Prevalence and prognostic significance of allelic imbalance by single-nucleotide polymorphism analysis in low-risk myelodysplastic syndromes

Azim Mohamedali,[1] Joop Gäken,[1] Natalie A. Twine,[1] Wendy Ingram,[1] Nigel Westwood,[1] Nicholas C. Lea,[1] Janet Hayden,[1] Nora Donaldson,[2] Carlo Aul,[3] Norbert Gattermann,[4] Aristotle Giagounidis,[4] Ulrich Germing,[4] Alan F. List,[5] and Ghulam J. Mufti[1]

[1]Department of Haematological Medicine, King's College London School of Medicine, London, United Kingdom; [2]Research and Development, King's College Hospital, London, United Kingdom; [3]St Johannes Hospital, Medizinische Klinik II, Duisberg, Germany; [4]Department of Haematology, Oncology, and Clinical Immunology, Heinrich-Heine-University, Düsseldorf, Germany; [5]H. Lee Moffitt Cancer Center and Research Institute, Department of Hematologic Malignancies, Tampa, FL

Low-risk myelodysplastic syndrome (MDS) with normal cytogenetics accounts for approximately 50% of MDS patients. There are no pathognomonic markers in these cases and the diagnosis rests on cytomorphologic abnormalities in bone marrow and/or peripheral blood. Affymetrix high-resolution single-nucleotide polymorphism (SNP) genotyping microarrays allow detection of cytogenetically cryptic genomic aberrations. We have studied 119 low-risk MDS patients (refractory anemia [RA] = 22; refractory cytopenia with multilineage dysplasia [RCMD] = 51; refractory anemia with ringed sideroblasts [RARS] = 12; refractory cytopenia with multilineage dysplasia with ringed sideroblasts [RCMD-RS] = 12; 5q− syndrome = 16; refractory anemia with excess blasts [RAEB] = 6) using SNP microarrays to seek chromosomal markers undetected by conventional cytogenetics. Loss of heterozygosity (LOH) detected by 50K arrays was verified using 250K and 500K arrays. We demonstrate the presence of uniparental disomy (UPD) in 46%, deletions in 10%, and amplifications in 8% of cases. Copy number (CN) changes were acquired, whereas UPDs were also detected in constitutional DNA. UPD on 4q was identified in 25% of RARS, 12% of RCMD with normal cytogenetics, 17% of RAEB, and 6% of 5q− syndrome cases. Univariate analysis showed deletions ($P = .04$) and International Prognostic Scoring System (IPSS; $P < .001$) scores correlated with overall survival; however, on multivariate analysis only IPSS scores retained prognostic significance ($P < .001$). We show, for the first time, that SNP microarray analysis in low-risk MDS patients reveals hitherto unrecognized UPD and CN changes that may allow stratification of these patients for early therapeutic interventions. (Blood. 2007;110:3365-3373)

Introduction

The myelodysplastic syndromes (MDSs) are clonal disorders of hematopoiesis characterized by morphologic dysplasia, ineffective hematopoiesis, and peripheral blood cytopenias, with progressive evolution to acute myeloid leukemia in approximately 25% of cases.[1] Based on the International Prognostic Scoring System (IPSS), patients with low-risk and intermediate-1 subgroups have a median survival of 5.7 years and 3.5 years, respectively.[2] However, within these low-/intermediate-risk patients there is a significant variation in outcome and attempts have been made to identify variables that may delineate subgroups with a poorer prognosis. The World Health Organization (WHO) classification of MDS has attempted to sharpen the distinction between subgroups of low-risk MDS.[3] Dysplasia within a hematopoietic lineage is now defined when at least 10% of cells in a specific lineage are affected; patients with unilineage dysplasia are classified as refractory anemia (RA), whereas patients with refractory cytopenia with multilineage dysplasia (RCMD) are identified as a separate category. Several groups have since reported on the utility of this classification with the observation that RCMD is associated with a significantly poorer survival than RA.[4,5] Most patients with RA or RCMD have normal cytogenetics and in the absence of definitive biologic markers, early diagnosis or prognostic stratification of patients is currently not possible.

Recent developments in microarray technology have enabled the exploitation of single-nucleotide polymorphisms (SNPs) for high-resolution genome-wide genotyping. This technology facilitates studies of acquired genetic imbalances, as it reveals areas of loss of heterozygosity (LOH). These copy number (CN) imbalances might be early events in the development of the neoplastic clone and predispose to the pathogenetic chromosomal aberrations that are detected by conventional cytogenetics and fluorescence in situ hybridization (FISH) and correlate with prognosis.[6] A recent example is the identification of chromosomal aberrations in 40% of B-cell acute lymphocytic leukemia (B-ALL) cases for genes affecting B-cell development and differentiation, with *PAX5* most frequently targeted by somatic mutation.[7] However, LOH can also occur by uniparental disomy (UPD), where an individual acquires a duplicated copy of an entire or partial chromosome derived from 1 parent through mitotic recombination, resulting in LOH without CN loss or gain. UPD cannot be detected by conventional cytogenetic analysis, as there is no net loss/gain of genetic material. Thus SNP microarrays have a distinct advantage over traditional techniques for the detection of chromosomal aberrations arising from UPD and CN change at higher resolution. The presence of somatically acquired UPD in 20% of acute myeloid leukemia

Submitted March 14, 2007; accepted July 16, 2007. Prepublished online as *Blood* First Edition paper, July 18, 2007; DOI 10.1182/blood-2007-03-079673.

The online version of this article contains a data supplement.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

Table 1. Clinical characteristics of 119 patients with myelodysplastic syndromes

Characteristics	Patients, n (%)
Sex	
Male	55 (46)
Female	64 (54)
Median age, y (range)	61 (20-90)
Median follow-up, mo (range)	46 (1-236)
WHO classification	
RA	22 (19)
RCMDnc	34 (29)
RCMDac	17 (14)
RARS	12 (10)
RCMD-RS	12 (10)
5q− syndrome	16 (13)
RAEB	6 (5)*
IPSS	
Low risk	70 (64)
Int-1	33 (30)
Int-2	3 (3)
High	3 (3)
Unavailable	10 (8)
Cytogenetics	
Normal	77 (65)
Isolated del(5q)	21 (18)
Del(5q) + additional cytogenetic changes	15 (12)
Unavailable	6 (5)

*Four patients had transformed from 5q− syndrome and 2 from RARS.

(AML) patients with normal karyotype was first demonstrated using the Affymetrix 10K SNP array (High Wycombe, United Kingdom).[8] Furthermore, in some AML patients the regions of UPD have been shown to correspond with areas of homozygous mutation in potentially leukemogenic genes.[9] More recently, Affymetrix 50K SNP analysis in chronic lymphocytic leukemia and myeloma has shown that UPD is a prominent feature of these diseases.[10,11] These genetic lesions can be further analyzed to identify regions harboring potential genes of pathogenetic importance,[10] thereby allowing stratification of different tumor types[12] and potentially assisting in treatment modalities.

