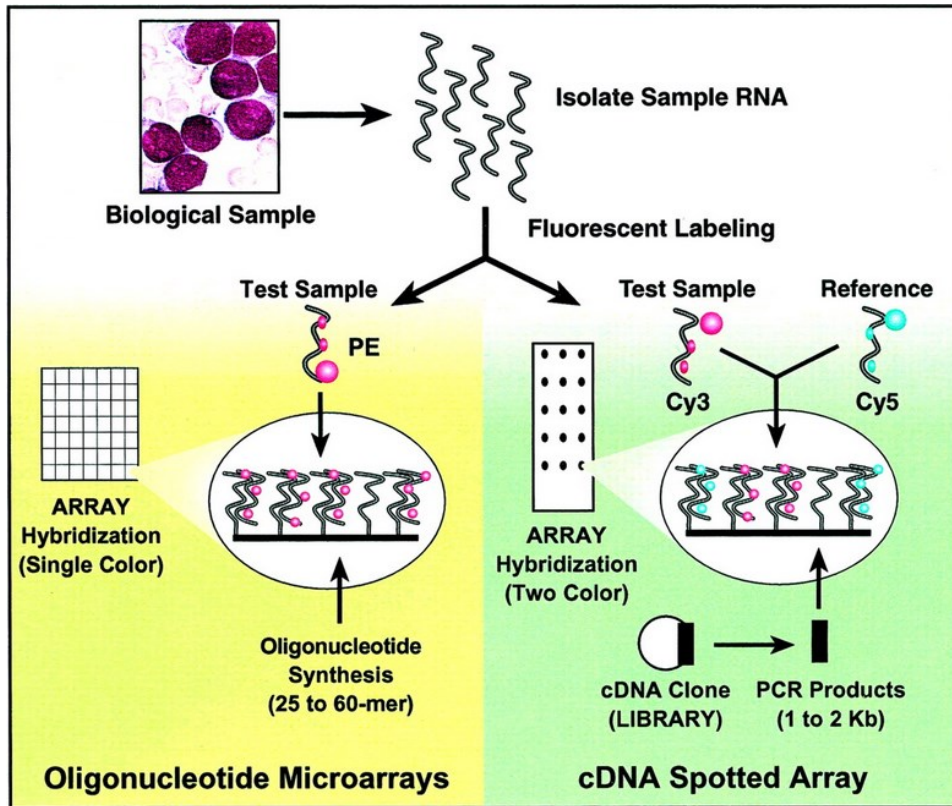


Figure 1.3: Comparaison entre les puces à oligonucléotides et les puces à polymorphisme nucléotidique.

Réédité avec permission. © 2002. American Society of Clinical Oncology. Tous droits réservés.⁹⁴

1.1. Séquençage à haut débit :

Le séquençage de l'ADN à haut-débit était restreint à la méthode de Sanger depuis les années 90. Après amplification clonage de l'ADN fragmenté, chaque cycle de séquençage inclus la dénaturation des fragments d'ADN, l'hybridation des amorces et l'extension des amorces avec une polymérase utilisant des didésoxynucléotides (ddNTP) fournis marqués par un fluorochrome différent pour chaque base (A,T,G,C) et leur incorporation arrête l'élongation. De cette manière, un signal correspondant à la dernière base de la séquence est détecté dans un capillaire d'électrophorèse et recueilli à chaque cycle de séquençage (**Figure 1.4a**). Malgré une grande précision, cette méthode est limitée par le

fait que l'électrophorèse ne puisse être réalisée que dans 96 ou 384 capillaires différents à la fois pour des fragments allant jusqu'à 1000 pb ¹⁰³.

En 2005, l'application du séquençage de « nouvelle génération » (SNG) cyclique sur puces à un génome bactérien, dans des prix compétitifs par rapport à la méthode de Sanger est démontrée par Shendure J. et son équipe ¹⁰⁴. Cette technique diffère du séquençage traditionnel (**Figure 1.4b**), car l'amplification clonale des fragments d'ADN se fait grâce à la ligation *in vitro* d'adaptateurs aux extrémités, celle-ci génère des amas (*clusters*) immobilisés sur des puces correspondant à chaque fragment et le séquençage s'effectue de manière cyclique par détection de l'ensemble des signaux fluorescents des puces ¹⁰³.

Ainsi, plusieurs méthodes d'amplification des fragments sont disponibles pour le SNG. La réaction en chaîne par polymérase (PCR) par émulsion (utilisée par les plateformes 454, Polonator et SOLiD) utilise des billes portant un des adaptateurs auquel s'hybride un type de fragment différent dans chaque émulsion, permettant l'amplification du fragment à partir de l'adaptateur libre. La PCR en ponts des fragments d'ADN (utilisée par la plateforme « Solexa » de Illumina) utilise des adaptateurs 5' et 3' immobilisés sur une surface solide. Les fragments simple-brin se lient à des endroits différents à la surface par leur adaptateur 5', leur adaptateur 3' libre s'hybride à son tour à un adaptateur proche du point d'attache pour former un pont d'ADN, lequel sert de matrice à l'amplification et le fragment double-brin est dénaturé pour former d'autres ponts ¹⁰³. D'autres types de plateformes existent, telles que la plateforme HeliScope (Helicos BioSciences) et la plateforme Pacific Biosciences ^{103,105}.

Suite à l'amplification clonale des fragments d'ADN, le séquençage consiste en la répétition de techniques biochimiques enzymatiques et d'acquisition par imagerie spécifiques à chaque type de plateforme, et permettant de reconstituer une séquence pour chaque puits d'une puce ¹⁰³. Par exemple, la technologie Solexa utilise la méthode cyclique à terminateurs réversibles, qui consiste successivement en l'incorporation de nucléotides, l'imagerie par fluorescence et le clivage ¹⁰⁵.

Une fois le séquençage effectué, de nombreux logiciels informatiques sont disponibles pour analyser les résultats bruts. L'analyse informatique vise à déduire les bases en fonction du signal et détecter les polymorphismes (comme avec PyroBayes et PbShort), aligner les éléments de lecture des fragments (*reads*) sur un génome de référence (avec BLAST et MAQ par exemple), assembler les lectures *de novo* (avec Edena par exemple), ou de visualiser le génome et annoter les séquences ¹⁰³.

Le SNG est applicable à la détection de mutations somatiques dans la LMA, telles que les variants d'un seul nucléotide (SNV), les petites insertions et délétions (indels), les réarrangements plus importants et les variations du nombre de copies géniques ¹⁰⁶. Trois méthodes expérimentales permettent la détection de mutations somatiques de la LMA : le séquençage du génome entier, le séquençage de l'exome et le séquençage du transcriptome (ou séquençage de l'ARN). Ces méthodes ont un coût variable, nécessitent une quantité de matériel différente et ont des applications différentes ¹⁰⁶.

Plusieurs études ont abouti à la détection de mutations somatiques non-synonymes de la LMA, par séquençage du génome entier des cellules leucémiques et des cellules normales d'un patient LMA-CN ^{62,107}. Les lectures provenant du séquençage des deux échantillons ont été alignés à un génome de référence, seuls les variants génétiques présents uniquement dans l'échantillon leucémique ont été gardés et ont été individuellement validés par un séquençage de Sanger. Une étude récente a permis d'identifier les mutations impliquées dans la rechute des patients LMA, avec le séquençage du génome entier de cellules leucémiques de 8 patients LMA avant la rechute et après la rechute ¹⁰⁸.

La majorité des études qui utilisent le SNG du génome entier se concentrent sur les mutations présentes au niveau des régions codantes, car leur effet sur la fonction protéique est plus facile à prédire. Le séquençage de l'exome cible spécifiquement les régions codantes (les exons) qui représentent seulement 1% du génome. Par conséquent, le séquençage de l'exome est moins coûteux que le séquençage du génome entier ¹⁰⁶. Les méthodes utilisant des puces à ADN à haute densité, ou des billes portant des sondes permettent de capturer n'importe quelle portion du génome, notamment les exons ^{109,110}.

Le séquençage de l'exome de 12 humains a permis de valider l'application de cette approche à la détection de mutations associées à des maladies humaines, telles que le syndrome de Freeman–Sheldon ¹¹¹. Par la suite, ce type de séquençage a pu être appliqué à la découverte de mutations somatiques de la LMA ¹¹².

Le séquençage du transcriptome est encore moins coûteux que celui de l'exome et nécessite une quantité de matériel plus faible. Cette approche est également applicable à la découverte de mutations somatiques liées à la LMA en comparant le séquençage de l'ARN provenant de la tumeur d'un patient avec celui de ces cellules normales ¹¹³. Le séquençage du transcriptome a aussi permis de détecter de nouveaux transcrits de fusion de la LMA-CN, grâce au séquençage de l'ARN de 45 cas de LMA ¹¹⁴. De plus, les données de séquençage du transcriptome d'un patient LMA ciblant spécifiquement les microARN, apporte des informations sur le profil d'expression des microARN et sur les variants génétiques détectés dans les microARN et dans leurs sites de liaison à l'ADN ¹¹⁵.

D'autres approches de séquençage sont possibles, telles que le séquençage du « méthylome » dans laquelle l'ADN génomique est traité au bisulfite pour différencier les cytosines méthylées de celles non méthylées avant le séquençage et qui apporte des information sur la méthylation de l'ADN à l'échelle du génome. De plus, les modifications de la chromatine peuvent être étudiées en réalisant le séquençage du génome associé aux histones qui consiste en une immunoprécipitation de la chromatine avec des anticorps contre les histones modifiés suivie du séquençage de la chromatine précipitée (ChIP-seq) ¹¹⁰. L'étude récente de 200 cas de LMA combine plusieurs techniques de SNG (séquençage du génome, de l'exome, du transcriptome, des microARN et du méthylome) pour apporter de l'information sur les variations du nombre de copies, les mutations codantes, les fusions géniques, la coopération entre les mutations, les PEG et la méthylation de l'ADN, impliqués dans la LMA ¹¹⁶. Ce travail constitue un début de réponse sur la façon dont ces événements génétiques participent à la pathogénèse de la LMA, et pourra être utilisée comme support à des études plus poussées sur la pathogénèse, la classification des LMA et la stratification des risques cliniques. Il permettrait aussi de développer des thérapies ciblées (personnalisées pour chaque patient).

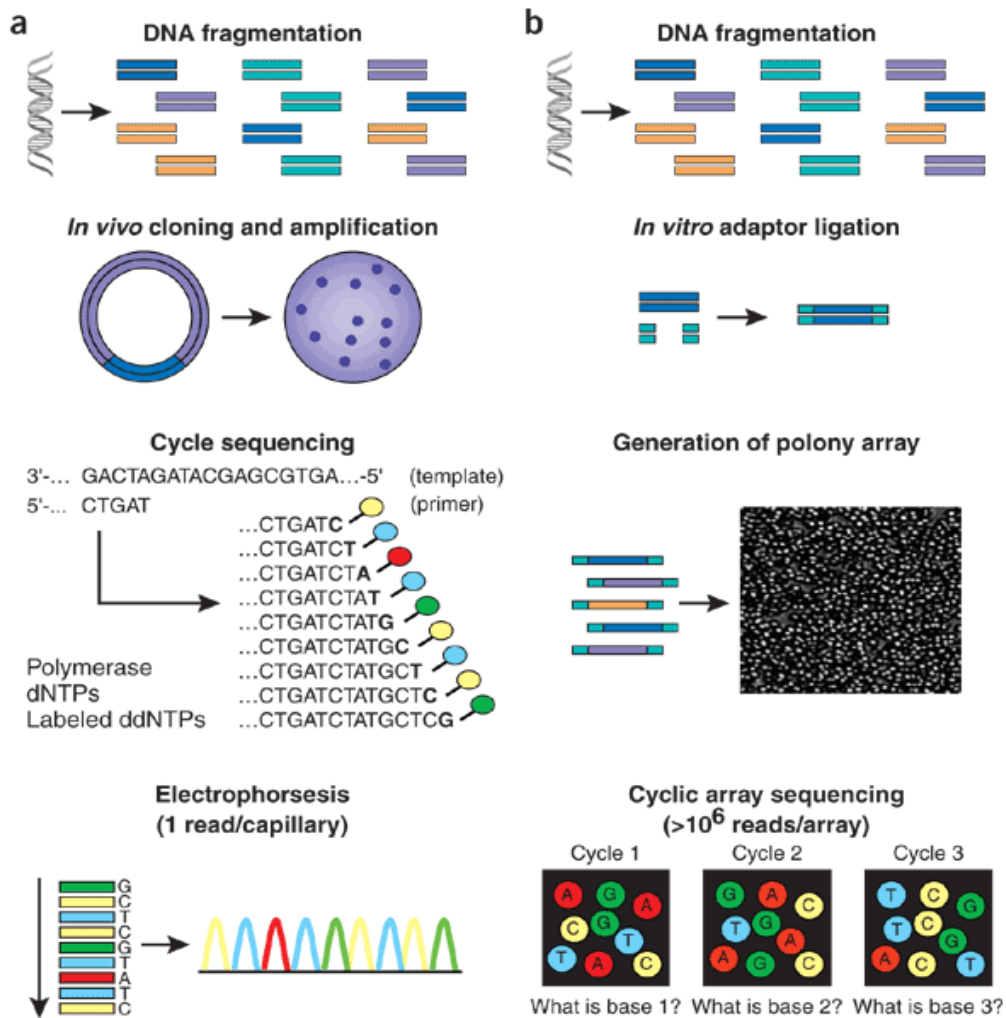


Figure 1.4 : Étapes successives du séquençage traditionnel versus le séquençage de seconde-génération.

Réédité avec la permission de Macmillan Publishers Ltd: *Nat Biotechnol* ¹⁰³ © 2008

4. Modèles d'étude de la leucémie myéloïde aigüe :

4.1. Échantillons primaires de patients pour l'étude *in vitro* :

Les échantillons primaires de patients LMA sont extraits de la moelle osseuse (ponction dans l'os de la hanche généralement) ou du sang périphérique des patients ¹¹⁷. Afin de sélectionner spécifiquement les blastes leucémiques, plusieurs techniques peuvent être appliquées successivement : une centrifugation sur gradient de densité (permettant d'isoler des cellules mononuclées), une déplétion des cellules matures, une sélection positive des blastes (enrichissement en cellules CD34+ par exemple). Les blastes LMA sont alors mis en culture cellulaire dans des milieux liquides variés en présence de sérum ou non, qui influence le comportement des blastes. De plus, la plupart des cellules de patients requièrent l'addition de cytokines (interleukines et facteurs de croissance) dans le milieu pour leur survie et leur prolifération *in vitro*. Malgré la sélection des blastes AML, il existe une grande hétérogénéité de cellules AML en culture, allant des blastes à long-terme à des progéniteurs plus différenciés.

L'étude *in vitro* des cellules LMA de patients a permis de mieux comprendre leur prolifération (autocrine ou spontanée) ¹¹⁸⁻¹²¹, leur différenciation ¹²²⁻¹²⁵ et leur apoptose (mort programmée) ¹²⁶⁻¹²⁸. De plus, le fait de travailler directement sur les cellules leucémiques humaines a permis d'identifier les mutations et les PEG caractéristiques de la LMA et d'associer les anomalies génétiques aux résultats cliniques des patients (voir section 2.6). Cela permet également de comparer des échantillons de patients à différents stades de la maladie (diagnostic comparé à la rechute par exemple) pour comprendre les changements associés ¹²⁹⁻¹³².

4.2. Lignées cellulaires leucémiques pour l'étude *in vitro* :

Les lignées cellulaires sont générées à partir d'un seul clone d'un échantillon de patient LMA, ce qui permet d'obtenir une population de cellules homogène, contrairement aux cellules primaires de patients ^{133,134}. Les cellules d'une lignée ont arrêté leur différenciation à un stade défini, leur prolifération en culture est maintenue indépendamment de facteurs de croissance (immortalisation), leur génotype et leur phénotype est relativement stable dans des conditions de culture optimales, et elles portent des variations génétiques et des patterns d'expression de gènes tumoraux spécifiques ¹³⁴. Les avantages des lignées cellulaires par rapport aux cellules de patients sont un apport infini en matériel cellulaire, la disponibilité d'un matériel cellulaire identique partout dans le monde, la possibilité d'entreposer ce matériel dans l'azote liquide et de le récupérer, ainsi que l'absence de cellules normales contaminantes. La création de banques de lignées cellulaires publiques a grandement facilité l'accès à ce modèle d'étude ¹³³. Cependant, l'établissement, l'authentification et la distribution des lignées cellulaires sont des processus très coûteux qui nécessitent généralement un financement extérieur. De plus, un désavantage majeur des lignées cellulaires est qu'elles ont de nombreuses anomalies chromosomiques, dont certains sont acquis pendant l'expansion *in vitro* ¹¹⁷.

Les lignées cellulaires ont facilité l'étude de la prolifération ¹³⁵⁻¹³⁸, la différenciation ¹³⁹⁻¹⁴¹ et l'apoptose ¹⁴²⁻¹⁴⁴ des cellules leucémiques. De plus, les effets de certains traitements actuels ou futurs ont été étudiés sur les lignées cellulaires LMA *in vitro* ¹⁴⁵⁻¹⁴⁹.

4.3. Modèles murin *in vivo* :

Trois stratégies sont utilisées dans les modèles murins des LMA humaines : la xénotransplantation, la transduction rétrovirale et la transgénèse ¹⁵⁰. La xénotransplantation consiste en la greffe de cellules LMA humaines primaires, dans des souris immunodéficientes (souris SCID, NOD/SCID, ou NOD/SCID/Il2rg^{-/-}). Les cellules de la moelle osseuse ou du sang périphérique de patients LMA sont injectées dans des

souris préalablement irradiées. Cependant, le succès de la greffe est variable et tous les effets phénotypiques de la LMA humaine ne sont pas reproductibles avec ce modèle. La transduction rétrovirale est une technique dans laquelle les cellules hématopoïétiques primitives de la moelle osseuse d'une souris donneuse sont génétiquement modifiées *ex vivo* grâce à un vecteur d'expression rétroviral dans lequel est cloné un gène d'intérêt. Les cellules génétiquement modifiées sont alors greffées à une souris. Le contexte génétique des cellules greffées, la dérégulation de l'expression du transgène et la mutagénèse due au site d'insertion du vecteur, peuvent influencer les expériences réalisées. La transgénèse permet de générer des souris porteuses de un ou plusieurs gènes exogènes. Il est possible par exemple de reproduire le produit de la fusion *RUNX1-RUNX1T1* humain en insérant *RUNX1T1* par recombinaison homologe dans le gène *Runx1* dans des embryons de souris. Avec ce modèle, il faut considérer l'effet de l'insertion sur la mutagénèse et sur l'expression du transgène, et l'effet du promoteur hétérogène sur l'expression endogène.

Les modèles murins ont permis de découvrir les gènes qui initient le développement de la LMA *in vivo* ¹⁵¹⁻¹⁵⁶. Ces modèles apportent de l'information sur la façon dont ces gènes participent à la leucémogénèse, en identifiant les voies signalétiques impliquées et les effets biologiques sur les cellules hématopoïétiques ¹⁵⁷⁻¹⁵⁹. Une expérience récente a permis de corrélérer des données de leucémies générées chez la souris avec les résultats cliniques des patients ¹⁶⁰. De plus, de potentiels et actuels traitements de la LMA ont été testés dans des modèles murins ^{161,162}. Cependant, le contexte génétique murin est un paramètre qui influence de manière significative la susceptibilité à la leucémie et son phénotype, que la communauté internationale tente de comprendre d'avantage.

5. Problématique :

Bien que certaines anomalies génétiques et certains PEG aient pu être détectés de façon récurrente dans la LMA, les mécanismes moléculaires et cellulaires menant à la

pathogénèse ne sont complètement élucidés. Cela est d'autant plus vrai pour les LMA-CN, représentant presque 50% des cas, qui n'ont pas de grands réarrangements chromosomiques apparaissant comme des causes génétiques évidentes de la maladie. C'est pourquoi l'utilisation de systèmes modèles reste indispensable à la compréhension des processus biologiques participant à la leucémogénèse. Les lignées cellulaires leucémiques sont des modèles d'étude de la LMA facilement accessibles grâce aux banques. Elles permettent de travailler avec une population de cellules homogène qu'il est facile de maintenir en culture au long terme, contrairement aux échantillons primaires de patients. Cependant, les lignées cellulaires disponibles aujourd'hui ont des réarrangements chromosomiques dont certains sont acquis pendant la culture cellulaire, ne permettant pas l'étude de la LMA-CN. Munker et son équipe ont résolu ces problèmes en 2009, en établissant une nouvelle lignée cellulaire LMA de type M5 : CG-SH, ayant la particularité d'avoir un caryotype normal ¹⁶³. Le phénotype de la lignée a été défini, et l'analyse moléculaire de quelques gènes LMA a révélé que seuls deux d'entre eux, *RUNX1* et *NRAS*, étaient porteur de mutations présumées. Nous nous sommes alors demandé quelles sont les autres caractéristiques moléculaires des cellules CG-SH.

6. Hypothèses et objectifs :

L'objectif principal de ce projet est de caractériser plus en détail le modèle d'étude de la LMA que représentent les cellules CG-SH. Pour cela, plusieurs hypothèses ont été émises à l'égard de ces cellules et des objectifs permettant de les vérifier ont été formulés.

- ❖ Hypothèse 1 : La lignée CG-SH possède d'autres variations génétiques que celles précédemment détectées qui participent à la pathogénèse.
 - Objectif 1 : Définir l'ensemble des variants génétiques présents.

