proportions of normal cell components, thus enabling reliable genomewide detection of LOH in a wide variety of primary cancer specimens.

Material and Methods

Samples and Microarray Analysis

Genomic DNA extracted from a lung cancer cell line (NCI-H2171) was intentionally mixed with DNA from its paired lymphoblastoid cell line (LCL) (NCI-BL2171) to generate a dilution series, in which tumor contents started at 10% and increased by 10% up to 90%. The ratios of admixture were validated using measurements of a microsatellite (*D3S1279*) within a UPD region on chromosome 3 (data not shown). The nine mixed samples, together with non-mixed original DNAs (0% and 100% tumor contents), were analyzed with GeneChip 50K Xba SNP arrays (Affymetrix). Microarray data corresponding to 5%, 15%, 25%,..., and 95% tumor content were interpolated by linearly superposing two adjacent microarray data sets after adjusting the mean array signals of the two sets. Both cell lines were obtained from the American Type Culture Collection (ATCC). Genomic DNA was also extracted from 85 primary leukemia samples, including 39 acute myeloid leukemia (AML [MIM #601626]) samples and 46 acute lymphoblastic leukemia (ALL) samples, and was subjected to analysis with 50K Xba SNP arrays. Of the 85 samples, 34 were analyzed with their matched complete-remission bone marrow samples. DNA from 53 MPD samples—13 polycythemia vera (PV [MIM #263300]), 21 essential thrombocythemia (ET [MIM #187950]), and 19 idiopathic myelofibrosis (IMF [MIM #254450])—43 of which had been studied for *JAK2* mutations,[8] were also analyzed with 50K Xba SNP arrays. Microarray analyses were performed according to the manufacturer's protocol,[15] except with the use of LA *Taq* (Takara) for adaptor-mediated PCR. Also, DNA from 96 normal volunteers was used for the analysis. All clinical specimens were made anonymous and were incorporated into this study in accordance with the approval of the institutional review boards of the University of Tokyo and Harvard Medical School.

AsCN Analyses Using Anonymous Control Samples (AsCNAR)

SNP typing on the GeneChip platform uses two discrete sets of SNP-specific probes, which are arbitrarily but consistently named "type *A*" and "type *B*" SNPs, at every SNP locus, each consisting of an equal number of perfectly matched probes (PM_As or PM_Bs) and mismatched probes (MM_As or MM_Bs). For AsCN analysis, the sums of perfectly matched probes (PM_As or PM_Bs) for the *i*th SNP locus in the tumor (tum) sample and reference samples (ref1, ref2,..., ref*N*),

$$S_{A,i}^{\mathrm{tum}} = \sum \mathrm{PM}_{A,i}^{\mathrm{tum}}\ ,\quad S_{B,i}^{\mathrm{tum}} = \sum \mathrm{PM}_{B,i}^{\mathrm{tum}}$$

and

$$S_{A,i}^{\mathrm{ref}I} = \sum \mathrm{PM}_{A,i}^{\mathrm{ref}I}\ ,\quad S_{B,i}^{\mathrm{ref}I} = \sum \mathrm{PM}_{B,i}^{\mathrm{ref}I}\ ,\quad (I = 1,2,3,\ldots,N)\ ,$$

are compared separately at each SNP locus, according to the concordance of the SNP calls in the tumor sample (O_i^{tum}) and the SNP calls in a given reference sample ($O_i^{\mathrm{ref}I}$),

$$R_{A,i}^{\mathrm{ref}I} = \frac{S_{A,i}^{\mathrm{tum}}}{S_{A,i}^{\mathrm{ref}I}} \qquad (\text{for } O_i^{\mathrm{tum}} = O_i^{\mathrm{ref}I}),$$

$$R_{B,i}^{\mathrm{ref}I} = \frac{S_{B,i}^{\mathrm{tum}}}{S_{B,i}^{\mathrm{ref}I}}$$

and the total CN ratio is calculated as follows:

