

Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN  
ONCOLOGIA E PATOLOGIA SPERIMENTALE

Ciclo XXIX

**Settore Concorsuale di afferenza: 06/G1**

**Settore Scientifico disciplinare: Med38**

EVOLUZIONE CLONALE NELLA RECIDIVA DELLA LEUCEMIA ACUTA  
MIELOIDE DELL'ETA' PEDIATRICA A CARIOTIPO NORMALE  
MEDIANTE IL SEQUENZIAMENTO DELL'ESOMA TRAMITE  
NEXT-GENERATION SEQUENCING

**Presentata da: Dr.ssa Tamara Belotti**

**Coordinatore Dottorato  
Prof. Pier Luigi Lollini**

**Relatore  
Prof. Andrea Pession**

**Esame finale anno 2017**

## Sommario

<b>INDICE .....</b>	<b>1</b>
<b>1 INTRODUZIONE .....</b>	<b>2</b>
1.1 LEUCEMIA ACUTA MIELOIDE IN ETA' PEDIATRICA .....	2
1.1.1 EPIDEMIOLOGIA E PROGNOSI .....	2
1.1.2 PRESENTAZIONE CLINICA .....	3
1.1.3 DIAGNOSI E CLASSIFICAZIONE .....	4
1.1.4 CARATTERIZZAZIONE CITOGENETICA E MOLECOLARE .....	7
<b>2 LO STUDIO .....</b>	<b>11</b>
2.1 OBIETTIVI .....	11
2.2 MATERIALI E METODI .....	11
2.2.1 PAZIENTI .....	11
2.2.2 SEQUENZIAMENTO MASSIVO PARALLELO DELL'ESOMA .....	12
2.2.3 ANALISI BIOINFORMATICA DEI DATI .....	13
2.2.3.1 FILTRAGGIO A BASSA ASTRINGENZA .....	14
2.2.3.2 FILTRAGGIO AD ALTA ASTRINGENZA .....	14
2.2.4 RISEQUENZIAMENTO MIRATO (VALIDAZIONE MEDIANTE SANGER) .....	15
2.2.4.1 DISEGNO DEI PRIMER .....	15
2.2.4.2 PCR .....	15
2.2.4.3 PURIFICAZIONE DEGLI AMPLIFICATI .....	16
2.2.4.4 SEQUENZIAMENTO SANGER .....	16
2.2.4.5 PURIFICAZIONE DELLE SEQUENZE .....	16
2.2.4.6 ANALISI DELLE SEQUENZE E VALIDAZIONE DELLE MUTAZIONI .....	17
2.2.5 SEQUENZIAMENTO MIRATO DELLE MUTAZIONI DRIVER .....	17
2.3 RISULTATI .....	17
2.4 DISCUSSIONE .....	26
2.5 CONCLUSIONI .....	31
<b>3 BIBLIOGRAFIA .....</b>	<b>33</b>

# 1 INTRODUZIONE

## 1.1 LEUCEMIA ACUTA MIELOIDE IN ETA' PEDIATRICA

Con il termine Leucemia Mieloide Acuta (LAM), altrimenti definita leucemia acuta non linfoblastica (LAnL), si intende un gruppo eterogeneo di leucemie che origina a partire dalla cellula staminale emopoietica orientata in senso mieloide: in relazione al tipo e al grado di differenziazione del clone cellulare interessato si possono riconoscere diverse forme (mieloblastica, promielocitica, mielomonocitica, monocitica, eritroblastica, megacarioblastica). In particolare, queste neoplasie derivano dalla trasformazione clonale dei precursori emopoietici indotta dall'acquisizione di riarrangiamenti cromosomici e mutazioni geniche multiple che sono in grado di alterare i meccanismi proliferativi e differenziativi delle cellule stesse.

### 1.1.1 EPIDEMIOLOGIA E PROGNOSI

La LAM rappresenta il 15-20% delle leucemie acute dell'età pediatrica ed è responsabile del 30% delle morti correlate a queste neoplasie [1]. L'incidenza annuale nei Paesi industrializzati è di circa 7-8 casi per milione di bambini/adolescenti di età inferiore a 18 anni [2], in Italia l'incidenza è di circa 65-70 nuovi casi l'anno. L'incidenza è maggiore nel sesso maschile (7,5 casi per milione) rispetto al sesso femminile (6.1 casi per milione)[2]. I sottotipi di LAM della classificazione *French-American-British* (FAB) sono equamente distribuiti nei gruppi etnici con l'eccezione della Leucemia Acuta Promielocitica (LAP), che mostra un'incidenza più alta nei bambini di origine ispanica o latina. La prognosi dei bambini con LAM è migliorata in modo significativo nel corso degli ultimi 30 anni [3]. Anche se la maggior parte dei pazienti raggiunge la remissione completa con la chemioterapia [4, 5, 6], circa il 30% dei bambini affetti da LAM recidiva [7, 8]. La prognosi di questi pazienti è nettamente peggiore rispetto all'esordio, con una probabilità di sopravvivenza globale (pOS) compresa tra il 29% e il 38% [7, 8, 9], facendo della recidiva la principale causa fallimento terapeutico e di morte in questi bambini. Alla luce dell'alta frequenza dei decessi correlati al trattamento (5-10%), sia con il trattamento di prima linea che con i protocolli per la recidiva di malattia, un'ulteriore intensificazione

della chemioterapia standard non sembra essere un'opzione per migliorare la prognosi di questi bambini [10].

### 1.1.2 PRESENTAZIONE CLINICA

La presentazione clinica della LAM dell'età pediatrica riflette i segni e i sintomi correlati all'infiltrazione leucemica del midollo osseo e dei siti extramidollari. L'insufficienza midollare che consegue alla sostituzione delle cellule emopoietiche da parte del clone neoplastico determina sviluppo di neutropenia, anemia e piastrinopenia. I bambini presentano comunemente le manifestazioni cliniche della pancitopenia, tra cui febbre, astenia, pallore, petecchie, ecchimosi, emorragie cutanee e mucose, dolori ossei e infezioni. In particolare il quadro di presentazione clinica di qualsiasi sottotipo di LAM si può associare allo sviluppo di una coagulazione intravascolare disseminata, ma questa risulta essere più frequentemente osservata nella leucemia acuta promielocitica (in più del 90% dei casi). Questa complicanza è innescata dall'attività procoagulante dei granuli dei promielociti ed è frequentemente aggravata dall'iniziale citolisi indotta dalla chemioterapia.

L'infiltrazione neoplastica di alcuni siti extramidollari può comportare sviluppo di linfadenopatia, epato-splenomegalia, noduli sottocutanei o lesioni definite a *blueberry muffin*, infiltrazione della gengiva, dell'orbita, dello spazio epidurale e raramente del testicolo, tumefazioni descritte come sarcoma granulocitico o cloroma (tipicamente associate a LAM con t(8;21) in assenza di apparente interessamento del midollo osseo) e interessamento del sistema nervoso centrale (che si verifica in circa il 15% dei casi alla diagnosi); la leucemia extramidollare è più comune nei pazienti con LAM sotto i due anni d'età. I bambini che presentano una leucocitosi marcata possono sviluppare segni e sintomi di leucostasi, più frequentemente localizzati a livello di polmone e cervello. Un esempio particolare è rappresentato dalla leucemia acuta megacariocitaria con t(1;22) che comunemente presenta aspetti clinici che richiamano quelli di un tumore solido con epato-splenomegalia e/o lesioni ossee (periostiti bilaterali e simmetriche e lesioni osteolitiche), a volte senza coinvolgimento del midollo osseo.

### 1.1.3 DIAGNOSI E CLASSIFICAZIONE

La diagnostica include la citomorfologia delle cellule leucemiche, l'immunofenotipo, il cariotipo e la genetica molecolare. L'eterogenicità della LAM si riflette nelle differenti presentazioni cliniche, nella morfologia, nell'immunofenotipo e nelle aberrazioni citogenetiche e molecolari.

Le LAM vengono distinte, sotto un profilo eziopatogenetico, in forme primarie (*de novo*) che compaiono in soggetti per i quali non è dimostrabile un'esposizione ad agenti leucemogeni, che rappresentano la maggior parte dei casi, e forme secondarie associate a fattori ambientali, condizioni ereditarie e disordini acquisiti predisponenti. In particolare tra le forme secondarie si annoverano:

- LAM che rappresentano l'evoluzione neoplastica di una precedente sindrome mielodisplastica;
- LAM associate all'esposizione ad agenti chemioterapici, radiazioni ionizzanti e prodotti chimici quali derivati del petrolio, solventi organici (benzene), erbicidi e pesticidi (organofosfati) [11, 12];
- LAM correlate a condizioni ereditarie: sindrome di Down, anemia di Fanconi, neutropenia congenita severa (sindrome di Kostmann), sindrome di Shwachman-Diamond, sindrome di Diamond-Blackfan, neurofibromatosi di tipo 1, sindrome di Noonan, discheratosi congenita, disordine familiare delle piastrine con predisposizione a LAM (FDP/AML, *Familial Platelet Disorder with a predisposition to Acute Myeloid Leukaemia*), trombocitopenia congenita amegacariocitica, atassia-telengectasia, sindrome di Klinefelter, sindrome di Li-Fraumeni e sindrome di Bloom [13, 14];
- LAM connesse a gravi condizioni acquisite comprendenti l'anemia aplastica, la trombocitopenia acquisita amegacariocitica e l'emoglobinuria parossistica notturna.

L'evoluzione della classificazione della LAM riflette una migliore comprensione della sua patogenesi. La classificazione del gruppo FAB fornisce una terminologia utile e comune dal punto di vista della comprensione e collaborazione internazionale, tuttavia appare limitata in termini di descrizione biologica, prognostica e terapeutica [8]. Secondo la classificazione FAB, le LAM si distinguono nei seguenti maggiori sottotipi: M0, M1, M2, M3, M3v, M4, M4eo, M5a, M5b, M6, M7 (tabella 1).

**Tabella 1.** Classificazione istochimico-morfologica FAB.

SIGLA	LAM	CRITERI MORFOLOGICI
M0*	<i>Indifferenziata</i>	MPO <3% NAE negativa
M1	<i>Mieloblastica senza maturazione</i>	Blasti >90% nel midollo esclusi eritroblasti; <10% componente monocitaria; <10% granulociti, MPO >3%
M2	<i>Mieloblastica con maturazione</i>	Blasti 30<90%; granulociti >10%; monociti <20% (NE), MPO >3%
M3**	<i>Promielocitica</i>	>20% promielociti anomali ipergranulari (corpi di Auer), MPO>3%
M3v**	<i>Promielocitica variante ipogranulare</i>	Nuclei reniformi, granulazioni fini nel citoplasma, poco visibili al MO, ma MPO +++, MPO >3%
M4	<i>Mielomonocitica</i>	Blasti >30%, NE componente granulocitaria 20<80%, componente monocitaria 20<80%, >5x10 <sup>9</sup> /l monociti nel sangue periferico o lisozima elevato, MPO >3%
M4 eo	<i>Mielomonocitica con ipereosinofilia</i>	Criteri per M4 associati a eosinofili anormali con granuli eosinofili e basofili, MPO >3%
M5a	<i>Monoblastica</i>	>80% NE componente monocitaria; monoblasti >80% della componente monocitaria, NAE positiva
M5b	<i>Monocitica</i>	>80% NE componente monocitaria; monoblasti <80% della componente monocitaria NAE positiva
M6	<i>Eritroleucemia</i>	Cellule eritroidi >50% delle cellule midollari; blasti >30% NE
M7***	<i>Megacarioblastica</i>	Blasti >30% NE; megacarioblasti, <i>blebs</i> , mielofibrosi
<p>*M0:</p> <ul style="list-style-type: none"> <li>• MPO e SBB &lt; del 3% in citochimica;</li> <li>• <i>markers</i> B e T-lineage negativi (ivi compresi CD3 e CD22 intracitoplasmatici);</li> </ul> <ul style="list-style-type: none"> <li>• CD13 e CD33 positivi in circa il 60% e 80% rispettivamente dei pazienti; co-espressione di CD13 e CD33 in circa il 40% dei casi; almeno uno dei due antigeni deve essere espresso per la diagnosi;</li> <li>• CD13 più sensibilmente dimostrato da tecniche immunocitochimiche (immunoperossidasi e APAAP) che dalla citometria a flusso; CD117, CD7 e TdT riscontrati nel 55, 41 e 46% dei casi rispettivamente.</li> </ul> <p>**M3:</p> <p>L'identificazione di questo sottotipo è critica per il rischio emorragico potenzialmente fatale che può comparire prima o durante la terapia di induzione.</p> <p>***M7:</p> <ul style="list-style-type: none"> <li>• Megacarioblasti &gt; 30% MO; <i>blebs</i> citoplasmatici e fibrosi midollare;</li> </ul> <ul style="list-style-type: none"> <li>• Positività per CD61, CD42, CD41 e antigene fattore VIII-relato;</li> <li>• Possono inoltre essere utili la dimostrazione ultrastrutturale di perossidasi piastrinica o del <i>subset</i> di cellule CD34+ è CD38+, e la biopsia ossea.</li> </ul>		

Il 50-60% delle LAM diagnosticate in età pediatrica è classificato come: M1, M2, M3, M6, o M7; nel restante 40% dei casi si tratta di forme M4 o M5. Circa l'80% dei bambini con una diagnosi di LAM al di sotto dei 2 anni d'età è affetto da un sottotipo M4 o M5.