- Objectif 2 : Déterminer les gènes mutés qui participent potentiellement à la pathogénèse.
- ❖ Hypothèse 2 : La lignée CG-SH est sensible à l'effet prolifératif des cytokines hématopoïétiques.
- Objectif 1 : Définir les changements d'expression des gènes induits par les cytokines.
 - Objectifs 2 : Déterminer les effets des cytokines sur la mort cellulaire des cellules CG-SH.

CHAPITRE 2 : APPROCHES EXPÉRIMENTALES

Culture Cellulaire

La lignée cellulaire CG-SH a été généreusement fournie par le laboratoire de Munker R. (2009) ¹⁶³. Les cellules ont été mises en culture à 37°C avec 5% de CO₂ dans des pétris contenant un milieu liquide Iscove's Modified Dulbecco's Medium (IMDM) avec 10% de sérum de veau fœtal (FBS), et 1% d'antibiotiques pénicilline/streptomycine (Invitrogen). Les CG-SH ont été cultivées dans un milieu standard ou dans un milieu additionné de cytokines : *stem cell factor* (SCF) à 166 ng/mL, interleukine-3 (IL3) à 20 ng/mL, *granulocyte colony-stimulating factor* (G-CSF) à 20 ng/mL et *Fms-related tyrosine kinase 3 ligand* (FL) à 100 ng/mL (Biotechnology). Afin de calculer le temps de doublement des cellules dans ces deux conditions, des duplicats biologiques (n=3) ont été maintenu en culture pendant 6 jours en présence et en absence de cytokines dans le milieu. Les comptes cellulaires ont été réalisés avec un hémocytomètre Bright-Line de chez Hausser Scientific, après marquage des cellules au bleu de trypan (Trypan Blue Stain 0.4% ; Invitrogen). Pour l'étude de l'expression génique au cours du temps, les cellules CG-H ont été récoltées pour que leur ARN soit collecté après un jour ("day_1"), trois jours ("day_3") et deux semaines d'exposition aux cytokines ("day_14").

Séquençage nouvelle génération de l'ARN et de l'ADN génomique (ADNg)

Les cellules en culture ont été collectées par centrifugation à 800 RPM pendant 10 minutes et les culots ont été resuspendus dans 1mL de réactif TRIzol (Invitrogen) pour 5x10⁶ de cellules (pour l'extraction de l'ARN), ou directement purifiés avec le kit PureLink Genomic DNA Mini Kit (Invitrogen) (pour l'extraction de l'ADNg). L'ARN a été extrait avec

le kit RNeasy Plus Mini (Invitrogen) et l'ADNc a été généré par transcription inverse de l'ARN avec le protocole Superscript III Reverse Transcriptase de chez Invitrogen. Les bibliothèques d'ADN ont été préparées à partir du protocole TruSeq DNA Sample Prep High Throughput (Illumina). Brièvement, l'ADNg a été fragmenté avec des plaques de fragmentation Covaris et les extrémités franches ont été générées avec le mélange End Repair Mix qui fournit l'activité 3' à 5' exonucléase. L'ADNg et l'ARN ont été adénylés à leurs extrémités 3' avant l'ajout d'adaptateurs de séquençage aux extrémités des fragments, par ligation suivie d'une PCR et d'une purification. Les fragments d'ADN de la bibliothèque ont été amplifiés par PCR en utilisant des amorces spécifiques des adaptateurs et normalisés à 10 nM pour former des *clusters* sur la Diluted Cluster Template. Les bibliothèques d'ADNc et d'ADNg ont été séquencées avec le séquenceur Illumina HiSeq 2000 à la plateforme génomique de l'IRIC.

Analyse des données de séquençage

La cartographie des lectures (*reads*) a été réalisée comme décrit précédemment par Simon et son équipe (2012) ¹⁶⁴. Une approche de filtration successive des données brutes a été utilisée pour l'analyse du très grand nombre de SNP détectés. L'analyse ciblait les variants génétiques non-synonymes et codant dans les gènes exprimés (avec un seuil d'expression en RPKM > 5) ayant une valeur minimale en nombre de lectures qui couvrent la région des variants (*depth* > 20). Les biais systématiques dus à la cartographie des lectures ont été éliminés en excluant les variants localisés dans des régions dupliquées ou des séquences répétitives. Enfin, seuls les variants apparaissant systématiquement dans l'ADNg et l'ARN ont été inclus dans la liste de variants utilisés pour l'annotation des termes de l'ontologie des gènes (*Gene Ontology*, GO). Tous les variants génétiques ont été classés comme connus (présents dans la base de données dbSNP Human Build 1.3.5; 4138 variants) ou nouveau (absent de dbSNP ; 60 variants). L'étude de l'effet délétère potentiel des SNP pour la protéine a été réalisée avec le logiciel Variant Effect Predictor (VEP) du

site internet de Ensembl (<http://useast.ensembl.org/tools.html>), qui combine les résultats de trois algorithmes publiés : PolyPhen, SIFT et Condel. Pour tous les variants connus ou nouveaux, seuls ceux qui étaient prédits comme délétères (« *damaging* » ou « *deleterious* ») par au moins deux des trois algorithmes de VEP ont été retenus.

Suivi de l'activité GATA2 avec un rapporteur luciférase

L'ADNc de référence *GATA2* hg19 amplifié par PCR à partir d'un ADNc LMA humain constitue la forme normale (*wild-type*) de *GATA2*, alors que la forme mutée de *GATA2* a été obtenue par amplification par PCR de l'ADNc des cellules CG-SH. *GATA2* normal et muté ont été clonés dans un plasmide pcDNA3.1 et pcDNA PU.1 publié, et le promoteur qui répond à *GATA2* (*LYL1*) dans PGL4.12 a été généreusement fournis par le laboratoire de Scott H.S ¹⁶⁵. Les cellules Cos-7 (lignée de fibroblastes rénaux de singes) ont été transfectées à 90% de confluence avec l'agent de transfection Lipofectamine 2000 (Invitrogen). Les cellules ont été récoltées après 20h et utilisées pour l'essai luciférase réalisé avec le protocole Luciferase Assay et le luminomètre Veritas Microplate Luminometer (Tunner Biosystem) de la compagnie Promega. Les résultats de l'essai luciférase ont été reportés en *relative light units* (RLU) de l'activité de la luciférase de la luciole. Les expériences ont été réalisées en triplicat.

Suivi des changements de l'expression des gènes dans le temps

Les valeurs d'expression (en RPKM) provenant du séquençage de l'ARN des cellules cultivées en présence ou en absence de cytokines ont été utilisées pour calculer un simple ratio (*fold change*) pour +cytokines/-cytokines ou pour être analysées dans R avec le paquet DESeq 2. Ces deux méthodes ont permis de mesurer l'expression différentielle entre les deux conditions, et un seuil pour déterminer si la différence est significative a

été utilisé pour chaque méthode : $p_{adj} < 0.05$ (paquet DESeq2), ou un ratio supérieur à 2 pour les gènes surexprimés et un ratio inférieur à 0.5 pour les gènes sous-exprimés (ratios directs en RPKM).

Analyse de la mort cellulaire

Les cellules CG-SH ont été cultivées tel que décrit ci-avant pendant 7 jours en présence ou en absence de cytokines. Elles ont été récoltées et marquées après 3 jours, 5 jours et 7 jours d'exposition aux cytokines. Les cellules ont été doublement marquées avec de l'annexinV couplée à l'isothiocyanate de fluorescéine (FITC) (BD Bioscience), et du iodure de propidium (PI) afin de distinguer les populations cellulaires en apoptose précoce (annexinV+/PI-), en apoptose tardive (annexinV+/PI+), et en nécrose (annexinV-/PI+). Les résultats ont été analysés à la plateforme cytométrie en flux de l'IRIC avec le cytomètre en flux BD FACSCanto II (BD Bioscience).

CHAPITRE 3: ARTICLE

1. Contributions:

Article soumis au journal *Experimental Hematology*.

Nature de la participation de **Géraldine Gosse** aux travaux de recherche :

- Conception, réalisation des expérimentations et analyse des résultats de toutes les expériences, excepté l'expérience pour la Figure 3.
- Élaboration de la Figure 1, la Figure 2, la Figure 3 (en partie), la Figure 4A, la Figure 4B, la Figure 5 et la Figure 8. Élaboration de tous les tableaux de l'article, ainsi que toutes les figures et tableaux supplémentaires.
- Écriture de l'article.

Nature de la contribution des co-auteurs à l'article :

Audrey Forest :

- Aide à la conception et à la réalisation des expériences.
- Réalisation des expérimentations pour la Figure 3.

Magalie Celton :

- Aide à l'analyse des résultats des expérimentations pour les Figures 1 à 7.
- Conception et analyse des expérimentations pour la Figure 3.
- Réalisation de la Figure 3 (en partie) et la Figure 6.
- Révision critique du manuscrit.

Brian T Wilhelm :

- Supervision du travail, conception et analyse de toutes les expériences.
- Réalisation de la Figure 7.
- Révision critique du manuscrit.

2. L'article:

Molecular characterization of a novel AML cell line with a normal karyotype

Géraldine Gosse^{1,2}, Magalie Celton^{1,2,3}, Audrey Forest^{1,2}, Brian T Wilhelm^{*1,2}

¹ Institute for Research in Immunology and Cancer, University of Montreal, Montreal, QC;

² Laboratory for high throughput genomics, Montreal, QC;

³INRA, UMR1083, Sciences Pour l'Oenologie.

***Correspondence to: Brian Wilhelm**

Institute for Research in Immunology and Cancer (IRIC), University of Montreal,
P.O. Box 6128, Station Centre-Ville, Montréal, Québec, Canada, H3C 3J7.

Hematological Malignancies

Abstract

Acute myeloid leukemia (AML) occurs when hematopoietic progenitor cells acquire genetic defects which interfere with their ability to regulate normal growth and division. Although many recurrent translocations have been identified in AML, almost half of all adult patients will present with a normal karyotype (NK-AML). While a variety of cell line models exist to study AML most of these have highly abnormal and unstable karyotypes, and primary cells from NK-AML patients are difficult to maintain *in vitro*. In the present study, we provide a thorough molecular characterization of a recently established cell line, CG-SH, which has normal cytogenetics and therefore represents a useful new model to study NK-AML. Using high-throughput DNA sequencing, we first defined the genetic background of this cell line. In addition to identifying potentially deleterious SNPs in a number of genes relevant to AML, we also found insertions in both *GATA2* and *EZH2*, two genes reported to be involved in AML. To improve its utility, we further characterized the growth of this model system *in vitro* with a cytokine mix (FL, IL3, SCF, G-CSF) which promotes a faster rate of cell growth in CG-SH. We assessed the gene expression changes due to cytokines and found differential expression in genes implicated in the proliferation, apoptosis and differentiation. Our results not only provide a molecular characterization of this interesting new cell line, but also elucidate the molecular defects in this cell line derived from an AML patient.

Introduction

Acute myeloid leukemia (AML) is a genetically heterogeneous disease, characterized by the accumulation of mutations in hematopoietic progenitor cells¹⁶⁶. In AML, recurrent mutations have been identified which affect the self-renewal (e.g. *RUNX1*), proliferation (e.g. *FLT3*) and differentiation (e.g. *CEBP α* , *GATA2*) capacity of progenitor cells in the bone marrow³⁷. In addition to point mutations, a variety of chromosomal translocations such as t(8;21)(q22;q22), inv(16)(p13.1q22), t(16;16)(p13.1;q22), t(15;17)(q22;q12) have also been identified in the disease^{48,167}. Despite these recurrent karyotypic anomalies, approximately 50% of acute myeloid leukemia cases present with a normal-karyotype (NK-AML), increasing the process of diagnosis, risk assessment and therapeutic guidance^{168,169}. In the past years, recurrently mutated genes in NK-AML have been identified, such as *FLT3*^{69,170,171}, *NPM1*^{172,173}, *CEBP α* ^{76,174}, *GATA2*¹⁷⁵, *IDH1/2*¹⁷⁶ and *DNMT3A*¹⁷⁷. Despite these discoveries, and the generation of completely sequenced AML genomes and transcriptomes, the complete genetic causes of NK-AML have yet to be determined. One limitation of current efforts to study the role of driver mutations in NK-AML is the lack of suitable model systems. Studies on primary AML patient samples are technically challenging because of the difficulties in manipulating primary AML blasts and maintaining them in their immature state in vitro for long periods of time. To overcome these issues, a variety of human myeloid cell lines, such as HL-60, THP1, or KG-1A have been used as model systems¹⁷⁸ although all of these lines carry multiple chromosomal rearrangements¹⁷⁹⁻¹⁸¹ and have unstable karyotypes which do not accurately reflect NK-AML biology.

The rarity of AML cell lines with normal karyotypes was recently addressed by publication of a new cell line which has a true normal karyotype¹⁸². Munker and colleagues recently established, CG-SH, an AML cell line with a normal karyotype which was derived from an M5 AML patient. This cell line has the phenotype of an immature blast cell (CD7+, CD34+, CD38+, CD117+ and CD10-, CD19-, CD20-, CD41-, CD42-) and expresses multiple myeloid markers (e.g. MPO, CD13, CD15, and CD65). Of the commonly recurrent mutations in AML (e.g. *FLT3*-ITD, *NPM1*, *MLL*-PTD), Munker and colleagues could only identify putative mutations in *RUNX1* and *NRAS*¹⁸² genes.

To identify molecular lesions in this cell line in an unbiased fashion, we performed a global analysis using whole-genome and whole-transcriptome sequencing data. Detailed analysis of these sequencing data for genes recurrently mutated in AML revealed no mutations in this cell line, confirming and extending the original observations of Munker and coworkers. We did however, identify novel sequence variants in 60 genes, including *GATA2*¹⁷⁵ and *USP42*¹⁸³, which have been found to be mutated in AML patients. In addition, we also identified novel insertions and deletions (indels) including one in the *EZH2* gene, which has also been reported to be mutated in AML^{184,185}. In addition, we investigated the biological pathways regulating the behaviour of CG-SH cells, through the addition of various cytokines into the growth media. We analyzed the differential gene expression profiles in presence or not of cytokines and observed enhanced proliferation and a decreased doubling time. RNA-seq analysis indicated that cytokines may play a role in controlling apoptosis and cell cycle processes, leading to a more rapid proliferation of leukemic cells. Together, these results highlight both the genetic defects in a novel NK-

AML cell line and its potential as a novel model system to study the genetic requirements of NK-AML.

Material and methods

Cell culture

The CG-SH cell line was provided by Munker et al. (2009) and cultured in IMDM medium with 10% FBS, 1% penicillin/streptomycin (Invitrogen). Cultures were either grown in standard conditions or with media supplemented with stem cell factor (SCF) at 166 ng/mL, interleukin 3 (IL3) at 20 ng/mL, granulocyte colony-stimulating factor (G-CSF) at 20 ng/mL and Fms-related tyrosine kinase 3 ligand (FL) at 100 ng/mL. To calculate the doubling-time, biological duplicates of cell aliquots (n=3) were grown for 6 days, with and without cytokines. Cell counts were done every 12 hours using a hemacytometer Bright-Line (Hausser Scientific) after staining the cells with Trypan Blue Stain (0.4%, Gibco by Life Technology). For time course experiments, CG-SH cells were harvested for RNA collection one day ("day_1"), three days ("day_3") and after two weeks of cytokine exposure ("day_14").

RNA and gDNA high throughput sequencing

Cultured cells were harvested through centrifugation at 800 RPM for 10 minutes and the pellets were either resuspended in 1mL TRIZOL reagent (Invitrogen) (for 5×10^6 cells) or directly purified with PureLink Genomic DNA Mini Kit (Invitrogen). RNA was extracted with the RNeasy Plus Mini Kit (Invitrogen) and cDNA was generated using the Superscript

III Reverse Transcriptase protocol from Invitrogen. DNA libraries were prepared according to standard TruSeq DNA Sample Prep High Throughput (HT) protocol (Illumina). Briefly, gDNA was fragmented using Covaris fragmentation plates and ends were blunted using an End Repair Mix with 3' to 5' exonuclease activity. Both gDNA and cDNA were adenylated at their 3' ends prior to sequencing adaptor addition to fragments ends by ligation followed by polymerase chain reaction (PCR) and purification. DNA fragments of the library were amplified by PCR using adapter-specific primers and normalized to 10 nM in the Diluted Cluster Template to form clusters. The cDNA and gDNA libraries were sequenced using Illumina HiSeq 2000 at the IRIC genomics core facility.

Data analysis

Read mapping was performed as previously described by Simon et al. (2012)¹⁶⁴. A step-wise approach of filtering was taken for the analysis of the large number of SNPs that were detected. The analysis was focused on non-synonymous, coding variants in expressed genes (minimum expression level in reads per kilo base per million (RPKM) > 5) that had a minimum coverage level (depth > 20 reads covering variant). Systematic read mapping biases were eliminated by excluding variants located within regions with duplicated or repetitive sequence. Lastly, only variants which appeared consistently in the gDNA and the RNA were included in the list of variants used for GO term annotation. All variants were categorized as being either known (present in dbSNP Human Build 1.3.5 database; 4138 variants) or novel (absent from dbSNP; 60 variants). The potential damaging effects of missense SNPs were accessed using Variant Effect Predictor (VEP) software from the Ensembl website (<http://useast.ensembl.org/tools.html>) which combines the results of

three published algorithms: PolyPhen, SIFT and Condel. For all known or novel SNPs, only variants that were predicted to be damaging or deleterious by at least two of the three predicting tools of the VEP were retained for further analysis.

Luciferase assay

An hg19 reference *GATA2* cDNA was PCR amplified from human AML cDNA while the CG-SH isoform of *GATA2* was PCR amplified from a CG-SH cDNA. All *GATA2* variants were cloned into pcDNA3.1 and published pcDNA PU.1, and the *GATA2*-responsive promoter (LYL1) in PGL4.12 was kindly provided by the Scott lab¹⁶⁵. Cos-7 cells were transfected at 90% confluence with Lipofectamine 2000 and were harvested and used for luciferase assays after 20h using a Veritas Microplate Luminometer (Tunner Biosystem). The luciferase assay results were reported as relative light units (RLU) of firefly luciferase activity. Experiments were performed in triplicate.

Time course and genes expression with cytokines

RNA-seq expression (RPKM) values from the samples treated with or without cytokines were used to calculate either a simple ratio (fold change for +cytokines/-cytokines) or were analyzed in R using the DESeq2 package. Threshold used for determining significantly differentially expressed genes were $\text{padj} < 0.05$ (DESeq2 package) or a fold change over 2 and under-expressed when the fold change was under 0.5 (direct RPKM ratios).

Cell death rate analysis

CG-SH cells were cultured at 37°C for 7 days in the presence or absence of cytokines in the media and were harvested and labeled after 3 days, 5 days and 7 days of cytokine exposure. Cells were stained with annexinV–fluorescein isothiocyanate (FITC) (BD Bioscience) and propidium iodide (PI) to distinguish populations of early apoptotic (annexinV+/PI-), late apoptotic (annexinV+/PI+), and necrotic (annexinV-/PI+) cells. The results were analyzed by flow cytometry using the BD FACSCanto II flow cytometer (BD Bioscience).

Results

Mutations in genes previously implicated in AML

Because the CG-SH cell line had recently been derived from an NK-AML patient and still maintained a normal karyotype in culture, we hypothesized that we might be able to use the whole genome and transcriptome sequence data to identify the initial causes of leukemogenesis in the patient. We specifically examined 14 genes known to be frequently mutated in NK-AML (including *CEBPA*, *DNMT3A*, *FLT3* and *NPM1*). In agreement with the initial targeting sequencing reported for CG-SH¹⁸², our results showed the presence of five known missense SNPs (annotated in dbSNP135 database) in four of these genes (**Table 2.1**), but no novel mutations in any of these genes. The five missense annotated SNPs in *FLT3*, *NRAS*, *RUNX1* and *TET2* were all predicted to be deleterious to the protein function with a high probability using the VEP software (http://useast.ensembl.org/info/docs/variation/vep/vep_script.html), including the two

variants in *NRAS* and *RUNX1* previously identified as “mutations”¹⁸². In addition to single nucleotide sequence variants, we also analysed small insertions and deletions (indels) in the CG-SH cell line. A homozygous single base pair insertion was detected at the beginning of the *EZH2* gene (**Figure 2.1**) leading to a shift of reading frame and a loss of the SET functional domain. In addition, while validating a novel mutation in *GATA2* we also identified a duplication and insertion of a 355 bp fragment of the *GATA2* gene, including part of the last intron and exon (**Figure 2.2**). This indel results in the loss of the last 20 amino acids of the protein due to a loss of reading frame.

GATA2 is mutationally and epigenetically inactivated in CG-SH cells

Our RNA-seq data and genomic DNA sequencing identified a novel mutation and rearrangement in the *GATA2* gene, however while the gDNA data clearly showed a heterozygous mutation, the transcriptome showed a homozygous mutation at the same position. A detailed examination of SNPs in the *GATA2* locus revealed that this gene exhibits allele specific expression (ASE) in CG-SH cells. The *GATA2* gene exhibits a high level of DNA sequence conservation between species and therefore the loss of the last exon suggested that this would have a deleterious impact on gene function. In order to verify this hypothesis, we performed in vitro studies of *GATA2* activity using a luciferase reporter assay. Results of the luciferase assays (**Figure 2.3**) verified that the mutated isoform of *GATA2* present in CG-SH has a complete loss of function, as the activity of the mutated form of *GATA2* from CG-SH cDNA is reduced by 7-fold compared to the wild-type. Because *GATA2* exhibits ASE, and the mutated and rearranged allele of *GATA2* has lost its activity, the cell line has the equivalent of a homozygous loss of function mutation. A more

detailed analysis of the molecular mechanisms of this regulation of *GATA2* and its implications for AML has recently been analyzed in a separate study ¹⁸⁶.