We have performed Affymetrix 250K SNP analysis on 119 MDS patients in order to determine the frequencies of CN changes and UPD in low-risk MDS with normal cytogenetics and in MDS associated with deletion of 5q. The results were subsequently correlated with known prognostic variables. We demonstrate the presence of novel LOH in 64% of these patients. The frequencies of UPD and CN changes are a prominent feature of MDS, distinguishing it clearly from healthy donors. In addition, the majority of UPDs in bone marrow samples from MDS patients were also present in paired constitutional DNA, whereas copy number aberrations were found to be acquired and indicative of poor prognosis.

Patients and methods

Patient samples

A diagnosis of MDS was made by 2 hematopathologists according to the WHO classification.[13] A total of 119 patients with a median age of 61 years (range, 20-90 y) were analyzed. Patient characteristics are summarized in Table 1. A diagnosis of RA was made in 19% (n = 22), RCMD with normal cytogenetics (RCMDnc) in 29% (n = 34), RCMD with abnormal cytogenetics (RCMDac) in 14% (n = 17), RA with ring sideroblasts (RARS) in 10% (n = 12), RCMD with ringed sideroblasts (RCMD-RS) in 10% (n = 12), 5q− syndrome (< 5% bone marrow [BM] blasts) in 13% (n = 16), and refractory anemia with excess blasts (RAEB) in 5% (n = 6). Within the RAEB subgroup, 4 patients had transformed from 5q− syndrome and 2 from RARS. Low-risk MDS (IPSS 0) was observed in 64% (n = 70) of evaluated patients, intermediate-1 (IPSS 0.5–1) in 30% (n = 33), intermediate-2 (IPSS 1.5-2) in 3% (n = 3), and high-risk MDS (IPSS ≥ 2.5) in 3% (n = 3). IPSS scores were unavailable in 10 patients. Cytogenetic analysis was normal in 65% (n = 77), isolated del(5q) in 18% (n = 21), and del(5q) plus additional cytogenetic changes in 12% (n = 15) of cases. Cytogenetics were unavailable in 5% (n = 6) of patients.

DNA from bone marrow films and fresh $CD34^+$ cells from patients were analyzed by SNP arrays. To determine LOH and CN change, DNA from healthy peripheral blood donor cells (n = 16; median age 41 y, range 17-59 y) and from age-matched subjects with no history of malignancies who were attending the anticoagulation clinic for lone atrial fibrillation (n = 17; median age 76 y, range 61-87 y) was analyzed. Furthermore, where available, constitutional DNA (n = 20) was also analyzed; the latter consisted of DNA from skin biopsy of 9 patients, buccal mucosa of 3 patients, and $CD3^+$ lymphocytes of 8 patients. Approval was obtained from King's College Hospital Local Research Ethics Committee prior to study commencement. Written informed consent was obtained in accordance with the Declaration of Helsinki prior to blood or bone marrow sample collection.

DNA extraction and amplification

Material from bone marrow films was recovered by PBS rehydration, transferred into 200 μL of PBS, and thoroughly resuspended. Freshly obtained bone marrow aspirate samples were enriched for $CD34^+$ cells using an AutoMACS cell sorter (Miltenyi Biotec, Surrey, United Kingdom) per the manufacturer's instructions and frozen as cell pellets at −80°C for batch processing. Genomic DNA was extracted using the Qiagen Blood Mini Kit (Qiagen, Crawley, United Kingdom) per manufacturer's instructions. The quality and quantity of genomic DNA was assessed by electrophoresis and spectrophotometry. Samples of good quality but insufficient quantity were amplified using phi29 polymerase (Qiagen) per the manufacturer's protocol. Phi29 polymerase–based whole-genome amplification has been shown to have a 99% concordance between amplified and unamplified genomic DNA samples.[14] These amplified samples were further tested by polymerase chain reaction (PCR) for either *ACTB* or *GAPDH* genes to evaluate the effectiveness of the amplification. Primers and PCR conditions are described in Document S1 (available on the *Blood* website; see the Supplemental Materials link at the top of the online article).

Genotyping assay

Genotyping was performed on DNA using either 50K *Hin*dIII (50K), 250K *Sty*I (250K), or 250K *Nsp*I + 250K *Sty*I (500K) SNP microarray chips and processed according to the manufacturer's instructions (Affymetrix) with minor modifications. Genomic DNA (375 ng) was digested with the appropriate restriction enzyme for 4 hours at 37°C. Digested DNA (125 ng) was electrophoresed to confirm complete digestion. All further downstream processes such as adaptor ligation, PCR amplification, purification, fragmentation, end labeling, and hybridization were done according to the manufacturer's protocols. The Affymetrix 450 fluidics station and the Affymetrix 3000 G7 gene scanner were used to wash, stain, and scan the arrays.

Expression array

Expression profiling of 2 RARS samples, with and without UPD on 4q, was performed using HGU133 Plus 2.0 arrays (Affymetrix). These samples were analyzed by a pairwise comparison and also compared with normal $CD34^+$ cells (n = 10). Details of the method and analysis are provided in Document S1.

Data analysis

DChipSNP software version 2006 (http://biosun1.harvard.edu/complab/dchip/) was used to determine LOH and CN. The HMM-considering haplotype method was used to determine LOH, and CN was inferred using

median smoothing and trimmed analysis (2% threshold; personal e-mail communication, Cheng Li, Harvard School of Public Health, October 2006). CN changes were verified using the Copy Number Analyzer for Affymetrix GeneChip Mapping (CNAG version 2.0) algorithm (http://www.genome.umin.jp/CNAGtop2.html).[15]

The median call rates were as follows: 50K *Hin*dIII arrays (n = 71) 98.7% (range, 91.7%-99.7%); 250K *Sty*I arrays (n = 145) 95.9% (range, 90.1%-98.8%); and 250K *Nsp*I arrays (n = 92) 95.8% (range, 92.4%-98.9%). The accuracy was verified by measuring the number of heterozygote calls on the X chromosome of male patients. The median of heterozygote calls for the 50K assay was 4 (range, 0-10) in 1159 SNPs, for the 250K *Nsp*I assay was 65 (range, 25-184) in 5710 SNPs, and for the 250K *Sty*I assay was 90 (range, 7-194) in 4826 SNPs. These levels indicate accuracies of 99.99%, 99.99%, and 99.98% for the 50K *Hin*dIII, 250K *Nsp*I, and 250K *Sty*I arrays, respectively.

Statistical analysis

Statistical analysis was performed using SPSS version 14.0 (SPSS, Chicago, IL). Patient characteristics were compared using chi-square analysis for categoric variables and the Mann-Whitney test for continuous variables. Correlation was calculated by using the Pearson product moment correlation coefficient. Overall survival curves were estimated using Kaplan-Meier methodology and the log-rank test was used to assess differences between groups. Univariate comparisons and multivariate analysis used the Cox proportional hazards regression model. Variables analyzed included WHO subgroup, IPSS score, cytogenetic status, UPD, and CN change. A *P* value of less than .05 was set as the threshold of significance.