La classificazione WHO del 2016 (tabella 2), come già la classificazione WHO del 2008, mantiene la terminologia FAB per la classificazione morfologica della LAM non altrimenti specificata (AML-NOS, *Acute Myeloid Leukaemia Not Otherwise Specified*,) ed entrambe, rispetto alla classificazione WHO del 2001, introducono tre anomalie cromosomiche bilanciate, più rare in termini epidemiologici [15].

**Tabella 2.** Classificazione WHO 2016.

<p><i>LAM con anomalie genetiche ricorrenti:</i></p> <ul style="list-style-type: none"> <li>• LAM con t(8;21)(q22;q22); (RUNX1-RUNX1T1)</li> <li>• LAM con inv(16)(p13.1q22) o t(16;16)(p13.1;q22); (CBFB-MYH11)</li> <li>• Leucemia acuta promielocitica (APL) con t(15;17)(q22;q12); (PML;RARA)</li> <li>• LAM con t(9;11)(p21.3;q23.3); (MLLT3-KMT2A)</li> <li>• LAM con t(6;9)(p23;q34); (DEK-NUP214)</li> <li>• LAM con inv(3)(q21.3;q26.2) o t(3;3)(q21.3;q26.2); (GATA2, MECOM)</li> <li>• LAM (megacariocitica) con t(1;22)(p13;q13); (RBM15-MKL1)</li> <li>• forma provvisoria: LAM con BCR-ABL1</li> <li>• LAM con gene NPM1 mutato</li> <li>• LAM con gene CEBPA mutato</li> <li>• forma provvisoria: LAM con gene RUNX1 mutato</li> </ul>
<p><i>LAM con alterazioni correlate a sindromi mielodisplastiche</i></p>
<p><i>Neoplasie mieloidi correlate a terapie</i></p>
<p><i>LAM non altrimenti specificate:</i></p> <ul style="list-style-type: none"> <li>• LAM con differenziazione minima</li> <li>• LAM senza maturazione</li> <li>• LAM con maturazione</li> <li>• Leucemia mielomonocitica acuta</li> <li>• Leucemia monoblastica/monocitica acuta</li> <li>• Leucemia acuta eritrocitica: leucemia eritrocitica pura</li> <li>• Eritroleucemia</li> <li>• Leucemia acuta megacariocitica</li> <li>• Leucemia acuta basofilica</li> <li>• Panmielosi acuta con mielofibrosi</li> </ul>

## Sarcoma mieloide

### *Proliferazioni mieloidi correlate a sindrome di Down:*

- Mielopoiesi anormale transitoria
- LAM associata a sindrome di Down

Gli altri aspetti che la classificazione WHO 2016 ha mantenuto dalla classificazione WHO 2008, all'epoca peculiari, sono: il mantenimento della LAM correlata alla terapia come categoria di importanza clinica e biologica [16], l'ulteriore rifinizione della diagnosi di mielodisplasia o di patologia correlata a sindrome mielodisplastica (MDS) [16] e il raggruppamento separato delle proliferazioni mieloidi correlate a sindrome di Down.

### 1.1.4 CARATTERIZZAZIONE CITOGENETICA E MOLECOLARE

L'analisi genetica riveste una notevole importanza nella diagnosi di LAM in quanto alcune anomalie cromosomiche rappresentano importanti marcatori diagnostici con notevole rilevanza prognostica. Anomalie cromosomiche clonali sono state identificate nei blasti del 70-80% dei bambini con LAM e si sono dimostrate utili nel definire i sottotipi per particolari associazioni pressoché costanti quali: t(8;21) e M2, t(15;17) e M3, inv16 e M4Eo, anomalità 11q23 con M4 e M5, t(1;22) e M7. Le leucemie con t(8;21) e inv16 sono anche definite *core-binding factor (CBF) leukemias*, ovvero leucemia con coinvolgimento del CBF, fattore di trascrizione coinvolto nella differenziazione staminale emopoietica. Anomalie citogenetiche ricorrenti sono riportate in tabella 3.

**Tabella 3.** Anomalie citogenetiche ricorrenti.

<i>Tipo</i>	<i>Sottotipo</i>	<i>Cromosoma-Gene</i>		<i>Note</i>
APL	LAM T(15;17)	15-PML	17-RAR $\alpha$	Ottima risposta all'ATRa
	LAM T(11;17)	11-PLZF	17-RAR $\alpha$	Ridotta sensibilità all'ATRA

LAM t(8;21)		21-AML1	8-ETO	Associate con sarcoma granulocitico. Prognosi favorevole
CBF leukemias	LAM inv(16)	cr.16q22-CBF $\beta$	16p13-MYH11	Prognosi favorevole
	LAM T(16;16)	16q22-CBF $\beta$	16p13-MYH11	Prognosi favorevole
LAM MLL	LAM 11q23	>50 geni partner descritti	11q23-MLL	Prognosi sfavorevole. Caratteristiche delle LAM secondarie a terapie antitumorali.
	t(9;11)	9 – AF9	11q23-MLL	Prognosi favorevole
LAM t(1;22)		1p13-OTT (RBM15)	1p13-MAL (MLK1)	Prognosi severa
Altre anomalie	-7	7q22		Rare. Prognosi sfavorevole.
	del(7q)			
	-5	13 bande da 5q11 a 5q35		
	del(5q)			
	inv(3)	3q21-?	3q26-EV11	
	t(3;3)			

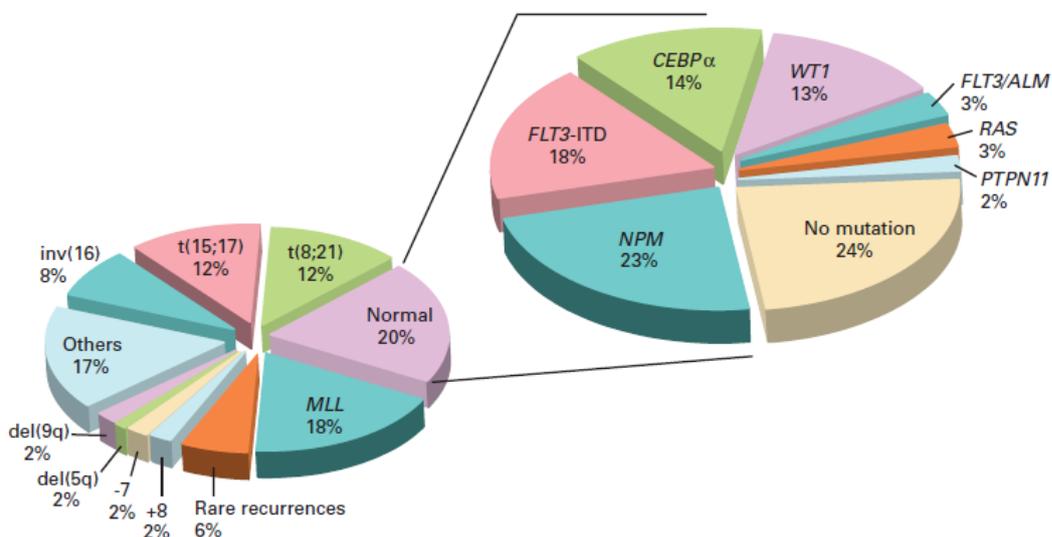
Altre anomalie molecolari ricorrenti di interesse non solo per il monitoraggio della malattia residua minima, ma anche potenziale ruolo di bersaglio terapeutico, sono rappresentate da:

- mutazione di FLT3: la presenza di una mutazione di FLT3 a tipo *internal-tandem duplication* (ITD) si associa a prognosi severa, anche se nell'età pediatrica è molto meno frequente (5-10%) che nell'adulto. Oggi sono disponibili inibitori di FLT3;
- mutazioni ras e del recettore per c-KIT: mutazioni dell'oncogene ras sono documentabili in circa il 25% dei casi di LAM, ma il loro significato prognostico non è noto. Mutazioni di c-KIT interessano meno del 5% delle LAM e frequentemente si associano ad anomalie CBF e prognosi severa;
- mutazioni di GATA1: le mutazioni di GATA1 sono caratteristiche della LAM M7 o TMD in sindrome di Down. Tali mutazioni

conferiscono ai mieloblasti un'aumentata sensibilità alla citarabina attraverso la *down*-regolazione dell'enzima citidina deaminasi. Questo meccanismo è alla base della buona risposta alle forme di LAM M7 in soggetti con sindrome di Down a regimi chemioterapico contenenti citarabina.

- mutazioni della Nucleofosmina (NPM1): NPM1 è una proteina essenziale per l'assemblaggio delle unità ribosomi ali con importanti funzioni di trasporto trans-membrana ed aggregazione proteica a livello nucleolare. Mutazioni di NPM1, peraltro rare in età pediatrica, riducono la localizzazione nucleare della proteina e si associano a forme CD34 negative. Nonostante la natura estremamente eterogenea della LAM, vari sottotipi sembrano condividere alcuni percorsi comuni che portano alla leucemogenesi. Diverse evidenze hanno portato ad un modello che mostra come la LAM nasce dalla cooperazione tra due classi di alterazioni genetiche che regolano auto-rinnovamento e differenziazione. Ad esempio, le mutazioni o alterazioni epigenetiche che coinvolgono CBF o i geni MLL, definite alterazioni di tipo II, giocano un ruolo chiave nel modificare la capacità dei precursori di differenziare, ma di solito non sono sufficienti da sole a generare la LAM. A sostegno di questo c'è l'evidenza che il trascritto di fusione AML1-ETO risultante dalla traslocazione t(8; 21) può essere rilevato nel sangue neonatale di adolescenti con LAM caratterizzata da questa traslocazione. Tali trascritti di fusione possono persistere nel midollo osseo di alcuni pazienti in remissione morfologica lungo termine.

**Figura 1.** Frequenza stimata delle alterazioni genetiche nella LAM pediatrica [17].



Il 20% delle LAM dell'infanzia è rappresentata della classe a cariotipo normale, cioè casi senza alterazioni citogenetiche ricorrenti. In questo gruppo di pazienti, un'adeguata caratterizzazione molecolare, la stratificazione del rischio e il monitoraggio della malattia rimangono compiti difficili con una conseguente grande variabilità in termini di risposta alla terapia e prognosi finale. Diventa quindi fondamentale la conoscenza delle alterazioni molecolari sottostanti la LAM, in particolare di quelle coinvolte nello sviluppo della recidiva, al fine di elaborare strategie di trattamento paziente-specifiche. Nell'ultimo decennio, diversi progressi tecnologici hanno notevolmente migliorato la capacità di studiare i genomi di cellule tumorali. Con lo sviluppo del sequenziamento massivo-parallelo (*Next Generation Sequencing*, NGS) si è arrivati ad oggi a poter sequenziare un intero genoma (*Whole-Genome Sequencing*), la sua componente codificante (*Whole-Exome Sequencing*, WES), oppure tutti i geni espressi (*Whole-Transcriptome Sequencing*), permettendo notevoli passi avanti nello studio della genomica del cancro [18, 19, 20]. Sulla base di recenti studi, la recidiva di malattia sembra essere associata con un'evoluzione clonale dall'esordio alla recidiva, promossa almeno in parte dalla stessa chemioterapia. Questo modello è supportato dal lavoro di Ding et al [21] negli adulti affetti da LAM, che solo recentemente è stato confermato da Farrar et al [22] nei bambini. Questi autori hanno individuato diversi subcloni all'interno dell'intera popolazione tumorale che sono stati caratterizzati dall'acquisizione di ulteriori mutazioni.

## 2 LO STUDIO

### 2.1 OBIETTIVI

La maggior parte dei bambini con LAM muore per una progressione di malattia dopo la recidiva. In base alle conoscenze attuali si ipotizza che il verificarsi della recidiva sia associato ad un'evoluzione clonale dello spettro mutazionale della malattia. Per meglio indagare le modalità secondo cui tale evoluzione avviene e poter individuare precisi pattern di evoluzione, si è scelto di sequenziare l'esoma dei blasti dei pazienti all'esordio e alla recidiva, al fine di confrontare lo spettro mutazionale durante le diverse fasi di malattia. Le conoscenze così acquisite permetteranno di identificare sia nuove mutazioni rilevanti nella patogenesi della malattia, che nella recidiva e valutare il loro contributo alla selezione clonale e alla resistenza alla chemioterapia. In accordo con quanto è già riportato in letteratura [18], i pattern di evoluzione clonale potrebbero delinarsi secondo due modalità:

1. il clone principale del tumore primitivo acquisisce nuove mutazioni e si espande dando luogo al clone della recidiva;
2. un subclone minoritario alla diagnosi, resistente alla terapia di prima linea, acquisisce ulteriori mutazioni e diventa il clone maggioritario alla recidiva. Attraverso la comprensione di tali meccanismi biologici, sarà possibile modulare la terapia in maniera sempre più adeguata e personalizzata.