Damaging Variants in other genes

In addition to the mutations in *EZH2* and *GATA2*, we hypothesized that other genes carrying novel mutations could also be involved in the leukemogenesis of CG-SH. Among the detected novel variants (not annotated in dbSNP135) we detected, 36 were predicted to be damaging by the VEP software (**Supplementary table 2.1**). In order to substantiate a potential role for these putatively damaging variants in AML we looked in the Catalogue Of Somatic Mutations In Cancer (COSMIC) database for evidence of mutations in other cancer types. This analysis revealed that among the 36 genes carrying novel SNVs, 9 are in genes implicated in AML or mutated in other types of cancer (**Table 2.2**). Because the combination of novel and known SNPs could participate to leukemogenesis in this case, we combined the list of 36 genes carrying damaging novel variants with the list of 296 non-AML genes carrying known SNPs predicted to be damaging and used the resulting list for characterization. Using a supervised Gene Ontology term (GO-term) analysis in AmiGO (<http://www.geneontology.org/GO.cite.shtml#data>), we compared our list of genes with 3 defined GO-term lists enriched in genes involved in three major biological functions associated with AML development: myeloid differentiation, proliferation and apoptosis. The genes which were in common between our list and the GO-terms lists are summarized in (**Table 2.3**). On the one hand, we found that a few novel variants are involved in rare AML cases or in other cancer cases. On the other hand, some of the genes with novel and known SNPs are implicated in important functions of leukemogenesis, suggesting that a

few genes, other than the ones frequently mutated in AML, could participate to AML development in the CG-SH cell line.

Normal expression levels of AML genes

Because the deregulation of normal gene expression can play an important role in the AML^{187,188}, we investigated whether abnormal expression, rather than mutations, of AML-associated genes could have been involved in the origin of the CG-SH leukemia. The expression levels (RPKM) of 10 genes recurrently mutated in AML were analyzed relative to all expressed genes in CG-SH. Our results suggest that none of these genes are silenced in this cell line (**Figure 2.4A**). We then hypothesised that if genes critical for AML were partially silenced in CG-SH, they might represent outliers in the distribution of gene expression values even if compared to other AML samples. We therefore searched for expression abnormalities in the AML genes by comparing the expression level of each of the 14 AML genes in CG-SH cells with their expression in 55 AML patients (**Figure 2.4B**). Once again, we saw no obvious difference in the expression level of these genes, suggesting that changes in gene expression levels are unlikely to be a part of the molecular mechanism of leukemogenesis in this cell line.

The effect of cytokines on gene expression

In the course of growing the CG-SH cells for study, we examined the effects of adding cytokines to the growth media. In particular, the combined action of FLT3-ligand (FL), stem cell factor (SCF), interleukine-3 (IL-3), and granulocyte-colony stimulating factor (G-CSF) was previously shown to increase the proliferation and long term survival of normal and leukemic precursor cells^{189,190}. We hypothesized that CG-SH cells, which grow slowly,

might also respond to this cytokine mix when grown *in vitro*. To test our hypothesis, we grew triplicate samples of CG-SH cells in the presence of the cytokine mix for 6 days, along with control cells grown in the absence of cytokines. Cell counts taken every 12 hours showed that there was a faster growth of the cells grown with cytokines compared to cells grown without cytokines during the exponential phase (**Figure 2.5**). Using the R software, we calculated the doubling-time of the cells using an exponential model of growth. These results showed a shortened doubling time of CG-SH cells grown with cytokines (~50 hours), in comparison with cells grown without cytokines (~75 hours).

We next sought to investigate the molecular effects of the cytokines on CG-SH cells at a molecular scale, since we reasoned that the gene expression changes caused by the addition of cytokines might provide some insight into the signalling pathways required for leukemic growth. We collected and sequenced total RNA from CG-SH cells grown with and without cytokines for more than two weeks. The expression data (RPKM values) was generated after read mapping and alignment with the Illumina Casava 1.8.1 package. As expected, we observed the differential expression of a large number of genes between the cells grown in the presence or absence of cytokines (**Figure 2.6**). To further understand the changes in gene expression triggered by cytokines, we compared their global gene expression profile with RNA-seq other AML cell lines (THP-1, HL60, and KG1a) generated in the lab and NK-AML patient samples. Through hierarchical clustering (using the “pvclust” R package), we observed that the expression profile of the CG-SH cells grown in normal media grouped with the NK-AML patients, whereas the expression profile with cytokines clustered more closely with the other cell lines (**Figure 2.7**).

Not only CG-SH cells proliferate faster in the presence of cytokines in their media, but also their expression profile is shifted towards other immortalized cell lines rather than the primary patient samples. These results suggest that cytokines promote an “immortalized cell line-like” behavior of the CG-SH cell line.

Go term enrichment analysis

In order to investigate the biological function of the genes differentially expressed after long term exposure to cytokine (more than two weeks), we performed a GO-term analysis of these genes using the DAVID analysis website (<http://david.abcc.ncifcrf.gov/>). We first generated a ranked gene list by measuring differential expression in presence/absence of cytokines (fold change) and then arbitrarily selected the first 150 upregulated genes (log2 fold change > 3.7) for analysis in DAVID. With this list we observed enrichment in AML-relevant GO-terms such as regulation of cell death (pval < 0.05) and cell cycle process (pval < 0.05). This analysis also revealed an enrichment in genes associated with regulation of myeloid cell differentiation (pval < 0.05) suggesting the genes whose expression increased did indeed have relevance for the behaviour of the cell line.

Time course RNA-seq analysis

Our analysis of the effects of cytokines on the growth CG-SH cells was performed after a long exposure period of two weeks. We hypothesized that the initial exposure to cytokines might trigger transient transactional effects that could be masked by secondary events after long-term exposure. We therefore studied the short term effects of cytokines exposure on gene expression levels by performing RNA-seq after one or three days of cytokine exposure. The RNA-seq data was processed using the DESeq2 R package. The

expression data of the early time points (day_1 and day_3) was compared to untreated cells as well as data from the long-term cytokine exposure (two weeks; day_14). As expected, the number of differentially genes increases with the time of cytokine exposure, and the number of down-regulated genes varies very little between day_3 and day_14 (**Supplementary figure 2.1A, Supplementary figure 2.1B**). We observed that 26 genes are constantly up-regulated (**Supplementary figure 2.1A**) and one gene was detected to be constantly down-regulated (**Supplementary figure 2.1B**) at all cytokine treated time points examined.

We again performed an unsupervised analysis using DAVID focusing specifically on clusters representing leukemia-promoting biological functions. The analysis of ranked list of genes up-regulated with cytokines revealed a significant enrichment (pval < 0.05) in the “negative regulation of apoptosis” GO-term at day1 time point. A similar enrichment was seen in the same cluster at the day_3 (not statistically significant) and day_14 (pval < 0.00005) timepoints. The genes constituting this GO-term cluster vary between the three time points with an increased total number of genes in the cluster with increasing time of cytokine exposure (**Supplementary figure 2.2**). We likewise found an enrichment in GO-terms associated with the differentiation of hematopoietic cells at every time point and an increasing large number of related terms as a function of time of exposure to cytokines (**Supplementary table 2.2**) and a similar pattern was seen for GO terms related to cell proliferation (**Supplementary table 2.3**). We also examined genes whose expression was down regulated when cytokines were added. In this case, we observed that the term hematopoietic differentiation was over-represented in gene lists from the day_1 and

day_3 time points (**Supplementary table 2.4**) however no enrichment was observed for the day_14 time point.

In conclusion, we found that cytokines induce changes in the expression of genes involved in leukemogenesis functions (essentially through the up-regulation of genes implicated in negative regulation of apoptosis, hematopoietic differentiation and proliferation) gradually over time.

Effects of cytokines on CG-SH apoptosis

Our experiments showed a reduced doubling time of CG-SH cells grown in culture with cytokines and our GO-term analysis of RNA-seq data suggested that cytokine treatment modulates the expression of genes involved in apoptosis. We therefore hypothesized that cytokine treatment reduces the rate of CG-SH cell death leading to a reduced doubling time. In order to test this hypothesis, we used a double staining with annexinV-FITC and PI. This labeling was designed to differentiate early apoptotic (annexinV+/PI-), late apoptotic (annexinV+/PI+), and necrotic (annexinV-/PI+) cells. To monitor the effects of cytokines on cell death, the CG-SH cells were stained at different time points (3 days, 5 days and 7 days after cytokine exposure) and examined by fluorescence-activated cell sorting (FACS) analysis. At day 3, our results showed 10% fewer PI positive cells with cytokine treatment compared to the control cells (**Figure 2.8a**). There was no significant difference in the frequency of CG-SH cells stained with AnnexinV-FITC between the two conditions, and the 10% difference in the unstained living cells seems to reflect the variation in the rate of PI stained cells (**Figure 2.8a**). At later time points the difference in necrosis rates that we observed in day 3 tended to diminish (**Figure 2.8b**, **Figure 2.8c**),

and no significant difference was observed at day 7 (**Figure 2.8c**). It seems that the decrease in necrosis rates rather than the apoptotic rates participates to the acceleration of CG-SH cells growth in the presence of cytokines.

Discussion

In this study, we provide a comprehensive description of the molecular genetics of a novel normal karyotype AML cell line, CG-SH. While previous studies aimed at defining the genetic background of the AML cell lines used either G banding and FISH to visualize karyotype abnormalities¹⁷⁹⁻¹⁸¹ our study uses next-generation whole genome sequencing to define the genetic background of an AML cell line. This technique has allowed the discovery of new genes which are recurrently mutated in NK-AML cases^{62,87,107}. We have focused our analysis specifically on two aspects: identifying genetic alterations which could participate in leukemogenesis and the ability of the cell line to respond to a mix of growth factors. To our knowledge, CG-SH is one of the only two AML cell lines which have normal karyotypes, the second being an acute erythroblastic leukemia cell line derived from a very rare form of AML (M6)¹⁹¹. Therefore, the CG-SH cell line represents a rare and valuable model system to study the biology of NK-AML which represents almost half of all patients' samples.

In order to understand the causes of AML in the case leading to the creation of this cell line, we initially examined genes which had previously been reported to be frequently mutated in AML (>20% of NK-AML cases for *NPM1*, *FLT3*, *DNMT3A*, *TET2*, 20-10% AML cases for *NRAS*, *RUNX1*, *CEBPA*, *IDH1/2*, *JAK2*, *MLL*, *TP53*, and 2% of NK-AML cases for *PTEN*). In confirmation of the previously published report on CG-SH, no mutations were found in recurrently mutated genes such as *FLT3*, *NPM1*, *DNMT3A* or *MLL*. We also examined the expression level of these genes to see if transcriptional silencing rather than mutations could impact their involvement, however the expression level of all of these

targets were within the range of other AML samples, suggesting this was not a mechanism involved in the origin of the disease.

We broadened our analysis to look at all missense/nonsense mutations in expressed genes, and specifically focused on those variants which were not in dbSNP135. These results identified a novel mutation in the *GATA2* gene, which has been implicated in familial MDS/AML and other hematopoietic diseases¹⁴. When we examined the *GATA2* gene more closely we observed that these novel mutations, in addition to all over SNPs within the locus, exhibited allele specific expression. Furthermore, the through cloning and sequencing of the *GATA2* cDNA expressed in CG-SH, we identified a 355 bp duplication and reinsertion into the last exon of the gene. The mutation and rearrangement of the last exon of *GATA2* combined with its allele specific expression would suggest that CG-SH has a complete loss of function of *GATA2*. We verified this hypothesis using luciferase reporter assays which showed that the version of *GATA2* expressed in CG-SH has no transactivational activity.

We further examined the NGS data for insertions and deletions and we identified a single basepair insertion in *EZH2*. The presence of this insertion in *EZH2* at the beginning of the genes results in the loss of the proper reading frame, leading to a highly truncated and likely defect protein being produced. The involvement of *EZH2* in AML through a similar mechanism has been previously reported ¹⁸⁵, suggesting that this heterozygous mutation is likely to be pathogenic.

Although the mutations in *GATA2* and *EZH2* might be sufficient to explain the leukemogenesis of CG-SH, we also analyzed all other coding SNPs which had the potential

to biologically impact cell. In order to identify those most likely to impact cellular function we used the Variant Effect Predictor software (VEP) to find SNPs which were likely to alter the proteins activity. By filtering the entire list of SNPs (165 356) for those which were predicted to be deleterious by two of the three algorithms of VEP, we were able to analyze a small subset of genetic variants (36 novel SNPs and 301 annotated SNPs). Using this filtered list of candidates, we looked for evidence in published studies for the implication of genes carrying novel SNPs in AML or cancer. Nine of the mutated genes were reported to be involved in some AML cases (more rarely than most AML-associated genes) and mutated in other cancers (present in the COSMIC database). These results suggest that these novel and known SNPs could potentially have participated in the development of this AML cell line. This hypothesis is supported by the fact that when performing a supervised GO-term analysis with AmiGO, the list of combined genes carrying damaging known and novel SNPs overlaps with GO-terms representing biological functions important for leukemia development. Although this overlap is not statistically significant, it is likely that only very few mutations in genes involved in key biological processes for leukemia trigger the development of the disease. Indeed it is well established that leukemic cells are affected in their self-renewal, apoptosis and myeloid differentiation processes ⁹. The exact combination of mutations detected in CG-SH cells that might have been required to generate this leukemia is not clear and cannot be deduced from this single sample. This study does, however, provide a useful overview of the magnitude of genetic variants that could influence this process in an NK-AML patient.

The majority of the newly established AML cell lines are characterized by their ability to respond to cytokines proliferation effects *in vitro*. For instance, the established cell line Kasumi-1 was reported to have increased colony formation in the presence of G-CSF, IL-3, or IL-6 in the media¹⁹². We initially sought to improve the proliferation of the CG-SH cell line using a cytokine mix, however we also postulated that any resulting change in growth might provide insight into factors regulating the growth of AML cells. For these experiments, we used a cytokine mix (IL3, SCF, FL and G-CSF) that has been reported to enhance proliferation and survival of myeloid cells^{189,190}. We demonstrated that CG-SH cells are sensitive to this cytokine mix, as their doubling-time is shortened by ~1.5 fold. Thus the CG-SH cell line could be used as a model for the study of the effects of cytokines or growth factors on NK-AML biology. To date, there have been a limited number of studies on the biological effects of cytokines, particularly using high-throughput assays to examine gene expression changes occurring after cytokine exposure. In the present study, we aimed to understand the effects of cytokines on the CG-SH cell line behavior by studying the changes in the global gene expression profile. We initially observed that hierarchical clustering of samples based on gene expression profiles showed that CG-SH grown without cytokines fell into the group of NK-AML samples, while after long term exposure to cytokines (more than two weeks) it fell closer to other cytogenetically abnormal AML cell lines. Interestingly, the growth of AML patient samples in culture is typically very poor compared to established cell lines¹⁷⁸, similar to CG-SH cells grown in the absence of cytokines. An unsupervised functional annotation analysis performed using DAVID with differentially expressed genes at day+14 revealed an up-regulation of genes involved in the proliferation, regulation of apoptosis, regulation of differentiation

and a down-regulation of genes involved in differentiation. These results were consistent with what we observed for the cytokine time course, where the GO-term analysis showed enrichment for the same biological functions and from earlier time points (day1 and day3). Our GO-term analysis of genes up-regulated after cytokine treatment suggested that cytokines might be affecting cellular behaviour by inactivating apoptotic pathways. In order to test this hypothesis, we stained the CG-SH cells with AnnexinV-FITC and PI after 3 days, 5 days and 7 days of growth with cytokines. Surprisingly, we did not see any significant change in the apoptotic rate (AnnexinV+/PI- and AnnexinV+/PI+) of the cells grown with cytokines compared to without cytokines. We did however observe a significant decrease of the necrosis rate (AnnexinV-/PI+) at day 3 of the cells with cytokines and this difference disappeared gradually until 7 days. The transient nature of the effects of cytokine treatment on cell death suggest that the modulation of expression of genes involved in proliferation rather than alterations in apoptosis or necrosis may not be involved in the increase growth of CG-SH cells with cytokines.

In summary, our work characterizing the precise molecular defects present in this novel NK-AML cell line will enable its use as a defined model system to study the biology of NK-AML *in vitro*. Moreover, this study has provided additional insight into the genetic and molecular determinants of NK-AML and on the effects of cytokine treatment on global gene expression profiles in the context of NK-AML.

Acknowledgements:

The authors thank Reinhold Munker (Louisiana State University Health Sciences Center, Shreveport, United States) for providing the CG-SH cells for the experiments. We also thank members of Guy Sauvageau's laboratory (IRIC, University of Montreal, Montreal, Canada) for technical support.

This work was supported by funding from the Cole Foundation (BTW) and Genome Quebec (BTW).

Authorship contributions:

B.T.W. designed and supervised the experimental work while G.G., A.F. and M.C. performed the experimental work presented. G.G., M.C. and B.T.W. analyzed the data. G.G. drafted the paper with revisions by B.T.W and M.C.

Disclosure of conflicts of interest:

The authors declare that they have no competing financial interests.

Figure legends:

Figure 2.1: Single base pair insertion in *EZH2* and loss of the SET domain. (A) Illustration of the *EZH2* gene structure and zoom in on the single base pair insertion occurring in the exon 2. The picture of the insertion represents a partial view of *EZH2* with the Integrative Genomics Viewer (IGV) tool including the gene portion (in blue at the top) total reads count (in grey at the top), individual mapped reads (horizontal grey bars), the nucleotide sequence (colored bases at the bottom) and the insertion (indicated by the black arrow). (B) Illustration of the wild-type *EZH2* protein with the functional domains, and the truncated protein below with the insertion point represented (black arrow).

Figure 2.2: Duplication/insertion in *GATA2* and loss of the protein end. (A) Illustration of the *GATA2* gene structure and zoom in on the duplication/insertion occurring in the last exon and intron. The picture represents a partial view of *GATA2* with the Integrative Genomics Viewer (IGV) tool including the gene portion (in blue at the top) total reads count (in grey at the top), individual mapped reads (horizontal grey and blue bars), the nucleotide sequence (colored bases at the bottom) and the insertion (indicated by the black arrow). The dark blue reads represent misplaced reads mapped on the genome. The duplicated portion of the gene is represented in a red frame and the location of its reinsertion in a red arrow. (B) Illustration of the wild-type *GATA2* protein with the functional domain, and the truncated protein below with the insertion point (black arrow) and the loss of the reading frame (red portion of the protein).

Figure 2.3: The activity of *GATA2* is altered in the CG-SH cells. Luciferase activity was measured in relative light units (RLU) 20h after cotransfection of Cos-7 cells with PGL3 promoter, PU.1 expression construct, and wild-type, or CG-SH mutant expression constructs or empty vector. We used Student's t-test to compare the wild-type and the CG-SH mutant. * Indicate a significant difference between two samples (p -value < 0.05, $n=3$).

Figure 2.4: Analysis of expression levels in the CG-SH cells of the genes recurrently mutated in AML. (A) Frequency plot showing the distribution of expression values (in RPKM) of all the genes in the CG-SH cells. On the Y axis are the frequencies and on the X

axis the log₂ values of expression in RPKM. The red lines represent the expression levels of the recurrently mutated genes in AML in which no genetic variant has been detected in the CG-SH cells. (B) Boxplot showing the expression value of the genes recurrently mutated in AML in the CG-SH cells (red crosses) compared to the distribution of expression values for 55 AML patients, with mean expression value (in box black line) and error bars. The values on the Y axis represent the log₂ of RPKM values. Gene names are mentioned on the X axis.

Figure 2.5: CG-SH cells grow faster in the presence of cytokines. CG-SH cells growth in the presence of cytokines is represented by the exponential model in a dashed line fitted to the cell counts (in triangle dots). The control cells growth without cytokines is represented by the exponential model in a solid line fitted to the cell counts (in round dots).

Figure 2.6: Differential expression in the CG-SH cells cultured with and without cytokines. MA plot of gene expression levels (in RPKM) in the CG-SH cells grown in the presence of cytokines compared to CG-SH grown in the absence of cytokines. The values on the Y axis represent the ratio of presence/absence of cytokines log₂ fold change. The values on the X axis represent the average log₂ intensity for RPKM value.

Figure 2.7: Global expression profiles of the CG-SH cells versus other AML cell lines and patient samples. This unsupervised hierarchical cluster dendrogram shows clustering of global gene expression patterns for AML cell lines: HL60, KG1a, Thp1 and CG-SH, and NK-AML patient samples (e.g. 11H009). It provides AU (approximately unbiased) p-values (in red) as well as BP (bootstrap probability) values (in green) computed via multiscale bootstrap resampling. The height of the branches represented by the left axis is proportional to the distances between samples. Indicated are the CG-SH cells grown with cytokines (CG-SH_plus) and without cytokines (CG-SH_minus).