Results

50K versus 250K versus 500K arrays

We investigated the presence or absence of LOH in 119 MDS patients using high-density SNP microarrays. Our initial experiments were performed on 50K SNP arrays followed by reanalyses on 250K and 500K SNP arrays. The regions of UPD greater than 2 megabases (Mb) and CN change were validated by the concordance of LOH found in the 50K, 250K, and 500K SNP arrays for the same MDS samples. All 119 samples were analyzed on 250K; 48 samples on 50K and 250K; and 27 samples on 50K, 250K, and 500K. There was 100% concordance between the 50K and 250K arrays in the detection of regions of UPD greater than 2 Mb and CN changes. Similarly, there was 100% concordance between the 250K and the 500K arrays for UPD regions greater than 2 Mb and CN changes. Of 48 cases identified as having abnormalities using the 50K array, reanalysis with the 250K array revealed an additional 22 regions of LOH (15 UPD > 2 Mb and 7 CN changes). However, the 500K array did not add any further regions of LOH over and above what was found in the 250K array; therefore, all further patients were analyzed using the 250K SNP array.

In all patients with del(5q) abnormality there was 100% concordance between the 250K array analyses and conventional cytogenetic findings. Cytogenetic analysis mapped the proximal breakpoint of del(5q) to 5q13 with the distal breakpoint ranging from q31 to q35. In del(5q) patients, SNP array analysis identified the deleted proximal breakpoint to 5q14.2 and the distal breakpoint to q34. Our results confirm previously identified common deleted regions.[16-18] There was only 1 patient from the del(5q) group with an additional copy number change that was not identified at presentation.

Table 2. Summary of the proportion of patients across the MDS subgroups presenting with LOH

MDS subtype	Patients, n = 119	Deletions, n (%)	Amplification, n (%)	UPD greater than 2 Mb, n (%)
5q− syndrome	16	1 (6.3)	0	6 (37.5)
RA + RCMDnc	56	4 (7.1)	6 (10.7)	27 (48.2)
RARS + RCMD RS	24	4 (16.7)	1 (4.2)	13 (54.1)
RCMDac	17	2 (11.8)	1 (5.9)	7 (41.1)
RAEB	6	1 (16.7)	1 (16.7)	2 (33.3)
Control subjects	33	0	0	4 (12)*

The UPD rate in 33 normal subjects was 12%, whereas it was at least 33% in various low-risk MDS subtypes. Deletions and amplification were only detected in MDS patients.

*The median size of UPD in healthy control is 1.21 Mb (range, 0.3-6.7 Mb) and the median size in MDS patients is 3.78 Mb (range, 2.03-138.57 Mb).

Loss of heterozygosity in MDS

In order to determine whether regions of LOH identified in patient samples were disease specific, DNA from 33 subjects with no history of malignancy was screened for LOH. These 33 subjects consisted of a group of 16 healthy blood donors and a further group of 17 older subjects who attend the anticoagulation clinic at King's College Hospital for lone atrial fibrillation. No regions of CN change were identified; therefore, the presence of CN changes of any size within the patients' DNA that were not previously identified by conventional cytogenetics was included in the analyses. However, 72% (n = 24) of control subjects had regions of UPD of median size of 1.21 Mb (range, 0.3-6.7 Mb) dispersed across the genome on 20 of 23 chromosomes. UPD of greater than 2 Mb was identified in 12% (n = 4) of our controls who had 7 regions of UPD (2.40 Mb, 2.66 Mb, 2.69 Mb, 2.74 Mb, 2.89 Mb, 3.05 Mb, and 6.70Mb). Examination of 250K SNP data from 60 healthy donors published by the HapMap project[19] showed 15% (n = 9) of individuals with UPD greater than 2 Mb. Therefore, all regions of UPD greater than 2 Mb in patient samples were excluded from subsequent analysis.

UPD was identified in bone marrow cells from 46% (n = 55/119) of MDS patients overall (Table 2). DNA of 54% (n = 13) of patients in the RARS + RCMD-RS group and 48% (n = 27) of patients in the RA + RCMDnc group contained regions of UPD. A total of 125 regions of UPD were identified. Thirty-eight MDS patients had 1 region of UPD greater than 2 Mb, 12 patients had 2 to 5 regions, and 5 patients had more than 5 regions of UPD. The median size of UPD was 3.78 Mb (range, 2.03-138.57 Mb). Paired constitutional DNA samples for UPD analysis were available for 13 patients: 5q− syndrome (n = 3), RA + RCMDnc (n = 6), RCMD-RS (n = 1), RCMDac (n = 2), and RAEB (n = 1). Twelve of 13 regions of UPD found in paired patient BM DNA were also identified in the respective constitutional DNA. Interestingly, 2 patients did not have any regions of UPD but had copy number aberrations in their bone marrow DNA. Analysis of the constitutional DNA from these patients also confirmed the absence of UPD and copy number aberrations. In addition, we did not observe any additional regions of UPD in constitutional DNA that were not present in the bone marrow DNA. None of the UPD regions identified in patient BM samples were present in any of the 33 healthy control subjects.

SNP array analysis identified 34 regions of novel CN change in 18% of MDS patients (n = 20), comprising deletions in 10% of patients (n = 12) and amplifications in 8% (n = 9), which were not identified on cytogenetic analysis (Table 2). One patient from the