Obiettivo di questo studio è stato definire:

1. lo spettro mutazionale del clone alla diagnosi e alla recidiva;
2. precisi pattern di evoluzione clonale, responsabili della recidiva;
3. l'identificazione di nuove mutazioni alla recidiva e il loro contributo alla selezione clonale e alla resistenza alla chemioterapia;
4. la selezione clonale chemioterapia-indotta.

### 2.2 MATERIALI E METODI

#### 2.2.1 PAZIENTI

I campioni analizzati sono stati raccolti al momento della diagnosi, della prima remissione completa e della recidiva di malattia da 4 bambini affetti da LAM non promielocitica di nuova insorgenza, arruolati nel protocollo AIEOP 2002/01, dopo aver ottenuto il consenso informato da

parte dei genitori secondo la dichiarazione di Helsinki. La diagnosi morfologica FAB e l'analisi immunofenotipica sono state centralizzate e revisionate presso il laboratorio di Ematologia Pediatrica dell'AOSP di Padova. L'analisi cromosomica è stata eseguita su sangue midollare utilizzando procedure di laboratorio standardizzate. I cariotipi sono stati riportati in base all'*International System for Human Cytogenetic Nomenclature* (ISCN 2005). Le caratteristiche cliniche dei pazienti analizzati sono rappresentate nella tabella 4.

**Tabella 4.** Caratteristiche cliniche dei pazienti dei quali è stato sequenziato l'esoma.

Paziente	Sesso	Età alla diagnosi (anni)	FAB	Leucociti alla diagnosi ( $\times 10^3/\text{ml}$ )	Blasti midollari alla diagnosi (%)	Leucociti alla recidiva ( $\times 10^3/\text{ml}$ )	Blasti midollari alla recidiva (%)	Intervallo Disease-free (mesi)	TCSE	Coinvolgimento extramidollare
*LAM#1	M	3.3	M2	12.010	70	3.210	60	24	SI <sup>1</sup> (AUTO)	NO
LAM#2	M	14.5	M5	114.040	92	53.010	80	14	SI <sup>2</sup> (MUD)	NO
LAM#3	F	7.4	M4	4.020	53	1.700	70	11	SI <sup>2</sup> (MUD)	NO
*LAM#4	M	12.6	M1	14.000	80	6.900	24	18	SI <sup>2</sup> (AUTO)	SI (Tonsille)

\* paziente vivo in remissione completa (RC); 1 TCSE in 1° RC; 2 TCSE in 2° RC.

### 2.2.2 SEQUENZIAMENTO MASSIVO PARALLELO DELL'ESOMA

Il DNA genomico è stato estratto mediante il kit QiaAmp Mini (Qiagen) a partire da blasti leucemici midollari/cellule mononucleate dei 4 pazienti con LAM. Le *library* di DNA e l'arricchimento esomico sono state effettuate mediante il kit Nextera Rapid Capture Exome Enrichment (Illumina, San Diego, CA). Con un processo di amplificazione a ponte (*bridge amplification*), le *library* sono state amplificate formando i *clusters* attraverso il sistema cBot cluster amplification system/TruSeq PE Cluster Kit v3-cBot-HS (Illumina). Quindi, le librerie sono state sequenziate con il kit TruSeq SBS Kit v3-HS, Illumina a 75 o 100 bp in *paired-end* (a seconda della dimensione media dei frammenti nella *library* finale) su piattaforma HiScanSQ di Illumina, presente all'interno del Laboratorio di Oncologia ed Ematologia Pediatrica del Policlinico S.Orsola-Malpighi.

La piattaforma HiScanSQ produce un massimo di 1.5 miliardi di *reads* totali per ogni corsa *paired-end* da 2x100 basi producendo fino a 150 Gb di *output* con una media di 60-70 milioni di *reads* per campione se caricato in pool da 3 campioni per *lane*.

### 2.2.3 ANALISI BIOINFORMATICA DEI DATI

Nella fase di *base-calling* del sequenziamento massivo *Whole Exome Sequencing* (WES) i segnali di fluorescenza sono stati acquisiti da un sistema interconnesso che converte le immagini in chiamate nucleotidiche generando i file FASTQ i quali contengono le *read* corrispondenti alle sequenze acquisite. I dati così ottenuti sono stati processati con il *software* Adapter Removal per il *trimming* delle basi a bassa qualità e la rimozione degli adattatori. Successivamente le sequenze ottenute sono state mappate con *Burrows-Wheeler Aligner* contro un genoma di riferimento contenente la sequenza genomica dei cromosomi umani 1-22, X, Y e il genoma mitocondriale derivante dalla *build* GRCh37/hg19 del genoma umano disponibile al sito UCSC Genome Browser. Dopo il passaggio di rimozione dei duplicati di PCR (effettuato con il *software* Samtools), è stato impiegato il *Genome Analysis Toolkit* (GATK) per la fase di ottimizzazione dell'allineamento in prossimità di inserzioni e delezioni (*InDel*) e per il processo di ricalibrazione delle qualità delle basi. Lo stesso GATK è stato adottato per la chiamata delle *Indel* (funzione *Haplotype Caller*) mentre le varianti a singolo nucleotide (SNV) sono state identificate con l'algoritmo *MuTect*.

L'annotazione delle varianti ottenute è stata effettuata mediante ANNOVAR. Questo software è in grado di distinguere quali delle SNV identificate siano proprie del paziente e quali riconducibili alla variabilità umana (frequenza nella popolazione > 1%). Tutte le SNV sono state cercate nel *database* dbSNP (<http://www.ncbi.nlm.nih.gov/project/SNP>), versione 138 e nel *database* 1000 Genomes (<http://www.1000genomes.org>), aggiornato nel febbraio del 2012 e scartate se segnalate come polimorfiche.

Per effettuare la prioritizzazione delle varianti ottenute sono state adottate due tipologie di filtraggio: filtraggio a bassa astringenza e filtraggio ad alta astringenza.

### 2.2.3.1 FILTRAGGIO A BASSA ASTRINGENZA

Inizialmente, per individuare tutte le possibili varianti somatiche coinvolgenti i geni implicati nella patogenesi della LAM, le varianti individuate sono state filtrate secondo criteri poco restrittivi per evitare di tralasciare dati potenzialmente rilevanti. Sono state quindi considerate solo le varianti che rispettavano i parametri seguenti:

- Effetto della variante: SNV non sinonima, SNV nonsense, SNV in prossimità di siti di splicing, delezione o inserzione *frameshift* e *non-frameshift*; sono state quindi scartate le SNV sinonime.
- Varianti valutabili: *coverage* totale  $\geq 5$  e *ratio* base alterata/base di riferimento  $\geq 0.1$  in almeno uno dei campioni;
- Varianti somatiche: assenti nel campione alla remissione o presenti in piccola percentuale potenzialmente ascrivibile ad errori del sequenziamento in tale campione o a persistenza del clone di malattia in maniera morfologicamente non evidenziabile. Per tale selezione si sono usati i criteri di una *coverage* totale alla remissione  $> 5$  e un *ratio* base alterata/base di riferimento alla remissione  $< 0.07$ . Sono state quindi scartate le varianti germinali, cioè presenti anche alla remissione, e le varianti indeterminate, cioè con *coverage* totale alla remissione  $\leq 5$ .

### 2.2.3.2 FILTRAGGIO AD ALTA ASTRINGENZA

Successivamente si è voluto focalizzare l'interesse sulle mutazioni coinvolgenti nuovi possibili "geni candidati" nella patogenesi o nella recidiva della malattia oppure potenzialmente rivelatori dell'evoluzione clonale durante la storia naturale della LAM. Tali varianti sono state selezionate in base ad una valutazione complessiva che teneva conto di uno o più dei seguenti criteri:

- valore minimo di *coverage* totale (numero di *read* che mappano nella posizione genomica della mutazione)  $\geq 15$  in almeno uno dei due campioni tra esordio e recidiva;
- effetto sulla proteina delle mutazioni nelle regioni codificanti valutato mediante l'utilizzo di diversi software di predizione (PROVEAN e SIFT);
- geni notoriamente mutati nell'ambito delle LAM;
- geni codificanti per proteine potenzialmente importanti nella patogenesi della LAM in base a quanto riportato in letteratura;
- geni mutati in due o più dei pazienti analizzati.

## 2.2.4 RISEQUENZIAMENTO MIRATO (VALIDAZIONE MEDIANTE SANGER)

Le mutazioni così individuate sono state validate mediante sequenziamento classico Sanger. La validazione è stata effettuata su tutti tre i campioni dello stesso paziente in cui tale mutazione era stata individuata mediante WES, per valutare la correttezza della chiamata e l'effettiva presenza o meno della mutazione nelle diverse fasi di malattia.

### 2.2.4.1 DISEGNO DEI PRIMER

I *primer* sono stati disegnati mediante il software NCBI/Primer BLAST (<http://www.ncbi.nlm.nih.gov/tools/primer-blast/>), in modo tale che i *primer Forward* e *Reverse* abbiano una temperatura di melting simile ed lunghezza tra le 20 e 30 basi. La lunghezza dell'amplicone invece è stata scelta tra le 151 e le 436 bp. Le sequenze geniche del frammento da amplificare sono state ottenute mediante la banca dati di Ensembl (<http://www.ensembl.org/index.html>). La tendenza delle coppie di *primer* a formare *hairpin*, eterodimeri o omodimeri è stata verificata tramite software OligoAnalyzer 3.1 disponibile al sito <http://eu.idtdna.com/calc/analyzer>.

### 2.2.4.2 PCR

L'amplificazione dei campioni è stata eseguita mediante l'utilizzo dell'AmpliTaq FastStart kit (Applied Biosystem). A 2 µl di DNA sono stati aggiunti 23 µl di mix, composta da:

- 5 µl di Buffer 5X
- 0.5 µl di dNTPs 10mM
- 1.25 µl di *primer Forward* 10µM
- 1.25 µl di *primer Reverse* 10 µM
- 0.1 µl di enzima Fast Start 5U/µl (FastStart Taq polimerasi, Roche)
- 15.9 µl di H<sub>2</sub>O

L'amplificazione è stata eseguita secondo le seguenti condizioni: 95°C per 2 minuti per l'attivazione della polimerasi, e 35 cicli di PCR composti da: denaturazione a 95°C per 15 secondi, *annealing* con una temperatura specifica per ogni primer per 15 secondi ed estensione a 72°C per 1 minuto.

#### **2.2.4.3 PURIFICAZIONE DEGLI AMPLIFICATI**

La purificazione degli amplificati è stata effettuata mediante l'utilizzo delle piastre filtranti Multiscreen PCR (Millipore) con l'ausilio di una pompa per il vuoto. Per la purificazione sono stati aggiunti 70 µl di H<sub>2</sub>O ad ogni prodotto di PCR e tutto il volume è stato trasferito su una piastra filtrante, precedentemente inserita nel collettore di aspirazione della macchina per il vuoto. Dopo aver applicato il vuoto con una pressione di circa 90 mbar per 10 minuti, il DNA adeso alla membrana è stato risospeso in 30 µl di H<sub>2</sub>O mediante agitazione meccanica per 10 minuti.

#### **2.2.4.4 SEQUENZIAMENTO SANGER**

I campioni purificati sono stati sequenziati mediante il Big Dye Terminator v 1.1 Cycle Sequencing Kit (Applied Biosystems). A 2 µl di DNA amplificato purificato è stata aggiunta una mix, costituita da:

- 0.32 µl di *primer* (Forward o Reverse)
- 2 µl di Buffer 5X
- 1µl di Big Dye 1.1
- H<sub>2</sub>O per arrivare ad un volume finale di 10 µl

Il sequenziamento è stato effettuato nelle seguenti condizioni: 96°C per 30 secondi e 25 cicli a 96°C per 10 secondi più 4 minuti a 60°C .

#### **2.2.4.5 PURIFICAZIONE DELLE SEQUENZE**

Le sequenze sono state purificate aggiungendo ai campioni 40 µl di mix composta da:

- 1.5 µl di sodio acetato
- 30.64 µl di etanolo assoluto
- 7.86 µl di H<sub>2</sub>O

I campioni sono stati quindi incubati a -20°C per 1 ora, poi centrifugati a massima velocità a 4°C per 30 minuti e infine è stato eliminato il surnatante. Dopo un lavaggio con etanolo 70%, i campioni sono stati centrifugati a massima velocità a 4°C per 15 minuti. L'etanolo rimanente è stato fatto evaporare in fornetto a 37°C.