$$R_{AB,i}^{\mathrm{ref}I} = \begin{cases} R_{A,i}^{\mathrm{ref}I} & \text{for } O_i^{\mathrm{tum}} = O_i^{\mathrm{ref}I} = AA \\ R_{B,i}^{\mathrm{ref}I} & \text{for } O_i^{\mathrm{tum}} = O_i^{\mathrm{ref}I} = BB \\ \frac{1}{2}(R_{A,i}^{\mathrm{ref}I} + R_{B,i}^{\mathrm{ref}I}) & \text{for } O_i^{\mathrm{tum}} = O_i^{\mathrm{ref}I} = AB \end{cases} (I = 1,2,3,\ldots,N)\ .$$

For CN estimations, however, $R_{AB,i}^{\mathrm{ref}I}$, $R_{A,i}^{\mathrm{ref}I}$, and $R_{B,i}^{\mathrm{ref}I}$ are biased by differences in mean array signals and different PCR conditions between the tumor sample and each reference sample and need to be compensated for these effects to obtain their adjusted values $\hat{R}_{AB,i}^{\mathrm{ref}I}$, $\hat{R}_{A,i}^{\mathrm{ref}I}$, and $\hat{R}_{B,i}^{\mathrm{ref}I}$, respectively (appendix A).[16]

These values are next averaged over the references that have a concordant genotype for each SNP in a given set of references (*K*), and we obtain $\tilde{R}_{AB,i}^{K}$, $\tilde{R}_{A,i}^{K}$, and $\tilde{R}_{B,i}^{K}$. Note that $\tilde{R}_{A,i}^{K}$ and $\tilde{R}_{B,i}^{K}$ are calculated only for heterozygous SNPs in the tumor sample (see appendix A for more details).

A provisional total CN profile Λ_K is provided by

$$\Lambda_K = \{\tilde{R}_{AB,i}^{K}\}\ ,$$

and provisional AsCN profiles are obtained by

$$\Lambda_K^{\mathrm{large}} = \{\max(\tilde{R}_{A,i}^{K}, \tilde{R}_{B,i}^{K})\}$$

$$\Lambda_K^{\mathrm{small}} = \{\min(\tilde{R}_{A,i}^{K}, \tilde{R}_{B,i}^{K})\}\ .$$

These provisional analyses, however, assume that the tumor genome is diploid and has no gross CN alterations, when the coefficients are calculated in regressions. In the next step, the regressions are iteratively performed using a diploid region that is truly or is expected to be diploid, to determine the coefficients on the basis of the provisional total CN, and then the CNs are recalculated.

Finally, the optimized set of references is selected that minimizes the SD of total CN at the diploid region by stepwise reference selection, as described in appendix A.

