OPEN

Citation: Blood Cancer Journal (2014) 4, e208; doi:10.1038/bcj.2014.27

www.nature.com/bcj

ORIGINAL ARTICLE

Copy-neutral loss of heterozygosity is prevalent and a late event in the pathogenesis of *FLT3*/ITD AML

DL Stirewalt[1], EL Pogosova-Agadjanyan[1], K Tsuchiya[1,2,3], J Joaquin[1] and S Meshinchi[1,4,5,6]

Patients with high *FLT3* internal tandem duplication allelic ratios (*FLT3*/ITD-ARs) have a poor prognosis. Single-nucleotide polymorphism/comparative genomic hybridization, single-cell PCR and colony-forming assays were used to evaluate genotypic evolution of high *FLT3*/ITD-ARs in 85 acute myeloid leukemia (AML) patients. Microarrays were used to examine molecular pathways disrupted in leukemic blasts with high *FLT3*/ITD-ARs. Copy-neutral loss of heterozygosity (CN-LOH) was identified at the *FLT3* locus in diagnostic samples with high *FLT3*/ITD-ARs ($N=11$), but not in samples with low *FLT3*/ITD-ARs ($N=24$), *FLT3*-activating loop mutations ($N=11$) or wild-type *FLT3* ($N=39$). Single-cell assays showed that homozygous *FLT3*/ITD genotype was present in subsets of leukemic blasts at diagnosis but became the dominant clone at relapse. Less differentiated $CD34^+/CD33^-$ progenitor colonies were heterozygous for *FLT3*/ITD, whereas more differentiated $CD34^+/CD33^+$ progenitor colonies were homozygous for *FLT3*/ITD. Expression profiling revealed that samples harboring high *FLT3*/ITD-ARs aberrantly expressed genes within the recombination/DNA repair pathway. Thus, the development of CN-LOH at the *FLT3* locus, which results in high *FLT3*/ITD-ARs, likely represents a late genomic event that occurs after the acquisition of the FLT3/ITD. Although the etiology underlying the development of CN-LOH remains to be clarified, the disruption in recombination/DNA repair pathway, which is present before the development of LOH, may have a role.

Blood Cancer Journal (2014) **4,** e208; doi:10.1038/bcj.2014.27; published online 2 May 2014

INTRODUCTION

Activating mutations of the *FLT3* gene occur as a result of an internal tandem duplication (*FLT3*/ITD) of the juxtamembrane domain-coding sequence or a point mutation within the activation loop domain (*FLT3*/ALM).[1–3] Although both mutations lead to the constitutive activation of the receptor, only presence of *FLT3*/ITD has consistently been associated with clinical outcome.[4–13] We and others have demonstrated that allelic ratio (AR) of *FLT3*/ITD has prognostic significance, such that those patients with a high *FLT3*/ITD-AR have an inferior outcome compared with those with a low *FLT3*/ITD-AR.[8,12–14] The central mechanism underpinning the variation in AR and its prognostic significance remains poorly understood.

Analyses of microsatellite markers have demonstrated that a small subset of patients with very high ARs actually harbor a loss of heterozygosity (LOH) of the *FLT3* gene at chr13q locus.[9,15,16] Recent reports suggest that copy-neutral LOH (CN-LOH), which is also known as acquired uniparental disomy, may have a role in the development of high *FLT3*/ITD-AR within leukemic blasts.[17,18] However, additional studies are warranted to better characterize the frequency of CN-LOH in patients with *FLT3*/ITD and how this process may contribute to the development of high *FLT3*/ITD-AR. In this study, single-nucleotide polymorphism/comparative genomic hybridization (SNP/CGH) profiling was used to better define the underlying mechanism and temporal evolution of allelic variation in AML blasts containing *FLT3* mutations.

MATERIALS AND METHODS

Patients and treatment

Pediatric patients with acute myeloid leukemia (AML) enrolled in the Children's Cancer Group, AML clinical protocols CCG 2941 and 2961, were candidates for this study. Details of the aforementioned protocols are described elsewhere.[19,20] Evaluated patients were limited to those with >80% leukemic blasts in the bone marrow (BM) or peripheral blood sample. All samples were further enriched for leukemic blast by ficoll preparation. DNA was extracted by Qiagen (Valencia, CA, USA) DNA extraction kit per manufacturer instructions.

Affymetrix mapping 250K SNP/CGH arrays and data processing

Samples were genotyped with Affymetrix 250K Nsp GeneChip Human Mapping arrays (Affymetrix, Inc., Santa Clara, CA, USA), which interrogated over 262 000 loci and allowed the evaluation of genomic copy number and LOH. DNA was digested with restriction enzymes, amplified by PCR, purified, labeled, fragmented and hybridized to the arrays according to the manufacturer's instructions. Briefly, 250 ng of DNA was digested with *Xba*I, *Hind*III or *Sty*I (New England Biolabs, Boston, MA, USA). Digested DNA was adaptor-ligated and PCR-amplified using AmpliTaq Gold (Applied Biosystems, Foster City, CA, USA) in 100 μl PCR reactions for each enzyme-digested sample. PCR products from each reaction were pooled, concentrated and fragmented using DNase I (Invitrogen, Carlsbad, CA, USA). Fragmented PCR products were then labeled, denatured and hybridized to the arrays. Arrays were washed using Affymetrix fluidics stations, and scanned using the GeneChip Scanner 3000 (Affymetrix, Inc.). CEL files were generated using Affymetrix GeneChip Genotyping Analysis Software (GTYPE) v4.0.[21]

SNP calls were generated using GTYPE by incorporating the Copy Number Analysis Tool (CNAT) 4.0.1 package.[22] Samples (85) passed the recommended minimum dynamic model call rate of 93%. Log2 signal ratios were estimated using the unpaired-copy-number workflow with a Gaussian smoothing window of 100 000 bp. Copy-number states were estimated by a Hidden Markov Model using CNAT as per the recommended parameters. LOH was estimated using the unpaired-LOH workflow on genotype files produced by the BRLMM algorithm at default settings. The reference set was comprised of 51 normal samples from the

[1]Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA; [2]Department of Pathology, Seattle Children's Hospital, Seattle, WA, USA; [3]Department of Laboratory Medicine, University of Washington Medical Center, Seattle, WA, USA; [4]Department of Pediatrics, University of Washington Medical Center, Seattle, WA, USA; [5]Children's Oncology Group, Arcadia, CA, USA and [6]Department of Hematology-Oncology, Seattle Children's Hospital, Seattle, WA, USA. Correspondence: Dr S Meshinchi, Clinical Research Division, Fred Hutchinson Cancer Research Center, D5-380, 1100 Fairview Ave. N, Seattle 98109, WA, USA. E-mail smeshinc@fhcrc.org
Received 23 February 2014; accepted 14 March 2014

International HapMap Project. Results were visualized with the Partek Genomics Suite, v. 6.07.0801. The algorithms utilized by CNAT 4.0 have been previously described.[21]

Fluorescence-activated cell sorting for colony-forming cells

Fluorescence-activated cell sorting purification for hematopoietic progenitor subpopulations for colony-forming cell assays on methylcellulose, as well as fluorescence-activated cell sorting-based single-cell sorts for PCR were performed using anti-CD34-fluorescein isothiocyanate and anti-CD33-phycoerythrin (Becton Dickinson, San Jose, CA, USA) with appropriate isotype controls as previously described.[23] Propidium iodide (Sigma-Aldrich, St. Louis, MO, USA) was used to select viable cells. Desired populations were selected on a Vantage flow cytometer (Becton Dickinson) and collected in Iscove's modified Dulbecco's medium (Invitrogen). Colonies were grown in methylcellulose culture supplemented with granulocyte colony-stimulating factor, granulocyte-macrophage colony-stimulating factor, stem cell factor, interleukin-3, interleukin-6 and erythropoietin (all from PeproTech, Rocky Hill, NJ, USA) as previously described.[23]