#### **2.2.4.6 ANALISI DELLE SEQUENZE E VALIDAZIONE DELLE MUTAZIONI**

Terminata la purificazione, i prodotti di reazione sono stati risospesi in Injection buffer e sequenziati mediante il sequenziatore automatico ABI

3730 Genetic Analyzer (Applied Biosystem), che sfrutta i segnali di fluorescenza dei terminatori di catena per fornire la sequenza delle basi del frammento amplificato. Per la lettura delle sequenze è stato utilizzato il software Sequencer 4.1 (Applied).

### 2.2.5 SEQUENZIAMENTO MIRATO DELLE MUTAZIONI DRIVER

Allo scopo di chiarire i meccanismi molecolari alla base della recidiva della malattia, tra le mutazioni rilevate tramite WES, abbiamo ulteriormente selezionato le mutazioni *driver* dei geni candidati per il sequenziamento mirato, sulla base di un'analisi della loro ricorrenza nella LAM o su un possibile ruolo patogenetico, in accordo con la letteratura. L'allestimento delle *library* di DNA è stato eseguito mediante il kit Nextera XT DNA Prep (Illumina). L'amplificazione dei campioni è stata eseguita mediante l'utilizzo dell'AmpliTaq FastStart kit (Applied Biosystem). Il sequenziamento mirato è stato eseguito su piattaforma MiSeq (Illumina) a 75 paia di basi in *paired end*, su flowcell v3.

## 2.3 RISULTATI

Al fine di indagare l'evoluzione clonale della LAM pediatrica ed individuare precisi modelli mutazionali responsabili della recidiva, sono stati sottoposti a sequenziamento massivo dell'esoma i campioni relativi alle condizioni di esordio, remissione e recidiva di quattro pazienti pediatrici affetti da LAM a cariotipo normale per un totale di 12 campioni. Le caratteristiche cliniche dei pazienti sono descritte nella tabella 4. Sono state sequenziate 93 miliardi di basi con una produzione totale di  $9,34 \times 10^8$  reads. La *coverage* media ottenuta nelle regioni target in ciascun campione è variabile tra un valore di 52X e un valore di 75X. L'allineamento delle sequenze contro il genoma umano e la successiva processazione dei dati ha permesso di identificare un numero totale di 4 340 023 variazioni a singolo nucleotide (SNVs) e 406 207 *InDel*. Di queste, 4 172 533, equivalenti a circa l'88%, sono state classificate come polimorfismi in base alla presenza con una frequenza >1% nelle banche dati di variabilità umana quali il database dbSNP (<http://www.ncbi.nlm.nih.gov/project/SNP>), versione 138 e il database 1000 Genomes (<http://www.1000genomes.org>), aggiornato nel febbraio del 2012. Tra le varianti classificate come *novel*, sono state distinte quelle identificate in regioni esoniche (che mappano in regioni codificanti) e

non esoniche (che mappano in 5'-3' UTR e introni). Le varianti non esoniche sono state scartate. Le varianti esoniche sono state analizzate e divise successivamente in SNV sinonime, in SNV non sinonime, intendendo come parte di quest'ultima categoria tutte le SNV che portano a modificazione amminoacidica e perdita/immissione di codoni di stop a livello della proteina corrispondente, SNV nelle regioni di splicing, *InDel* frameshift e *InDel* non frameshift.

Sono state ulteriormente considerate solo le SNV non sinonime, le SNV nelle regioni di splicing e le *InDel* sia frameshift che non frameshift, con coverage totale >5X e ratio>0.1 (rapporto tra coverage della base alterata e coverage totale) in almeno uno dei campioni. In ciascuno dei pazienti sono state individuate le seguenti varianti: 2090 SNV e 542 *InDel* in LAM#1, 2073 SNV e 549 *InDel* in LAM#2, 1992 SNV e 557 *InDel* nel paziente LAM#3, 2094 SNV e 590 *InDel* in LAM #4.

Nelle fasi successive, sono state considerate le sole varianti somatiche, cioè assenti nel campione di remissione. Sono state quindi scartate le varianti germinali o indeterminate, pari all'85% delle varianti fino ad ora considerate. Sono state così individuate un totale di 716 mutazioni nei quattro pazienti, di cui 666 SNV e 50 *InDel*. Queste mutazioni sono così ripartite tra i quattro pazienti: 132 SNV e 8 *InDel* in LAM#1, 171 SNV e 20 *InDel* in LAM#2, 179 SNV e 15 *InDel* in LAM#3, 184 SNV e 7 *InDel* in LAM#4. Successivamente si è voluto focalizzare l'interesse sulle mutazioni coinvolgenti nuovi possibili "geni candidati" nella patogenesi o nella recidiva della malattia oppure potenzialmente rivelatori dell'evoluzione clonale durante la storia naturale della LAM.

Le mutazioni considerate nelle fasi successive sono state individuate sulla base di una valutazione complessiva che tenesse conto dell'affidabilità dei dati del sequenziamento NGS oppure di un potenziale ruolo del gene mutato nella patogenesi e progressione della LAM in base a quanto riportato in letteratura. Le mutazioni così individuate sono state ricontrollate mediante il software *Integrative Genomics Viewer* (scaricabile al sito internet <http://www.broadinstitute.org/software/igv/home>) per verificare la correttezza dell'allineamento. In tal modo sono state selezionate il seguente numero di mutazioni ad alta qualità nei quattro pazienti: 6 mutazioni in LAM#1, 41 mutazioni in LAM#2, 11 mutazioni in LAM#3, 7 mutazioni in LAM#4. Questi risultati sono in accordo con l'evidenza che

il genoma della LAM generalmente ha un basso numero di mutazioni somatiche [23], con l'eccezione di LAM#2.

L'analisi di *copy number* a partire dai dati di *exome sequencing* ha mostrato la presenza di pochissime regioni di *gain* o *loss*, in accordo con i dati della letteratura [23, 24], ad eccezione del caso LAM#1, che era portatore di una delezione del braccio corto (17:0-18965000) e *gain* (17:18965000-81188000) del cromosoma 17 alla diagnosi. Questa alterazione è stata persa in recidiva, quando non era più rilevabile attraverso la sensibilità del test, probabilmente a causa della sua natura subclonale. L'elenco di tutte le SNVs e *InDel* somatiche di ogni paziente sono mostrate nelle tabelle 5.1-5.4 e riassunte in figura 2.

Il WES ci ha permesso di identificare le mutazioni diagnosi-specifiche, le mutazioni recidiva-specifiche e le mutazioni presenti sia all'esordio che alla recidiva in ogni paziente. Questi risultati hanno confermato che la LAM pediatrica è una malattia policlonale, con un'architettura complessa dell'intera popolazione leucemica, che si modifica evolvendosi dalla diagnosi alla recidiva. Con l'intento di tracciare l'origine di tali cloni, è stato effettuato un sequenziamento mirato delle mutazioni somatiche con possibile ruolo patogenetico, in accordo con la letteratura. È stata ottenuta una *coverage* media di ~ 7000X, con una sensibilità stimata nel rilevare subcloni minori di 0,12%.

Questa analisi ci ha permesso di calcolare una dimensione percentuale del clone portatore di mutazione, basando il calcolo sulla *Mutation Frequency* (MF), che rappresenta la proporzione di *reads* con l'allele mutato rispetto al numero di *reads* totali, aggiustato per il numero di copie locali del cromosoma. I risultati sono mostrati in tabella 6.

**Tabella 5.1.** SNVs e *InDel* somatiche identificate tramite WES nel paziente LAM#1.

Gene	Exon	cDNA	Tot Cov at diagnosis	MF at diagnosis	Tot Cov at relapse	MF at relapse	Provean Prediction	Sift Prediction
ATXN7L1	exon10	c.1977_1979del	38	0,1316	36	0,0000	Neutral	Deletion
NAP1L2	exon1	c.663_665del	26	0,1923	24	0,1250	Neutral	Deletion
KIF1A	exon29	c.G3010A	46	0,1957	34	0,0588	Neutral	Tolerated
TYK2	exon18	c.T2597A	95	0,3368	107	0,1495	Deleterious	Damaging
ISOC2	exon4	c.378delG	26	0,0000	18	0,2778	Frameshift	Frameshift
H2AFV	exon5	c.T356C	50	0,0400	28	0,2143	Deleterious	Tolerated

Tot Cov: *coverage* totale, MF: *mutation frequency*.

Tabella 5.2. SNVs e *InDel* somatiche identificate tramite WES nel paziente LAM#2.

Gene	Exon	cDNA	Tot Cov at diagnosis	MF at diagnosis	Tot Cov at relapse	MF at relapse	Provean Prediction	Sift Prediction
EMX1	exon1	c.C454A	15	0,3333	39	0,0000	Neutral	Damaging
KMT2D	exon39	c.11580_11582del	25	0,2400	35	0,0000	Neutral	Deletion
GLRB	exon10	c.C1414T	36	0,5556	45	0,3556	Nonsense	Nonsense
SETD2	exon15	c.6306_6307insCACC	23	0,3478	42	0,4048	Frameshift	Frameshift
SRSF7	exon8	c.G677A	71	0,1972	71	0,2394	Neutral	Damaging
LRP1B	exon46	c.C7535T	44	0,2500	34	0,5588	Neutral	Tolerated
CEBPA	exon1	c.937_938insCAG	99	0,8990	103	0,7961	Deleterious	Insertion
DOCK8	exon2	c.C141A	112	0,4018	109	0,2018	Neutral	Tolerated
ZNF503	exon1	c.78_79insAGGCGGAGGCGG	51	0,2745	64	0,1563	Neutral	Insertion
FLRT1	exon2	c.G446A	70	0,4857	95	0,5474	Neutral	Damaging
RREB1	exon10	c.G3088A	13	0,8462	21	0,5714	Neutral	Damaging
BAIAP2	exon8	c.G805A	14	0,7143	36	0,4167	Neutral	Damaging
FAM46A	exon2	c.117_131del	13	0,3846	20	0,3000	Neutral	Deletion
UFL1	exon17	c.G1985A	24	0,0000	25	0,5600	Neutral	Tolerated
MAP1A	exon4	c.G163C	61	0,0000	95	0,1895	Deleterious	Damaging
DNAH14	exon75	c.11920_11923del	75	0,0000	87	0,3333	Frameshift	Frameshift
WISP1	exon5	c.G1040A	32	0,0000	28	0,3214	Neutral	Tolerated
FLVCR2	exon1	c.G103T	100	0,0000	160	0,3188	Neutral	Tolerated
LDLR	exon15	c.C2174G	144	0,0000	175	0,2571	Deleterious	Damaging
FAM73A	exon10	c.G1175T	75	0,0000	109	0,2294	Deleterious	Damaging
SSPO	exon2	c.G215A	20	0,0000	41	0,4634	NA	NA
SELP	exon3	c.G170A	88	0,0000	56	0,1429	Deleterious	Damaging
KCNAB3	exon1	c.G179A	37	0,0000	39	0,3590	Neutral	Tolerated
MIB1	exon13	c.G1862C	26	0,0000	26	0,3077	Deleterious	Damaging
TAS1R2	exon2	c.G406A	63	0,0000	49	0,1429	Neutral	Damaging
WT1	exon5	c.G962C	75	0,0000	15	0,4000	Neutral	Tolerated
ASXL3	exon12	c.C4409T	44	0,0000	49	0,2449	Deleterious	Damaging
ZNF671	exon4	c.C681A	44	0,0000	53	0,2642	Neutral	Tolerated
ZNF630	exon5	c.T346A	20	0,0000	51	0,1765	Neutral	Tolerated
TEX11	exon12	c.C881A	8	0,0000	26	0,3462	Deleterious	Damaging
IGSF21	exon4	c.T346A	45	0,0000	84	0,3333	Deleterious	Damaging
PSIP1	exon11	c.C1003A	35	0,0000	27	0,3704	Neutral	Tolerated
FAM155B	exon1	c.54_56del	39	0,0256	27	0,1852	Neutral	Deletion
BHLHE41	exon5	c.C893T	14	0,0000	19	0,2632	Neutral	Tolerated
PCBP2	exon15	c.G1077C	76	0,0000	65	0,1538	Neutral	Tolerated
CHST6	exon3	c.C1019T	158	0,0000	114	0,5702	Neutral	Tolerated
SCAPER	exon10	c.T309A	37	0,0000	41	0,2439	Neutral	Tolerated
XKR9	exon4	c.G343A	50	0,0000	54	0,2407	Neutral	Tolerated
COL6A2	exon26	c.2410_2411insT	24	0,0000	43	0,2558	Frameshift	Frameshift
FAM193A	exon17	c.G2671T	23	0,0000	23	0,2609	Neutral	Tolerated
KRBOX4	exon6	c.G315A	40	0,0000	29	0,5862	Nonsense	Nonsense

Tot Cov: coverage totale, MF: mutation frequency, NA: not available.