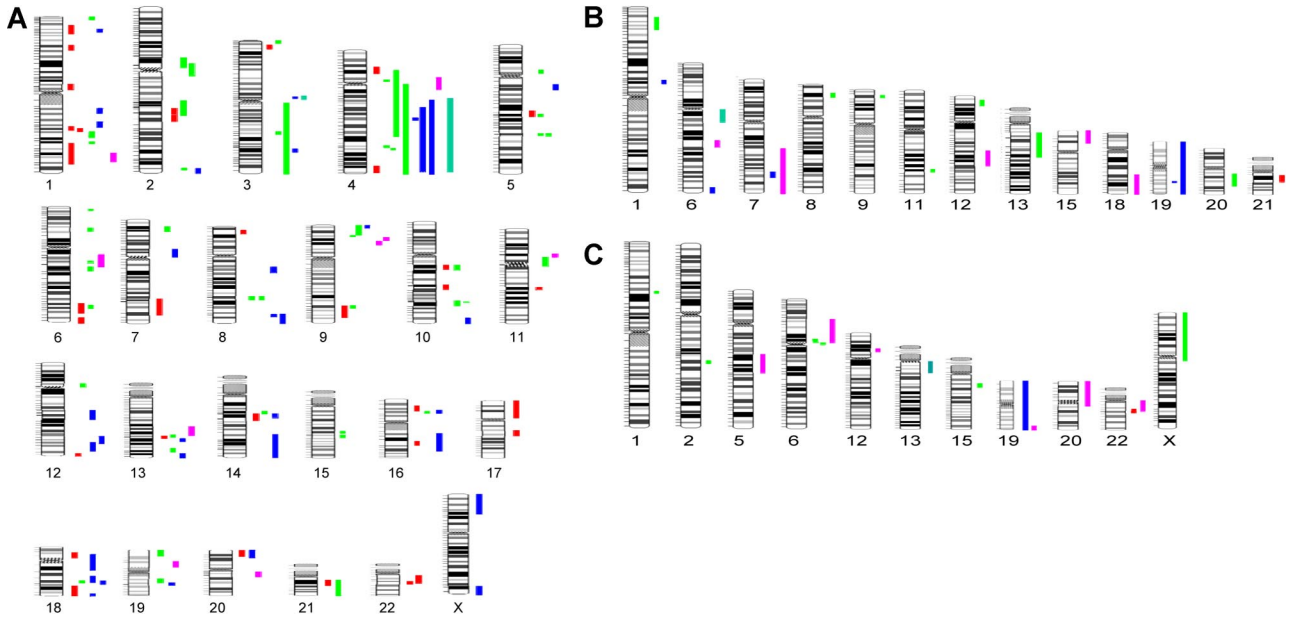

Figure 1. Genomic ideogram showing the distribution of LOH in low-risk MDS patients detected using the 250K SNP assay. Chromosome numbers are below each ideogram. (A) Distribution of UPD greater than 2Mb is shown. One hundred twenty-five regions of UPD were detected: red 5q− Syndrome (n = 33 regions); green RA+RCMDnc (n = 46 regions); blue RARS+RCMD-RS (n = 34 regions); pink RCMDac (n = 10 regions); and green RAEB (n = 2 regions). (B,C) Regions of deletions (del) and amplifications (amp) are shown. CN changes detected: red 5q− syndrome (del, n = 1 region); green RA+RCMDnc (del, n = 7 regions; amp, n = 7 regions); blue RARS+RCMD-RS (del, n = 5 regions; amp, n = 1 region); pink RCMDac (del, n = 5 regions; amp, n = 6 regions); and green RAEB (del, n = 1 region; amp, n = 1 region).

RCMDac group had both multiple deletions and amplifications in the DNA. There were 2 patients in the RAEB group with either a deletion (17%; n = 1) or an amplification (17%; n = 1). The RARS + RCMD-RS group had 17% (n = 4) with deletions but only 4% (n = 1) had an amplification. The frequencies of CN changes in patients were observed as follows: 1 region (n = 16), 2 regions (n = 2), and more than 5 regions (n = 2). Copy number change consisted of 19 regions of deletions of median size of 2.86 Mb (range, 0.14-63.54 Mb) and 15 amplifications with a median size of 4.13 Mb (range, 0.001-68.11 Mb). Paired constitutional DNA was available for 8 patients who had CN changes detected by SNP analysis of their bone marrow cells (4 with and 4 without cytogenetic abnormalities). In all 8 cases, constitutional tissue DNA was found to be normal by SNP array analysis, confirming that CN changes in the bone marrow are somatically acquired.

Genomic spread of LOH

Figure 1A through C summarizes the genomic location of UPD and CN change in all our MDS patient samples. UPD was distributed across the entire genome with no specific pattern, and none of these regions were shared by more than 3 MDS patients. Six chromosomes had 0 to 2 regions of UPD, a further 6 had 3 to 5 regions of UPD, 9 chromosomes had 6 to 9 regions of UPD, and 2 chromosomes had more than 10 regions of UPD. The most notable chromosomes with frequent regions of UPD in MDS samples were chromosome 4 (n = 12 regions), chromosome 1 (n = 10 regions), chromosome 6 (n = 8 regions), and chromosomes 2 and 3 (n = 7 regions). All other chromosomes had fewer than 6 regions of UPD. Chromosomes most frequently associated with CN changes were as follows: chromosome 6 (n = 5 regions), chromosome 19 (n = 4 regions), and chromosomes 1 and 12 (n = 3 regions). CN changes were not detected on chromosomes 3, 4, 10, 14, and 16. The remaining chromosomes had 2 or fewer regions of CN change. There was no significant correlation between chromosome size and frequency of UPD (r = 0.69) or CN change (r = −0.01).

UPD on chromosome 4

UPD on chromosome 4q was identified in 3 of 12 RARS patients with a median size of 91 Mb (range, 2.9-108.89 Mb), 4 of 34 RCMDnc patients ranging from 2.6 to 138.5 Mb with a median size of 49.9 Mb, 1 of 6 patients from the RAEB group (107 Mb), and 1 of 16 patients with 5q− syndrome as shown in Figure 2. Constitutional DNA from 1 RARS patient with a UPD of 90.9 Mb (range, q21.5-q35.1) was available and this did not show the identical UPD on 4q. However, this patient had additional UPD present on 3q26.1 (2.6 Mb) and 13q31.1 (3.1 Mb), which were present in both the BM and paired constitutional DNA.

The gene-expression profile of $CD34^+$ cells from 2 RARS patients with and without UPD on 4q, respectively, were compared with $CD34^+$ cells from 10 healthy donors (see Document S1 for methods and data analysis). Of the 334 genes that were expressed on the region of UPD (4q21.5-q35.1), there were 118 genes from the patient with 4q UPD and 45 genes from the patient without 4q UPD that changed by at least 3-fold in comparison to expression from normal $CD34^+$ cells. Pairwise comparison of gene expression between the RARS samples identified 105 differentially expressed genes, of which 64 were down-regulated and 41 were up-regulated by 3-fold (Table S1). Interestingly, erythroid genes such as glycophorin A (*GYPA,* 6.5-fold), glycophorin B (*GYPB,* 3-fold), and the mitochondrial uncoupling protein 1 (*UCP1,* 4.7-fold) were up-regulated in the patient with 4q UPD. Functional comparison of the 105 differentially expressed genes using ingenuity pathway analysis (IPA) identified the top 3 biofunctional groups as genetic disorders (4 genes: *FGG, FGB, GYPA,* and *NR3C2;* significance = 2.9×10^{-5} to 1.2×10^{-2}); hematologic disease (3 genes: *FGB, FGG,* and *EDNRA*; significance = 2.9×10^{5} to 6.3×10^{-3}); and cardiovascular disease (8 genes: *EDNRA, FGB, FGG, GUCY1A3, NR3C2, PDE5A, CCNA2,* and *GRIA2*; significance = 1.6×10^{-4} to 1.9×10^{-2}).