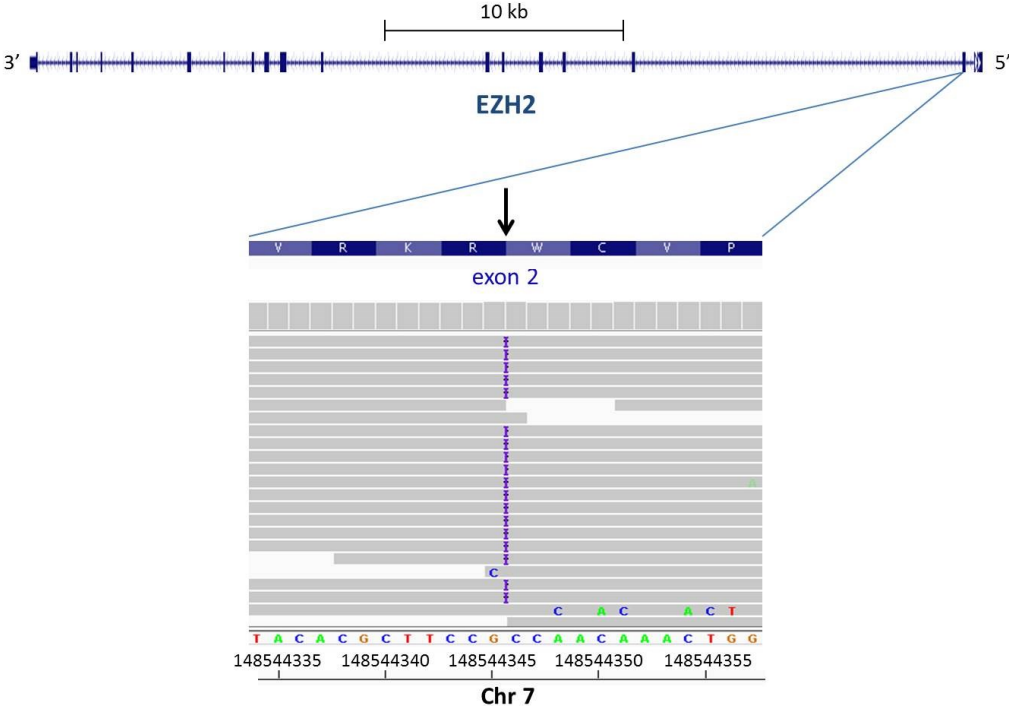
Figure 2.8: Flow cytometry analysis of cytokines effects on the CG-SH cells death. CG-SH cells were cultured for 7 days in a liquid media supplemented with cytokines (presence of cytokines) or not (absence of cytokines) and harvested at different time points. Staining with annexinV-FITC and PI was performed to distinguish early apoptotic

(annexinV), necrotic (PI), and late apoptotic (annexinV+PI) cell populations. (a) The cells were harvested at day 3. (b) The cells were harvested at day 5. (c) The cells were harvested at day 7.

Figures

Figure 2.1.

A



B

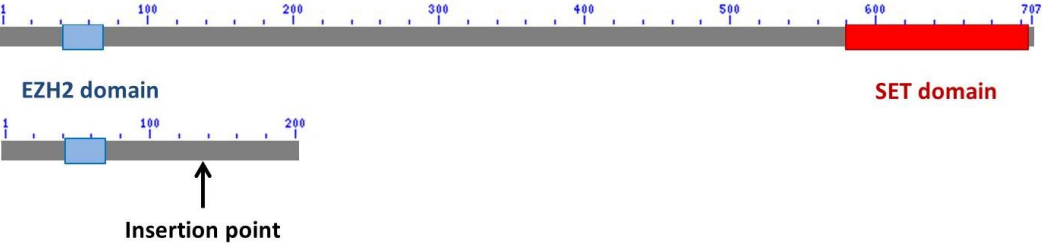
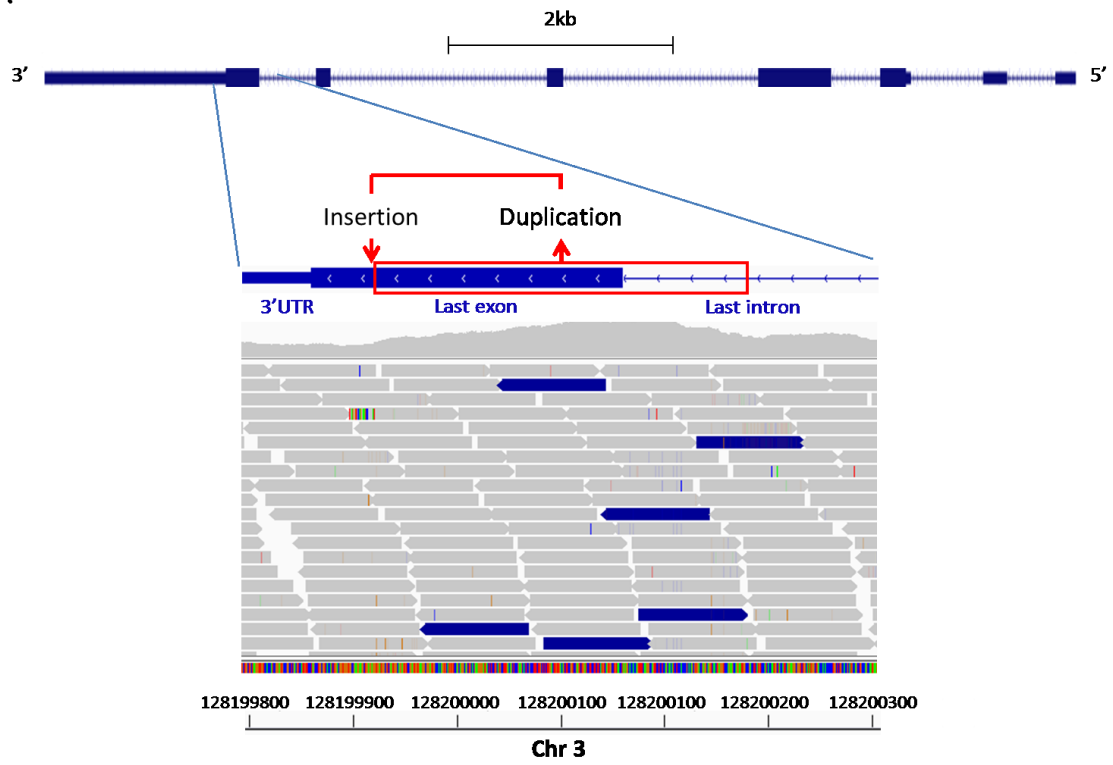


Figure 2.2

A



B

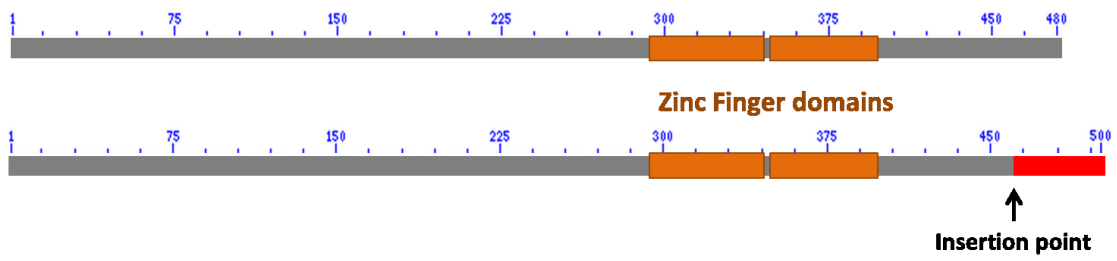


Figure 2.3

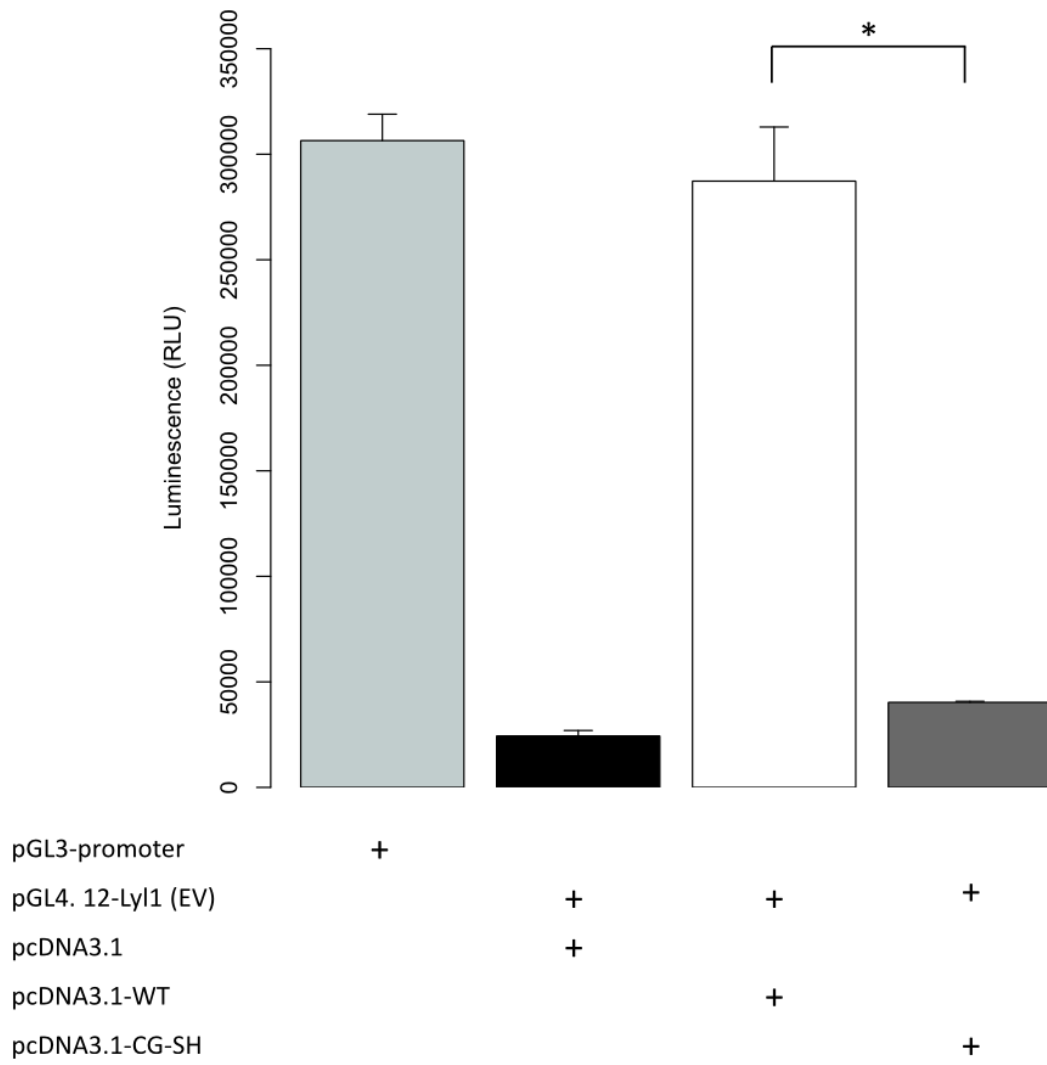


Figure 2.4

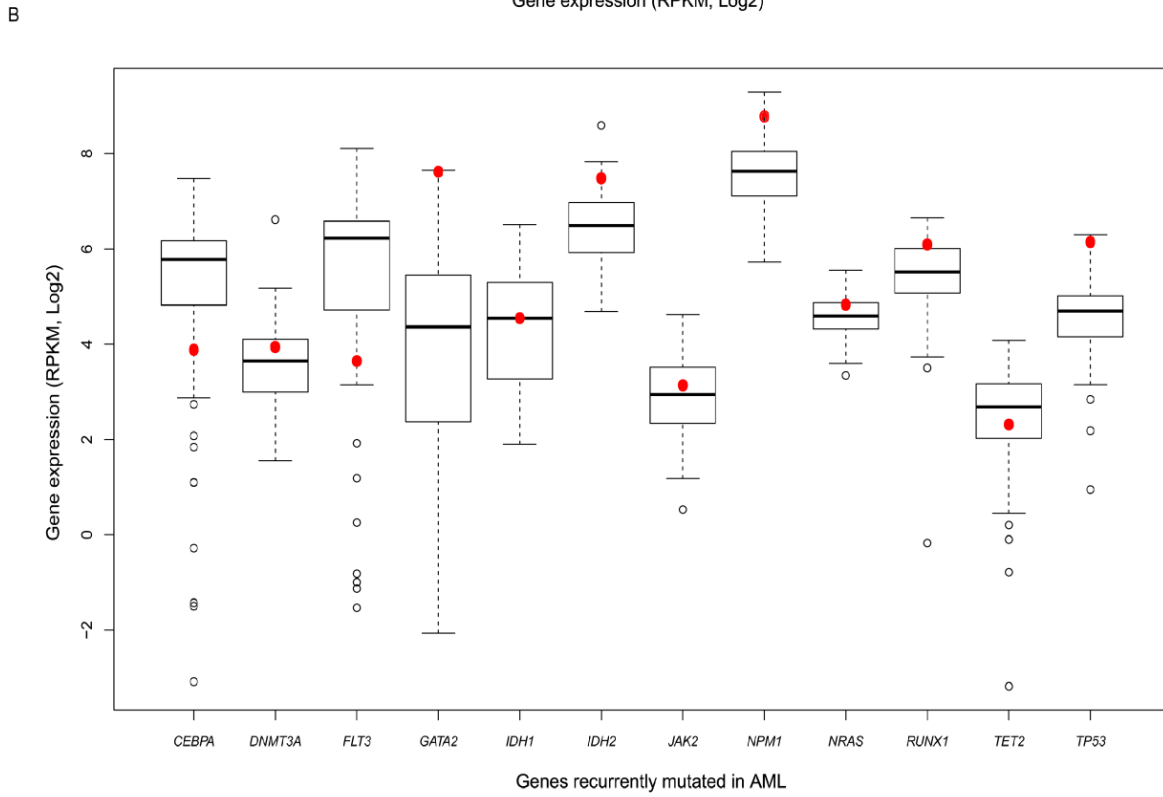
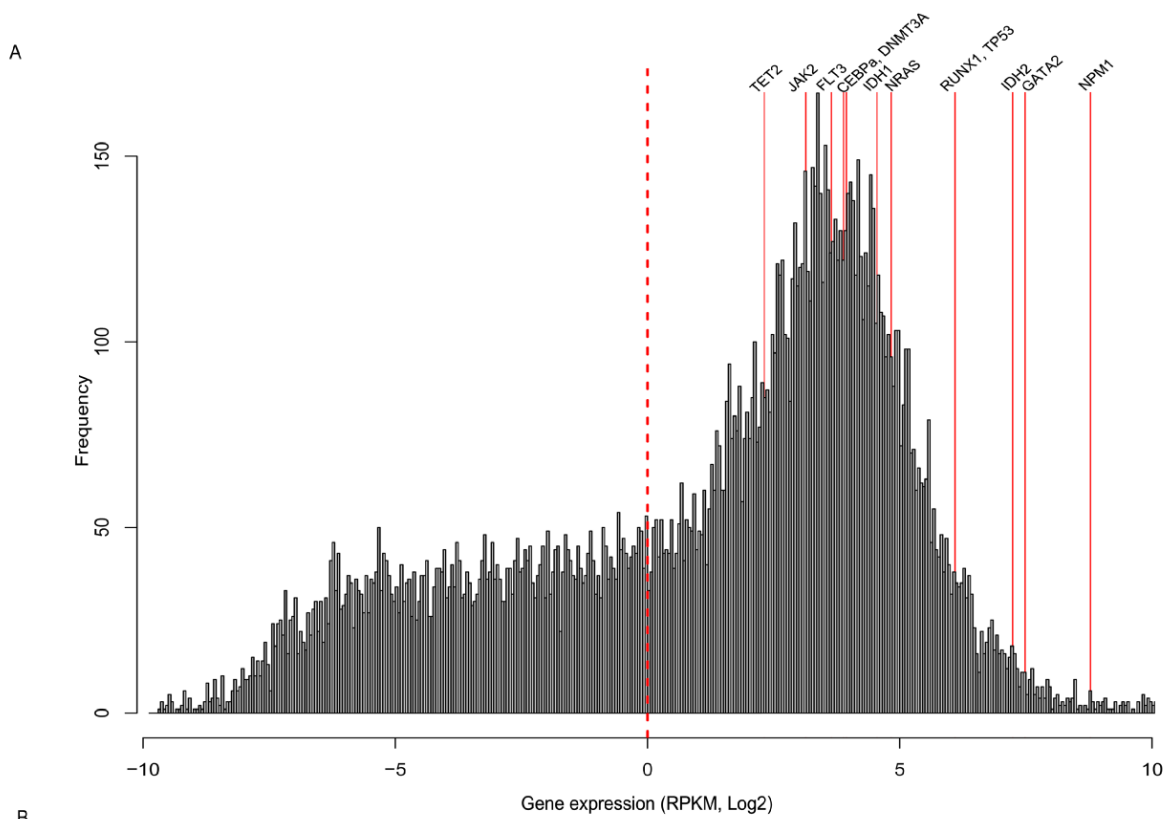