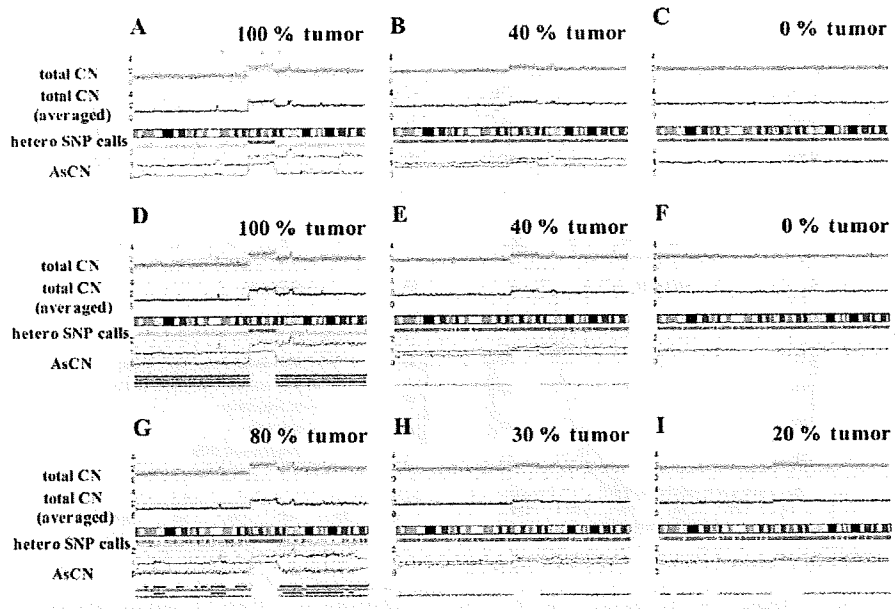

Figure 1. AsCN analysis with or without paired DNA. DNA from a lung cancer cell line (NCI-H2171) was mixed with DNA from an LCL (NCI-BL2171) established from the same patient at the indicated percentages and was analyzed with GeneChip 50K Xba SNP arrays. AsCNs, as well as total CNs, were analyzed using either the paired reference sample (NCI-BL2171) (*upper panels, A–C*) or samples from unrelated individuals simultaneously processed with the tumor samples (*middle and lower panels, D–I*). On each panel, the upper two graphs represent total CNs and their moving averages for the adjacent 10 SNPs, whereas moving averages of AsCNs for the adjacent 10 SNPs are shown below (*red and green lines*). Green and pink bars in the middle are heterozygous (hetero) calls and discordant SNP calls between the tumor and its paired reference, respectively. At the bottom of each panel, LOH regions inferred from AsCNAR (*orange*), SNP call–based LOH inference of CNAG (*blue*), dChip (*purple*), and PLASQ (*light green*) are depicted. Asterisks (*) indicate the loci at which total CNs were confirmed by FISH analysis (data not shown). The calibrations of CN graphs are linearly adjusted so that the mean CNs of null and single alleles should be 0 and 1, respectively.

Allele-specific analysis using a constitutive reference, refSelf, is provided by

$$\Lambda^{\text{large}} = \{\max(\hat{R}_{A,i}^{\text{refSelf}}, \hat{R}_{B,i}^{\text{refSelf}})\}$$

and

$$\Lambda^{\text{small}} = \{\min(\hat{R}_{A,i}^{\text{refSelf}}, \hat{R}_{B,i}^{\text{refSelf}})\} .$$

Computational details of AsCNAR are provided in appendix A.

Comparison with Other Algorithms

dChip[17] and PLASQ[18] were downloaded from their sites, and the identical microarray data were analyzed using these programs. Since PLASQ requires both Xba and Hind array data, microarray data of mixed tumor contents for Hind arrays were simulated by linearly superimposing the tumor cell line (NCI-H2171) and LCL (NCI-BL2171) data at indicated proportions.

Statistical Analysis

Significance of the presence of allelic imbalance (AI) in a given region, Γ, called as having AI by the hidden Markov model (HMM), was statistically tested by calculating *t* statistics for the difference in AsCNs, $|\log_2 \tilde{R}_{A,i}^{\kappa} - \log_2 \tilde{R}_{B,i}^{\kappa}|$, between Γ and a normal diploid region, where the tests were unilateral. Significance between the numbers of UPDs detected by the SNP call–based method and by AsCNAR was tested by one-tailed binominal tests. *P* values for AI detection by allele-specific PCR were calculated by one-tailed *t* tests, comparing triplicates of the target sample and triplicates of five normal samples that have heterozygous alleles in the SNP.

Detection of the JAK2 *Mutation and Measurements of Relative Allele Doses*

The *JAK2* V617F mutation was examined by a restriction enzyme–based analysis, in which PCR-amplified *JAK2* exon 12 fragments were digested with *Bsa*XI, and the presence of the undigested fragment was examined by gel electrophoresis.[5] Relative allele dose between wild-type and mutated *JAK2* was determined by measuring allele-specific PCR products for wild-type and mutated *JAK2* alleles by