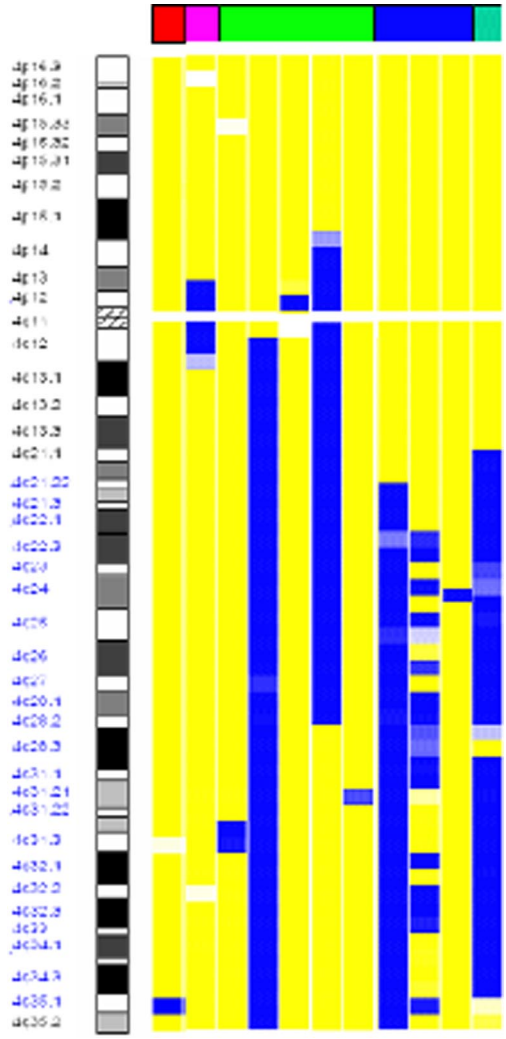

Figure 2. Detailed ideogram of UPD on chromosome 4. The yellow areas indicate regions of heterozygosity and the blue areas indicate regions of LOH. All regions of LOH were copy neutral: red 5q− syndrome (n = 2); pink RCMDac (n = 1); RA+RCMDnc (n = 5); blue RARS (n = 3); and green RAEB (n = 1).

Copy number changes

There were 13 chromosomes carrying a deletion and 11 chromosomes with an amplification detected by SNP arrays that were not detected by conventional cytogenetics summarized in Figure 1B and C. Cases from the RA + RCMDnc (41%) and RCMDac (32%) subgroups contributed the most CN changes overall from all of the MDS categories. Indeed 1 patient from the RCMDac group had monosomy 7 and an additional 8 CN abnormalities that were not detected by conventional cytogenetic analysis at sampling. This patient progressed to the RAEB subtype. Among the most frequent copy number abnormalities previously described, we also found cases of deletion 20q11.23-13.13 (12.91 Mb) in an RCMDnc patient and deletion 7q33.3-qterm (0.35 Mb) in a RCMD-RS patient who transformed to RAEB. Furthermore, 1 RCMDac patient and 3 RARS patients also showed CN aberrations in chromosome 19: RCMDac, amplification at 19q13.42-13.43 (4.13 Mb); and RARS, deletions at q13.31 (0.37 Mb), p13.3-q13.43 (63.55 Mb), and trisomy 19 (63.81 Mb).

Correlation of LOH with clinical variables

The median duration of follow-up from time of diagnosis of the study group was 46 months (range, 1-236 months) with a median time to disease progression of 36 months (range, 6-200 months). There was no correlation between the presence of either UPD or CN change with age (UPD, $P = .248$; CN, $P = .586$), sex (UPD, $P = .443$; CN, $P = .387$), or disease progression (UPD, $P = .446$; CN, $P = .449$). Disease progression was defined as either progression to RAEB (n = 19) or transformation to AML (n = 5). At the time of disease progression, 4 cases had evidence of cytogenetic evolution. On further analysis of those patients with CN change, the presence of a deletion was associated with a higher proportion of patients showing disease progression (37.5%, 3/8) when compared with the rate of progression in those without a deletion (13%, 9/69), although this was not significant ($P = .19$).

Furthermore, there was no correlation between the presence of UPD and increasing IPSS score. There was however a correlation between CN aberrations and higher IPSS score ($P = .01$). This correlation was more pronounced when deletions/amplifications were analyzed separately. A higher frequency of deletions was observed in those with high IPSS scores ($P = .007$) compared with amplifications ($P = .17$).

The overall survival of patients with or without UPD or CN aberrations was analyzed by Kaplan-Meier plots. No difference in the overall survival was seen on comparison of patients with or without UPD greater than 2 Mb ($P = .51$; Figure 3A). To account for cases of UPD greater than 2 Mb in healthy subjects, survival was also analyzed for UPD greater than 3.1 Mb and no significant difference was observed. Univariate analysis of factors affecting overall survival identified the presence of CN change ($P = .03$; Figure 3B) and IPSS score ($P < .001$) as the most significant variables affecting survival (Table 3). In particular, the presence of deletions ($P = .04$) compared with amplifications ($P = .44$) adversely affected outcome (Figure 3C,D). The WHO subgroups of MDS ($P = .06$), presence of cytogenetic abnormalities ($P = .99$), and the presence of UPD ($P = .23$) did not affect the overall survival. Although CN change correlated with increasing IPSS score, on multivariate analysis IPSS score ($P < .001$) was the only predictive variable for overall survival (Table 4).

Discussion

Low-risk MDS subgroups accounts for at least 50% of all MDS patients, and up to 80% of these patients have a normal karyotype.[5] In the absence of clonal cytogenetic aberrations, the diagnosis depends on the presence of morphologic abnormalities defined in the WHO classification.[13] Due to the poor in vitro growth of the dysplastic clone, it is often difficult to obtain useful cytogenetic data from MDS patients with a "normal" karyotype. The advent of high-resolution SNP microarrays has provided an opportunity to identify genome-wide cytogenetically cryptic genomic aberrations in low-risk MDS patients, which may prove to be of diagnostic or prognostic significance. Furthermore, these aberrations may help identify genes that lead to progression of low-risk MDS subtypes. To our knowledge this is the first report using high-resolution SNP genotyping to ascertain the incidence, spread, and prognostic consequence of subcytogenetic changes in MDS.

Most studies to date have used 10K and 50K Affymetrix arrays that have an SNP intermarker median distance of 470 kb