Figure 2.5

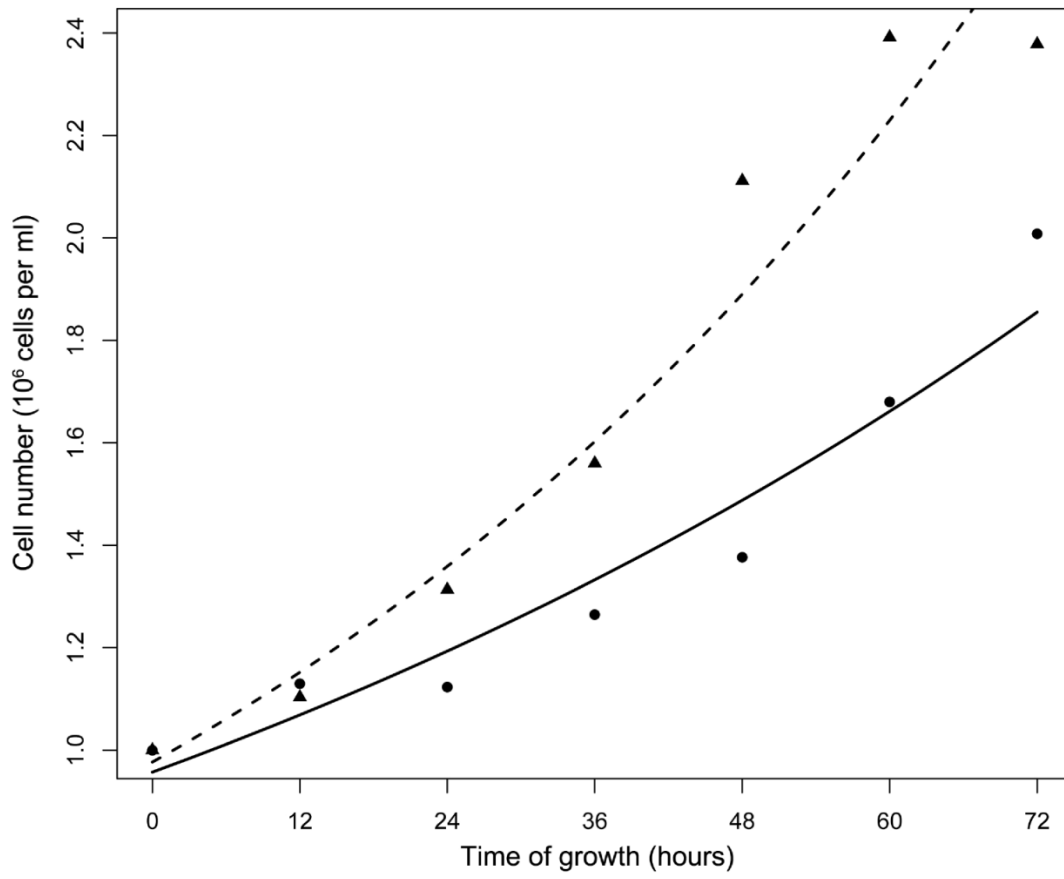


Figure 2.6

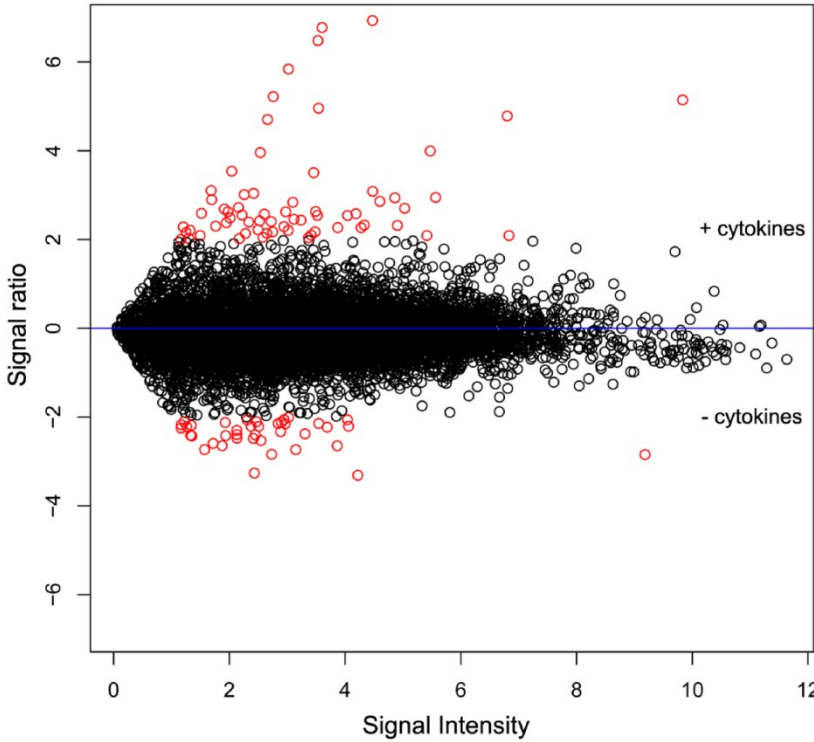


Figure 2.7

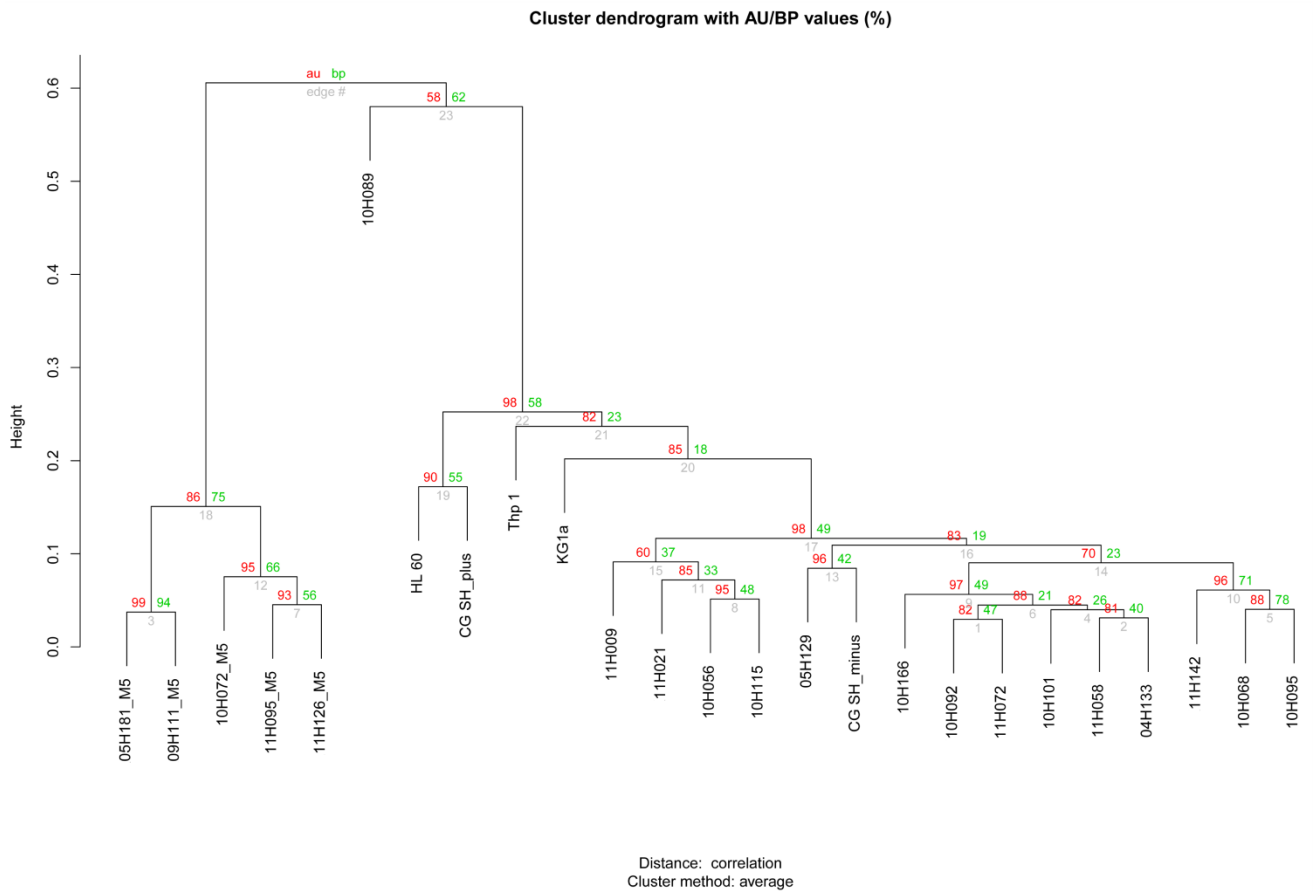
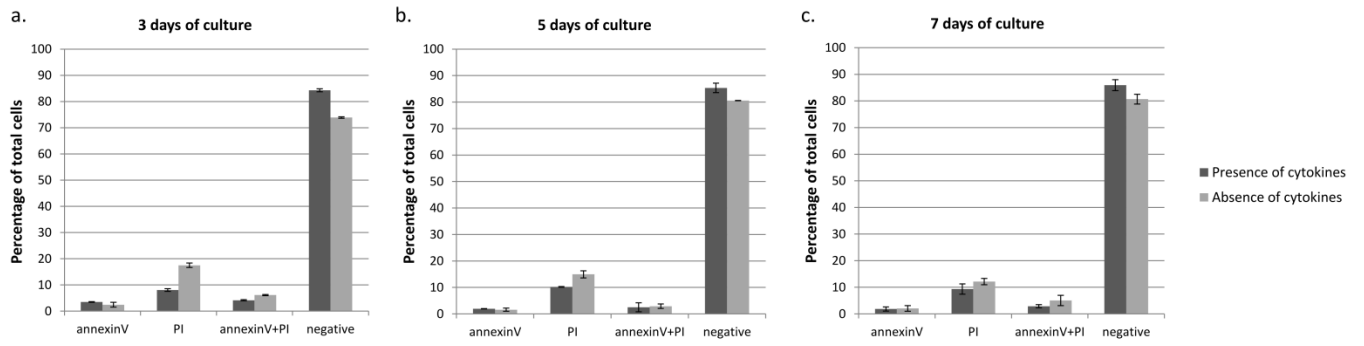


Figure 2.8



Tables

Table 2.1: Genes recurrently mutated in AML carrying damaging annotated SNPs in the CG-SH cells.

Gene symbol	Mutation frequency in NK-AML *	gene expression level (RPKM)	SNPs annotated in dbSNP135	Transcript feature†	VEP prediction		
					PolyPhen‡	Condel‡	SIFT‡
<i>FLT3</i>	36-50% ¹⁹³	12,5	rs1933437	ENST00000380987	probably_damaging	deleterious	deleterious
<i>NRAS</i>	6% ¹⁹³	28,5	rs12143459 5	ENST00000369535	probably_damaging	deleterious	deleterious
<i>RUNX1</i>	14.8% ¹⁹³	68,3	rs11152773 8	ENST00000300305	probably_damaging	deleterious	tolerated
<i>TET2</i>	23-26% ^{193,194}	5,0	rs17253672	ENST00000545826	probably_damaging	deleterious	tolerated
			rs34402524	ENST00000540549	possibly_damaging	deleterious	deleterious

* Column including the numbered references in which are published the frequencies.

† Ensembl feature used for Variant Effect Predictor (VEP) analysis.

‡ Qualitative predictions of the impact of missense SNPs from the three algorithms in VEP

Table 2.2: Genes with novel damaging variants in the CG-SH cells reported to be implicated in AML and cancer in the literature.

Gene symbol	Biological function (from NCBI)	References for potential implication of the gene in AML	COSMIC cancer mutations (Mutated samples/ total samples)*
<i>CCNB1</i>	cell cycle regulatory protein involved in mitosis	195	1/7 skin
<i>CCND1</i>	cell cycle regulatory protein involved in G1/S transition	196,197	2/2 central nervous system
<i>SETDB1</i>	histone methyltransferase which regulates histone methylation, gene silencing and transcriptional repression	198	3/3 ovary, 1/1 nervous system, 1/3 upper aerodigestive tract, 2/8 skin
<i>SS18L1</i>	gene originally described in synovial sarcoma to translocate and fuse with SSX genes (function not well known)	199	1/7 skin
<i>TMEM2</i>	—	200	4/4 prostate, 2/2 ovary, 1/1 skin, 1/1 stomach, 1/1 kidney
<i>TRAF2</i>	mediator of the anti-apoptotic signals from TNF receptors	201,202	1/4 upper aerodigestive tract
<i>WWOX</i>	protein degradation, transcription and RNA splicing	203	1/1 upper aerodigestive tract
<i>GATA2</i>	regulates transcription of genes involved in the development and proliferation of hematopoietic and endocrine cell lineages	175,204,205	3/3 prostate, 1/6 skin
<i>USP42</i>	—	183	3/7 skin

* Results from the Catalogue Of Somatic Mutations In Cancer (COSMIC) including the ratio of mutated samples on total samples tested. Are showed only the results for which the ratio is above 1/7.

— indicates unknown function.

Table 2.3: Mutated genes in the CG-SH cells are associated with biological functions involved in leukemia development.

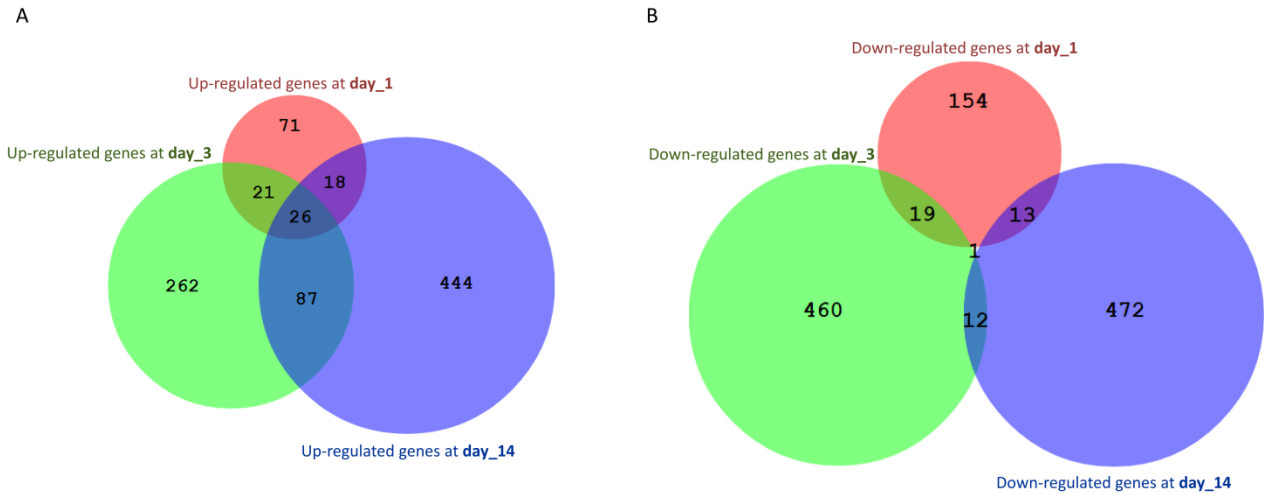
GO terms*	Gene symbol	Type of variant†	Gene biological function (from NCBI)
myeloid differentiation	<i>FLVCR1</i>	known SNP	heme transporter that may play a critical role in erythropoiesis
	<i>INPP5D</i>	known SNP	negative regulator of myeloid cell proliferation and survival
apoptotic signaling pathway	<i>RIPK1</i>	novel SNP	—
	<i>TRAF2</i>	novel SNP	mediator of the anti-apoptotic signals from TNF receptors
	<i>CHEK2</i>	known SNP	cell cycle checkpoint regulator and putative tumor suppressor
	<i>MLH1</i>	known SNP	locus frequently mutated in hereditary nonpolyposis colon cancer
	<i>FNIP2</i>	known SNP	—
positive regulation of cell proliferation	<i>CCNB1</i>	novel SNP	cell cycle regulatory protein involved in mitosis
	<i>CCND1</i>	novel SNP	cell cycle regulatory protein involved in G1/S transition
	<i>CNOT6L</i>	novel SNP	—
	<i>AKR1C3</i>	known SNP	aldo/keto reductase utilizing NADH and/or NADPH, may play a role in the pathogenesis of allergic diseases and in controlling cell growth and/or differentiation
	<i>DDX11</i>	known SNP	enzyme that possesses both ATPase and DNA helicase activities
	<i>IQGAP3</i>	known SNP	—
negative regulation of cell proliferation	<i>UTP20</i>	known SNP	U3 small nucleolar RNA (snoRNA) involved in 18S rRNA processing
	<i>ETV4</i>	known SNP	—
	<i>STRN</i>	known SNP	—
	<i>INPP5D</i>	known SNP	negative regulator of myeloid cell proliferation and survival
	<i>STK3</i>	known SNP	serine/threonine protein kinase activated by proapoptotic molecules, promotes the condensation of chromatin

* Annotation clusters from the AmiGO database and representing the biological functions reported to be important for leukemia development in the literature.

† “Known SNP” corresponds to a genetic variant annotated in the dbSNP135 database and “novel SNP” corresponds to a variant not annotated in this database.

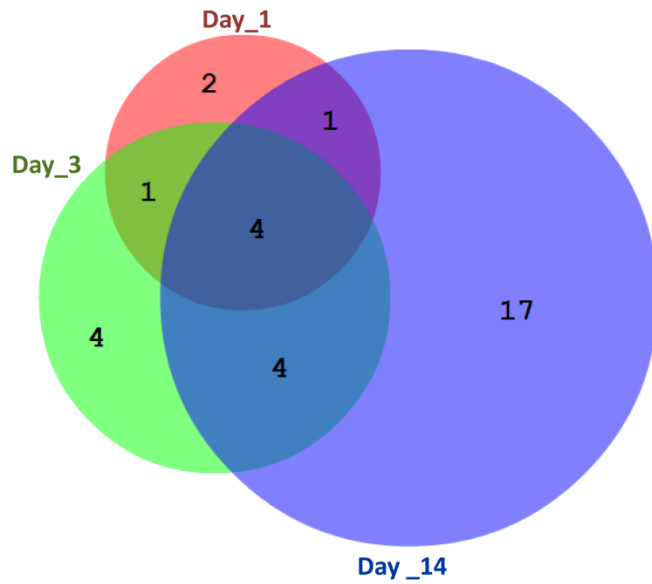
— indicates unknown function.

Supplementary figures



Supplementary figure 2.1: Time course analysis of differentially expressed genes in the presence of cytokines. The lists of differentially expressed genes are shown at day 1 (in red), at day 3 (in yellow) and at day +14 (in blue) of cytokine exposure. (A) Venn diagram including the lists of up-regulated genes with cytokines. (B) Venn diagram including the list of down-regulated genes with cytokines.

Genes involved in the negative regulation of apoptosis



Supplementary figure 2.2: Time course analysis of up-regulated genes involved in negative regulation of apoptosis. This figure is a Venn diagram showing genes present in the GO-term called “negative regulation of apoptosis” which results from a functional annotation analysis in DAVID performed with genes up-regulated in the presence of cytokines. The lists of genes are shown at day 1 (in red), at day 3 (in yellow) and at day +14 (in blue) of cytokine exposure.

Supplementary tables

Supplementary table 2.1: Identification of novel variants in the CG-SH cells in genes unknown to be frequently mutated in NK-AML.

Gene symbol	Mutation position	reference nucleotide	variant nucleotide	strand	VEP prediction			
					Transcript	PolyPhen*	Condel*	SIFT*
<i>ACLY</i>	chr17:40035155	C	T	-	ENST00000537919	probably_damaging	deleterious	deleterious
<i>ANKRD40</i>	chr17:48774393	C	A	-	ENST00000315273	benign	neutral	tolerated
<i>ASAP2</i>	chr2:9533755	C	T	+	ENST00000464683	possibly_damaging	deleterious	deleterious
<i>ATP6V0E2</i>	chr7:149571219	T	C	+	ENST00000340611	benign	neutral	tolerated
<i>BAAT1</i>	chr7:2577945	A	C	-	ENST00000239899	probably_damaging	deleterious	deleterious
<i>BRD8</i>	chr5:137504358	C	T	-	ENST00000239899	probably_damaging	deleterious	deleterious
<i>C10orf47</i>	chr10:11912164	C	T	+	ENST00000277570	probably_damaging	deleterious	deleterious
<i>C16orf48</i>	chr16:67697373	C	A	-	ENST00000243878	probably_damaging	deleterious	tolerated
<i>C16orf59</i>	chr16:2510363	C	T	+	ENST00000569496	probably_damaging	deleterious	deleterious
<i>C16orf70</i>	chr16:67173935	G	A	+	ENST00000219139	probably_damaging	deleterious	deleterious
<i>CCNB1</i>	chr5:68473121	T	C	+	ENST00000256442	possibly_damaging	deleterious	deleterious
<i>CCND1</i>	chr11:69466019	C	G	+	ENST00000227507	probably_damaging	deleterious	deleterious
<i>CEP70</i>	chr3:138289912	T	C	-	ENST00000481834	possibly_damaging	deleterious	deleterious
<i>CNOT6L</i>	chr4:78647442	T	C	-	ENST00000505983	benign	neutral	tolerated
<i>COL24A1</i>	chr1:86361735	A	G	-	ENST00000436319	benign	neutral	tolerated
<i>COL6A3</i>	chr2:238280836	T	G	-	ENST00000392003	probably_damaging	deleterious	deleterious
<i>CSGALNACT2</i>	chr10:43654185	T	G	+	ENST00000374464	probably_damaging	deleterious	tolerated
<i>DBF4B</i>	chr17:42807309	A	G	+	ENST00000525011	benign	neutral	tolerated
<i>DOM3Z</i>	chr6:31938749	C	A	-	ENST00000480240	benign	neutral	tolerated
<i>E4F1</i>	chr16:2283048	C	T	+	ENST00000565090	probably_damaging	neutral	tolerated

<i>FBXL17</i>	chr5:107716411	A	G	-	ENST0000054226 7	benign	neutral	tolerated
<i>FYN</i>	chr6:112041092	G	A	-	ENST0000052431 0	benign	neutral	tolerated
<i>GATA2</i>	chr3:128200682	G	C	-	ENST0000034110 5	probably_damaging	deleterious	deleterious
<i>GON4L</i>	chr1:155720488	C	G	-	ENST0000043780 9	probably_damaging	deleterious	deleterious
<i>HNRNPA2B1</i>	chr7:26235512	C	T	-	ENST0000035466 7	unknown	neutral	tolerated
<i>KLHL13</i>	chrX:117054206	T	G	-	ENST0000037187 8	probably_damaging	deleterious	deleterious
<i>LSM14B</i>	chr20:60705718	A	G	+	ENST0000036167 0	probably_damaging	deleterious	tolerated
<i>MAN2A2</i>	chr15:91449174	T	C	+	ENST0000056045 1	possibly_damaging	deleterious	deleterious
<i>MAP7D3</i>	chrX:135307012	A	G	-	ENST0000037066 3	probably_damaging	deleterious	deleterious
<i>MAST4</i>	chr5:66462494	C	A	+	ENST0000044380 8	probably_damaging	deleterious	deleterious
<i>MCCC1</i>	chr3:182789073	C	G	-	ENST0000039261 6	benign	neutral	tolerated
<i>MYADM</i>	chr19:54377700	C	T	+	ENST0000039177 1	possibly_damaging	neutral	tolerated
<i>NLN</i>	chr5:65058841	G	A	+	ENST0000050246 4	probably_damaging	deleterious	deleterious
<i>PCF11</i>	chr11:82880438	G	C	+	ENST0000029828 1	possibly_damaging	neutral	tolerated
<i>PHRF1</i>	chr11:610300	A	G	+	ENST0000053432 0	benign	neutral	tolerated
<i>PITHD1</i>	chr1:24112221	C	G	+	ENST0000041537 2	benign	neutral	tolerated
<i>PMPCB</i>	chr7:102949502	C	T	+	ENST0000042815 4	possibly_damaging	deleterious	deleterious
<i>POLR3A</i>	chr10:79759846	A	G	-	ENST0000054084 2	probably_damaging	deleterious	deleterious
<i>PPP1R14B</i>	chr11:64014048	C	T	-	ENST0000030931 8	probably_damaging	deleterious	deleterious
<i>PPP1R7</i>	chr2:242098741	G	C	+	ENST0000045036 7	benign	neutral	tolerated
<i>RIPK1</i>	chr6:3078060	G	C	+	ENST0000054179 1	benign	deleterious	deleterious
<i>RSPRY1</i>	chr16:57255319	C	G	+	ENST0000053786 6	benign	neutral	tolerated
<i>SETDB1</i>	chr1:150900359	C	T	+	ENST0000044802 9	probably_damaging	deleterious	deleterious
<i>SS18L1</i>	chr20:60738674	G	C	+	ENST0000042156 4	probably_damaging	deleterious	tolerated
<i>SUMF1</i>	chr3:4452568	T	C	-	ENST0000045846 5	benign	neutral	tolerated
<i>TBC1D14</i>	chr4:6925708	C	T	+	ENST0000044850 7	benign	neutral	tolerated