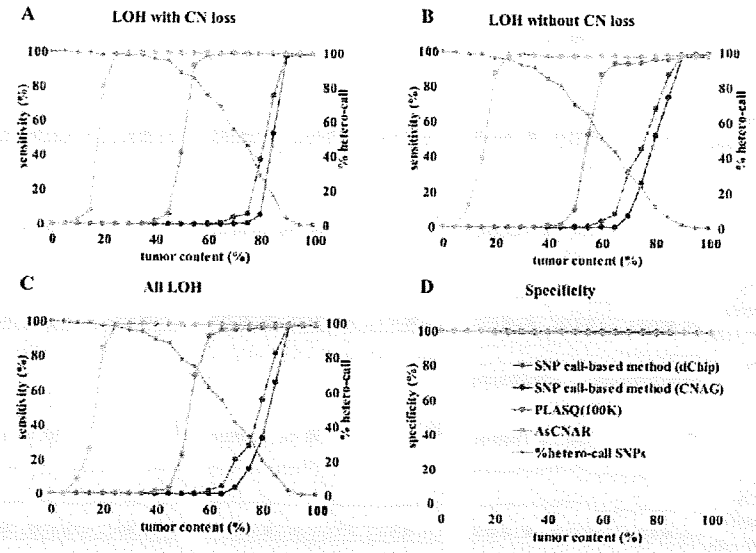

Figure 2. Sensitivity and specificity of LOH detection for intentionally mixed tumor samples. Sensitivity of detection of LOH with or without CN loss (*A* and *B*) in different algorithms were compared using a mixture of the tumor sample (NCI-H2171) and the paired LCL sample (NCI-BL2171). The results for all LOH regions are shown in panel C, and the specificities of LOH detection are depicted in panel D. For precise estimation of sensitivity and specificity, we defined the SNPs truly positive and negative for LOH as follows. The tumor sample and the paired LCL sample were genotyped on the array three times independently, and we considered only SNPs that showed the identical genotype in the three experiments. SNPs that were heterozygous in the paired LCL sample and were homozygous in the tumor sample were considered to be truly positive for LOH, and SNPs that were heterozygous both in the paired LCL sample and in the tumor sample were considered to be truly negative. Proportions of heterozygous SNP calls (%hetero-call) that remained in LOH regions of each sample are also shown in panels A–C.

capillary electrophoresis by use of the 3100 Genetic Analyzer (Applied Biosystems), as described in the literature.[19] Likewise, the fraction of tumor components having 9p and other UPDs was measured by either allele-specific PCR or STR PCR,[7,19] by use of the primers provided in appendix B [online only]. The percentage of UPD-positive cells (%UPD(+)) was also estimated as the mean difference of AsCNs for heterozygous SNPs within the UPD region divided by that for homozygous SNPs within an arbitrary selected normal region:

$$\%\mathrm{UPD}(+) = \frac{E(|\tilde{R}^K_{A,i} - \tilde{R}^K_{B,i}|_{i\in\text{hetero SNPs in UPD region}})}{E(|\tilde{R}^K_{A,i} - \tilde{R}^K_{B,i}|_{i\in\text{homo SNPs with normal CN}})},$$

where AsCNs for the denominator were calculated as if the homozygous SNPs were heterozygous. However, in those samples with a high percentage of UPD-positive components, the heterozygous SNP rate in the UPD region decreased. For such regions, we calculated the percentage of UPD-positive cells by randomly selecting 30% (the mean heterozygous SNP call rate for this array) of all the SNPs therein and by assuming that they were heterozygous SNPs. Cellular composition of *JAK2* wild-type (wt) and mutant (mt) homozygotes (wt/wt and mt/mt) and heterozygotes (wt/mt) in each MPD specimen was estimated assuming that all UPD components are homozygous for the *JAK2* mutation. The fractions of the wt/mt heterozygotes in cases with a 9p gain were estimated assuming that the duplicated 9p alleles had the *JAK2* mutation. Throughout the calculations, small negative values for wt/mt were disregarded.

FISH

FISH analysis was performed according to the previously published method, to confirm the absolute total CNs in NCI-H2171.[20] The genomic probes were generated by whole-genome amplification of FISH-confirmed RP11 BAC clones 169N13 (3q13; CN = 2), 227F7 (8q24; CN = 2), 196H14 (12q14; CN = 2), 25E13 (13q33; CN = 2), 84E24 (17q24; CN = 2), 12C9 (19q13; CN = 2), 153K19 (3q13; CN = 3), 94D19 (3p14; CN = 1), 80P10 (8q22; CN = 1), and 64C21 (13q12-13; CN = 1), which were obtained from the BACPAC Resources Center at the Children's Hospital Oakland Research Institute in Oakland, California.