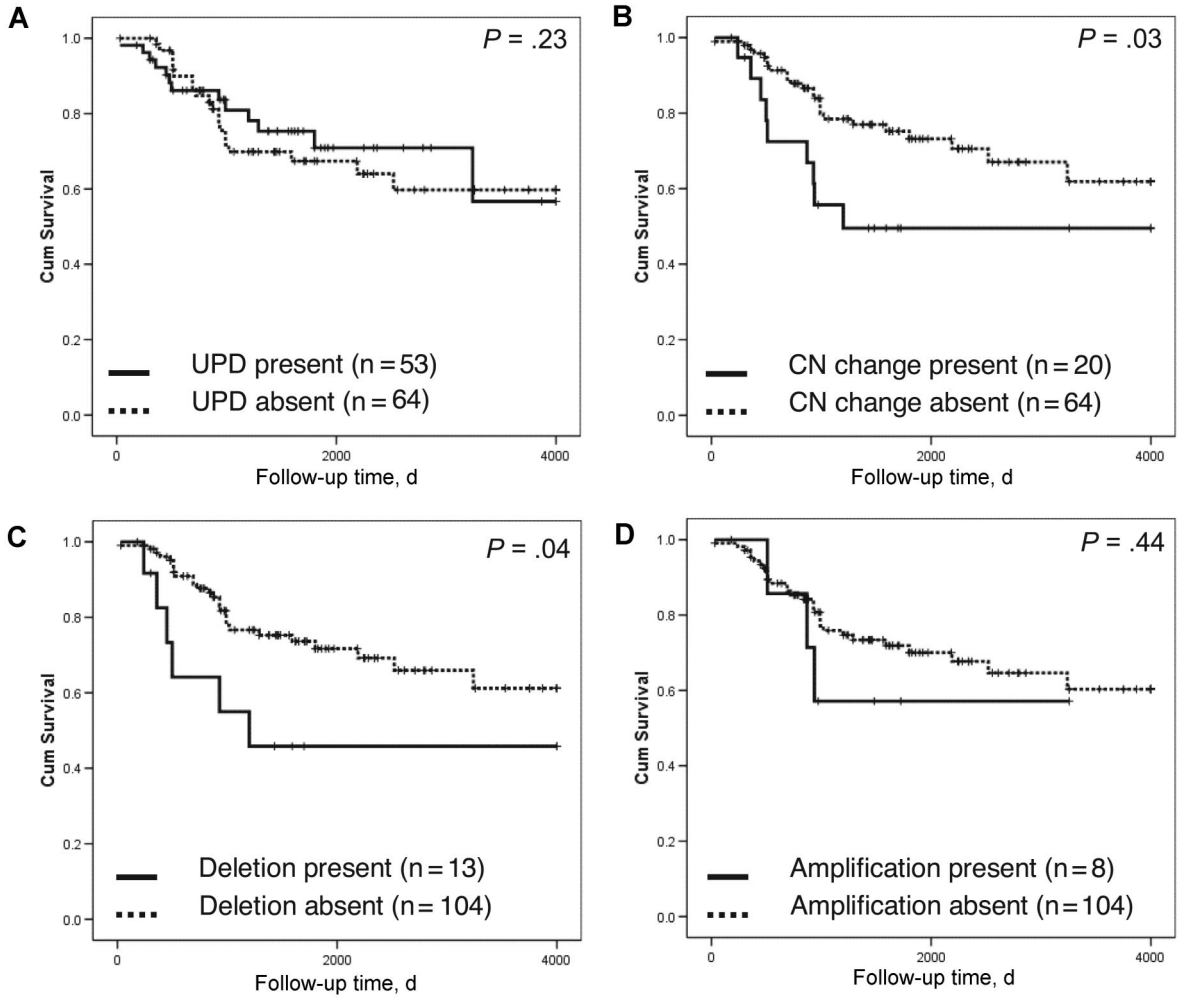

Figure 3. Kaplan-Meier plots demonstrating the effect of LOH on patient survival. (A,B) The effect of UPD and copy number change on overall survival is shown. UPD present (n = 53), UPD absent (n = 64), CN present (n = 20), CN absent (n = 97). (C,D) The overall survival of MDS patients according to the presence of deletions or amplifications, respectively, is shown. Deletion present (n = 13), deletion absent (n = 104), amplification present (n = 8), amplification absent (n = 109). Two patients did not have any follow-up data.

and 8.5 kb, respectively, but recent improvements in photolithography techniques have led to the advent of 500K SNP arrays with an improved intermarker resolution of 2.5 kb. The procedure has been refined to provide a precise, practical, and high-resolution technique that can be applied for diagnostic purposes. In comparison to the standard genotyping technique of microsatellite marker analysis, SNP microarrays have a clear advantage for detection of LOH at a much higher resolution. However, our results show that increased resolution provided by the 500K SNP arrays did not add any further information over what was identified using the 250K SNP arrays for low-risk MDS patients. We observed a 100% concordance in the regions of LOH detected with 50K arrays and verified on the 250K arrays for 48 samples.

Uniparental disomy arises from the duplication of one of the parental chromosomes and loss of the other, thereby conferring homozygosity across the affected region. This consequence of mitotic recombination can contribute to pathogenicity by "activating" potential recessive oncogenes[9] or inactivating tumor suppressor genes[20] and also plays a major role in imprinting[21] as outlined in the 2-hit Knudson model[22] describing the inactivation of tumor suppressor genes. Two population studies used SNP genotyping technology to document long contiguous stretches of homozygosity in white samples obtained from the Center d'Etude du

Table 3. Univariate analysis of factors affecting overall survival

	Hazard ratio	95% CI	*P*
WHO subgroup	1.2	0.99-1.81	.06
IPSS, IPSS 0 vs greater than 0	2.67	1.53-4.6	<.001
Cytogenetic abnormalities	0.98	0.61-1.65	.99
UPD	0.62	0.28-1.4	.23
CN change	2.3	1.08-4.8	.03
Deletion	2.4	1.04-5.62	.04
Amplification	1.6	0.49-5.26	.44

CI indicates confidence interval.

Table 4. Multiple Cox regression analysis following adjustment for age and CN change

	Hazard ratio	95% CI	*P*
WHO subgroup	0.84	0.56-1.26	.40
IPSS, IPSS 0 vs greater than 0	2.66	1.53-4.60	<.001
CN change	1.13	0.41-3.11	.81

CI indicates confidence interval.

Polymorphissme Humain (CEPH).[23,24] Both studies showed the presence of homozygous regions in the genome; however, data from the CEPH collection may be biased due to autozygosity that occurs when 2 copies of an ancestral haplotype come together in an individual. Of the 8 reference families from CEPH, 2 showed clear signs of consanguinity, whereas the remaining 6 families, all from Utah, had been subject to some inbreeding in the past.[25] These factors might explain the overestimation of CN-neutral LOH. In another report, only 1 region of UPD (4.95 Mb) was found in 36 DNA samples from healthy bone marrow analyzed using the 250K SNP array.[26] We show the occurrence of small regions of UPD ranging from 0.3 to 6.7 Mb in our control subjects. It needs emphasizing that only 12% of our control subjects had UPD greater than 2 Mb; therefore, we excluded all UPDs below 2 Mb from our analysis.

UPD has been shown to be present in high-risk MDS,[27] AML with normal karyotype,[28] and polycythemia vera,[29] but to our knowledge this is the first report showing UPD greater than 2 Mb (range, 2.03-138.57 Mb) in 46% of patients with low-risk MDS. More significantly, regions of UPD found in BM DNA were also present in their respective constitutional DNA, where available, except for 1 RARS patient who had 3 regions of UPD in the BM DNA, of which the largest UPD (90.9 Mb) was not present in the paired constitutive DNA. Constitutive DNA from a patient with RCMDnc had 7 regions of UPD ranging from 2.1 to 96 Mb, with all being present in the patient's constitutive DNA. We cannot exclude the possibility that large regions of UPD are either somatic or germ-line aberrations.

The high incidence of constitutional UPD found in MDS may suggest an underlying predisposition of genomic instability. Fragile sites, which are known sites of frequent genomic instability, were mapped from The Genome Database[30] to the regions of UPD.[31] We found that 43.2% of the UPD regions were localized within or as part of a known fragile site. In the absence of a reliable marker for low-risk MDS patients with a normal karyotype, the high frequency of UPD may provide a surrogate disease marker. Even though there was a high incidence of UPD in our MDS patients, the size or their frequency did not affect progression to a high-risk MDS subtype or prognosis. The full relevance of this has yet to be ascertained, but it may have an important consequence in the inheritance of homozygous recessive mutations or aberrant genomic imprinting resulting in loss of a functional allele.