**Tabella 5.3.** SNVs e *InDel* somatiche identificate tramite WES nel paziente LAM#3.

Gene	Exon	cDNA	Tot Cov at diagnosis	MF at diagnosis	Tot Cov at relapse	MF at relapse	Provean Prediction	Sift Prediction
RIMS1	exon1	c.G89A	27	0,2963	21	0,0000	Neutral	Tolerated
OR10G2	exon1	c.C508T	18	0,3889	14	0,1429	Deleterious	Damaging
SCAF4	exon20	c.2831_2833del	46	0,1304	27	0,1111	Neutral	Deletion
WT1	exon7	c.1139_1140insTCTTGAC	80	0,1625	95	0,0316	Frameshift	Frameshift
ENAH	exon5	c.651_668del	79	0,1519	97	0,1237	Neutral	Deletion
FLT3	exon20	c.T2505A	48	0,0000	67	0,1791	Deleterious	Damaging
SUPT6H	exon34	c.4573_4575del	38	0,0000	40	0,1500	Neutral	Deletion
SALL1	exon2	c.T1517C	97	0,0000	103	0,2621	Deleterious	Damaging
PTPN11	exon3	c.C215T	51	0,0000	40	0,2750	Deleterious	Damaging
UBE2D3	exon6	c.A211G	16	0,0000	24	0,3333	Deleterious	Damaging
SOX11	exon1	c.673_675del	31	0,0323	45	0,1778	Neutral	Deletion

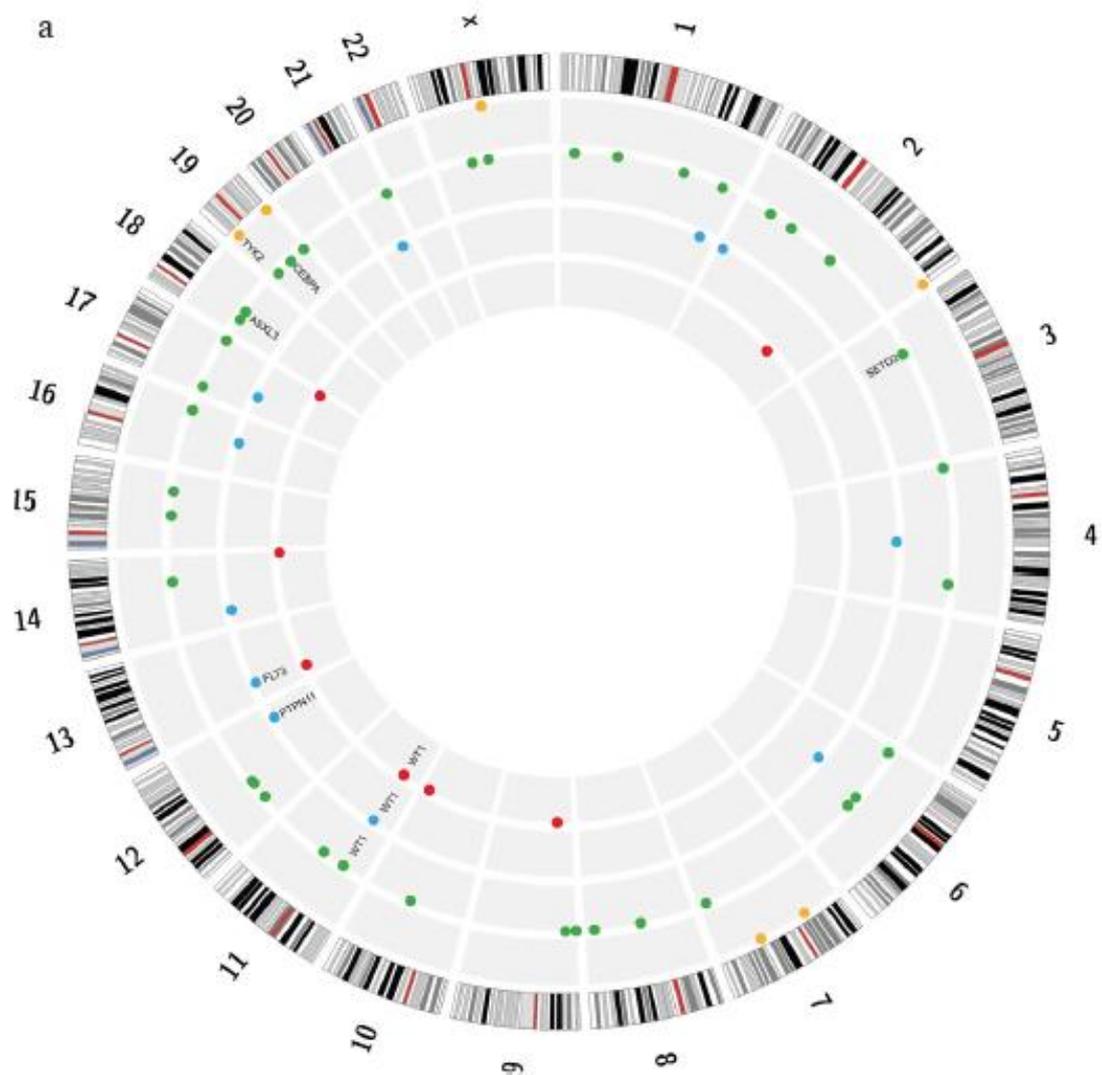
Tot Cov: *coverage* totale, MF: *mutation frequency*.

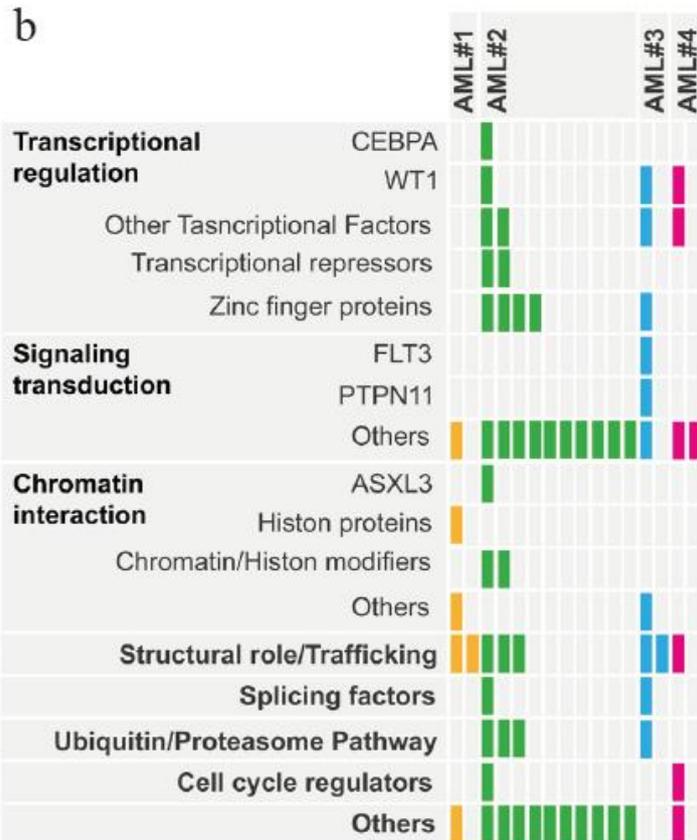
**Tabella 5.4.** SNVs e *InDel* somatiche identificate tramite WES nel paziente LAM#4.

Gene	Exon	cDNA	Tot Cov at diagnosis	MF at diagnosis	Tot Cov at relapse	MF at relapse	Provean Prediction	Sift Prediction
CUZD1	exon2	c.A214T	25	0,3600	29	0,0000	Neutral	Damaging
WT1	exon7	c.1145_1146insTCGG	107	0,2150	141	0,0000	Frameshift	Frameshift
PYCR1	exon3	c.C285G	104	0,3654	144	0,0000	Deleterious	Damaging
TUBA3C	exon4	c.G1040T	15	0,3333	19	0,1053	Deleterious	NA
GPR132	exon5	c.A817G	50	0,0000	73	0,1918	Neutral	Tolerated
TEK	exon17	c.G2849A	35	0,0000	60	0,2167	Neutral	Tolerated
DLX2	exon3	c.G817T	16	0,0625	30	0,2000	Neutral	Tolerated

Tot Cov: *coverage* totale, MF: *mutation frequency*, NA: *not available*.

**Figura 2.** Mutazioni somatiche non sinonime identificate tramite WES in 4 LAM pediatriche. **Figura 2a.** La figura è stata creata con l'uso del software *Circos* (<http://circos.ca/>), mostrando tutte le mutazioni somatiche rilevate in ciascun paziente. I cromosomi sono disposti in senso orario dal cromosoma 1 al cromosoma X, ogni cerchio grigio rappresenta un singolo paziente, procedendo da LAM#1 al cerchio esterno al cerchio interno, ciascun punto rappresenta una mutazione. **Figura 2b.** Le mutazioni sono raggruppate in categorie funzionali di geni coinvolti in base alle annotazioni Pubmed. Ogni casella rappresenta una singola mutazione, ogni colore rappresenta un distinto paziente.





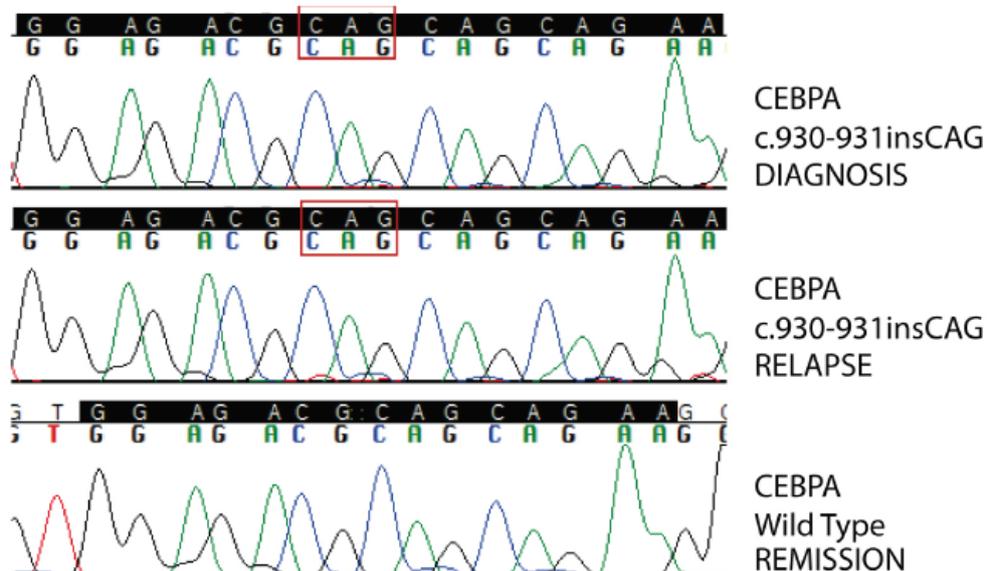
**Tabella 6.** Risultato del sequenziamento mirato delle mutazioni *driver*.

Gene	Patient	Chr	Start	Ref	Alt	cDNA	Protein	MF at diagnosis	Inferred clone size at diagnosis*	MF at remission	MF at relapse	Inferred clone size at relapse*
TYK2	AML#1	19	10467264	A	T	c.T2597A	p.L866H	43,0%	80-90%	0,40%	14,9%	30%
RREB1	AML#2	6	7231420	G	A	c.G3088A	p.V1030M	57,6%	100%	ND	53,1%	100%
PSIP1	AML#2	9	15469966	G	T	c.C1003A	p.P335T	ND	0%	ND	43,9%	40-50%
ASXL3	AML#2	18	31324221	C	T	c.C4409T	p.P1470L	0,3%	<1%	ND	29,7%	60%
WT1	AML#2	11	32438075	C	G	c.G962C	p.G321A	ND	0%	ND	40,0%	40%
CEBPA	AML#2	19	33792384	-	CTG	c.937_938 insCAG	p.K313 delinsQK	88,6%	100%	0,08%	80,8%	100%
SETD2	AML#2	3	47098968	-	GGTG	c.6306_6307 insCACC	p.P2102fs	32,5%	60-70%	0,07%	31,7%	60-70%
WISP1	AML#2	8	134239889	G	A	c.G1040A	p.R347K	ND	0%	ND	34,0%	60-70%
FLT3	AML#3	13	28592640	A	T	c.T2505A	p.D835E	3,4%	<10%	ND	13,3%	25-30%
WT1	AML#3	11	32417913	-	GTACAAGA	c.1139_1140 insTCTGTAC	p.R380fs	13,9%	25-30%	ND	4,2%	<10%
SALL1	AML#3	16	51174616	A	G	c.T1517C	p.I506T	ND	0%	ND	28,6%	50-60%
UBE2D3	AML#3	4	103722704	T	C	c.A211G	p.T71A	ND	0%	ND	27,7%	50-60%
PTPN11	AML#3	12	112888199	C	T	c.C215T	p.A72V	ND	0%	ND	31,9%	60-70%
TEK	AML#4	9	27212867	G	A	c.G2849A	p.R950Q	ND	0%	ND	21,0%	40-50%
WT1	AML#4	11	32417907	-	CCGA	c.1145_1146 insTCGG	p.A382fs	27,6%	55%	ND	ND	0%

### Le mutazioni di biCEBP $\alpha$ sono altamente persistenti, contrariamente alle mutazioni di WT1

È stata rilevata una mutazione biallelica di CEBP $\alpha$  (biCEBP $\alpha$ ) altamente penetrante. Nel paziente LAM#2, sia il WES che il sequenziamento mirato hanno mostrato un'inserzione non-frameshift omozigote (c.937\_938insCAG, p.K313delinsQK) di CEBP $\alpha$ , coinvolgente il dominio bZIP, nella maggior parte della popolazione di cellule tumorali, come rivelato da MF>80%, sia alla diagnosi che alla recidiva. La validazione mediante sequenziamento Sanger è mostrato in figura 3.