<i>TMEM2</i>	chr9:74319689	C	G	-	ENST0000037704 4	probably_damaging	deleterious	tolerated
<i>TMPO</i>	chr12:98921673	A	G	+	ENST0000055602 9	probably_damaging	deleterious	deleterious
<i>TRAF2</i>	chr9:139802586	G	A	+	ENST0000037164 5	probably_damaging	deleterious	tolerated
<i>TRIAP1</i>	chr12:12088414 5	A	G	-	ENST0000054695 4	benign	neutral	deleterious
<i>TSPAN32</i>	chr11:2323389	T	C	+	ENST0000038112 1	possibly_damaging	deleterious	tolerated
<i>TST</i>	chr22:37414343	A	C	-	ENST0000024904 2	possibly_damaging	neutral	tolerated
<i>TUBGCP6</i>	chr22:50671897	A	G	-	ENST0000043930 8	benign	neutral	tolerated
<i>UBE2E2</i>	chr3:23258779	C	G	+	ENST0000045289 4	probably_damaging	deleterious	deleterious
<i>USP42</i>	chr7:6194711	A	G	+	ENST0000042624 6	probably_damaging	deleterious	deleterious
<i>WWOX</i>	chr16:79245590	G	A	+	ENST0000056678 0	probably_damaging	deleterious	tolerated
<i>ZBTB7B</i>	chr1:154989003	A	T	+	ENST0000053542 0	probably_damaging	deleterious	deleterious
<i>ZNF627</i>	chr19:11728466	G	A	+	ENST0000036111 3	probably_damaging	deleterious	tolerated
<i>ZNF841</i>	chr19:52569729	G	A	-	ENST0000042639 1	benign	neutral	tolerated

* Qualitative prediction of the damaging effect resulting from the three algorithms used in VEP (PolyPhen, Condel and SIFT)

Supplementary table 2.2: GO-terms involving myeloid differentiation in the functional analysis of up-regulated genes with cytokines:

Time point	GO-term	Number of genes	% of genes	PValue
Day_1	B cell activation	4	2,96	0,019
	B cell differentiation	3	2,22	0,049
Day_3	hemopoietic or lymphoid organ development	13	3,33	0,010
	positive regulation of cell differentiation	11	2,82	0,024
	hemopoiesis	11	2,82	0,029
	negative regulation of cell differentiation	10	2,56	0,041
	myeloid cell differentiation	6	1,54	0,048
Day_14	negative regulation of CD4-positive, alpha beta T cell differentiation	3	0,54	0,008
	negative regulation of T-helper cell differentiation	3	0,54	0,008
	negative regulation of lymphocyte differentiation	4	0,73	0,010
	regulation of T-helper 2 cell differentiation	3	0,54	0,012
	negative regulation of alpha-beta T cell differentiation	3	0,54	0,012
	positive regulation of cell differentiation	14	2,54	0,017
	regulation of T-helper cell differentiation	3	0,54	0,033
	regulation of lymphocyte differentiation	6	1,09	0,037
	negative regulation of T cell differentiation	3	0,54	0,039
	hemopoiesis	13	2,36	0,045

Supplementary table 2.3: GO-terms involving cell proliferation in the functional analysis of up-regulated genes with cytokines:

Time point	GO-term	Number of genes	% of genes	PValue
Day_1	<i>no significant outcome</i>			
	cell proliferation	20	5,13	0,002
	regulation of growth	16	4,10	0,006
	regulation of cell proliferation	28	7,18	0,009
	positive regulation of cell proliferation	17	4,36	0,015
Day_3	regulation of cell growth	10	2,56	0,023
	regulation of mononuclear cell proliferation	6	1,54	0,033
	regulation of leukocyte proliferation	6	1,54	0,033
	negative regulation of cell growth	6	1,54	0,046
	positive regulation of cell division	4	1,03	0,049
	cell proliferation	25	4,54	0,002
	regulation of cell proliferation	36	6,53	0,009
	regulation of growth	19	3,45	0,012
Day_14	G1/S transition of mitotic cell cycle	6	1,09	0,024
	regulation of cell growth	12	2,18	0,027
	regulation of meiotic cell cycle	3	0,54	0,033

Supplementary table 2.4: GO-terms involving myeloid differentiation in the functional analysis of down-regulated genes with cytokines:

Time point	GO-term	Number of genes	% of genes	PValue
day 1	leukocyte differentiation	6	3,21	0,006
	lymphocyte differentiation	5	2,67	0,013
	regulation of lymphocyte differentiation	4	2,14	0,018
day 3	leukocyte differentiation	9	1,86	0,020
	B cell differentiation	5	1,03	0,034

CHAPITRE 4: DISCUSSION

À travers cette étude, nous apportons une description détaillée de la génétique moléculaire d'une nouvelle lignée cellulaire LMA, CG-SH, qui présente un caryotype normal. À notre connaissance, CG-SH est une des deux seules lignées cellulaires LMA à caryotype normal avec EEB, une lignée cellulaire à LMA érythroblastique (M6)¹⁹¹. La LMA-M6 est une forme rare de LMA ne représentant que 4% des cas²⁰⁶, contre 9-17% pour la LMA-M5²⁰⁷ dont fait partie CG-SH. Ainsi, la lignée cellulaire CG-SH apparaît comme un modèle rare et précieux pour l'étude de la LMA-CN.

Nous nous sommes intéressés particulièrement à deux aspects des CG-SH, qui sont ses variants génétiques participant à la leucémogénèse et sa capacité à répondre aux cytokines. Les études précédentes visant à caractériser génétiquement les lignées cellulaires LMA utilisent le marquage des bandes G au Giemsa et l'hybridation *in situ* en fluorescence (FISH) qui permettent de détecter les anomalies génétiques à l'échelle des chromosomes¹⁷⁹⁻¹⁸¹, ou la PCR après transcription inverse (RT-PCR) et le séquençage direct pour détecter des mutations plus petites dans des gènes spécifiques^{208,209}. Nous avons choisi le SNG dans cette étude car contrairement aux méthodes précédentes, le SNG permet la détection simultanée de nombreuses variations génétiques sur l'ensemble du génome, telles que des SNV et des indels. De plus, il nécessite peu de matériel biologique de départ, et est disponible à un coût relativement bas. Le SNG est une méthode de choix dans la détection de petites mutations dans le génome de patients LMA-CN^{62,63,107,113}.

Afin d'identifier l'ensemble des variations génétiques qui définissent le génome de CG-SH, nous avons utilisé le SNG du génome entier et du transcriptome, en considérant uniquement les variants détectés dans les régions codantes. Dans les données brutes du séquençage nous avons détecté une très grande majorité de SNV (environ 99%). Il a récemment été montré que la détection des indels avec le SNG est déficiente, en partie à cause du glissement de la polymérase (*polymerase slippage*) généré par les répétitions

en tandem ²¹⁰. Le SNG ciblé permettrait une détection efficace des indels, mais s'appliquerait à des gènes déjà connus dans la maladie ^{211,212}. Puisque la lignée cellulaire CG-SH a récemment été établie et qu'aucun réarrangement chromosomique n'est survenu après la mise en culture de la lignée, nous avons entrepris de découvrir quelles variations génétiques sont susceptibles d'être à l'origine du développement de la leucémie chez le patient dont la lignée est issue. Étant donné que le patient est décédé en 2006 ¹⁶³, nous n'avons accès ni à ces CSH normales, ni à ces cellules leucémiques primaires, qui nous auraient servi de contrôles positifs pour identifier respectivement les mutations participant à la leucémogénèse et celles acquises pendant la culture cellulaire. Par conséquent, notre étude est fondée sur l'hypothèse que les anomalies génétiques ayant lieu dans des gènes déjà reportés comme mutés ou dérégulés dans la LMA, sont ceux qui participent probablement à la leucémogénèse. Afin d'éliminer les variants potentiellement bénins parmi les nombreux SNV détectés dans la lignée cellulaire CG-SH, nous avons éliminé ceux dont l'effet prédit par le logiciel VEP (Variant Effect Predictor) n'était pas délétère pour la protéine. Les SNV trouvés dans quatre gènes connus pour être mutés de façon récurrente dans la LMA (*FLT3*, *NRAS*, *RUNX1* et *TET2*) sont des polymorphismes connus (annotés dans dbSNP135) qui ont par conséquent peu de chance d'être impliqué dans la pathogénicité. En utilisant le séquençage du génome et du transcriptome, nous avons pu associer le statut mutationnel des gènes fréquemment mutés dans la LMA avec leur niveau d'expression dans la lignée cellulaire. Ainsi nous avons éliminé l'hypothèse d'une anomalie de l'expression de ces gènes qui aurait pu participer à la leucémogénèse. En revanche, des SNV inconnus ont été détectés dans des gènes mutés dans de rares cas de LMA et dans d'autres types de cancer d'après la littérature. Par conséquent, ces mutations pourraient jouer un rôle dans la pathogénèse. L'hypothèse de l'implication des SNV dans la leucémogénèse a été supportée par l'étude des termes GO porteurs des SNV connus ou nouveaux. En effet, les analyses GO supervisées et non supervisées révèlent que les gènes mutés sont impliqués dans des fonctions biologiques essentielles au développement de la leucémie.

De plus, le séquençage a permis la détection d'une insertion d'une paire de base dans *EZH2*, provoquant la perte du domaine fonctionnel SET de la protéine, site de son

activité méthyltransférase ²¹³. Cette mutation est homozygote, ce qui suggère une perte totale de fonction du gène. EZH2 est un régulateur de l'hématopoïèse faisant partie du complexe PRC2 muté dans de rares cas de LMA ¹⁸⁴. D'autres mutations de type indels résultant en un décalage du cadre de lecture et en une perte de l'activité méthyltransférase ont été découvertes tous au long du gène *EZH2*, chez des patients atteints de pathologies myéloïdes malignes ²¹⁴. Il a été montré que EZH2 accélère la progression de la leucémie *in vivo* initiée par *MLL-AF9*, notamment en participant au blocage de la différenciation ^{215,216}. EZH2 aurait donc pu participer à la progression de la leucémie en partenariat avec d'autres gènes dans ce cas de LMA-CN.

Bien qu'un SNV inconnu ait été détecté dans *GATA2* par le SNG dans CG-SH, l'insertion/délétion hétérozygote de 355 pb a été identifiée par RT-PCR et séquençage de Sanger. En effet, les algorithmes actuellement utilisés pour le SNG ne sont pas appropriés à la détection d'insertions de taille moyenne ²¹⁷. Cependant, l'utilisation du SNG ciblé avec des lectures de taille plus courte (~100 pb) aux extrémités couplées (*paired-end reads*) a permis la détection d'insertions en tandem de 15 à 300 pb dans le gène *FLT3*. Puisque l'insertion dans *GATA2* est présente à la fin du gène, nous avons vérifié qu'elle avait mené à une perte de fonction en testant l'activité du facteur de transcription grâce à un gène rapporteur luciférase. De plus, le séquençage de l'ARN a révélé que seul l'allèle muté portant l'insertion était exprimé dans la lignée cellulaire (ASE), ce qui conduit à une perte de fonction totale du gène. Les mutations de perte de fonction de *GATA2* sont associées aux formes familiales et somatiques de la LMA, et sont très fortement corrélées aux mutations du gène *CEBPA* ^{165,175,218,219}. *GATA2* est associé à la « reprogrammation » du réseau de transcription des gènes *in vivo* engendré lors du développement de la LMA ²²⁰. Les souris *GATA2*-déficientes présentent une hématopoïèse anormale, avec des défauts de prolifération, de survie, de différenciation, d'apoptose et de quiescence des CSH ²²¹. Ces informations suggèrent que l'insertion/délétion détectée dans le gène *GATA2* des cellules CG-SH participe très probablement à la pathogénèse, peut-être même à l'initiation de la maladie. Cependant, l'étude que nous avons menée ne nous permet pas de connaître la combinaison exacte de mutations qui ont participé au développement de la LMA chez le patient dont est dérivée la lignée cellulaire CG-SH.

Afin d'améliorer notre modèle d'étude *in vitro*, nous avons voulu accélérer la croissance des CG-SH en culture dont le temps de doublement étaient relativement long (supérieur à 72h). Dans de nombreux cas, la prolifération *in vitro* des cellules leucémiques primaires de patients et de lignées cellulaires est dépendante des cytokines ^{222,223}. La réponse aux cytokines et leur niveau d'expression chez les patients LMA confèrent une valeur pronostique ²²⁴⁻²²⁷. De plus, les cytokines sont utilisées dans les thérapies ciblées de la LMA ²²⁸. Ainsi, la compréhension de l'effet des cytokines sur le comportement des CG-SH apporterait de l'information sur la régulation de la prolifération des cellules LMA-CN par des facteurs de croissance, pouvant être utilisée à des fins thérapeutiques. Dans cette étude, nous avons choisi une combinaison de quatre cytokines : FL, IL3, G-CSF et SCF, pour son effet connu sur la prolifération et la survie des cellules LMA ^{189,190,222}. Il est important de noter que l'effet synergique des cytokines sur la prolifération des cellules leucémiques est plus significatif que l'effet individuel de chaque cytokine ¹⁸⁹. Bien que le milieu de culture des cellules leucémiques soit connu pour influencer l'expression génique ¹¹⁷, nous n'avons pas trouvé d'étude portant sur l'analyse à grande échelle des changements d'expression géniques provoqués par les cytokines. Afin de comprendre comment les cytokines entraînent une diminution (de facteur ~1,5) du temps de doublement de CG-SH, nous avons étudié les changements dans l'expression des gènes engendrés par les cytokines grâce au séquençage du transcriptome. Premièrement, l'analyse globale de l'expression des gènes montre que le profil de CG-SH en l'absence de cytokines est proche de celui des patients et qu'en présence de cytokines il tend vers un profile proche de la lignée cellulaire HL-60 et des autres lignées LMA avec un caryotype anormal. Ces résultats suggèrent que les changements de l'expression des gènes mis en jeu dans les CG-SH en présence cytokines, sont similaires à ceux qui participent à l'avantage prolifératif des lignées cellulaires par rapport aux cellules primaires de patients LMA. De façon inattendue, le profil de CG-SH en l'absence de cytokines est proche de celui des patients LMA M1, alors que cette lignée est dérivée d'un patient M5. L'analyse des termes GO pour les gènes différentiellement exprimés dans les cellules CG-SH exposées aux cytokines sur le long terme et le court terme, suggère que les gènes qui participent aux changements du profile d'expression sont impliqués dans la prolifération,

la régulation de l'apoptose et la régulation de la différenciation. Ces fonctions cellulaires sont typiquement dérégulées lors de la leucémogénèse ⁹, suggérant que les cytokines augmentent la pathogénicité de CG-SH. De plus, l'analyse GO au court terme révèle un enrichissement dans la fonction de régulation négative de l'apoptose pour les gènes surexprimés en présence de cytokines. Ces résultats suggèrent l'inactivation de la voie de l'apoptose par les cytokines. Cette hypothèse a été testée par cytométrie en flux des CG-SH avec et sans cytokines, à l'aide d'un double-marquage AnnexinV/PI. Étonnamment cette analyse ne révèle qu'une diminution significative du taux de nécrose (cellules marquées au PI seul) des CG-SH après 3 jours d'exposition aux cytokines. Cette différence tend à diminuer avec le temps jusqu'à disparaître au bout de 7 jours. Ces résultats suggèrent que les cytokines ont un effet inhibiteur sur la nécrose (mort non-physiologique) des CG-SH uniquement sur le court terme. Nous n'avons pas trouvé d'étude faisant état du rôle des cytokines dans la nécrose dans la littérature. De plus, la nécrose des cellules de la moelle osseuse des patients leucémiques est un phénomène très rare qui confère un pronostic défavorable ²²⁹. Il se pourrait donc que les cytokines favorisent la survie des CG-SH au court terme en diminuant le taux de nécrose, et activent la prolifération de CG-SH sur le long-terme, ce qui expliquerait la diminution du temps de doublement observée en présence de cytokines.

CHAPITRE 5 : PERSPECTIVES

Étant donné le peu d'insertions et de délétions détectées avec le SNG dans CG-SH; un problème déjà connu de cette approche; nous pourrions analyser le génome de CG-SH en utilisant le SNG ciblé générant des lectures courtes aux extrémités couplées, tel que décrit par Spencer et son équipe ²¹⁷. Cette approche s'appliquerait à la détection d'indels (jusqu'à 300 pb) dans certains gènes connus pour être impliqués dans la LMA ou le cancer de façon récurrente.

Afin d'approfondir nos connaissances sur l'effet des cytokines sur les CG-SH, nous pourrions étudier les signatures d'expression génique pour chaque cytokine indépendamment. Tout comme nous l'avons fait dans la présente étude, nous pourrions utiliser le séquençage du transcriptome et comparer les profils d'expression des CG-SH en présence et en absence de chacune des cytokines dans le milieu de culture. De cette manière, nous pourrions identifier les voies biologiques activées ou réprimées par chaque cytokine. Ceci apporterait de l'information précieuse sur les voies biologiques impliquées dans la régulation de la survie et la prolifération dans la LMA-CN, aboutissant potentiellement à la découverte de cibles thérapeutiques pour ce type fréquent de leucémie.

La lignée cellulaire LMA SHI-1 se caractérise par sa capacité à générer une tumeur *in vivo* après injection dans des modèles murins ²⁰⁸. Ainsi, nous pourrions tester le potentiel tumorigénique de CG-SH dans des modèles murins tel que décrit dans l'étude de Chen et son équipe ²⁰⁸. Les lignées cellulaires sont utilisées pour les xénogreffes dans des modèles de souris pour l'étude des mécanismes biologiques qui participent à la leucémogénèse ^{160,230,231}. Par conséquent, nous pourrions dans un objectif au long terme, utiliser la lignée CG-SH dans les xénogreffes pour étudier la LMA-CN *in vivo*.

CHAPITRE 6 : CONCLUSION

La présente étude a permis d'approfondir la caractérisation moléculaire de la nouvelle lignée cellulaire, par la découverte des mutations qui participent probablement la pathogénicité de la lignée CG-SH et l'identification des changements d'expression des gènes impliqués dans les voies biologiques de réponse aux cytokines. Ces informations permettront sûrement d'utiliser la lignée CG-SH par la suite comme modèle d'étude *in vitro* et de la greffer dans des modèles murins, dans le but de faciliter la compréhension des mécanismes biologiques participant au développement de LMA-CN. De plus, notre étude a apporté un aperçu des causes génétiques associées à ce cas de LMA-CN, et une vue d'ensemble des changements d'expression géniques associés au traitement avec les cytokines dans un contexte de LMA-CN.

CHAPITRE 7: BIBLIOGRAPHIE

1. Orkin, S.H. & Zon, L.I. Hematopoiesis: an evolving paradigm for stem cell biology. *Cell* **132**, 631-44 (2008).
2. Sashida, G. & Iwama, A. Epigenetic regulation of hematopoiesis. *Int J Hematol* **96**, 405-12 (2012).
3. Orkin, S.H. Diversification of haematopoietic stem cells to specific lineages. *Nat Rev Genet* **1**, 57-64 (2000).
4. Huang, X., Cho, S. & Spangrude, G.J. Hematopoietic stem cells: generation and self-renewal. *Cell Death Differ* **14**, 1851-9 (2007).
5. Jin, H., Xu, J. & Wen, Z. Migratory path of definitive hematopoietic stem/progenitor cells during zebrafish development. *Blood* **109**, 5208-14 (2007).
6. Murayama, E. *et al.* Tracing hematopoietic precursor migration to successive hematopoietic organs during zebrafish development. *Immunity* **25**, 963-75 (2006).
7. Ciriza, J., Thompson, H., Petrosian, R., Manilay, J.O. & Garcia-Ojeda, M.E. The migration of hematopoietic progenitors from the fetal liver to the fetal bone marrow: Lessons learned and possible clinical applications. *Exp Hematol* **41**, 411-23 (2013).
8. Niu, W. *et al.* Characterization of hemangioblast in umbilical arteries of mid-gestation mouse embryos. *Int J Hematol* **95**, 632-9 (2012).
9. Cellot, S. & Sauvageau, G. Zfx: at the crossroads of survival and self-renewal. *Cell* **129**, 239-41 (2007).

10. 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).
11. Antonchuk, J., Sauvageau, G. & Humphries, R.K. HOXB4-induced expansion of adult hematopoietic stem cells *ex vivo*. *Cell* **109**, 39-45 (2002).
12. Buske, C. *et al.* Deregulated expression of HOXB4 enhances the primitive growth activity of human hematopoietic cells. *Blood* **100**, 862-8 (2002).
13. Roeder, I. & Glauche, I. Towards an understanding of lineage specification in hematopoietic stem cells: a mathematical model for the interaction of transcription factors GATA-1 and PU.1. *J Theor Biol* **241**, 852-65 (2006).
14. Rice, K.L., Hormaeche, I. & Licht, J.D. Epigenetic regulation of normal and malignant hematopoiesis. *Oncogene* **26**, 6697-714 (2007).
15. Letting, D.L., Rakowski, C., Weiss, M.J. & Blobel, G.A. Formation of a tissue-specific histone acetylation pattern by the hematopoietic transcription factor GATA-1. *Mol Cell Biol* **23**, 1334-40 (2003).
16. Perez-Campo, F.M., Borrow, J., Kouskoff, V. & Lacaud, G. The histone acetyl transferase activity of monocytic leukemia zinc finger is critical for the proliferation of hematopoietic precursors. *Blood* **113**, 4866-74 (2009).
17. Shilatifard, A. Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. *Annu Rev Biochem* **75**, 243-69 (2006).
18. Konuma, T., Oguro, H. & Iwama, A. Role of the polycomb group proteins in hematopoietic stem cells. *Dev Growth Differ* **52**, 505-16 (2010).