Single-cell *FLT3* genotyping

Individual colonies were harvested from methylcellulose and sorted as individual cells into 20 μl of RNase-free water supplemented with 10 U RNaseOUT RNase Inhibitor and 1 mM dithiothreitol (Invitrogen) on Vantage flow cytometer. *FLT3* genotyping was carried out using nested touch-down PCR. First-round amplification was performed using primers ITDA-F (5′-CCTGATTGTCTGTGGGGAGT)-3′ and ITDB-R (5′-ACCCCAATCCACTCCAT TTT-3′) in a final reaction volume of 50 μl, consisting of 25 μl 2 × Failsafe PCR Premix D (Epicentre, Madison, WI, USA), 2.5 μl PCR-grade water, 2.5 U Platinum Taq (Invitrogen), 0.2 μM of each primer and 20 μl single-cell template. Cycling conditions were as follows: initial denaturation, 94 °C for 5 min, followed by 15 cycles of 94 °C for 30 s denaturation, 70–56 °C (Δ − 1 °/cycle) for 45 s annealing and 72 °C for 1 min extension; this was followed by another 25 cycles with the same parameters, but with annealing temperature of 56 °C and completed with a final extension of 7 min at 72 °C. Second-round amplification was performed using primers CG13F-6FAM (5′-[6FAM]TGCCTATTCCTAACTGACTCATCA-3′) and CG14R (5′-TCTTTGTTGCTGTCCTTCCA-3′) and a 25-μl reaction mix consisting of 12.5 μl 2 × Failsafe PCR Premix B (Epicentre), 6.25 μl PCR-grade water, 2.5 U Platinum Taq (Invitrogen), 0.2 μM of each primer and 5 μl template DNA from the first-round amplification described above. Touch-down cycling conditions were as listed above, with the exception of the annealing and extension times were reduced to 30 s each.

Fluorescence *in situ* hybridization (FISH) evaluation of colony-forming cell colonies

FISH was performed using bacterial artificial chromosome (BAC) RP11-136G6 (13q12.2) that contains the *FLT3* gene, and BAC RP11-408E5 (13q12.11, near the centromere) as a control for chromosome 13 copy number. BAC DNA was isolated from a single-colony culture using the Large-Construct Kit according to the manufacturer's protocol (Qiagen). DNA was labeled by nick translation using SpectrumOrange dUTP (RP11-136G6) or SpectrumGreen dUTP (RP11-408E5) following the manufacturer protocol (Vysis Inc., Downers Grove, IL, USA). Labeled BAC DNA was precipitated, resuspended in LSI/WCP hybridization buffer (Vysis Inc.), and hybridized to slides according to the manufacturer's protocol. Labeled BAC DNA was initially hybridized to normal human metaphase cells to verify probe localization to the correct band on chromosome 13. At least two and up to 20 independent colonies from each cell population (30–323 total interphase nuclei per population) were scored.

Microarray expression analyses

Details of subjects included in the microarray studies are provided in Supplementary Table 1. RNA was extracted from BM and peripheral blood samples, and quality of RNA was analyzed on HP 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA).[24] DNA from available samples was evaluated for NPM1, DNMT3A, IDH1 and IDH2 mutations as previously described.[25–27] Five micrograms of total RNA underwent biotin-labeling using the Enzo BioArray HighYield RNA Transcript Labeling Kit (Enzo Life Technologies, Farmingdale, NY, USA) as per manufacturer's recommendations. The biotin-labeled product was fragmented as per the standard Affymetrix protocol.[24] Fifteen micrograms of fragmented antisense cRNA was then hybridized to the HG-U133A arrays (Affymetrix, Inc.). DAT, CEL and CHP files were generated using GCOS 1.2.1 and MAS 5.0 software (Affymetrix, Inc.) as previously described.[24] All microarray data met previously determined quality controls.[28] CEL files were normalized using robust multi-array average (gcRMA).[29,30]

Log2 expression values were analyzed using GenePlus software (Enodar Biologic, Seattle, WA, USA).[24,31] Two separate analyses were performed. The first analysis directly compared the expression profile between eight AML samples without *FLT3*/ITDs (*FLT3*/wild type (WT)) and five AML samples with high *FLT3*/ITD-AR (*FLT3*/$ITD_{H\text{-}AR}$).[24,31] Expression changes with a *Z*-score $\geqslant 4.75$ or $\leqslant -4.75$ were deemed to be statistically significant.[32,33] The second analysis examined expression changes across the three different groups: normal CD34 cells from BM (BM34), *FLT3*/WT and *FLT3*/$ITD_{H\text{-}AR}$. For this analysis, the three groups were assigned a group number (BM34 = 0, *FLT3*/WT = 1 and *FLT3*/$ITD_{H\text{-}AR}$ = 2). A linear regression analysis examined the expression changes across the three groups.[24,31] Expression changes with a *Z*-score $\geqslant 4.75$ or $\leqslant -4.75$ were deemed to be statistically significant.[32,33] Detailed clinical, cytogenetic and *FLT3*/ITD-AR information, along with the CEL files for each of the normal and AML samples, are available at the Gene Expression Omnibus.[34] All normal BM34 samples and a portion of the AML samples have previously been used for other microarray analyses.[24,31]

Network generation and pathway analyses

Affymetrix IDs and *Z*-scores for significant genes were analyzed using Ingenuity Pathways Analysis (IPA) software (Qiagen, Redwood City, CA, USA). Using the IPA software, we determined the significance of functional and canonical pathways as previously described.[24] All relationships were supported by at least one literature reference and/or information from the Ingenuity Pathways Knowledge Base.[35] Functions and canonical pathways with a *P*-value $\leqslant 0.05$ were considered to be statistically significant.[24,36] Networks were generated using genes identified from the microarray analyses and IPA Network algorithm. A Network Score was based on hypergeometric distribution and calculated using the right-tailed Fisher Exact Test, taking into account total number of network eligible genes in the Knowledge Base, the number of genes analyzed and network size: the higher the score, the less likely that the observed fit between the analyzed genes and the network was due to chance.[24,36]

RESULTS

SNP/CGH genomic profiling identifies CN-LOH at 13q locus

Diagnostic specimens from 85 patients with and without *FLT3*/ITDs were analyzed by SNP/CGH arrays. To control for the potential impact of blast variability, all specimens had >80% blasts at diagnosis, and leukemic blast populations were further enriched before the study by ficoll processing. The 85 patients displayed the following genotypes: *FLT3*/ITD 35/85 (41%), *FLT3*/ALM 11/85 (13%) and *FLT3*/WT 39/85 (46%). Patients with *FLT3*/ITD had ARs (that is, ratio of ITD to WT product) ranging from 0.2 to 7.05.