The most prominent region for UPD was on chromosome 4q in patients with RCMDnc, RARS, and RAEB. UPD on 4q has previously been shown to occur in patients with A-beta Lipoproteinemia.[32] Interestingly, the RAEB patient with UPD on 4q had transformed from a previous RARS condition. A recent study on 4 MDS and AML patients also identified a common microdeletion on 4q24, suggesting the possible involvement of a tumor suppressor gene within the region.[33] Our expression analyses did not detect any overtly down-regulated genes in a patient with 4q UPD. We performed an ontologic survey using OntoExpress[34] for genes localized to the large region of UPD4q and found 2 genes of interest: guanylate cyclase soluble beta 3 (*GUCY1B3*)-4q32.1 and electron-transferring lavoprotein dehydrogenase (*ETFDH*)-4q32.1, which are involved in heme binding. However, there was no difference in gene expression between the patients with and without 4q UPD for *GUCY1B3* and *ETFDH.* The up-regulation of erythroid-specific genes *GYPA*, *GYPB,* and *GYPE* and a mitochondrial transporter protein *UCP1* in the patient with 4q UPD relative to normal controls and also by pair-wise analysis with the patient without 4q UPD suggests a possible role in dyserythropoeisis associated with RARS. These genes were also up-regulated in 2 of 19 RARS patients from a previously published gene-expression study[34] and it would be interesting to determine if these patients also had UPD on 4q. However, because of the small sample size, this observation needs to be validated in a larger cohort of MDS patients where genotyping data are available to confirm the 4q UPD.

In addition, we also analyzed differential gene expression from $CD34^+$ cells from patients with RARS and del(5q) patients[35] in the context of our LOH data and found no significant correlation (Table S2 and Table 3). None of the up-regulated interferon-stimulated genes mapped onto any of our identified regions of LOH. Apart from a common theme of UPD on 4q identified in a small percentage of our patients, there were no common regions of LOH between any of the subtypes of MDS or between any of the patients within a MDS subtype.

Copy number aberrations in low-risk MDS cases were not as prevalent as UPDs. Previous reports have highlighted the presence of LOH in up to 20% of cases with AML with normal cytogenetics using SNP arrays.[28] In addition, studies have investigated LOH in the transformation of MDS to AML using informative microsatellite analysis and found LOH, but not deletions, in greater than 20% of cases on chromosome arms 6q, 7p, 10p, 11q, 14q, and 20q, suggesting the presence of unidentified tumor suppressor genes in these regions.[27] Therefore, the use of SNP arrays may allow detection of additional CN changes that may aid in the prognostic stratification of patients and identify those most likely to transform.

To confirm that the CN changes we observed were acquired, we analyzed matched constitutional DNA where available and found that CN changes were present only in BM DNA and not in the respective constitutional DNA. These CN changes were compared with the genomic variants in the general population[36] and most of the CN changes we observed, including confirmed acquired CN changes, were present at low frequencies within the general population. This highlights the importance of analyzing paired samples to confirm acquired copy number changes. Common regions of CN aberrations were not observed for the different MDS subgroups. Chromosome 19 harbored 4 CN changes (2 deletions and 2 amplifications) from RARS and RCMDac patients. A shared deleted (19p13.31) region was identified in 2 RARS patients. This region is also susceptible to recurrent chromosomal translocation and is highlighted by the fusion of E2A-PBX1 and E2A-HLF, which plays a role in the pathogenesis of lymphoblastic leukemias.[36] Gain of chromosome 19 has been shown to be a feature of chronic myelomonocytic[38] and megakaryoblastic leukemias.[39]

A large study in patients with AML with an abnormal karyotype has shown that patients with copy number changes, especially deletions, had a poorer outcome.[6] Our results extend this observation to MDS, where we found that copy number aberrations play a significant role in stratifying prognosis. The presence of deletions ($P = .05$) resulted in a significantly adverse prognosis for low-risk MDS patients (Figure 3C,D). It is highly relevant to note that our results were concordant with the IPSS score. Patients with a high IPSS score had a significantly ($P = .007$) higher incidence of deletions than patients with a low IPSS score, reflecting poor prognosis. Given that cytogenetics is an important factor in the derivation of the IPSS score, the present findings that SNP microarray results correlate with IPSS score in cytogenetically normal cases strongly indicate that molecular methods for the detection of LOH could be an important addendum to the already

established prognosticators in MDS. This finding needs further investigation.

In conclusion, this study has shown that UPD is an important feature of low-risk MDS. This UPD was found to be present both in BM DNA and in paired constitutional DNA, where available. Although there was no prognostic significance of UPD in low-risk MDS patients, the frequency of UPD may provide a marker of inherent genomic instability in these patients. Moreover, UPD on 4q in the DNA of a subset of patients may have pathogenic significance, especially since an increase in erythroid gene expression was detected. CN changes were detected by SNP microarrays in cytogenetically normal MDS cases, confirming the greater sensitivity and reliability of SNP microarrays over metaphase-dependent conventional cytogenetics. However, there was no difference in LOH detected between 250K and 500K SNP arrays in our cases. Finally, we show that patients with CN changes, especially those who harbor deletions, are those with higher IPSS scores who have a substantially poorer prognosis.

Acknowledgments

We would like to acknowledge the technical assistance provided by Lee Macpherson and Ashutosh Chauhan with SNP array experiments.

This work was supported by grants from the Wolfson Charitable Trust, Leukemia Research Fund (LRF), Bundesministerium für Bildung und Forschung (BMBF), and Elimination of Leukemia Fund (ELF). W.I. is funded by the LRF.

Authorship

Contribution: A.M. and J.G. conducted research, analyzed data, and prepared the manuscript; N.T and W.I. collected and analyzed clinical data and cowrote the manuscript; N.W. and N.C.L. processed clinical samples and provided intellectual contribution; J.H. coordinated collection of clinical samples; N.D. performed and verified statistical analysis; C.A., N.G., A.G., U.G., and A.L. provided clinical samples and intellectual contribution; G.J.M. conceived the project and was involved in manuscript preparation.

A.M. and J.G. contributed equally to this study.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

Correspondence: G. J. Mufti, Department of Haematological Medicine, Kings College London and Kings College Hospital, Denmark Hill, London, United Kingdom SE5 9RS; e-mail: ghulam.mufti@kcl.ac.uk.