Results

SNP Call–Based Genomewide LOH Detection by Use of SNP Arrays

When a pure tumor sample is analyzed with a paired constitutive reference on a GeneChip Xba 50K array, LOH is easily detected as homozygous SNP loci in the tumor spec-

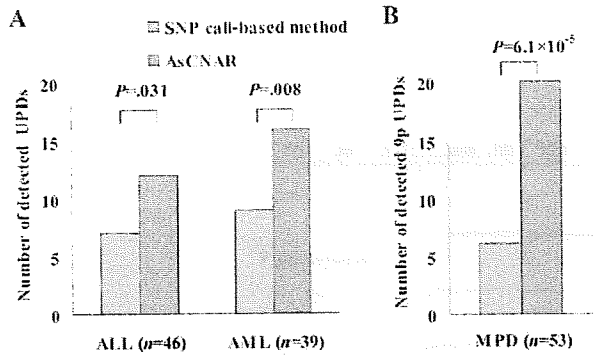

Figure 3. The number of UPD regions for acute leukemia and MPD samples detected by either the SNP call–based method or AsCNAR. The number of UPD regions for ALL and AML samples detected by the SNP call–based method or by AsCNAR is shown in panel A, and the number of 9p UPDs for MPD samples detected by the two methods is shown in panel B. Some samples have more than one UPD region. Details of UPD regions are given in table 1. Significance between the numbers of UPDs detected by the SNP call–based method and by the AsCNAR method was tested by one-tailed binomial tests.

imen that are heterozygous in the constitutive DNA (fig. 1*A*, pink bars). In addition, given a large number of SNPs to be genotyped, the presence of LOH is also inferred from the grossly decreased heterozygous SNP calls, even in the absence of a paired reference (fig. 1*D*). The accuracy of the LOH inference would depend partly on the algorithm used but more strongly on the tumor content of the specimens. Thus, our SNP call–based LOH inference algorithm in CNAG (appendix C), as well as that of dChip,[17] show almost 100% sensitivity and specificity for pure tumor specimens. But, as the tumor content decreases, the LOH detection rate steeply declines (fig. 1*G*), and, with <50% tumor cells, no LOH can be detected, even when complete genotype information for both tumor and paired constitutive DNA is obtained (fig. 1*B*, 1*E*, 1*H*, and 1*I*).

LOH Detection Based on AsCN Analysis

On the other hand, the capability of allele-specific measurements of CN alterations in cancer genomes is an excellent feature of the SNP array–based CN-detection system that uses a large number of SNP-specific probe sets.[16,18,21] When constitutive DNA is used as a reference, AsCN analysis is accomplished by separately comparing the SNP-specific array signals from the two parental alleles at the heterozygous SNP loci in the constitutive genomic DNA.[16] It determines not only the total CN changes but also the alterations of allelic compositions in cancer genomes, which are captured as the split lines in the two AsCN graphs (fig. 1*A* and 1*B*). In this mode of analysis, the presence of LOH can be detected as loss of one parental allele, even in specimens showing almost no discordant calls (fig. 1*B*).