Using data from the Affymetrix Nsp arrays, chromosome 13 was evaluated for copy-number and LOH alterations, and analyses compared with samples with and without *FLT3* mutations. Unclustered analysis of the entire cohort demonstrated large region of LOH in specimens from 12/85 patients (Figure 1a). Clustering analyses revealed that all samples with LOH_{13q} contained *FLT3*/ITDs (Figures 1b and c), and the ITD-AR was substantially higher in those *FLT3*/ITD samples with LOH_{13q} (median = 2.5, range 1.21–8.88) as compared with *FLT3*/ITD samples without LOH_{13q} (median = 0.65, range 0.20–0.88). Those with LOH_{13q} were further analyzed for copy-number alterations to determine whether observed LOH was due to large segmental deletion or CN-LOH. The observed LOH_{13q} was not due to deletion, thus indicating the presence of CN-LOH at 13q ($CN\text{-}LOH_{13q}$) in 12/35 specimens with *FLT3*/ITD. Furthermore, the region of CN-LOH conversion always involved the *FLT3* gene (Figure 1d), and the LOH extended from a locus that was centromeric to *FLT3* and encompassed the entire distal part of 13q. To define the transition point for the CN-LOH conversion, a detailed evaluation of the centromeric locus was performed. In 11/12 cases with $CN\text{-}LOH_{13q}$, the CN-LOH transition occurred at

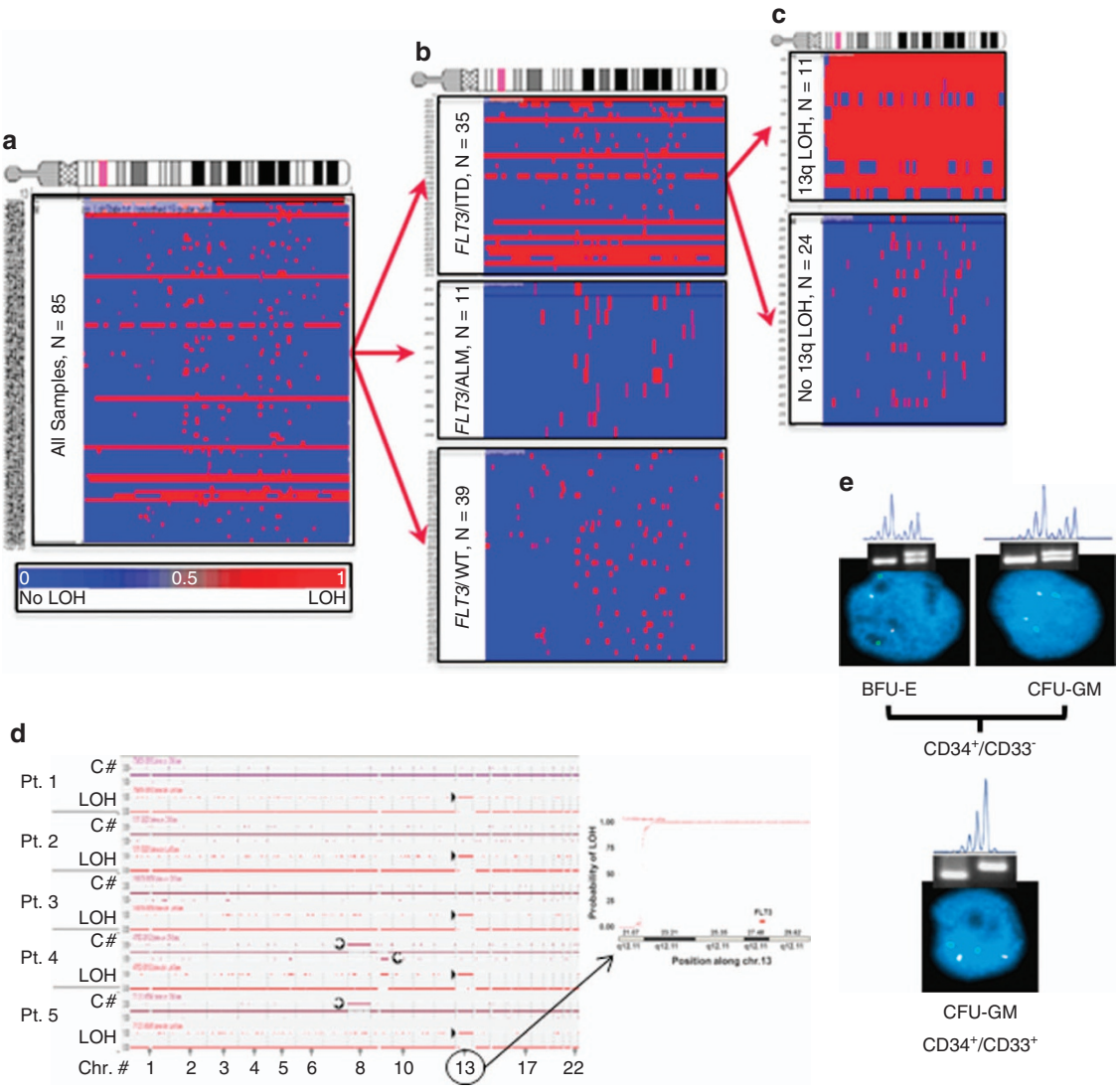

Figure 1. CN-LOH in FLT3/ITD samples with high allelic ratio. (**a**) Unclustered profile for chr13 for the entire cohort demonstrates segmental LOH (red color) in a subset of samples. (**b**) Clustering based on *FLT3* genotype shows that a large segmental LOH segregates with *FLT3*/ITD. (**c**) Clustering of *FLT3*/ITD AML by LOH_{13q}, segregates samples based on allelic ratio. (**d**) Simultaneous copy number (C#) and LOH alteration in five representative patients with LOH_{13q} are shown (left panel). C# state represents patients with or without any copy-number changes. CN state of 1 represents no loss of copy number. Deletions are represented as a downward deflection (for example, patient 4, del9q, ⟲) and any duplication or trisomy are represented with a upward deflection from baseline (for example, patients 4 and 5, trisomy 8, ⟳). LOH state is represented as either 0 (no LOH, baseline) or 1 (LOH, upward deflection, patients 1–5, chr. 13, ▶). Inset is a close up of Chr13 for patient number 5, demonstrating the region of transition into LOH (right panel). (**e**) Figure shows BFU-E or CFU-GM colonies derived from $CD34^+/CD33^-$ myeloid progenitors and CFU-GM colonies derived from more differentiated $CD34^+/CD33^+$ cells. The cells were subjected to *FLT3* genotyping by PCR and LOH determination by STR analysis and FISH. Colonies derived from $CD34^+/CD33^-$ progenitors were heterozygous for *FLT3*/ITD (gel), with no LOH detected by STR and no deletions by FISH (green signals—control probe near chromosome 13 centromere, red signals—BAC probe that includes *FLT3*). Cultured colonies from more differentiated $CD34^+/CD33^+$ progenitors were homozygous for *FLT3*/ITD (gel), had LOH by STR, but showed no evidence of deletion by FISH, suggesting CN-LOH.

13q12.1, and in the remaining case, the entire 13q had undergone conversion.

Single-cell analysis to detect subpopulations of leukemic blasts with potential $CN\text{-}LOH_{13q}$

Given the intra-individual genomic heterogeneity of AML, we examined whether subpopulations of leukemic blasts may harbor $CN\text{-}LOH_{13q}$, which would be undetected by standard SNP/CGH analyses. Individual leukemic blasts from five patients with known *FLT3*/ITD without previously identifiable $CN\text{-}LOH_{13q}$ by SNP/CGH analyses (see Figure 1d) were sorted into PCR tubes, and single-cell PCR assays were utilized to define *FLT3*/ITD status. *FLT3* signal recovery for the single-cell assays ranged from 50 to 85% in each case. In each of the five diagnostic samples, three genotypic subpopulations were identified: (1) heterozygous *FLT3*/ITD, (2) homozygous *FLT3*/ITD and (3) *FLT3*/WT (Table 1). The homozygous *FLT3*/ITD genotype, likely owing to $CN\text{-}LOH_{13q}$, was demonstrated in 13–34% cells of the five samples tested. These findings suggest that patients with low *FLT3*/ITD-AR at diagnosis may harbor $CN\text{-}LOH_{13q}$ in minor subpopulations of leukemic blasts. Single-cell analysis of paired diagnostic and relapsed samples from two patients demonstrated that the homozygous *FLT3*/ITD genotype became the dominant genotype at relapse (Table 1), indicating a clonal evolution to a more uniform population of leukemic blasts harboring $CN\text{-}LOH_{13q}$ upon relapse.