**Figura 3.** Validazione mediante sequenziamento Sanger della mutazione di biCEBP $\alpha$  nel paziente LAM#2.



Al contrario, le mutazioni di WT1 appaiono altamente instabili. WT1 codifica per un fattore trascrizionale mutato in maniera ricorrente nelle LAM [25], ma con un ruolo non ancora chiaro nello sviluppo della leucemia [26]. Sono state trovate le seguenti mutazioni di WT1: inserzione frameshift c.1145\_1146insTCGG (p.A382fs) in AML#4 (identificata solo alla diagnosi con MF di 27.6%) e inserzione frameshift c.1139\_1140insTCTTG TAC (p.R380fs) in AML#3 (identificata sia alla diagnosi (MF 13.9%) che in recidiva (MF 4,2%)). Queste inserzioni coinvolgono un'area dell'esone 7, associata con LAM, risultante nella perdita del dominio *zinc-finger* di legame al DNA [25]. Inoltre, SNV c.G962C (p.P1470L) è stato rilevato solo nella recidiva di LAM#2 con un MF del 40%, associato con la perdita di una copia di allele *wild type* alla

recidiva. La mutazione coinvolge l'esone 5 che non è un sito comune di mutazioni nelle LAM [25].

### **Identificazione di meccanismi genomici sottostanti alla recidiva di LAM**

Il sequenziamento delle coppie esordio-recidiva ci ha consentito di individuare modelli di mutazione specifici associati alla recidiva. Nel paziente LAM#3, è stata rilevata l'acquisizione di una mutazione attivante (p.A72V) di PTPN11 in recidiva (MF 31,9%). Nello stesso paziente, il sequenziamento mirato ha rivelato un subclone minore con mutazione FLT3-TKD (p.D835E), già presente alla diagnosi (MF 3,4%), anche se inferiore alla sensibilità di WES, e incrementato alla recidiva (MF 13,3%). Visto l'incremento di MF di PTPN11 e FLT3-TKD, è ragionevole ipotizzare che ciascuno caratterizzi un clone diverso.

Nel paziente LAM#1, è stata evidenziata una mutazione puntiforme (C.T2597A, p.L866H) che coinvolge il dominio pseudochinasico di TYK2, sia alla diagnosi (MF del 43%) che in recidiva (MF del 14,9%). Il dominio pseudochinasico di questa proteina svolge un ruolo nell'inibizione dell'attività chinasica e molte mutazioni descritte associate a tumori si trovano vicino o nell'interfaccia tra questi due domini, determinando una maggiore attività chinasica in vitro [27, 28]. La mutazione che abbiamo identificato non è mai stata riportata in precedenza, ma, considerando il suo coinvolgimento del dominio pseudochinasico, possiamo ipotizzare che costituisca una mutazione attivante.

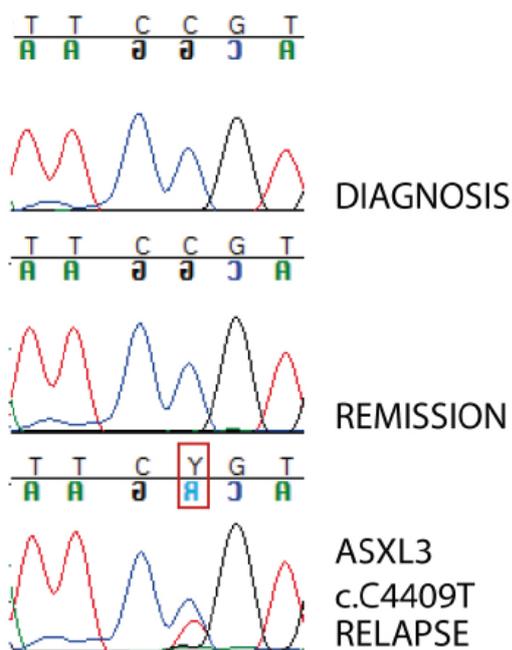
Nel paziente LAM#2, è stata rilevata un'inserzione frameshift (c.6306\_6307insCACC, p.P2102fs) del gene SETD2 in una frazione considerevole della popolazione tumorale sia alla diagnosi (MF 32,5%) che alla recidiva (MF 31,7%). SETD2 è una metiltransferasi responsabile della trimetilazione di H3K36 (H3K36me3), a sua volta, responsabile del reclutamento del complesso *mismatch repair* (MMR) del DNA. Questa mutazione coinvolge il dominio Set2-Rpb1 *interacting* (SRI) nella porzione C-terminale della proteina, che a sua volta prende contatto con il dominio PCTD della RNA polimerasi II e che determina il reclutamento di SETD2 sui geni target [29]. La perdita di questo dominio dovuta a una mutazione missenso o *loss of function* (nonsenso o

frameshift) è stata precedentemente segnalata come causa della perdita di funzione dell'attività metiltransferasica [30].

### Identificazione di una nuova mutazione di ASXL3 come evento tardivo

Nel paziente LAM#2, la mutazione puntiforme c.C4409T (P.P1470L) di ASXL3 è stata individuata alla recidiva (MF del 29,7%) e in un piccolo sub clone nel campione d'esordio (MF del 0,3%). Questa mutazione è descritta essere deleteria a livello proteico. La validazione con sequenziamento Sanger è mostrata in figura 4. Questo rappresenta uno dei pochi esempi di mutazioni ASXL3 descritte in LAM [31, 23], al contrario di mutazioni di ASXL1 e ASXL2, gli altri due membri della famiglia *Additional Sex combs (ASX)-Like*, che sembrano essere comunemente mutati in LAM [31].

**Figura 4.** Validazione mediante sequenziamento Sanger della mutazione di ASXL3 nel paziente LAM#2.



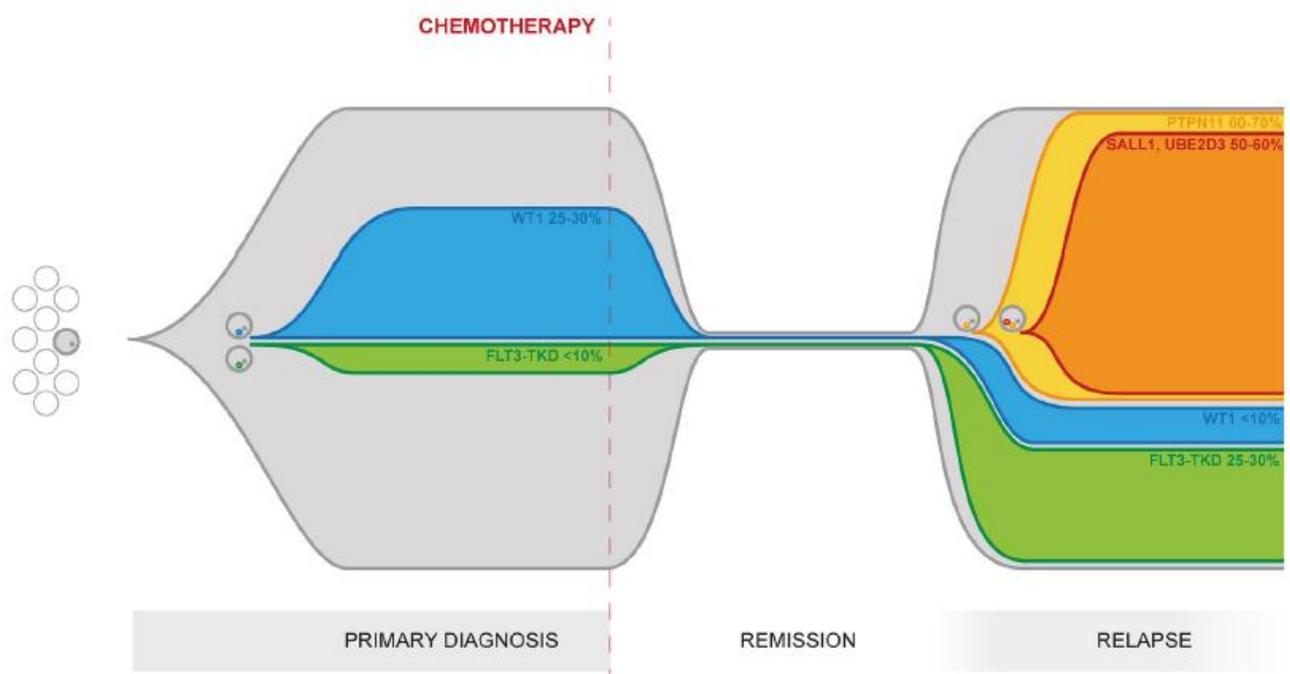
## 2.4 DISCUSSIONE

Dato che le LAM sono caratterizzate da un'alta frequenza di recidiva che a sua volta è associata a prognosi infausta [7, 8], diventa fondamentale indagare più in profondità i meccanismi responsabili della ricaduta di malattia per migliorare la risposta al trattamento e la

sopravvivenza. Un punto importante per comprendere la genomica sottostante alla leucemia è identificare le mutazioni che sottostanno all'evento primario leucemogeno. Nonostante i grandi progressi degli ultimi decenni, nella maggior parte dei casi questa identificazione rimane ancora una questione irrisolta. Tuttavia, un compito altrettanto importante è capire come la leucemia si evolva una volta che questa trasformazione è avvenuta. Una miglior conoscenza di questo processo ci permetterà di affinare il controllo della malattia e di scegliere le terapie ottimali, sia in termini di bersaglio molecolare che di programma terapeutico. Il termine "evoluzione clonale" si riferisce alla progressione del tumore attraverso l'acquisizione di nuove mutazioni che forniscono la diversità genetica all'interno di una linea cellulare. La dinamica di questo processo dipende dall'interazione tra l'effetto specifico di nuove mutazioni e condizioni microambientali, come la chemioterapia. Questo si traduce nella selezione e nell'espansione del subclone più idoneo e nell'eliminazione o nell'auto-estinzione di subcloni meno adatti [32]. I nostri risultati, come mostrato in figura 5 e 6, dimostrano chiaramente l'evoluzione clonale dalla diagnosi alla recidiva, un'ulteriore conferma delle recentissime scoperte di Farrar et al [22]. Infatti, mentre le alterazioni genetiche condivise tra tumore primitivo e recidiva dimostrano un'origine comune da un clone ancestrale, l'evidenza di mutazioni diagnosi-specifiche e recidiva-specifiche dimostrano il modello di evoluzione clonale. Gli eventi primari responsabili dell'origine della LAM sono altamente penetranti nella popolazione tumorale e sono stabili durante il corso della malattia, essendo continuamente presenti nella leucemia di un paziente dalla diagnosi alla recidiva. Dall'altra parte, gli eventi secondari si verificano successivamente e solo in una parte di cellule, conferendo una maggior complessità al profilo genomico della malattia [33]. Per chiarire se i cloni che danno luogo alla recidiva erano già presenti alla diagnosi, ma al di sotto della soglia di rilevazione di WES, o se si sono presentati successivamente nel corso della malattia, abbiamo eseguito il sequenziamento mirato delle mutazioni. È interessante notare che il sequenziamento mirato è stato anche in grado di rilevare la persistenza di alcune mutazioni leucemia-relate (biCEBP $\alpha$ , TYK2 e SETD2) a bassa MF (0,08%, 0,4% e 0,07%, rispettivamente) durante la remissione, supportando la potenziale applicazione del sequenziamento di nuova

generazione per il monitoraggio della malattia minima residua [34, 35, 36]. Al contrario, la grande instabilità delle mutazioni WT1, sia in termini di perdita che di acquisizione, in accordo con la letteratura [22, 25, 37], suggerisce cautela nell'adottare tecniche di monitoraggio basate su WT1. Secondo il modello descritto precedentemente, la mutazione di biCEBP $\alpha$ , definita come la perdita di entrambi gli alleli CEBP $\alpha$  [38], è stata rilevata nella stragrande maggioranza della popolazione tumorale di uno dei nostri pazienti, sia alla diagnosi che alla recidiva. Questo risultato rispecchia il ruolo patogenetico dell'evento: CEBP $\alpha$ , infatti, sembra giocare vari ruoli nella normale emopoiesi, sia nella regolazione della differenziazione che della proliferazione cellulare, e la sua sola alterazione risulta nella formazione di blasti. La stabilità di biCEBP $\alpha$  supporta pertanto il suo ruolo nel definire un sottotipo distinto, molecolare e clinico di LAM, come riportato nella classificazione WHO del 2016 [39]. Lo scopo del nostro lavoro è stato identificare modelli mutazionali alla base della recidiva. Il primo modello riportato è correlato ad un incremento del segnale di proliferazione che dà un vantaggio significativo ad un clone specifico. Questo modello è ben illustrato in figura 5, tracciando le dinamiche di evoluzione clonale del paziente LAM#3.