19. Iwama, A. *et al.* Enhanced self-renewal of hematopoietic stem cells mediated by the polycomb gene product Bmi-1. *Immunity* **21**, 843-51 (2004).
20. Chou, R.H., Yu, Y.L. & Hung, M.C. The roles of EZH2 in cell lineage commitment. *Am J Transl Res* **3**, 243-50 (2011).
21. Kamminga, L.M. *et al.* The Polycomb group gene Ezh2 prevents hematopoietic stem cell exhaustion. *Blood* **107**, 2170-9 (2006).
22. Trowbridge, J.J., Snow, J.W., Kim, J. & Orkin, S.H. DNA methyltransferase 1 is essential for and uniquely regulates hematopoietic stem and progenitor cells. *Cell Stem Cell* **5**, 442-9 (2009).
23. Calvi, L.M. *et al.* Osteoblastic cells regulate the haematopoietic stem cell niche. *Nature* **425**, 841-6 (2003).
24. Zhang, J. *et al.* Identification of the haematopoietic stem cell niche and control of the niche size. *Nature* **425**, 836-41 (2003).
25. Arai, F. *et al.* Tie2/angiopoietin-1 signaling regulates hematopoietic stem cell quiescence in the bone marrow niche. *Cell* **118**, 149-61 (2004).
26. Ehninger, A. & Trumpp, A. The bone marrow stem cell niche grows up: mesenchymal stem cells and macrophages move in. *J Exp Med* **208**, 421-8 (2011).
27. Winkler, I.G. *et al.* Bone marrow macrophages maintain hematopoietic stem cell (HSC) niches and their depletion mobilizes HSCs. *Blood* **116**, 4815-28 (2010).
28. Kiel, M.J. *et al.* SLAM family receptors distinguish hematopoietic stem and progenitor cells and reveal endothelial niches for stem cells. *Cell* **121**, 1109-21 (2005).

29. Wright, D.E., Bowman, E.P., Wagers, A.J., Butcher, E.C. & Weissman, I.L. Hematopoietic stem cells are uniquely selective in their migratory response to chemokines. *J Exp Med* **195**, 1145-54 (2002).
30. Sugiyama, T., Kohara, H., Noda, M. & Nagasawa, T. Maintenance of the hematopoietic stem cell pool by CXCL12-CXCR4 chemokine signaling in bone marrow stromal cell niches. *Immunity* **25**, 977-88 (2006).
31. Mendez-Ferrer, S. *et al.* Mesenchymal and haematopoietic stem cells form a unique bone marrow niche. *Nature* **466**, 829-34 (2010).
32. Société Canadienne du cancer.
33. Comité directeur de la Société canadienne du cancer, Institut National du cancer du Canada, Statistique Canada, Registres du cancer des provinces et des territoires, Agence de la santé publique du Canada. in *Statistiques canadiennes sur le cancer 2010* (Editor, 2010).
34. American Cancer Society. (2013).
35. Howlader N, N.A., Krapcho M, Neyman N, Aminou R, Waldron W, Altekruse SF, , Kosary CL, R.J., Tatalovich Z, Cho H, Mariotto A, Eisner MP, Lewis DR, Chen HS, Feuer & EJ, C.K., Edwards BK, editors. SEER cancer statistics review, 1975–2008. (National Cancer Institute, Bethesda., based on November 2010 SEER data Submission, posted to the SEER web site, 2011).
36. Estey, E.H. Acute Myeloid Leukemia (AML). in *Leukemia and Related Disorders : Integrated Treatment Approaches* (ed. York, S.N.) 1-36 (Springer eBooks, 2012).
37. Lowenberg, B., Downing, J.R. & Burnett, A. Acute myeloid leukemia. *N Engl J Med* **341**, 1051-62 (1999).

38. Vardiman, J.W., Harris, N.L. & Brunning, R.D. The World Health Organization (WHO) classification of the myeloid neoplasms. *Blood* **100**, 2292-302 (2002).
39. Roboz, G.J. Current treatment of acute myeloid leukemia. *Curr Opin Oncol* **24**, 711-9 (2012).
40. Dohner, H. *et al.* Diagnosis and management of acute myeloid leukemia in adults: recommendations from an international expert panel, on behalf of the European LeukemiaNet. *Blood* **115**, 453-74 (2010).
41. Khan, C., Pathe, N., Fazal, S., Lister, J. & Rossetti, J.M. Azacitidine in the management of patients with myelodysplastic syndromes. *Ther Adv Hematol* **3**, 355-73 (2012).
42. Fenaux, P. *et al.* Efficacy of azacitidine compared with that of conventional care regimens in the treatment of higher-risk myelodysplastic syndromes: a randomised, open-label, phase III study. *Lancet Oncol* **10**, 223-32 (2009).
43. Kindler, T., Lipka, D.B. & Fischer, T. FLT3 as a therapeutic target in AML: still challenging after all these years. *Blood* **116**, 5089-102 (2010).
44. Sievers, E.L. *et al.* Efficacy and safety of gemtuzumab ozogamicin in patients with CD33-positive acute myeloid leukemia in first relapse. *J Clin Oncol* **19**, 3244-54 (2001).
45. Bennett, J.M. *et al.* Proposals for the classification of the acute leukaemias. French-American-British (FAB) co-operative group. *Br J Haematol* **33**, 451-8 (1976).
46. Bennett, J.M. *et al.* Proposed revised criteria for the classification of acute myeloid leukemia. A report of the French-American-British Cooperative Group. *Ann Intern Med* **103**, 620-5 (1985).

47. Vardiman, J.W. The World Health Organization (WHO) classification of tumors of the hematopoietic and lymphoid tissues: an overview with emphasis on the myeloid neoplasms. *Chem Biol Interact* **184**, 16-20 (2010).
48. Vardiman, J.W. *et al.* The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. *Blood* **114**, 937-51 (2009).
49. Gilliland, D.G. Molecular genetics of human leukemias: new insights into therapy. *Semin Hematol* **39**, 6-11 (2002).
50. Grimwade, D. & Mrozek, K. Diagnostic and prognostic value of cytogenetics in acute myeloid leukemia. *Hematol Oncol Clin North Am* **25**, 1135-61, vii (2011).
51. Grimwade, D. *et al.* Refinement of cytogenetic classification in acute myeloid leukemia: determination of prognostic significance of rare recurring chromosomal abnormalities among 5876 younger adult patients treated in the United Kingdom Medical Research Council trials. *Blood* **116**, 354-65 (2010).
52. Brown, N.J. *et al.* PML nuclear bodies in the pathogenesis of acute promyelocytic leukemia: active players or innocent bystanders? *Front Biosci* **14**, 1684-707 (2009).
53. Pabst, T. *et al.* AML1-ETO downregulates the granulocytic differentiation factor C/EBPalpha in t(8;21) myeloid leukemia. *Nat Med* **7**, 444-51 (2001).
54. Helbling, D. *et al.* CFBF-SMMHC is correlated with increased calreticulin expression and suppresses the granulocytic differentiation factor CEBPA in AML with inv(16). *Blood* **106**, 1369-75 (2005).
55. Goyama, S. & Mulloy, J.C. Molecular pathogenesis of core binding factor leukemia: current knowledge and future prospects. *Int J Hematol* **94**, 126-33 (2011).

56. Kuhn, A. & Grimwade, D. Molecular markers in acute myeloid leukaemia. *Int J Hematol* **96**, 153-63 (2012).
57. Meyer, C. *et al.* New insights to the MLL recombinome of acute leukemias. *Leukemia* **23**, 1490-9 (2009).
58. Mrozek, K. *et al.* Adult patients with de novo acute myeloid leukemia and t(9;11)(p22; q23) have a superior outcome to patients with other translocations involving band 11q23: a cancer and leukemia group B study. *Blood* **90**, 4532-8 (1997).
59. Blum, W. *et al.* Adult de novo acute myeloid leukemia with t(6;11)(q27;q23): results from Cancer and Leukemia Group B Study 8461 and review of the literature. *Cancer* **101**, 1420-7 (2004).
60. Soekarman, D. *et al.* The translocation (6;9) (p23;q34) shows consistent rearrangement of two genes and defines a myeloproliferative disorder with specific clinical features. *Blood* **79**, 2990-7 (1992).
61. Oancea, C., Ruster, B., Henschler, R., Puccetti, E. & Ruthardt, M. The t(6;9) associated DEK/CAN fusion protein targets a population of long-term repopulating hematopoietic stem cells for leukemogenic transformation. *Leukemia* **24**, 1910-9 (2010).
62. Mardis, E.R. *et al.* Recurring mutations found by sequencing an acute myeloid leukemia genome. *N Engl J Med* **361**, 1058-66 (2009).
63. Heo, S.G., Hong, E.P. & Park, J.W. Genetic risk prediction for normal-karyotype acute myeloid leukemia using whole-exome sequencing. *Genomics Inform* **11**, 46-51 (2013).
64. Grisendi, S., Mecucci, C., Falini, B. & Pandolfi, P.P. Nucleophosmin and cancer. *Nat Rev Cancer* **6**, 493-505 (2006).

65. Becker, H. *et al.* Favorable prognostic impact of NPM1 mutations in older patients with cytogenetically normal de novo acute myeloid leukemia and associated gene- and microRNA-expression signatures: a Cancer and Leukemia Group B study. *J Clin Oncol* **28**, 596-604 (2010).
66. Dohner, K. *et al.* Mutant nucleophosmin (NPM1) predicts favorable prognosis in younger adults with acute myeloid leukemia and normal cytogenetics: interaction with other gene mutations. *Blood* **106**, 3740-6 (2005).
67. Haferlach, C. *et al.* AML with mutated NPM1 carrying a normal or aberrant karyotype show overlapping biologic, pathologic, immunophenotypic, and prognostic features. *Blood* **114**, 3024-32 (2009).
68. Thiede, C. *et al.* Prevalence and prognostic impact of NPM1 mutations in 1485 adult patients with acute myeloid leukemia (AML). *Blood* **107**, 4011-20 (2006).
69. Whitman, S.P. *et al.* Absence of the wild-type allele predicts poor prognosis in adult de novo acute myeloid leukemia with normal cytogenetics and the internal tandem duplication of FLT3: a cancer and leukemia group B study. *Cancer Res* **61**, 7233-9 (2001).
70. Singh, H. *et al.* Outcome of older adults with cytogenetically normal AML (CN-AML) and FLT3 mutations. *Leuk Res* **35**, 1611-5 (2011).
71. Park, S.H. *et al.* Prognostic significance of the FLT3 ITD mutation in patients with normal-karyotype acute myeloid leukemia in relapse. *Korean J Hematol* **46**, 88-95 (2011).
72. Noguera, N.I. *et al.* Alterations of the FLT3 gene in acute promyelocytic leukemia: association with diagnostic characteristics and analysis of clinical outcome in patients treated with the Italian AIDA protocol. *Leukemia* **16**, 2185-9 (2002).

73. Gale, R.E. *et al.* Relationship between FLT3 mutation status, biologic characteristics, and response to targeted therapy in acute promyelocytic leukemia. *Blood* **106**, 3768-76 (2005).
74. Pabst, T. *et al.* Dominant-negative mutations of CEBPA, encoding CCAAT/enhancer binding protein-alpha (C/EBPalpha), in acute myeloid leukemia. *Nat Genet* **27**, 263-70 (2001).
75. Preudhomme, C. *et al.* Favorable prognostic significance of CEBPA mutations in patients with de novo acute myeloid leukemia: a study from the Acute Leukemia French Association (ALFA). *Blood* **100**, 2717-23 (2002).
76. Fos, J., Pabst, T., Petkovic, V., Ratschiller, D. & Mueller, B.U. Deficient CEBPA DNA binding function in normal karyotype AML patients is associated with favorable prognosis. *Blood* **117**, 4881-4 (2011).
77. Gaidzik, V.I. *et al.* TET2 mutations in acute myeloid leukemia (AML): results from a comprehensive genetic and clinical analysis of the AML study group. *J Clin Oncol* **30**, 1350-7 (2012).
78. Chou, W.C. *et al.* TET2 mutation is an unfavorable prognostic factor in acute myeloid leukemia patients with intermediate-risk cytogenetics. *Blood* **118**, 3803-10 (2011).
79. Shen, Y. *et al.* Gene mutation patterns and their prognostic impact in a cohort of 1185 patients with acute myeloid leukemia. *Blood* **118**, 5593-603 (2011).
80. Zhou, K.G. *et al.* Potential application of IDH1 and IDH2 mutations as prognostic indicators in non-promyelocytic acute myeloid leukemia: a meta-analysis. *Leuk Lymphoma* **53**, 2423-9 (2012).
81. Nomdedeu, J. *et al.* Adverse impact of IDH1 and IDH2 mutations in primary AML: experience of the Spanish CETLAM group. *Leuk Res* **36**, 990-7 (2012).

82. Chotirat, S., Thongnoppakhun, W., Promsuwicha, O., Boonthimat, C. & Auewarakul, C.U. Molecular alterations of isocitrate dehydrogenase 1 and 2 (IDH1 and IDH2) metabolic genes and additional genetic mutations in newly diagnosed acute myeloid leukemia patients. *J Hematol Oncol* **5**, 5 (2012).
83. Ravandi, F. *et al.* Prognostic significance of alterations in IDH enzyme isoforms in patients with AML treated with high-dose cytarabine and idarubicin. *Cancer* **118**, 2665-73 (2012).
84. Ibanez, M. *et al.* Rapid screening of ASXL1, IDH1, IDH2, and c-CBL mutations in de novo acute myeloid leukemia by high-resolution melting. *J Mol Diagn* **14**, 594-601 (2012).
85. Gelsi-Boyer, V. *et al.* Mutations in ASXL1 are associated with poor prognosis across the spectrum of malignant myeloid diseases. *J Hematol Oncol* **5**, 12 (2012).
86. Schnittger, S. *et al.* ASXL1 exon 12 mutations are frequent in AML with intermediate risk karyotype and are independently associated with an adverse outcome. *Leukemia* **27**, 82-91 (2013).
87. Ley, T.J. *et al.* DNMT3A mutations in acute myeloid leukemia. *N Engl J Med* **363**, 2424-33 (2010).
88. Thol, F. *et al.* Incidence and prognostic influence of DNMT3A mutations in acute myeloid leukemia. *J Clin Oncol* **29**, 2889-96 (2011).
89. de Jonge, H.J., Huls, G. & de Bont, E.S. Gene expression profiling in acute myeloid leukaemia. *Neth J Med* **69**, 167-76 (2011).
90. Tholouli, E., MacDermott, S., Hoyland, J., Yin, J.L. & Byers, R. Quantitative multiplex quantum dot in-situ hybridisation based gene expression profiling in tissue microarrays identifies prognostic genes in acute myeloid leukaemia. *Biochem Biophys Res Commun* **425**, 333-9 (2012).

91. Babashah, S., Sadeghizadeh, M., Tavirani, M.R., Farivar, S. & Soleimani, M. Aberrant microRNA expression and its implications in the pathogenesis of leukemias. *Cell Oncol (Dordr)* **35**, 317-34 (2012).
92. Jongen-Lavrencic, M., Sun, S.M., Dijkstra, M.K., Valk, P.J. & Lowenberg, B. MicroRNA expression profiling in relation to the genetic heterogeneity of acute myeloid leukemia. *Blood* **111**, 5078-85 (2008).
93. Garzon, R. *et al.* MicroRNA signatures associated with cytogenetics and prognosis in acute myeloid leukemia. *Blood* **111**, 3183-9 (2008).
94. Ramaswamy, S. & Golub, T.R. DNA microarrays in clinical oncology. *J Clin Oncol* **20**, 1932-41 (2002).
95. Maciejewski, J.P., Tiu, R.V. & O'Keefe, C. Application of array-based whole genome scanning technologies as a cytogenetic tool in haematological malignancies. *Br J Haematol* **146**, 479-88 (2009).
96. Duggan, D.J., Bittner, M., Chen, Y., Meltzer, P. & Trent, J.M. Expression profiling using cDNA microarrays. *Nat Genet* **21**, 10-4 (1999).
97. Tang, Y.C. & Amon, A. Gene copy-number alterations: a cost-benefit analysis. *Cell* **152**, 394-405 (2013).
98. Costa, A.R., Belangero, S.I., Melaragno, M.I. & Chauffaille Mde, L. Additional chromosomal abnormalities detected by array comparative genomic hybridization in AML. *Med Oncol* **29**, 2083-7 (2012).
99. Kolquist, K.A. *et al.* Microarray-based comparative genomic hybridization of cancer targets reveals novel, recurrent genetic aberrations in the myelodysplastic syndromes. *Cancer Genet* **204**, 603-28 (2011).

100. Stirewalt, D.L. *et al.* Identification of genes with abnormal expression changes in acute myeloid leukemia. *Genes Chromosomes Cancer* **47**, 8-20 (2008).
101. Valk, P.J. *et al.* Prognostically useful gene-expression profiles in acute myeloid leukemia. *N Engl J Med* **350**, 1617-28 (2004).
102. Tan, D.S., Lambros, M.B., Natrajan, R. & Reis-Filho, J.S. Getting it right: designing microarray (and not 'microawry') comparative genomic hybridization studies for cancer research. *Lab Invest* **87**, 737-54 (2007).
103. Shendure, J. & Ji, H. Next-generation DNA sequencing. *Nat Biotechnol* **26**, 1135-45 (2008).
104. Shendure, J. *et al.* Accurate multiplex polony sequencing of an evolved bacterial genome. *Science* **309**, 1728-32 (2005).
105. Metzker, M.L. Sequencing technologies - the next generation. *Nat Rev Genet* **11**, 31-46 (2010).
106. Riva, L., Luzi, L. & Pelicci, P.G. Genomics of acute myeloid leukemia: the next generation. *Front Oncol* **2**, 40 (2012).
107. Ley, T.J. *et al.* DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. *Nature* **456**, 66-72 (2008).
108. Ding, L. *et al.* Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature* **481**, 506-10 (2012).
109. Hodges, E. *et al.* Genome-wide in situ exon capture for selective resequencing. *Nat Genet* **39**, 1522-7 (2007).
110. Welch, J.S. & Link, D.C. Genomics of AML: clinical applications of next-generation sequencing. *Hematology Am Soc Hematol Educ Program* **2011**, 30-5 (2011).

111. Ng, S.B. *et al.* Targeted capture and massively parallel sequencing of 12 human exomes. *Nature* **461**, 272-6 (2009).
112. Yan, X.J. *et al.* Exome sequencing identifies somatic mutations of DNA methyltransferase gene DNMT3A in acute monocytic leukemia. *Nat Genet* **43**, 309-15 (2011).
113. Greif, P.A. *et al.* Identification of recurring tumor-specific somatic mutations in acute myeloid leukemia by transcriptome sequencing. *Leukemia* **25**, 821-7 (2011).
114. Wen, H. *et al.* New fusion transcripts identified in normal karyotype acute myeloid leukemia. *PLoS One* **7**, e51203 (2012).
115. Ramsingh, G. *et al.* Complete characterization of the microRNAome in a patient with acute myeloid leukemia. *Blood* **116**, 5316-26 (2010).
116. Cancer Genome Atlas Research, N. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N Engl J Med* **368**, 2059-74 (2013).
117. Bruserud, O., Gjertsen, B.T., Foss, B. & Huang, T.S. New strategies in the treatment of acute myelogenous leukemia (AML): in vitro culture of aml cells--the present use in experimental studies and the possible importance for future therapeutic approaches. *Stem Cells* **19**, 1-11 (2001).
118. Bruserud, O. *et al.* Effects of interleukin 10 on blast cells derived from patients with acute myelogenous leukemia. *Leukemia* **9**, 1910-20 (1995).
119. Blair, A., Hogge, D.E. & Sutherland, H.J. Most acute myeloid leukemia progenitor cells with long-term proliferative ability in vitro and in vivo have the phenotype CD34(+)/CD71(-)/HLA-DR. *Blood* **92**, 4325-35 (1998).

120. Nara, N. & McCulloch, E.A. The proliferation in suspension of the progenitors of the blast cells in acute myeloblastic leukemia. *Blood* **65**, 1484-93 (1985).
121. Delwel, R. *et al.* Growth regulation of human acute myeloid leukemia: effects of five recombinant hematopoietic factors in a serum-free culture system. *Blood* **72**, 1944-9 (1988).
122. Geissler, K., Tricot, G., Leemhuis, T., Walker, E. & Broxmeyer, H.E. Differentiation-inducing effect of recombinant human tumor necrosis factor alpha and gamma-interferon in vitro on blast cells from patients with acute myeloid leukemia and myeloid blast crisis of chronic myeloid leukemia. *Cancer Res* **49**, 3057-62 (1989).
123. Charbonnier, A., Gaugler, B., Sainty, D., Lafage-Pochitaloff, M. & Olive, D. Human acute myeloblastic leukemia cells differentiate in vitro into mature dendritic cells and induce the differentiation of cytotoxic T cells against autologous leukemias. *Eur J Immunol* **29**, 2567-78 (1999).
124. Santini, V. *et al.* Induction of granulocytic maturation in acute myeloid leukemia by G-CSF and retinoic acid. *Leuk Res* **15**, 341-50 (1991).
125. Howell, A.L., Stukel, T.A., Bloomfield, C.D., Davey, F.R. & Ball, E.D. Induction of differentiation in blast cells and leukemia colony-forming cells from patients with acute myeloid leukemia. *Blood* **75**, 721-9 (1990).
126. Lotem, J. & Sachs, L. Control of apoptosis in hematopoiesis and leukemia by cytokines, tumor suppressor and oncogenes. *Leukemia* **10**, 925-31 (1996).
127. Chant, I.D., Rose, P.E. & Morris, A.G. Susceptibility of AML cells to in vitro apoptosis correlates with heat shock protein 70 (hsp 70) expression. *Br J Haematol* **93**, 898-902 (1996).
128. Ryningen, A. *et al.* Stress-induced in vitro apoptosis of native human acute myelogenous leukemia (AML) cells shows a wide variation between patients and

is associated with low BCL-2:Bax ratio and low levels of heat shock protein 70 and 90. *Leuk Res* **30**, 1531-40 (2006).