Table 1. Distribution of molecular subtypes of *FLT3* gene in a single-cell assay

	Diagnostic			*Relapse*		
	Heterozygous ITD	*Homozygous ITD*	*FLT3/WT*	*Heterozygous ITD*	*Homozygous ITD*	*FLT3/WT*
Pt 1	27%	21%	62%	10%	83%	7%
Pt 2	32%	32%	37%	12%	77%	11%
Pt 3	3%	13%	83%	ND	ND	ND
Pt 4	21%	26%	53%	ND	ND	ND
Pt 5	16%	34%	50%	ND	ND	ND

Abbreviations: ITD, internal tandem duplication; ND, not determined; WT, wild type.

Table 2. Top functions for genes with significant expression differences between *FLT3*/WT and *FLT3*/$ITD_{H\text{-}AR}$ samples

Function	P-*value*	*Number of genes associated with function*
[a]Cancer	4.50E-04–3.70E-02	28
[a]Cellular growth and proliferation	8.66E-05–3.57E-02	23
[a]Hematological system development and function	3.82E-05–3.67E-02	16
[a]Immune response	1.86E-04–3.67E-02	15
Immune and lymphocyte development and function	1.86E-04–3.57E-02	14
Small molecule biochemistry	9.31E-05–3.53E-02	12
[a]Hematologic disease	4.61E-04–3.70E-02	12
Cellular development	3.82E-05–3.81E-02	11
[a]Immunologic disease	4.50E-04–3.33E-02	11
Neurological disease	1.09E-03–3.33E-02	9
[a]Cell cycle	1.31E-03–3.56E-02	9
Molecular transport	9.31E-05–3.53E-02	8
Tissue development	1.86E-04–3.33E-02	7
[a]Cell morphology	4.61E-04–3.33E-02	7
Lipid metabolism	9.31E-05–3.53E-02	6
Cellular compromise	4.61E-04–3.33E-02	5
Genetic disorders	1.09E-03–3.33E-02	5
Gene expression	3.08E-04–2.78E-02	4
Nucleic acid metabolism	1.09E-03–3.33E-02	4
Hepatic system disease	1.66E-03–1.66E-03	2

Abbreviations: ITD, internal tandem duplication; WT, wild type. [a]Represents an overlap of top functions between the two analyses evaluating (a) *FLT3*/WT and *FLT3*/$ITD_{H\text{-}AR}$ and (b) BM34, *FLT3*/WT and *FLT3*/$ITD_{H\text{-}AR}$.

$CN\text{-}LOH_{13q}$ is a late event in AML pathogenesis

In order to evaluate the sequence of events leading to $CN\text{-}LOH_{13q}$, $CD34^+/CD33^-$ and $CD34^+/CD33^+$ progenitor cells from two patient samples with $CN\text{-}LOH_{13q}$ (high *FLT3*/ITD-AR) were isolated and subjected to colony-forming cell culture. Erythroid burst-forming unit (BFU-E) and granulocyte, monocyte colony-forming unit (CFU-GM) colonies were sequentially evaluated by PCR for *FLT3*/ITD status, short tandem repeat analyses (STR) for LOH, and FISH for genetic loss. BFU-E and CFU-GM colonies grown from $CD34^+/CD33^-$ leukemic blasts (less differentiated population) demonstrated heterozygous ITD and no LOH by STR, indicating that the less differentiated progenitors within the colony harbored *FLT3*/ITDs but had not yet developed LOH. As expected, the FISH results for this population uniformly demonstrated two copies of chromosome 13 and no deletion of *FLT3*. In contrast, CFU-GM colonies cultured from more differentiated $CD34^+/CD33^+$ leukemic blasts uniformly demonstrated homozygous ITD and LOH_{13q} by STR evaluation but still showed no evidence of chromosome 13 loss or deletion of *FLT3* by FISH (Figure 1e). This data demonstrates that *FLT3*/ITD evolves initially in the progenitor population as a heterozygous clone (that is, no $CN\text{-}LOH_{13q}$), and subsequent events likely lead to the evolution of $CN\text{-}LOH_{13q}$ in the more differentiated population (Figure 1e).

Gene expression changes associated with $CN\text{-}LOH_{13q}$

To identify expression changes potentially associated with the development of homozygous ITD status and subsequent $CN\text{-}LOH_{13}$, microarray studies were performed using leukemic samples from AML patients without *FLT3*/ITD (*FLT3*/WT, $N = 8$) and those with very high *FLT3*/ITD ARs (*FLT3*/$ITD_{H\text{-}AR}$, $N = 5$). To reduce the potential impact of normal hematopoietic cells and leukemic maturation stage on the expression profile, only AML samples with high blast counts ($>80\%$) that expressed CD34 were selected for the microarray studies (Supplementary Table 1). A direct comparison of expression profiles between *FLT3*/WT and *FLT3*/$ITD_{H\text{-}AR}$ identified 84 genes and 2 unique probe sets (that is, sequences without an annotated gene) with significant expression differences between the two sample groups (Supplementary Table 2). Using *FLT3*/WT samples as the baseline, 43 genes and 1 unique probe set displayed increased expression in the *FLT3*/$ITD_{H\text{-}AR}$ samples, four of which were HOX family members, whereas 41 genes and 1 unique probe set displayed a decreased expression (Supplementary Figure 1). The uniformity in expression of top candidate genes suggests that presence of other known mutations (for example, *NPM1* and *IDH1*) may not be a major contributor to expression differences observed between *FLT3*/$ITD_{H\text{-}AR}$ and *FLT3*/WT samples (Supplementary Table 1).

IPA software was used to identify significant functions, networks and canonical pathways associated with the 84 identified genes. The top 20 functions associated with the 84 genes are listed in Table 2, and the top 12 networks are provided in Supplementary Table 3. The most significant functions were[1] cancer,[2] cellular growth and proliferation and[3] hematologic system function and development. Not surprisingly, the network with the highest score was comprised of genes associated with similar functions: cellular

proliferation, hematologic development and function, and immune response. The only canonical pathway associated with the 84 genes was the cell cycle pathway that regulates G2/M checkpoint for DNA damage (that is, cell cycle: G2/M DNA damage checkpoint regulation).