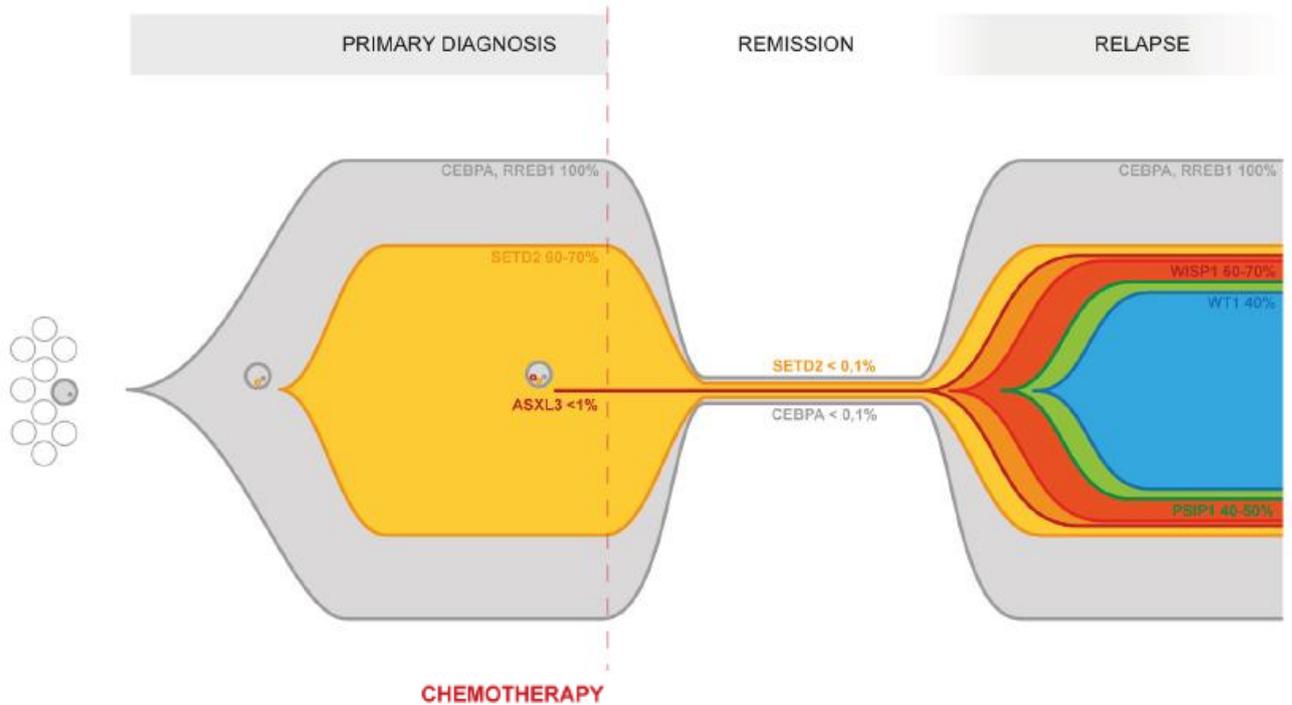
**Figura 5.** Rappresentazione grafica dell'evoluzione clonale dalla diagnosi alla recidiva del paziente LAM#3.



Un piccolo sub clone FLT3-TKD-mutato (<10%) della leucemia primitiva è sopravvissuto alla chemioterapia ed è andato incontro ad espansione nella recidiva (25-30%). Un clone supplementare, caratterizzato dall'acquisizione da una mutazione di PTPN11, è comparso successivamente, diventando il clone predominante in recidiva (60-70%). È ben noto come questi due tipi di eventi conducano ad un aumento della proliferazione cellulare e/o della sopravvivenza [40, 41]. Anche se questi eventi tardivi possono non essere necessari per la leucemogenesi, giocano chiaramente un ruolo importante nella progressione della malattia conferendo uno specifico vantaggio proliferativo. Ciò è particolarmente rilevante se si considera che queste mutazioni potrebbero potenzialmente essere il bersaglio di una terapia mirata. Una precisa caratterizzazione molecolare della malattia anche alla recidiva può guidare la scelta delle terapie ottimali, anche mirando i vari subcloni attraverso più agenti contemporaneamente. Un vantaggio può anche essere espresso come una resistenza anormale all'apoptosi che causa una sopravvivenza clonale alla terapia. Questa caratteristica è stata trovata nel paziente LAM#1, dove un clone con una mutazione puntiforme deleteria di TYK2, un membro della famiglia Janus chinasi (JAK), ha contribuito alla recidiva. Questa chinasi citoplasmatica media l'attivazione intracellulare di una varietà di vie di crescita, differenziazione e sopravvivenza cellulare [42]. L'iperattivazione della via di TYK2, attraverso mutazioni o altri meccanismi, ha dimostrato di contribuire alla sopravvivenza delle cellule aberranti attraverso una sovraregolazione della proteina anti-apoptotica BCL2 in diverse LAL T [43]. C'è anche una crescente evidenza di un ruolo dei membri anti-apoptotici della famiglia BCL2 nel promuovere la resistenza alla terapia e la sopravvivenza sia nelle LAM che nelle sindromi mielodisplastiche [44]. Allo stesso modo, nel nostro paziente una mutazione attivante TYK2 può aver promosso la sopravvivenza clonale alla terapia. Questo clone rappresenta circa un terzo di tutta la popolazione alla recidiva; pertanto eventi sconosciuti devono aver promosso la sopravvivenza di altri cloni. Tuttavia, un singolo campione di recidiva è un'istantanea dove la proporzione relativa dei vari cloni è fortemente correlata alla cinetica di divisione cellulare, come precedentemente descritto. A lungo termine, invece, la persistenza di una riserva di cellule che hanno affrontato la terapia può anche essere correlato ad uno stato più stabile o

addirittura quiescente. Un terzo ed inaspettato modello di evoluzione clonale è l'acquisizione di un fenotipo caratterizzato da un accumulo di un gran numero di mutazioni subclonali a causa di un'alterata riparazione del DNA. Questo è già stato descritto nei casi di recidiva di LAL [45, 46], mentre è un sorprendentemente nuova nelle LAM, tradizionalmente note come caratterizzate da un numero molto basso di mutazioni somatiche e dall'assenza di instabilità genomica [23]. Infatti, nel paziente LAM#2, sia alla diagnosi che alla recidiva, abbiamo rilevato un carico di mutazioni molto più elevato che negli altri tre pazienti. Questo fenotipo è associato ad una mutazione inattivante di SETD2, in grado di interferire con la sua funzione nel reclutamento di riparazione del DNA di tipo MMR [47]. Oltre ad essere state identificate nel carcinoma renale a cellule chiare [48, 49], le mutazioni di SETD2 sono state recentemente identificate da Zhu et al [30] nel 6,2% di 241 pazienti con leucemia acuta, sia mieloide che linfoblastica. Le mutazioni di SETD2 sono state anche descritte come acquisite nella recidiva di LAL pediatrica [50]. I nostri risultati supportano ulteriormente l'idea di un ruolo delle mutazioni di SETD2 nella LAM, in particolare contribuendo alla selezione clonale e alla sopravvivenza. Infatti, l'accumulo di ulteriori mutazioni, aumentando notevolmente la plasticità e l'adattabilità delle cellule leucemiche, conduce alla fine ad una maggiore probabilità di chemioresistenza. Una rappresentazione grafica di evoluzione clonale del paziente LAM#2 dalla diagnosi alla ricaduta è mostrata in figura 6. Infine, abbiamo anche identificato una nuova mutazione puntiforme di ASXL3. È stata rilevata in un piccolo subclone del campione d'esordio (<1%), ma ha subito un'espansione alla recidiva (60%). Mentre le mutazioni degli altri due membri della famiglia *Additional Sex combs (Asx)-Like*, ASXL1 e ASXL2, sono correntemente mutate nella LAM, finora mutazioni di ASXL3 sono state descritte solo due volte in questa patologia [23, 31]. Inoltre, mentre le mutazioni di ASXL1, in base al loro ruolo nel modulare l'espressione genica attraverso la regolazione epigenetica, sono considerate come eventi causativi del processo di leucemogenesi precoce [51, 52], la mutazione di ASXL3 è stata identificata essere un evento tardivo.

**Figura 6.** Rappresentazione grafica dell'evoluzione clonale dalla diagnosi alla recidiva del paziente LAM#2.



## 2.5 CONCLUSIONI

In sintesi, questo studio descrive la struttura policlonale della LAM pediatrica, rivelando un'evoluzione globale dello spettro mutazionale dalla diagnosi alla recidiva. I nostri risultati confermano ulteriormente le recenti evidenze [10] sull'evoluzione clonale nella LAM dell'infanzia e mettono in evidenza una notevole e finora sconosciuta complessità genomica della malattia. I possibili modelli di evoluzione clonale sono vari ed eterogenei in ogni singolo paziente e questo supporta ulteriormente la necessità di strategie diagnostiche e terapeutiche personalizzate. Inoltre, il monitoraggio del clone dominante rilevato al momento della diagnosi può non essere un indicatore affidabile di recidiva, mentre può essere più informativo identificare e tracciare i subcloni emergenti. Attraverso i notevoli miglioramenti tecnologici delle piattaforme di sequenziamento, sia in termini di tempo necessario per eseguire l'analisi che dei costi, si può prevedere che un approccio simile sarà presto disponibile per uso clinico. Questi risultati forniscono una maggior conoscenza dell'architettura clonale sottostante alla recidiva. Più precisamente, considerando i diversi tipi di geni coinvolti abbiamo dimostrato che, una volta che la trasformazione leucemica si è instaurata, l'evoluzione clonale risulta da una complessa interazione di

lesioni *driver*, come ad esempio le mutazioni che forniscono un vantaggio proliferativo (cioè FLT3 o PTPN11) o in termini di aumentata sopravvivenza (cioè mutazioni Tyk2), così come da lesioni deleterie che vengono perse durante l'evoluzione della malattia e da lesioni *passenger*. Abbiamo anche identificato una lesione "mutatore" (inattivazione di SETD2) come un modo non precedentemente descritto di aumentare il tasso di altri cambiamenti genetici che promuovono quindi la perpetuazione della malattia.

In conclusione, la LAM pediatrica è caratterizzata da una notevole complessità genomica e da un'evoluzione clonale. Nello sviluppo della recidiva contribuiscono diversi *pathway* molecolari che causano aumentata proliferazione, resistenza all'apoptosi e ipermutabilità somatica e si configurano come possibili bersagli di terapie mirate.

### 3 BIBLIOGRAFIA

1. Kaspers GJ, Zwaan CM. Pediatric acute myeloid leukaemia: towards high-quality cure of all patients. *Haematologica* 2007; 92(11):1519-32.
2. Stevens RF, Hann IM, Wheatley K, et al. Marked improvements in outcome with chemotherapy alone in paediatric acute myeloid leukemia: results of the United Kingdom Medical Research Council's 10th AML trial. MRC Childhood Leukaemia Working Party. *Br J Haematol* 1998; 101:130-40.
3. Zwaan CM, Kolb EA, Reinhardt D, Abrahamsson J, Adachi S, Aplenc R, De Bont ESJM, De Moerloose B, Dworzak M, Gibson BES, Hasle H, Leverger G, Locatelli F, et al. Collaborative efforts driving progress in pediatric acute myeloid leukemia. *J. Clin. Oncol.* 2015; 33:2949-62.
4. Gibson BES, Webb DKH, Howman AJ, De Graaf SSN, Harrison CJ, Wheatley K. Results of a randomized trial in children with Acute Myeloid Leukaemia: medical research council AML12 trial. *Br. J. Haematol.* 2011; 155:366-76.
5. Pession A, Masetti R, Rizzari C, Putti MC, Casale F, Fagioli F, Luciani M, Lo Nigro L, Menna G, Micalizzi C, Santoro N, Testi AM, Zecca M, et al. Results of the AIEOP AML 2002/01 multicenter, prospective trial for treatment of children with acute myeloid leukemia. *Blood.* 2013; 122:170-8.
6. Creutzig U, van den Heuvel-Eibrink MM, Gibson B, Dworzak MN, Adachi S, de Bont E, Harbott J, Hasle H, Johnston D, Kinoshita A, Lehrnbecher T, Leverger G, Mejstrikova E, et al. Diagnosis and management of acute myeloid leukemia in children and adolescents: recommendations from an international expert panel. *Blood.* 2012; 120:3187-205.
7. Sander A, Zimmermann M, Dworzak M, Fleischhack G, von Neuhoff C, Reinhardt D, Kaspers GJL, Creutzig U. Consequent and intensified relapse therapy improved survival in pediatric AML: results of relapse treatment in 379 patients of three consecutive AML-BFM trials. *Leukemia.* 2010; 24:1422-8.
8. Kaspers GJL, Zimmermann M, Reinhardt D, Gibson BES, Tamminga RYJ, Aleinikova O, Armendariz H, Dworzak M, Ha SY, Hasle H, Hovi L, Maschan A, Bertrand Y, et al. Improved outcome in pediatric relapsed acute myeloid leukemia: results of a randomized trial on liposomal daunorubicin by the international BFM study group. *J. Clin. Oncol.* 2013; 31:599-607.
9. Gorman MF, Ji L, Ko RH, Barnette P, Bostrom B, Hutchinson, Raymond Raetz E, Seibel NL, Twist CJ, Eckroth E, Sposto R, Gaynon PS, Loh ML. Outcome for children treated for relapsed or refractory Acute Myelogenous Leukemia (rAML): a therapeutic advances in childhood leukemia (TACL) consortium study. *J. Clin. Oncol.* 2010; 55:421-9.
10. de Rooij J, Zwaan C, van den Heuvel-Eibrink M. Pediatric AML: from biology to clinical management. *J. Clin. Med.* 2015; 4:127-49.
11. Korte JE, Hertz-Picciotto I, Schultz MR et al. The contribution of benzene to smoking-induced leukaemia. *Environ Health Perspect* 2000;108(4):333-339.