129. Kronke, J. *et al.* Clonal evolution in relapsed NPM1-mutated acute myeloid leukemia. *Blood* **122**, 100-8 (2013).
130. Parkin, B. *et al.* Clonal evolution and devolution after chemotherapy in adult acute myelogenous leukemia. *Blood* **121**, 369-77 (2013).
131. Waterhouse, M. *et al.* Genome-wide profiling in AML patients relapsing after allogeneic hematopoietic cell transplantation. *Biol Blood Marrow Transplant* **17**, 1450-1459 e1 (2011).
132. Cloos, J. *et al.* Stability and prognostic influence of FLT3 mutations in paired initial and relapsed AML samples. *Leukemia* **20**, 1217-20 (2006).
133. Drexler, H.G. & Matsuo, Y. Guidelines for the characterization and publication of human malignant hematopoietic cell lines. *Leukemia* **13**, 835-42 (1999).
134. Drexler, H.G. & Macleod, R.A. History of leukemia-lymphoma cell lines. *Hum Cell* **23**, 75-82 (2010).
135. Xiu, B. *et al.* Genetic inhibition of vascular endothelial growth factor receptor-1 significantly inhibits the migration and proliferation of leukemia cells and increases their sensitivity to chemotherapy. *Oncol Rep* **29**, 2030-8 (2013).
136. Oellerich, T. *et al.* beta2 integrin-derived signals induce cell survival and proliferation of AML blasts by activating a Syk/STAT signaling axis. *Blood* **121**, 3889-99, S1-66 (2013).
137. Anfossi, G., Gewirtz, A.M. & Calabretta, B. An oligomer complementary to c-myc-encoded mRNA inhibits proliferation of human myeloid leukemia cell lines. *Proc Natl Acad Sci U S A* **86**, 3379-83 (1989).

138. Ortlepp, C. *et al.* Autotaxin is expressed in FLT3-ITD positive acute myeloid leukemia and hematopoietic stem cells and promotes cell migration and proliferation. *Exp Hematol* **41**, 444-461 e4 (2013).
139. Ojima, Y., Duncan, M.T., Nurhayati, R.W., Taya, M. & Miller, W.M. Synergistic effect of hydrogen peroxide on polyploidization during the megakaryocytic differentiation of K562 leukemia cells by PMA. *Exp Cell Res* (2013).
140. Breitman, T.R., Selonick, S.E. & Collins, S.J. Induction of differentiation of the human promyelocytic leukemia cell line (HL-60) by retinoic acid. *Proc Natl Acad Sci U S A* **77**, 2936-40 (1980).
141. Auwerx, J. The human leukemia cell line, THP-1: a multifaceted model for the study of monocyte-macrophage differentiation. *Experientia* **47**, 22-31 (1991).
142. Konopleva, M. *et al.* Stromal cells prevent apoptosis of AML cells by up-regulation of anti-apoptotic proteins. *Leukemia* **16**, 1713-24 (2002).
143. Kremer, K.N. *et al.* SDF-1/CXCR4 signaling induces apoptosis in acute myeloid leukemia cells via regulation of the Bcl-2 family members Bcl-XL, Noxa, and Bak. *J Biol Chem* (2013).
144. Su, X. *et al.* Involvement of MAPK activation and ROS generation in human leukemia U937 cells undergoing apoptosis in response to sonodynamic therapy. *Int J Radiat Biol* (2013).
145. Teng, C.L. *et al.* Effector mechanisms of sunitinib-induced G1 cell cycle arrest, differentiation, and apoptosis in human acute myeloid leukaemia HL60 and KG-1 cells. *Ann Hematol* **92**, 301-13 (2013).
146. Sanchez, P.V., Glantz, S.T., Scotland, S., Kasner, M.T. & Carroll, M. Induced differentiation of acute myeloid leukemia cells by activation of Retinoid X and Liver X Receptors. *Leukemia* (2013).

147. Fang, Y. *et al.* MEK/ERK Dependent Activation of STAT1 Mediates Dasatinib-Induced Differentiation of Acute Myeloid Leukemia. *PLoS One* **8**, e66915 (2013).
148. Wang, L. *et al.* Pharmacologic inhibition of CDK4/6: mechanistic evidence for selective activity or acquired resistance in acute myeloid leukemia. *Blood* **110**, 2075-83 (2007).
149. Levis, M., Pham, R., Smith, B.D. & Small, D. In vitro studies of a FLT3 inhibitor combined with chemotherapy: sequence of administration is important to achieve synergistic cytotoxic effects. *Blood* **104**, 1145-50 (2004).
150. Fortier, J.M. & Graubert, T.A. Murine models of human acute myeloid leukemia. *Cancer Treat Res* **145**, 183-96 (2010).
151. Li, Z. *et al.* Deletion of Tet2 in mice leads to dysregulated hematopoietic stem cells and subsequent development of myeloid malignancies. *Blood* **118**, 4509-18 (2011).
152. Lamothe, B. *et al.* Deletion of TAK1 in the myeloid lineage results in the spontaneous development of myelomonocytic leukemia in mice. *PLoS One* **7**, e51228 (2012).
153. Greenblatt, S. *et al.* Knock-in of a FLT3/ITD mutation cooperates with a NUP98-HOXD13 fusion to generate acute myeloid leukemia in a mouse model. *Blood* **119**, 2883-94 (2012).
154. Wang, Y.Y. *et al.* C-KIT mutation cooperates with full-length AML1-ETO to induce acute myeloid leukemia in mice. *Proc Natl Acad Sci U S A* **108**, 2450-5 (2011).
155. Watanabe-Okochi, N. *et al.* The shortest isoform of C/EBPbeta, liver inhibitory protein (LIP), collaborates with Evi1 to induce AML in a mouse BMT model. *Blood* **121**, 4142-55 (2013).

156. Miller, P.G. *et al.* In Vivo RNAi Screening Identifies a Leukemia-Specific Dependence on Integrin Beta 3 Signaling. *Cancer Cell* (2013).
157. Zhang, Y. *et al.* AML1-ETO mediates hematopoietic self-renewal and leukemogenesis through a COX/beta-catenin signaling pathway. *Blood* **121**, 4906-16 (2013).
158. Lobry, C. *et al.* Notch pathway activation targets AML-initiating cell homeostasis and differentiation. *J Exp Med* **210**, 301-19 (2013).
159. Sportoletti, P. *et al.* The human NPM1 mutation A perturbs megakaryopoiesis in a conditional mouse model. *Blood* **121**, 3447-58 (2013).
160. Woiterski, J. *et al.* Engraftment of low numbers of pediatric acute lymphoid and myeloid leukemias into NOD/SCID/IL2Rgamma null mice reflects individual leukemogenicity and highly correlates with clinical outcome. *Int J Cancer* (2013).
161. Zhang, Y. *et al.* CXCR4 inhibitors selectively eliminate CXCR4-expressing human acute myeloid leukemia cells in NOG mouse model. *Cell Death Dis* **3**, e396 (2012).
162. Wunderlich, M. *et al.* AML cells are differentially sensitive to chemotherapy treatment in a human xenograft model. *Blood* **121**, e90-7 (2013).
163. Munker, R. *et al.* Characterization of a new myeloid leukemia cell line with normal cytogenetics (CG-SH). *Leuk Res* **33**, 1405-8 (2009).
164. Simon, C. *et al.* A key role for EZH2 and associated genes in mouse and human adult T-cell acute leukemia. *Genes Dev* **26**, 651-6 (2012).
165. Hahn, C.N. *et al.* Heritable GATA2 mutations associated with familial myelodysplastic syndrome and acute myeloid leukemia. *Nat Genet* **43**, 1012-7 (2011).

166. Lowenberg, B. Acute myeloid leukemia: the challenge of capturing disease variety. *Hematology Am Soc Hematol Educ Program*, 1-11 (2008).
167. Tefferi, A., Skoda, R. & Vardiman, J.W. Myeloproliferative neoplasms: contemporary diagnosis using histology and genetics. *Nat Rev Clin Oncol* **6**, 627-37 (2009).
168. Marcucci, G., Haferlach, T. & Dohner, H. Molecular genetics of adult acute myeloid leukemia: prognostic and therapeutic implications. *J Clin Oncol* **29**, 475-86 (2011).
169. Schlenk, R.F. *et al.* Mutations and treatment outcome in cytogenetically normal acute myeloid leukemia. *N Engl J Med* **358**, 1909-18 (2008).
170. Santos, F.P. *et al.* Prognostic value of FLT3 mutations among different cytogenetic subgroups in acute myeloid leukemia. *Cancer* **117**, 2145-55 (2011).
171. Frohling, S. *et al.* Prognostic significance of activating FLT3 mutations in younger adults (16 to 60 years) with acute myeloid leukemia and normal cytogenetics: a study of the AML Study Group Ulm. *Blood* **100**, 4372-80 (2002).
172. Schneider, F. *et al.* NPM1 but not FLT3-ITD mutations predict early blast cell clearance and CR rate in patients with normal karyotype AML (NK-AML) or high-risk myelodysplastic syndrome (MDS). *Blood* **113**, 5250-3 (2009).
173. Falini, B. *et al.* Cytoplasmic nucleophosmin in acute myelogenous leukemia with a normal karyotype. *N Engl J Med* **352**, 254-66 (2005).
174. Dufour, A. *et al.* Monoallelic CEBPA mutations in normal karyotype acute myeloid leukemia: independent favorable prognostic factor within NPM1 mutated patients. *Ann Hematol* **91**, 1051-1063 (2012).

175. Greif, P.A. *et al.* GATA2 zinc finger 1 mutations associated with biallelic CEBPA mutations define a unique genetic entity of acute myeloid leukemia. *Blood* **120**, 395-403 (2012).
176. Paschka, P. *et al.* IDH1 and IDH2 mutations are frequent genetic alterations in acute myeloid leukemia and confer adverse prognosis in cytogenetically normal acute myeloid leukemia with NPM1 mutation without FLT3 internal tandem duplication. *J Clin Oncol* **28**, 3636-43 (2010).
177. Ribeiro, A.F. *et al.* Mutant DNMT3A: a marker of poor prognosis in acute myeloid leukemia. *Blood* **119**, 5824-31 (2012).
178. Montesinos, J.J. *et al.* Deficient proliferation and expansion in vitro of two bone marrow cell populations from patients with acute myeloid leukemia in response to hematopoietic cytokines. *Leuk Lymphoma* **47**, 1379-86 (2006).
179. Gallagher, R. *et al.* Characterization of the continuous, differentiating myeloid cell line (HL-60) from a patient with acute promyelocytic leukemia. *Blood* **54**, 713-33 (1979).
180. Odero, M.D., Zeleznik-Le, N.J., Chinwalla, V. & Rowley, J.D. Cytogenetic and molecular analysis of the acute monocytic leukemia cell line THP-1 with an MLL-AF9 translocation. *Genes Chromosomes Cancer* **29**, 333-8 (2000).
181. Mrozek, K., Tanner, S.M., Heinonen, K. & Bloomfield, C.D. Molecular cytogenetic characterization of the KG-1 and KG-1a acute myeloid leukemia cell lines by use of spectral karyotyping and fluorescence in situ hybridization. *Genes Chromosomes Cancer* **38**, 249-52 (2003).
182. Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F. & Richmond, T.J. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**, 251-60 (1997).

183. Paulsson, K. *et al.* A novel and cytogenetically cryptic t(7;21)(p22;q22) in acute myeloid leukemia results in fusion of RUNX1 with the ubiquitin-specific protease gene USP42. *Leukemia* **20**, 224-9 (2006).
184. Wang, X. *et al.* EZH2 mutations are related to low blast percentage in bone marrow and -7/del(7q) in de novo acute myeloid leukemia. *PLoS One* **8**, e61341 (2013).
185. Ernst, T. *et al.* A somatic EZH2 mutation in childhood acute myeloid leukemia. *Leukemia* **26**, 1701-3 (2012).
186. Celton, M. *et al.* Allele-specific expression of GATA2 in Normal Karyotype Acute Myeloid Leukemia. *unpublished observations*.
187. de Jonge, H.J. *et al.* Gene expression profiling in the leukemic stem cell-enriched CD34+ fraction identifies target genes that predict prognosis in normal karyotype AML. *Leukemia* **25**, 1825-33 (2011).
188. Mawad, R. & Estey, E.H. Acute myeloid leukemia with normal cytogenetics. *Curr Oncol Rep* **14**, 359-68 (2012).
189. Piacibello, W. *et al.* Effects of human FLT3 ligand on myeloid leukemia cell growth: heterogeneity in response and synergy with other hematopoietic growth factors. *Blood* **86**, 4105-14 (1995).
190. Woost, P.G. *et al.* High-resolution kinetics of cytokine signaling in human CD34/CD117-positive cells in unfractionated bone marrow. *Blood* **117**, e131-41 (2011).
191. Kawano-Yamamoto, C. *et al.* Establishment and characterization of a new erythroblastic leukemia cell line, EEB: phosphatidylglucoside-mediated erythroid differentiation and apoptosis. *Leuk Res* **30**, 829-39 (2006).

192. Asou, H. *et al.* Establishment of a human acute myeloid leukemia cell line (Kasumi-1) with 8;21 chromosome translocation. *Blood* **77**, 2031-6 (1991).
193. Fernandez-Mercado, M. *et al.* Mutation patterns of 16 genes in primary and secondary acute myeloid leukemia (AML) with normal cytogenetics. *PLoS One* **7**, e42334 (2012).
194. Metzeler, K.H. *et al.* TET2 mutations improve the new European LeukemiaNet risk classification of acute myeloid leukemia: a Cancer and Leukemia Group B study. *J Clin Oncol* **29**, 1373-81 (2011).
195. Ersvaer, E. *et al.* Cyclin B1 is commonly expressed in the cytoplasm of primary human acute myelogenous leukemia cells and serves as a leukemia-associated antigen associated with autoantibody response in a subset of patients. *Eur J Haematol* **79**, 210-25 (2007).
196. Aref, S., Mabed, M., El-Sherbiny, M., Selim, T. & Metwaly, A. Cyclin D1 expression in acute leukemia. *Hematology* **11**, 31-4 (2006).
197. Wang, C.M. *et al.* Effect of small interfering RNA targeting wild-type FLT3 in acute myeloid leukaemia cells in vitro and in vivo. *J Int Med Res* **39**, 1661-74 (2011).
198. Martens, J.H. *et al.* PML-RARalpha/RXR Alters the Epigenetic Landscape in Acute Promyelocytic Leukemia. *Cancer Cell* **17**, 173-85 (2010).
199. Kim, K.I. *et al.* Copy Number Variations in Normal Karyotype Acute Myeloid Leukaemia and their Association with Treatment Response. *Basic Clin Pharmacol Toxicol* **111**, 317-24 (2012).
200. Koren-Michowitz, M. *et al.* Older patients with normal karyotype acute myeloid leukemia have a higher rate of genomic changes compared to young patients as determined by SNP array analysis. *Leuk Res* **36**, 467-73 (2012).

201. Sawanobori, M. *et al.* Expression of TNF receptors and related signaling molecules in the bone marrow from patients with myelodysplastic syndromes. *Leuk Res* **27**, 583-91 (2003).
202. Yagi, T. *et al.* Identification of a gene expression signature associated with pediatric AML prognosis. *Blood* **102**, 1849-56 (2003).
203. Ozar, D. The characteristics of a valid "empirical" slippery-slope argument. *J Clin Ethics* **3**, 301-2 (1992).
204. Luesink, M. *et al.* High GATA2 expression is a poor prognostic marker in pediatric acute myeloid leukemia. *Blood* **120**, 2064-75 (2012).
205. Hyde, R.K. & Liu, P.P. GATA2 mutations lead to MDS and AML. *Nat Genet* **43**, 926-7 (2011).
206. Olopade, O.I. *et al.* Clinical, morphologic, and cytogenetic characteristics of 26 patients with acute erythroblastic leukemia. *Blood* **80**, 2873-82 (1992).
207. Haferlach, T. *et al.* Distinct genetic patterns can be identified in acute monoblastic and acute monocytic leukaemia (FAB AML M5a and M5b): a study of 124 patients. *Br J Haematol* **118**, 426-31 (2002).
208. Chen, S. *et al.* A new human acute monocytic leukemia cell line SHI-1 with t(6;11)(q27;q23), p53 gene alterations and high tumorigenicity in nude mice. *Haematologica* **90**, 766-75 (2005).
209. Hayashi, M. *et al.* Establishment of a novel childhood acute myeloid leukaemia cell line, KOPM-88, containing partial tandem duplication of the MLL gene and an in vivo model for childhood acute myeloid leukaemia using NOD/SCID mice. *Br J Haematol* **137**, 221-32 (2007).

210. Boland, J.F. *et al.* The new sequencer on the block: comparison of Life Technology's Proton sequencer to an Illumina HiSeq for whole-exome sequencing. *Hum Genet* (2013).
211. Chin, E.L., da Silva, C. & Hegde, M. Assessment of clinical analytical sensitivity and specificity of next-generation sequencing for detection of simple and complex mutations. *BMC Genet* **14**, 6 (2013).
212. Sikkema-Raddatz, B. *et al.* Targeted Next-Generation Sequencing can Replace Sanger Sequencing in Clinical Diagnostics. *Hum Mutat* **34**, 1035-42 (2013).
213. Rea, S. *et al.* Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* **406**, 593-9 (2000).
214. Ernst, T. *et al.* Inactivating mutations of the histone methyltransferase gene EZH2 in myeloid disorders. *Nat Genet* **42**, 722-6 (2010).
215. Tanaka, S. *et al.* Ezh2 augments leukemogenicity by reinforcing differentiation blockage in acute myeloid leukemia. *Blood* **120**, 1107-17 (2012).
216. Neff, T. *et al.* Polycomb repressive complex 2 is required for MLL-AF9 leukemia. *Proc Natl Acad Sci U S A* **109**, 5028-33 (2012).
217. Spencer, D.H. *et al.* Detection of FLT3 internal tandem duplication in targeted, short-read-length, next-generation sequencing data. *J Mol Diagn* **15**, 81-93 (2013).
218. Fasan, A. *et al.* GATA2 mutations are frequent in intermediate-risk karyotype AML with biallelic CEBPA mutations and are associated with favorable prognosis. *Leukemia* **27**, 482-5 (2013).
219. Green, C.L. *et al.* GATA2 mutations in sporadic and familial acute myeloid leukaemia patients with CEBPA mutations. *Br J Haematol* **161**, 701-5 (2013).

220. Bonadies, N. *et al.* Genome-wide analysis of transcriptional reprogramming in mouse models of acute myeloid leukaemia. *PLoS One* **6**, e16330 (2011).
221. Rodrigues, N.P., Tipping, A.J., Wang, Z. & Enver, T. GATA-2 mediated regulation of normal hematopoietic stem/progenitor cell function, myelodysplasia and myeloid leukemia. *Int J Biochem Cell Biol* **44**, 457-60 (2012).
222. Bruserud, O., Gjertsen, B.T. & von Volkman, H.L. In vitro culture of human acute myelogenous leukemia (AML) cells in serum-free media: studies of native AML blasts and AML cell lines. *J Hematother Stem Cell Res* **9**, 923-32 (2000).
223. Bruserud, O., Frostad, S. & Foss, B. In vitro culture of acute myelogenous leukemia blasts: a comparison of four different culture media. *J Hematother* **8**, 63-73 (1999).
224. Sanchez-Correa, B. *et al.* Cytokine profiles in acute myeloid leukemia patients at diagnosis: survival is inversely correlated with IL-6 and directly correlated with IL-10 levels. *Cytokine* **61**, 885-91 (2013).
225. Redell, M.S. *et al.* FACS analysis of Stat3/5 signaling reveals sensitivity to G-CSF and IL-6 as a significant prognostic factor in pediatric AML: a Children's Oncology Group report. *Blood* **121**, 1083-93 (2013).
226. Bruserud, O. *et al.* Subclassification of patients with acute myelogenous leukemia based on chemokine responsiveness and constitutive chemokine release by their leukemic cells. *Haematologica* **92**, 332-41 (2007).
227. Rombouts, W.J., Lowenberg, B., van Putten, W.L. & Ploemacher, R.E. Improved prognostic significance of cytokine-induced proliferation in vitro in patients with de novo acute myeloid leukemia of intermediate risk: impact of internal tandem duplications in the Flt3 gene. *Leukemia* **15**, 1046-53 (2001).

228. Ferretti, E., Cocco, C., Airoidi, I. & Pistoia, V. Targeting acute myeloid leukemia cells with cytokines. *J Leukoc Biol* **92**, 567-75 (2012).
229. Forrest, D.L. *et al.* Bone marrow necrosis in adult acute leukemia and non-Hodgkin's lymphoma. *Leuk Lymphoma* **38**, 627-32 (2000).
230. Gao, X.N. *et al.* A histone acetyltransferase p300 inhibitor C646 induces cell cycle arrest and apoptosis selectively in AML1-ETO-positive AML cells. *PLoS One* **8**, e55481 (2013).
231. Kannan, S. *et al.* Notch activation inhibits AML growth and survival: a potential therapeutic approach. *J Exp Med* **210**, 321-37 (2013).