AML samples can harbor CN-LOH at loci other than 13q.[17,18,37–40] Thus, AML samples with CN-LOH at other loci may display the same or similar expression changes as those with CN-LOH$_{13q}$. To further examine potential expression changes associated with CN-LOH, we incorporated the expression profiles of normal CD34 cells from the BM (BM34, $N=8$) into our analyses, which are unlikely to harbor CN-LOH.[41] For these analyses, samples were divided into three groups based on the likelihood of harboring CN-LOH: normal BM34 (unlikely to harbor CN-LOH), *FLT3*/WT (unlikely to have CN-LOH$_{13}$ but potentially harboring CN-LOH at other regions) and *FLT3*/ITD$_{H\text{-}AR}$ (CN-LOH$_{13}$ likely in the majority of blasts and potential CN-LOH at other loci). The normal BM34 were designated as the 'control' and given the covariate of 0, whereas *FLT3*/WT and *FLT3*/ITD$_{H\text{-}AR}$ were designated as having covariates of 1 and 2, respectively. Linear regression analysis identified 642 genes and 3 unique probe sets with significant expression changes (Supplementary Table 4). The analysis identified genes with relatively heterogeneous expression changes within the *FLT3*/WT group, which became more homogenous in the *FLT3*/ITD$_{H\text{-}AR}$ samples (Figures 2a and b). The presence of other mutations did not appear to explain these mutation differences. Approximately 55% of the genes/unique probe sets ($N=364$) displayed a linear increase in expression with increasing certainty of CN-LOH presence, while 45% ($N=281$) displayed decreasing expression with increasing certainty of CN-LOH presence (Figure 2c). Thirty-three of these genes were also identified in the previous analysis comparing AML patients with *FLT3*/WT vs *FLT3*/ITD$_{H\text{-}AR}$.

Similar to the first analysis, IPA software was used to identify significant functions, networks and canonical pathways associated with the 642 significant genes. The top 20 functions are provided in Table 3, and as expected, there was considerable overlap among the significant functions associated with 84 and 642 gene sets (Tables 2 and 3). The top five networks are provided in Supplementary Table 5. The network with the highest score was comprised of genes associated with cell cycle, cellular movement and cancer. When considering the top 20 networks (Supplementary Table 6), the three most frequently represented functions were (1) cancer, (2) cell cycle and (3) DNA replication, recombination and repair. As DNA replication, recombination and repair likely have a central role in the development of CN-LOH, we examined the genes associated with these functions and their relationship in more detail (Table 4). All the genes within this functional category displayed decreasing expression with increasing likelihood of CN-LOH being present (that is, BM34 to *FLT3*/WT to *FLT3*/ITD$_{H\text{-}AR}$).

DISCUSSION

In this study, we demonstrate that not only CN-LOH$_{13q}$ is detected in a subset of *FLT3*/ITD patients with very high AR (AR >1), but we show lack of CN-LOH$_{13q}$ in patients without *FLT3*/ITD, suggesting a potential causal link between *FLT3*/ITD and evolution of CN-LOH$_{13q}$. We also demonstrate that presence of CN-LOH$_{13q}$ is

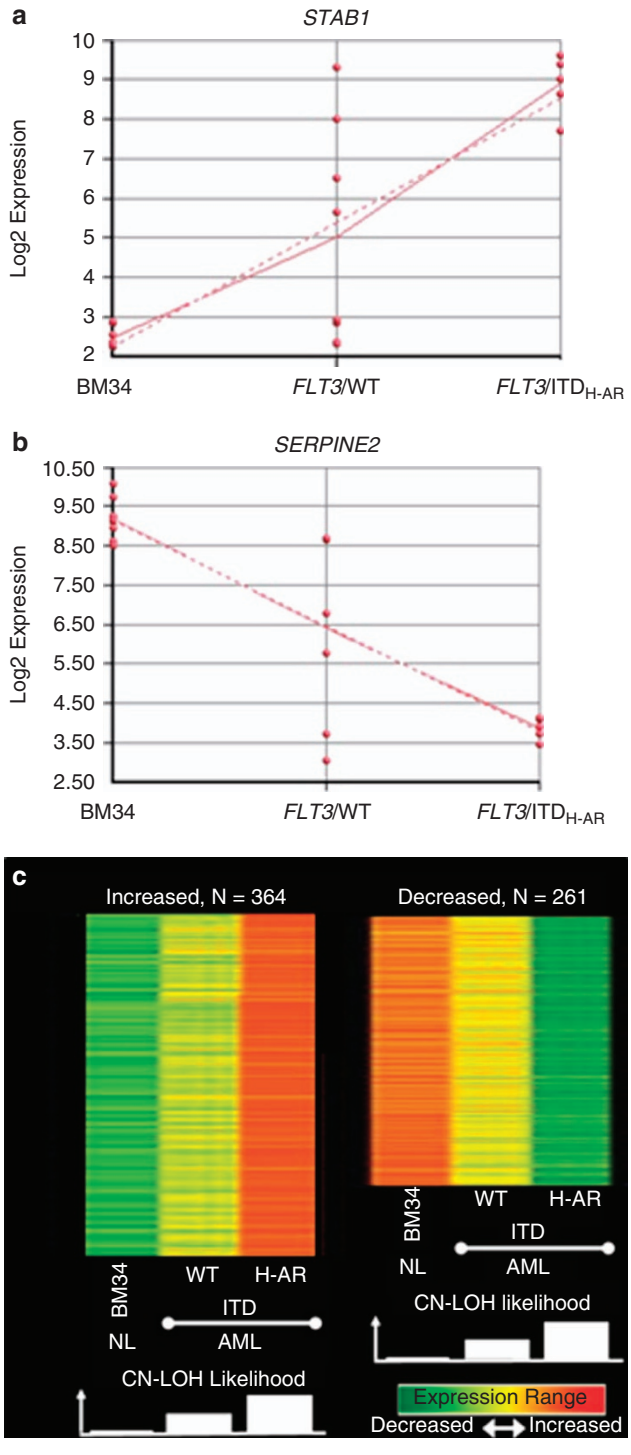

Figure 2. Linear expression changes associated with likelihood of CN-LOH. (**a**) Figure shows the log2 expression (y-axis) of stabilin 1 (*STAB1*, potential role in regulating lipid metabolism and angiogenesis) across normal BM34 ($N=8$), *FLT3*/WT ($N=8$) and *FLT3*/ITD$_{H\text{-}AR}$ ($N=5$) samples. As depicted in the figure, the expression significantly increases ($Z=17.09$, $P<0.01$), yet there is considerable heterogeneity in its expression within the *FLT3*/WT group. (**b**) Figure shows the log2 expression (y-axis) of serpin pepitidase inhibitor, clade B (*SERPINE2*, potential role in regulating apoptosis and differentiation) across normal BM34, *FLT3*/WT and *FLT3*/ITD$_{H\text{-}AR}$ samples. As depicted in the figure, the expression significantly increases ($Z=-20.90$, $P<0.01$), and as with *STAB1*, there is heterogeneity in its expression within the *FLT3*/WT group. (**c**) Heatmaps show increasing and decreasing expression (y-axis) of 364 and 281 genes from two analyses, respectively, across the three different subpopulations of cells (x-axis). Expression range is provided in the lower right corner. Expression for each gene within a group represents the average expression of the gene across all samples within the group. Overall, as the likelihood of CN-LOH increases, so does the uniformity of the expression and differences in the expression within the normal and malignant cell populations.