References

1. Delforge M. Understanding the pathogenesis of myelodysplastic syndromes. Hematol J. 2003;4:303-309.
2. Greenberg P, Cox C, LeBeau MM, et al. International scoring system for evaluating prognosis in myelodysplastic syndromes. Blood. 1997;89:2079-2088.
3. Harris NL, Jaffe ES, Diebold J, et al. The World Health Organization classification of neoplastic diseases of the haematopoietic and lymphoid tissues: Report of the Clinical Advisory Committee Meeting, Airlie House, Virginia, November 1997. Histopathology. 2000;36:69-86.
4. Dunkley SM, Manoharan A, Kwan YL. Myelodysplastic syndromes: prognostic significance of multilineage dysplasia in patients with refractory anemia or refractory anemia with ringed sideroblasts. Blood. 2002;99:3870-3871; author reply 3871.
5. Haase D, Germing U, Schanz J, et al. New and comprehensive cytogenetic prognostication and categorization in MDS [abstract]. Blood. 2006;108:248. Abstract 248.
6. Schoch C, Haferlach T, Bursch S, et al. Loss of genetic material is more common than gain in acute myeloid leukemia with complex aberrant karyotype: a detailed analysis of 125 cases using conventional chromosome analysis and fluorescence in situ hybridization including 24-color FISH. Genes Chromosomes Cancer. 2002;35:20-29.
7. Mullighan CG, Goorha S, Radtke I, et al. Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. Nature. 2007;446:758-764.
8. Raghavan M, Lillington DM, Skoulakis S, 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.
9. Fitzgibbon J, Smith LL, Raghavan M, et al. Association between acquired uniparental disomy and homozygous gene mutation in acute myeloid leukemias. Cancer Res. 2005;65:9152-9154.
10. Pfeifer D, Pantic M, Skatulla I, et al. Genome-wide analysis of DNA copy number changes and LOH in CLL using high-density SNP arrays. Blood. 2007;109:1202-1210.
11. Walker BA, Leone PE, Jenner MW, et al. Integration of global SNP-based mapping and expression arrays reveals key regions, mechanisms, and genes important in the pathogenesis of multiple myeloma. Blood. 2006;108:1733-1743.
12. Janne PA, Li C, Zhao X, et al. High-resolution single-nucleotide polymorphism array and clustering analysis of loss of heterozygosity in human lung cancer cell lines. Oncogene. 2004;23:2716-2726.
13. Jaffe ESH, Stein H, Vardiman JW. WHO Classification of Tumors, Pathology and Genetics of Tumours of Haematopoetic and Lymphoid Tissues. Lyon, France: IARC Press; 2001.
14. Paez JG, Lin M, Beroukhim R, et al. Genome coverage and sequence fidelity of phi29 polymerase-based multiple strand displacement whole genome amplification. Nucleic Acids Res. 2004;32:e71.
15. Nannya Y, Sanada M, Nakazaki K, et al. A robust algorithm for copy number detection using high-density oligonucleotide single nucleotide polymorphism genotyping arrays. Cancer Res. 2005;65:6071-6079.
16. Horrigan SK, Arbieva ZH, Xie HY, et al. Delineation of a minimal interval and identification of 9 candidates for a tumor suppressor gene in malignant myeloid disorders on 5q31. Blood. 2000;95:2372-2377.
17. Boultwood J, Fidler C, Strickson AJ, et al. Narrowing and genomic annotation of the commonly deleted region of the 5q- syndrome. Blood. 2002;99:4638-4641.
18. Zhao N, Stoffel A, Wang PW, et al. Molecular delineation of the smallest commonly deleted region of chromosome 5 in malignant myeloid diseases to 1–1.5 Mb and preparation of a PAC-based physical map. Proc Natl Acad Sci U S A. 1997;94:6948-6953.
19. International HapMap Project. http://www.hapmap.org. Accessed February 26, 2007.
20. Stephens K, Weaver M, Leppig KA, et al. Interstitial uniparental isodisomy at clustered breakpoint intervals is a frequent mechanism of NF1 inactivation in myeloid malignancies. Blood. 2006;108:1684-1689.
21. Ledbetter DH, Engel E. Uniparental disomy in humans: development of an imprinting map and its implications for prenatal diagnosis. Hum Mol Genet. 1995;4 Spec No:1757-1764.
22. Knudson AG. Two genetic hits (more or less) to cancer. Nat Rev Cancer. 2001;1:157-162.
23. Gibson J, Morton NE, Collins A. Extended tracts of homozygosity in outbred human populations. Hum Mol Genet. 2006;15:789-795.
24. Li LH, Ho SF, Chen CH, et al. Long contiguous stretches of homozygosity in the human genome. Hum Mutat. 2006;27:1115-1121.
25. Broman KW, Weber JL. Long homozygous chromosomal segments in reference families from the centre d'Etude du polymorphisme humain. Am J Hum Genet. 1999;65:1493-1500.
26. O'Keefe CL, Gondek LP, Tiu R, et al. Identification of chromosomal abnormalities in healthy bone marrow using 250K SNP arrays [abstract]. Blood. 2006;108:2076. Abstract 2076.
27. Mori N, Morosetti R, Hoflehner E, Lubbert M, Mizoguchi H, Koeffler HP. Allelic loss in the progression of myelodysplastic syndrome. Cancer Res. 2000;60:3039-3042.
28. Gorletta TA, Gasparini P, D'Elios MM, Trubia M, Pelicci PG, Di Fiore PP. Frequent loss of heterozygosity without loss of genetic material in acute myeloid leukemia with a normal karyotype. Genes Chromosomes Cancer. 2005;44:334-337.
29. Kralovics R, Guan Y, Prchal JT. Acquired uniparental disomy of chromosome 9p is a frequent stem cell defect in polycythemia vera. Exp Hematol. 2002;30:229-236.
30. The GDB Human Genome Database. http://www.gdb.org/gdb. Accessed February 10, 2007.

31. Gumus G, Sunguroglu A, Tukun A, Sayin DB, Bokesoy I. Common fragile sites associated with the breakpoints of chromosomal aberrations in hematologic neoplasms. Cancer Genet Cytogenet. 2002;133:168-171.
32. Yang XP, Inazu A, Yagi K, Kajinami K, Koizumi J, Mabuchi H. Abetalipoproteinemia caused by maternal isodisomy of chromosome 4q containing an intron 9 splice acceptor mutation in the microsomal triglyceride transfer protein gene. Arterioscler Thromb Vasc Biol. 1999;19:1950-1955.
33. Viguie F, Aboura A, Bouscary D, et al. Common 4q24 deletion in four cases of hematopoietic malignancy: early stem cell involvement? Leukemia. 2005;19:1411-1415.
34. Intelligent Systems and Bioinformatics Laboratory, Wayne State University, Detroit, MI. Accessed November 28, 2006.
35. Pellagatti A, Cazzola M, Giagounidis AAN, et al. Gene expression profiles of CD34+ cells in myelodysplastic syndromes: involvement of interferon-stimulated genes and correlation to FAB subtype and karyotype. Blood. 2006;108:337-345.
36. The Centre for Applied Genomics. Database of Genomic Variants. http://projects.tcag.ca/variation. Accessed January 19, 2007.
37. LeBrun DP. E2A basic helix-loop-helix transcription factors in human leukemia. Front Biosci. 2003;8:s206-222.
38. Daskalakis M, Mauritzson N, Johansson B, et al. Trisomy 19 as the sole chromosomal abnormality in proliferative chronic myelomonocytic leukemia. Leuk Res. 2006;30:1043-1047.
39. Alvarez S, MacGrogan D, Calasanz MJ, Nimer SD, Jhanwar SC. Frequent gain of chromosome 19 in megakaryoblastic leukemias detected by comparative genomic hybridization. Genes Chromosomes Cancer. 2001;32:285-293.