12. Mills PK, Zahm SH. Organophosphate pesticide residues in urine of farmworkers and their children in Fresno Country, California. *Am J Ind Med* 2001;40(5):571-577.
13. Rosenberg PS, Greene MH, Alter BP. Cancer incidence in persons with Fanconi anemia. *Blood* 2003;101(3):822-826.
14. Bader-Meunier B, Tchernia G, Mielot F et al. Occurrence of myeloproliferative disorder in patients with Noonan syndrome. *J PEdiatr* 1997;130(6):885-889.
15. McKenney AH, Arber DA et al. Acute Myeloid Leukemia. *Hematol Oncol Clin N Am* 2009;23:633-654.
16. Swerdlow SH, Campo E, Harris NL et al. WHO classification of tumors of hematopoietic and lymphoid tissues. Lyon:IARC Press 2008.
17. Pui CH, Carroll WL, Meshinchi S, Arceci RJ. Biology, risk stratification, and therapy of pediatric acute leukemias: an update. *J Clin Oncol*. 2011 Feb 10;29(5):551-65.
18. Meyerson M, Gabriel S and Getz G. Advances in understanding cancer genomes through second generation sequencing. *Nature Reviews Genetics* 2010; 11: 685-96.
19. Schuback HL, Arceci RJ, Meshinchi S. Somatic characterization of pediatric acute myeloid leukemia using next-generation sequencing. *Semin Hematol*. 2013; 50(4):325-32.
20. Oszolak F and Milos MP. RNA sequencing: advances, challenges and opportunities. *Nature Reviews Genetics* 2011; 12:87-98.
21. Ding L, Ley TJ, Larson DE, Miller C a, Koboldt DC, Welch JS, Ritchey JK, Young M a, Lamprecht T, McLellan MD, McMichael JF, Wallis JW, Lu C, et al. Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature*.
22. Farrar JE, Schuback HL, Ries RE, Wai D, Hampton OA, Trevino LR, Alonzo TA, Guidry Auvil JM, Davidsen TM, Gesuwan P, Hermida L, Muzny DM, Dewal N, et al. Genomic profiling of pediatric acute myeloid leukemia reveals a changing mutational landscape from disease diagnosis to relapse. *Cancer Res*. 2016; 76:2197-205.
23. Cancer Genome Atlas Research Network. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N. Engl. J. Med*. 2013; 368:2059–74.
24. Walter MJ, Payton JE, Ries RE, Shannon WD, Deshmukh H, Zhao Y, Baty J, Heath S, Westervelt P, Watson MA, Tomasson MH, Nagarajan R, O’Gara BP, et al. Acquired copy number alterations in adult acute myeloid leukemia genomes. *Proc. Natl. Acad. Sci*. 2009; 106:12950–5.
25. Hollink I, Van den Heuvel-Eibrink MM, Zimmermann M, Balgobind B V, Arentsen-Peters STCJM, Alders M, Willasch A, Kaspers GJL, Trka J, Baruchel A, de Graaf SSN, Creutzig U, Pieters R, et al. Clinical relevance of Wilms tumor 1 gene mutations in childhood acute myeloid leukemia. *Blood*. 2009; 113:5951–60.

26. Ho PA, Zeng R, Alonzo TA, Gerbing RB, Miller KL, Pollard JA, Stirewalt DL, Heerema NA, Raimondi SC, Hirsch B, Franklin JL, Lange B, Meshinchi S. Prevalence and prognostic implications of WT1 mutations in pediatric acute myeloid leukemia (AML): a report from the Children's Oncology Group. *Blood*. 2010; 116:702–10.
27. Min X, Ungureanu D, Maxwell S, Hammarén H, Thibault S, Hillert E-K, Ayres M, Greenfield B, Eksterowicz J, Gabel C, Walker N, Silvennoinen O, Wang Z. Structural and functional characterization of the JH2 pseudokinase domain of JAK Family Tyrosine Kinase 2 (TYK2). *J. Biol. Chem.* 2015; 290:27261–70.
28. Lupardus PJ, Ultsch M, Wallweber H, Bir Kohli P, Johnson AR, Eigenbrot C. Structure of the pseudokinase-kinase domains from protein kinase TYK2 reveals a mechanism for Janus kinase (JAK) autoinhibition. *Proc. Natl. Acad. Sci. U.S.A.* 2014; 111:8025–30.
29. Li M, Phatnani HP, Guan Z, Sage H, Greenleaf AL, Zhou P. Solution structure of the Set2-Rpb1 interacting domain of human Set2 and its interaction with the hyperphosphorylated C-terminal domain of Rpb1. *Proc. Natl. Acad. Sci. U. S. A.* 2005; 102:17636–41.
30. Zhu X, He F, Zeng H, Ling S, Chen A, Wang Y, Yan X, Wei W, Pang Y, Cheng H, Hua C, Zhang Y, Yang X, et al. Identification of functional cooperative mutations of SETD2 in human acute leukemia. *Nat Genet.* 2014; 46:287–93.
31. Duployez N, Micol JB, Boissel N, Petit A, Geffroy S, Bucci M, Lapillonne H, Renneville A, Leverger G, Ifrah N, Dombret H, Abdel-Wahab O, Jourdan E, et al. Unlike ASXL1 and ASXL2 mutations, ASXL3 mutations are rare events in acute myeloid leukemia with t(8;21). *Leuk. Lymphoma.* 2016; 57:199–200.
32. Jan M, Majeti R. Clonal evolution of acute leukemia genomes. *Oncogene.* 2013; 32:135–40.
33. Grove CS, Vassiliou GS. Acute myeloid leukaemia: a paradigm for the clonal evolution of cancer? *Dis. Model. Mech.* 2014; 7:941–51.
34. Bibault J, Figeac M, Hélevaut N, Quief S, Sebda S, Renneville A, Rousselot P, Gruson B, Dombret H. Nextgeneration sequencing of FLT3 internal tandem duplications for minimal residual disease monitoring in acute myeloid leukemia. *Oncotarget.* 2015; 6:22812–21.
35. Kohlmann A, Nadarajah N, Alpermann T, Grossmann V, Schindela S, Dicker F, Roller A, Kern W, Haferlach C, Schnittger S, Haferlach T. Monitoring of residual disease by next-generation deep-sequencing of RUNX1 mutations can identify acute myeloid leukemia patients with resistant disease. *Leukemia.* 2014; 28:129–37.
36. Klcó JM, Miller CA, Griffith M, Petti A, Spencer DH, Ketkar-kulkarni S, Wartman LD, Christopher M, Lamprecht TL, Helton NM, Duncavage EJ, Payton JE, Baty J, et al. Association between mutation clearance after induction therapy and outcomes in Acute Myeloid Leukemia. *JAMA.* 2015; 314:811–22.
37. Hou HA, Huang TC, Lin LI, Liu CY, Chen CY, Chou WC, Tang JL, Tseng MH, Huang CF, Chiang YC, Lee FY, Liu MC, Yao M, et al. WT1 mutation in

470 adult patients with acute myeloid leukemia: stability during disease evolution and implication of its incorporation into a survival scoring system. *Blood*. 2010; 115:5222–31.

38. Ho PA, Alonzo TA, Gerbing RB, Pollard J, Stirewalt DL, Hurwitz C, Heerema NA, Hirsch B, Raimondi SC, Lange B, Franklin JL, Radich JP, Meshinchi S. Prevalence and prognostic implications of CEBPA mutations in pediatric acute myeloid leukemia (AML): a report from the Children's Oncology Group. *Blood*. 2009; 113:6558–66.

39. Arber DA, Orazi A, Hasserjian R, Thiele J, Borowitz MJ, Le Beau MM, Bloomfield CD, Cazzola M, Vardiman JW. The 2016 revision to the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia. *Blood*. 2016; 127:2391–405.

40. Tartaglia M, Martinelli S, Iavarone I, Cazzaniga G, Spinelli M, Giarin E, Petrangeli V, Carta C, Masetti R, Aricò M, Locatelli F, Basso G, Sorcini M, et al. Somatic PTPN11 mutations in childhood acute myeloid leukaemia. *Br. J. Haematol*. 2005; 129:333–9.

41. Liang D-C, Shih L-Y, Hung I-J, Yang C-P, Chen S-H, Jaing T-H, Liu H-C, Wang L-Y, Chang W-H. FLT3-TKD mutation in childhood acute myeloid leukemia. *Leukemia*. 2003; 17:883–6.

42. Meyer S, Levine R. Molecular pathways: molecular basis for sensitivity and resistance to JAK kinase inhibitors. *Clin Cancer Res*. 2014; 20:2051–9.

43. Sanda T, Tyner JW, Gutierrez A, Ngo VN, Glover J, Chang BH, Yost A, Ma W, Fleischman AG, Zhou W, Yang Y, Kleppe M, Ahn Y, et al. TYK2-STAT1-BCL2 Pathway Dependence in T-Cell Acute Lymphoblastic Leukemia. *Cancer Discov*. 2013; 3:564–77.

44. Chan SM, Thomas D, Corces-Zimmerman MR, Xavy S, Rastogi S, Hong W-J, Zhao F, Medeiros BC, Tyvoll DA, Majeti R. Isocitrate dehydrogenase 1 and 2 mutations induce BCL-2 dependence in acute myeloid leukemia. *Nat. Med*. 2015; 21:178–84.

45. Kunz JB, Rausch T, Bandapalli OR, Eilers J, Pechanska P, Schuessle S, Assenov Y, Stütz AM, Kirschner-Schwabe R, Hof J, Eckert C, von Stackelberg A, Schrappe M, et al. Pediatric T-lymphoblastic leukemia evolves into relapse by clonal selection, acquisition of mutations and promoter hypomethylation. *Haematologica*. 2015; 100:1442–50.

46. Ma X, Edmonson M, Yergeau D, Muzny DM, Hampton OA, Rusch M, Song G, Easton J, Harvey RC, Wheeler DA, Ma J, Doddapaneni HV, Vadodaria B, et al. Rise and fall of subclones from diagnosis to relapse in pediatric B-acute lymphoblastic leukaemia. *Nat. Commun*. 2015; 6:6604.

47. Li F, Mao G, Tong D, Huang J, Gu L, Yang W, Li G. The histone mark H3K36me3 regulates human DNA Mismatch Repair through its interaction with MutS $\alpha$ . *Cell*. 2013; 153:590–600.

48. Xiang W, He J, Huang C, Chen L, Tao D, Wu X. miR-106b-5p targets tumor suppressor gene SETD2 to inactivate its function in clear cell renal cell carcinoma. *Oncotarget*. 2015; 6:4066–79.

49. Tiedemann RL, Hlady RA, Hanavan PD, Lake DF, Tibes R, Lee J, Choi J, Ho TH, Robertson KD. Dynamic reprogramming of DNA methylation in SETD2-deregulated renal cell carcinoma. *Oncotarget*. 2015; 7:1927–46.
50. Mar BG, Bullinger LB, Mclean KM, Grauman P V, Harris MH, Stevenson K, Neuberg DS, Sinha AU, Sallan SE, Silverman LB, Kung AL, Lo Nigro L, Ebert BL, et al. Mutations in epigenetic regulators including SETD2 are gained during relapse in pediatric acute lymphoblastic leukemia. *Nat Commun*. 2014; 5:3469.
51. Corces-Zimmerman MR, Hong W-J, Weissman IL, Medeiros BC, Majeti R. Preleukemic mutations in human acute myeloid leukemia affect epigenetic regulators and persist in remission. *Proc. Natl. Acad. Sci. U. S. A.* 2014; 111:2548–53.
52. Majeti R. Clonal evolution of preleukemic hematopoietic stem cells precedes human acute myeloid leukemia. *Best Pract. Res. Clin. Haematol.* 2014; 27:229–34.