Table 3. Top functions for genes with significant linear expression changes in BM34, FLT3/WT and FLT3/ITD$_{H\text{-}AR}$ samples

Function	P-*value*	*Number of genes associated with function*
[a]Cancer	1.67E-08–1.36E-02	196
[a]Cellular growth and proliferation	5.91E-06–1.33E-02	173
Cell death	1.37E-06–1.39E-02	158
[a]Hematologic disease	2.40E-06–1.36E-02	104
[a]Cell cycle	9.28E-07–1.41E-02	92
[a]Immunologic disease	1.61E-05–1.36E-02	91
Cellular movement	8.03E-07–1.36E-02	91
[a]Hematological system development and function	8.03E-07–1.36E-02	86
Inflammatory disease	3.51E-07–1.41E-02	85
[a]Immune response	8.03E-07–1.41E-02	75
DNA replication, recombination and repair	2.79E-07–1.29E-02	73
Gastrointestinal disease	3.53E-06–1.36E-02	68
Skeletal and muscular disorders	4.24E-06–1.36E-02	60
[a]Cell morphology	4.81E-06–1.36E-02	58
Connective tissue disorders	9.86E-07–3.79E-05	53
Reproductive system diseases	1.70E-06–1.03E-02	46
Cell-to-cell signaling and interaction	1.35E-05–1.36E-02	39
Cellular assembly and organization	2.79E-07–1.36E-02	29
Cell function and maintenance	1.35E-05–1.36E-02	23
Renal and urological disease	4.24E-05–1.36E-02	18

Abbreviations: ITD, internal tandem duplication; WT, wild type. [a]Represents an overlap of top functions between the two analyses evaluating (a) *FLT3*/WT and *FLT3*/ITD$_{H\text{-}AR}$ and (b) BM34, *FLT3*/WT and *FLT3*/ITD$_{H\text{-}AR}$.

Table 4. Downregulated genes within DNA replication, recombination and repair pathways associated with FLT3/ITD$_{H\text{-}AR}$

Process annotation	P-*value*	*Genes/molecules*	*Number of molecules*
Recombination of DNA	2.50E-03	*CHEK1*, *DKFZP762E1312*, *HMGB1*, *KPNA2*, *MSH6*, *REC8*, *RUVBL2*, *XRCC5*	8
Checkpoint control	3.67E-03	*BUB3*, *BUB1*, *CDC6*, *CDC20*, *CDK2*, *CHEK1*, *MCM7*, *NDC80*	8
Double-stranded DNA break repair	4.96E-03	*FEN1*, *KPNA2*, *POLA1*, *SOD1*, *SOD2*, *XRCC5*	6

Abbreviation: ITD, internal tandem duplication. Process associated with the overall function category (first column), *P*-value for significance of the process (column 2), genes within the process (column 3, all with decreasing expression from BM34 to FLT3/WT to FLT3/ITDH-AR), and number of genes within the network (column 4) are provided. Overall, the data suggest that the expression of genes is downregulated in these functional pathways as the likelihood of harboring CN-LOH increases.

not limited to those with very high ITD-AR, but in fact, as demonstrated by single-cell *FLT3* genotyping, CN-LOH$_{13q}$ can be present as a minor clone in *FLT3*/ITD cases with low AR, bringing into question what are the additional events that may lead to clonal dominance of this homozygous clone. The results from our studies indicate that the evolution of CN-LOH$_{13q}$ likely represents a late event in the pathogenesis of high *FLT3*/ITD-ARs, such that *FLT3*/ITD initially develops as a heterozygous mutation with the corresponding WT *FLT3* allele. Over time, the heterozygous *FLT3*/ITD blasts acquire CN-LOH$_{13q}$, resulting in homozygous mutated alleles and a high *FLT3*/ITD-AR.

Previous genome-wide studies have identified CN-LOH at other loci in AML and in a variety of different hematopoietic malignancies.[17,18,37–40,42] These other loci also frequently contain oncogenes or tumor suppressors (for example, *TP53*, *MPL*, *NRAS* and so on).[39] As with *FLT3*/ITD and CN-LOH$_{13q}$, the development of CN-LOH at these loci also results in homozygous mutant alleles and the loss of expression of the WT protein. The relative expression of mutant vs WT proteins (that is, 'mutant dosage effect') likely has a critical role in malignant transformation and disease progression,[43] and a mutant dosage effect has been demonstrated for *RAS* mutations.[44–50] Like *FLT3*/ITDs, *RAS* mutations enhance proliferation and likely provide a growth advantage.[44,48–50] For example, hematopoietic progenitors with homozygous *NRAS* mutations develop a phenotypically different hematopoietic malignancy than similar progenitors harboring heterozygous *NRAS* mutations.[50] In the case of *FLT3*/ITDs, CN-LOH$_{13q}$ also likely provides a growth and/or survival advantage through the establishment of a homozygous *FLT3*/ITD state.[43] Higher *FLT3*/ITD-AR (that is, likely to harbor CN-LOH$_{13q}$) are associated with higher relapse rates (that is, more resistant disease),[14] and are found more frequently at relapse.

Few studies have examined the temporal relationship between the development of mutations and CN-LOH. Wang *et al.*[50,51] used a Cre–Lox murine model to introduce a single *NRAS* mutant allele into hematopoietic stem cells. Using this model, they showed that ~40% of mice harboring a heterozygous *NRAS* mutation subsequently acquired CN-LOH at the *NRAS* locus, resulting in a homozygous *NRAS* mutant state. Recently, similar murine models have been used to examine the cooperative nature between *FLT3*/ITDs and *NPM1* mutations.[43] In these experiments, transgenic mice harboring heterozygous mutations for both *FLT3*/ITD and *NPM1* rapidly developed an AML-like disease. Interestingly, LOH of the WT *FLT3* locus was observed in a very high percentage of the mice with the most rapid disease progression.[43] Thus, the available data from *NRAS* and *FLT3*/ITD murine models are in keeping with our findings, which indicate that *FLT3*/ITD mutations in humans likely precede the development of LOH.

Several theories have been proposed to explain the potential mechanism linking genomic mutations and CN-LOH.[39,52] Expanding upon previous mechanistic models, it can be hypothesized that some mutations, particularly those with extensive deletions or insertions, may destabilize the DNA structure in ways that directly promote CN-LOH during mitotic recombination or during failed attempts to correct the mutant allele. According to this model, CN-LOH develops in a

'deterministic' manner, occurring initially at the mutated locus. Alternatively, CN-LOH may develop in a more 'stochastic' manner. According to this model, CN-LOH randomly develops throughout the genome of leukemic blasts. Clones developing CN-LOH at loci containing advantageous mutations are perpetuated, while those clones developing CN-LOH in normal allelic regions will have no survival advantage and eventually perish. This stochastic model does not exclude the possibility that the mutations have a role in the initial development of CN-LOH. For example, certain oncogenic mutations may disrupt key regulatory pathways that control genomic fidelity, apoptosis and proliferation. In the case of *FLT3*/ITD mutations, the fact that CN-LOH$_{13q}$ is not observed in *FLT3*/WT cases is highly suggestive of causal link between *FLT3*/ITD and evolution of CN-LOH$_{13q}$. In addition, the expression data from others suggest that many pathways are disrupted in leukemic blasts harboring ITD mutations.[53–55] In the current study, we identified a number of expression differences between *FLT3*/WT and *FLT3*/ITD$_{H\text{-}AR}$ (Figure 2; Tables 2–4). Some of these expression changes involve the DNA repair/recombination pathway, which, if antedate the development of CN-LOH, may have a role in its development. Hence, the disruption of these critical functions may predispose cells to the development of CN-LOH throughout the genome, eventually leading to the emergence of a dominant clone with *FLT3*/ITD$_{H\text{-}AR}$. However, the number of samples evaluated for expression studies were limited, and we cannot rule out that other potential confounders may be contributing to the observed expression changes. Additional studies are needed to better dissect out the molecular mechanisms underlying these expression differences and determine the potential etiology driving the development of CN-LOH in *FLT3*/ITD blasts.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

This work was supported in part by National Institute of Health grant numbers CA114563 (SM) and CA160872 (DLS).

AUTHOR CONTRIBUTIONS

Conception and design was by DLS and SM. Development of methodology (FISH, SNP/CGH, PCR and expression arrays) was by KT, JJ and ELP-A. Analysis and interpretation of data was done by DLS, KT, ELP-A, JJ and SM. Writing, review and/or revision of the manuscript was done by DLS, KT, ELP-A, JJ and SM.

REFERENCES

1 Nakao M, Yokota S, Iwai T, Kaneko H, Horiike S, Kashima K *et al.* Internal tandem duplication of the flt3 gene found in acute myeloid leukemia. *Leukemia* 1996; **10**: 1911–1918.
2 Yamamoto Y, Kiyoi H, Nakano Y, Suzuki R, Kodera Y, Miyawaki S *et al.* Activating mutation of D835 within the activation loop of FLT3 in human hematologic malignancies. *Blood* 2001; **97**: 2434–2439.
3 Matsuno N, Nanri T, Kawakita T, Mitsuya H, Asou N. A novel FLT3 activation loop mutation N841K in acute myeloblastic leukemia. *Leukemia* 2005; **19**: 480–481.
4 Kiyoi H, Towatari M, Yokota S, Hamaguchi M, Ohno R, Saito H *et al.* Internal tandem duplication of the FLT3 gene is a novel modality of elongation mutation which causes constitutive activation of the product. *Leukemia* 1998; **12**: 1333–1337.
5 Lavagna-Sevenier C, Marchetto S, Birnbaum D, Rosnet O. FLT3 signaling in hematopoietic cells involves CBL, SHC and an unknown P115 as prominent tyrosine-phosphorylated substrates. *Leukemia* 1998; **12**: 301–310.
6 Kiyoi H, Naoe T, Nakano Y, Yokota S, Minami S, Miyawaki S *et al.* Prognostic implication of FLT3 and N-RAS gene mutations in acute myeloid leukemia. *Blood* 1999; **93**: 3074–3080.
7 Abu-Duhier FM, Goodeve AC, Wilson GA, Gari MA, Peake IR, Rees DC *et al.* FLT3 internal tandem duplication mutations in adult acute myeloid leukaemia define a high-risk group. *Br J Haematol* 2000; **111**: 190–195.
8 Kottaridis PD, Gale RE, Frew ME, Harrison G, Langabeer SE, Belton AA *et al.* The presence of a FLT3 internal tandem duplication in patients with acute myeloid leukemia (AML) adds important prognostic information to cytogenetic risk group and response to the first cycle of chemotherapy: analysis of 854 patients from the United Kingdom Medical Research Council AML 10 and 12 trials. *Blood* 2001; **98**: 1752–1759.
9 Meshinchi S, Woods WG, Stirewalt DL, Sweetser DA, Buckley JD, Tjoa TK *et al.* Prevalence and prognostic significance of FLT3 internal tandem duplication in pediatric acute myeloid leukemia. *Blood* 2001; **97**: 89–94.
10 Schnittger S, Schoch C, Dugas M, Kern W, Staib P, Wuchter C *et al.* Analysis of FLT3 length mutations in 1003 patients with acute myeloid leukemia: correlation to cytogenetics, FAB subtype, and prognosis in the AMLCG study and usefulness as a marker for the detection of minimal residual disease. *Blood* 2002; **100**: 59–66.
11 Thiede C, Steudel C, Mohr B, Schaich M, Schakel U, Platzbecker U *et al.* Analysis of FLT3-activating mutations in 979 patients with acute myelogenous leukemia: association with FAB subtypes and identification of subgroups with poor prognosis. *Blood* 2002; **99**: 4326–4335.
12 Meshinchi S, Alonzo TA, Stirewalt DL, Zwaan M, Zimmerman M, Reinhardt D *et al.* Clinical implications of FLT3 mutations in pediatric AML. *Blood* 2006; **108**: 3654–3661.
13 Gale RE, Green C, Allen C, Mead AJ, Burnett AK, Hills RK *et al.* The impact of FLT3 internal tandem duplication mutant level, number, size, and interaction with NPM1 mutations in a large cohort of young adult patients with acute myeloid leukemia. *Blood* 2008; **111**: 2776–2784.
14 Schnittger S, Bacher U, Kern W, Alpermann T, Haferlach C, Haferlach T. Prognostic impact of FLT3-ITD load in NPM1 mutated acute myeloid leukemia. *Leukemia* 2011; **25**: 1297–1304.
15 Frohling S, Schlenk RF, Breitruck J, Benner A, Kreitmeier S, Tobis K *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* 2002; **100**: 4372–4380.
16 Griffiths M, Mason J, Rindl M, Akiki S, McMullan D, Stinton V *et al.* Acquired isodisomy for chromosome 13 is common in AML, and associated with FLT3-itd mutations. *Leukemia* 2005; **19**: 2355–2358.
17 Raghavan M, Lillington DM, Skoulakis S, Debernardi S, Chaplin T, Foot NJ *et al.* Genome-wide single nucleotide polymorphism analysis reveals frequent partial uniparental disomy due to somatic recombination in acute myeloid leukemias. *Cancer Res* 2005; **65**: 375–378.
18 Dunbar AJ, Gondek LP, O'Keefe CL, Makishima H, Rataul MS, Szpurka H *et al.* 250K single nucleotide polymorphism array karyotyping identifies acquired uniparental disomy and homozygous mutations, including novel missense substitutions of c-Cbl, in myeloid malignancies. *Cancer Res* 2008; **68**: 10349–10357.
19 Sievers EL, Larson RA, Stadtmauer EA, Estey E, Lowenberg B, Dombret H *et al.* Efficacy and safety of gemtuzumab ozogamicin in patients with CD33-positive acute myeloid leukemia in first relapse. *J Clin Oncol* 2001; **19**: 3244–3254.
20 Lange BJ, Dinndorf P, Smith FO, Arndt C, Barnard D, Feig S *et al.* Pilot study of idarubicin-based intensive-timing induction therapy for children with previously untreated acute myeloid leukemia: Children's Cancer Group Study 2941. *J Clin Oncol* 2004; **22**: 150–156.
21 Affymetrix I. CNAT 4.0: Copy Number and Loss of Heterozygosity Estimation Algorithms for the GeneChip® Human Mapping 10/50/100/250/500 K Array Set, 2007 (cited 2013 0716/2013); revision version 1.2:(available from http://www.affymetrix.com/esearch/search.jsp?N=4000097&Ntt=cnat + 4.0&st=doc.
22 Konings P, Vanneste E, Jackmaert S, Ampe M, Verbeke G, Moreau Y *et al.* Microarray analysis of copy number variation in single cells. *Nat Protoc* 2012; **7**: 281–310.
23 Pollard JA, Alonzo TA, Gerbing RB, Woods WG, Lange BJ, Sweetser DA *et al.* FLT3 internal tandem duplication in CD34 + /CD33 − precursors predicts poor outcome in acute myeloid leukemia. *Blood* 2006; **108**: 2764–2769.
24 Stirewalt DL, Meshinchi S, Kopecky KJ, Fan W, Pogosova-Agadjanyan EL, Engel JH *et al.* Identification of genes with abnormal expression changes in acute myeloid leukemia. *Genes Chromosomes Cancer* 2008; **47**: 8–20.
25 Ho PA, Kutny MA, Alonzo TA, Gerbing RB, Joaquin J, Raimondi SC *et al.* Leukemic mutations in the methylation-associated genes DNMT3A and IDH2 are rare events in pediatric AML: a report from the Children's Oncology Group. *Pediatr Blood Cancer* 2011; **57**: 204–209.
26 Ho PA, Kopecky KJ, Alonzo TA, Gerbing RB, Miller KL, Kuhn J *et al.* Prognostic implications of the IDH1 synonymous SNP rs11554137 in pediatric and adult AML: a report from the Children's Oncology Group and SWOG. *Blood* 2011; **118**: 4561–4566.
27 Ostronoff F, Othus M, Ho PA, Kutny M, Geraghty DE, Petersdorf SH *et al.* Mutations in the DNMT3A exon 23 independently predict poor outcome in older patients with acute myeloid leukemia: a SWOG report. *Leukemia* 2013; **27**: 238–241.
28 Brazma A, Parkinson H, Sarkans U, Shojatalab M, Vilo J, Abeygunawardena N *et al.* ArrayExpress--a public repository for microarray gene expression data at the EBI. *Nucleic Acids Res* 2003; **31**: 68–71.

29 Bolstad BM, Irizarry RA, Astrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* 2003; **19**: 185–193.
30 Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U *et al.* Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 2003; **4**: 249–264.
31 Stirewalt DL, Choi YE, Sharpless NE, Pogosova-Agadjanyan EL, Cronk MR, Yukawa M *et al.* Decreased IRF8 expression found in aging hematopoietic progenitor/stem cells. *Leukemia* 2009; **23**: 391–393.
32 Storey JD, Tibshirani R. Statistical significance for genomewide studies. *Proc Natl Acad Sci USA* 2003; **100**: 9440–9445.
33 Xu XL, Olson JM, Zhao LP. A regression-based method to identify differentially expressed genes in microarray time course studies and its application in an inducible Huntington's disease transgenic model. *Hum Mol Genet* 2002; **11**: 1977–1985.
34 Gene Expression Omnibus (database on the internet), 2007. Available from: http://www.ncbi.nlm.nih.gov/geo/.
35 Ingenuity S. FAQs about statistical considerations, 2006 (cited 2013 7/15/2013). Available from: http://ingenuity.force.com/ipa/IPATutorials?id=kA250000000TNACCA4.
36 Stirewalt DL, Mhyre AJ, Marcondes M, Pogosova-Agadjanyan E, Abbasi N, Radich JP *et al.* Tumour necrosis factor-induced gene expression in human marrow stroma: clues to the pathophysiology of MDS? *Br J Haematol* 2008; **140**: 444–453.
37 Akagi T, Ogawa S, Dugas M, Kawamata N, Yamamoto G, Nannya Y *et al.* Frequent genomic abnormalities in acute myeloid leukemia/myelodysplastic syndrome with normal karyotype. *Haematologica* 2009; **94**: 213–223.
38 Akagi T, Shih LY, Kato M, Kawamata N, Yamamoto G, Sanada M *et al.* Hidden abnormalities and novel classification of t(15;17) acute promyelocytic leukemia (APL) based on genomic alterations. *Blood* 2009; **113**: 1741–1748.
39 O'Keefe C, McDevitt MA, Maciejewski JP. Copy neutral loss of heterozygosity: a novel chromosomal lesion in myeloid malignancies. *Blood* 2010; **115**: 2731–2739.
40 Nowak D, Klaumuenzer M, Hanfstein B, Mossner M, Nolte F, Nowak V *et al.* SNP array analysis of acute promyelocytic leukemia may be of prognostic relevance and identifies a potential high risk group with recurrent deletions on chromosomal subband 1q31.3. *Genes Chromosomes Cancer* 2012; **51**: 756–767.
41 Liehr T. Cytogenetic contribution to uniparental disomy (UPD). *Mol Cytogenet* 2010; **3**: 8.
42 Yi JH, Huh J, Kim HJ, Kim SH, Kim HJ, Kim YK *et al.* Adverse prognostic impact of abnormal lesions detected by genome-wide single nucleotide polymorphism array-based karyotyping analysis in acute myeloid leukemia with normal karyotype. *J Clin Oncol* 2011; **29**: 4702–4708.
43 Mupo A, Celani L, Dovey O, Cooper JL, Grove C, Rad R *et al.* A powerful molecular synergy between mutant Nucleophosmin and Flt3-ITD drives acute myeloid leukemia in mice. *Leukemia* 2013; **27**: 1917–1920.
44 Sistonen L, Keski-Oja J, Ulmanen I, Holtta E, Wikgren BJ, Alitalo K. Dose effects of transfected c-Ha-rasVal 12 oncogene in transformed cell clones. *Exp Cell Res* 1987; **168**: 518–530.
45 Venkatachalam S, Shi YP, Jones SN, Vogel H, Bradley A, Pinkel D *et al.* Retention of wild-type p53 in tumors from p53 heterozygous mice: reduction of p53 dosage can promote cancer formation. *EMBO J* 1998; **17**: 4657–4667.
46 Araki T, Mohi MG, Ismat FA, Bronson RT, Williams IR, Kutok JL *et al.* Mouse model of Noonan syndrome reveals cell type- and gene dosage-dependent effects of Ptpn11 mutation. *Nat Med* 2004; **10**: 849–857.
47 Rosenbauer F, Wagner K, Kutok JL, Iwasaki H, Le Beau MM, Okuno Y *et al.* Acute myeloid leukemia induced by graded reduction of a lineage-specific transcription factor, PU.1. *Nat Genet* 2004; **36**: 624–630.
48 Chan IT, Kutok JL, Williams IR, Cohen S, Kelly L, Shigematsu H *et al.* Conditional expression of oncogenic K-ras from its endogenous promoter induces a myeloproliferative disease. *J Clin Invest* 2004; **113**: 528–538.
49 Sarkisian CJ, Keister BA, Stairs DB, Boxer RB, Moody SE, Chodosh LA. Dose-dependent oncogene-induced senescence *in vivo* and its evasion during mammary tumorigenesis. *Nat Cell Biol* 2007; **9**: 493–505.
50 Wang J, Liu Y, Li Z, Wang Z, Tan LX, Ryu MJ *et al.* Endogenous oncogenic Nras mutation initiates hematopoietic malignancies in a dose- and cell type-dependent manner. *Blood* 2011; **118**: 368–379.
51 Wang J, Liu Y, Li Z, Du J, Ryu MJ, Taylor PR *et al.* Endogenous oncogenic Nras mutation promotes aberrant GM-CSF signaling in granulocytic/monocytic precursors in a murine model of chronic myelomonocytic leukemia. *Blood* 2010; **116**: 5991–6002.
52 Tuna M, Knuutila S, Mills GB. Uniparental disomy in cancer. *Trends Mol Med* 2009; **15**: 120–128.
53 Lacayo NJ, Meshinchi S, Kinnunen P, Yu R, Wang Y, Stuber CM *et al.* Gene expression profiles at diagnosis in *de novo* childhood AML patients identify FLT3 mutations with good clinical outcomes. *Blood* 2004; **104**: 2646–2654.
54 Radmacher MD, Marcucci G, Ruppert AS, Mrozek K, Whitman SP, Vardiman JW *et al.* Independent confirmation of a prognostic gene-expression signature in adult acute myeloid leukemia with a normal karyotype: a Cancer and Leukemia Group B study. *Blood* 2006; **108**: 1677–1683.
55 Balgobind BV, Van den Heuvel-Eibrink MM, De Menezes RX, Reinhardt D, Hollink IH, Arentsen-Peters ST *et al.* Evaluation of gene expression signatures predictive of cytogenetic and molecular subtypes of pediatric acute myeloid leukemia. *Haematologica* 2011; **96**: 221–230.

Supplementary Information accompanies this paper on Blood Cancer Journal website (http://www.nature.com/bcj)