AsCNAR

The previous method for AsCN analysis, however, essentially depends on the availability of constitutive DNA, since AsCNs are calculated only at the heterozygous SNP loci in constitutive DNA.[16] Alternatively, allele-specific signals can be compared with those in anonymous references on the basis of the heterozygous SNP calls in the tumor specimen. In the latter case, the concordance of heterozygous SNP calls between the tumor and the unrelated sample is expected to be only 37% with a single reference. However, the use of multiple references overcomes the low concordance rate with a single reference, and the expected overall concordance rate for heterozygous SNPs and for all SNPs increases to 86% and 92%, respectively, with five unrelated references (appendix D [online only]). Thus, for AsCNAR, allele-specific signal ratios are calculated at all the concordant heterozygous SNP loci for individual references, and then the signal ratios for the identical SNPs are averaged across different references over the entire genome. For the analysis of total CNs, all the concordant SNPs, both homozygous and heterozygous, are included in the calculations, and the two allele-specific signal ratios for heterozygous SNP loci are summed together. Since AsCNAR computes AsCNs only for heterozygous SNP loci in tumors, difficulty may arise on analysis of an LOH region in highly pure tumor samples, in which little or no heterozygous SNP calls are expected. However, as shown above, such LOH regions can be easily detected by the SNP call–based algorithm, where AsCNAR is formally calculated assuming all the SNPs therein are heterozygous. Thus, the AsCNAR provides an essentially equivalent result to that from AsCN analysis using constitutional DNA, with similar sensitivity in detecting AI and LOH (compare fig. 1*A* with 1*D* and 1*B* with 1*E*).

As expected from its principle, AsCNAR is more robust in the presence of normal cell contaminations than are SNP call–based algorithms. To evaluate this quantitatively, we analyzed tumor DNA that was intentionally mixed with its paired normal DNA at varying ratios in 50K Xba SNP arrays, and the array data were analyzed with AsCNAR. To preclude subjectivity, LOH regions were detected by an HMM-based algorithm, which evaluates difference in AsCNs in both parental alleles (appendix E).[22] As the tumor content decreases, the SNP call–based LOH inference fails to detect LOH because of the appearance of heterozygous SNP calls from the contaminated normal cell component (fig. 1*E* and 1*G*–1*I*), but these heterozygous SNP calls, in turn, make AsCNAR operate effectively.

Table 1. CN-Neutral LOH in Primary Acute Leukemia

The table is available in its entirety in the online edition of *The American Journal of Human Genetics*.

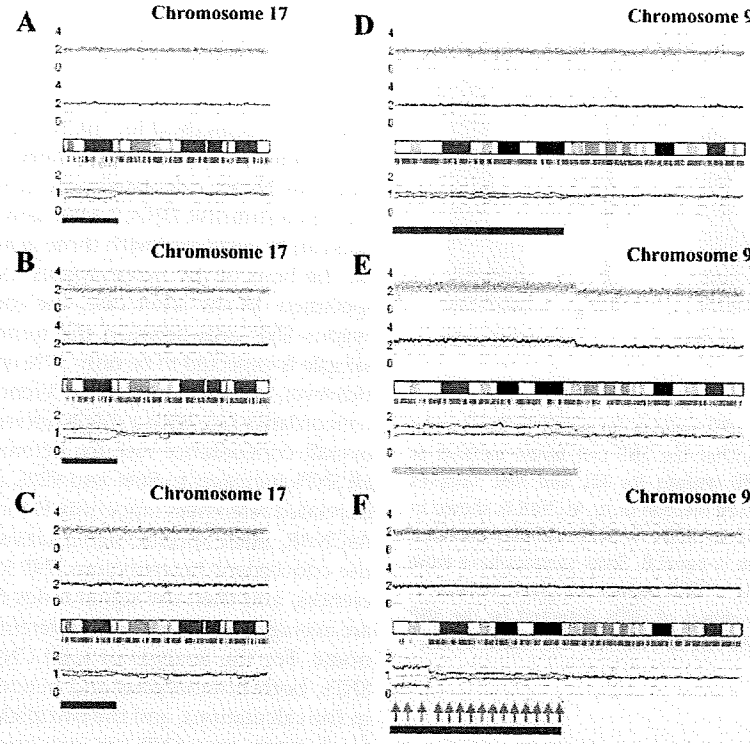

Figure 4. Detection of AI in samples of primary AML and MPD. AsCN analyses disclosed the presence of a small population with 17p UPD in a primary AML specimen (W150673) (93% blasts in microscopic examination) with either a paired sample (*A*) or anonymous reference samples (*B*). The difference of the mean CNs of the two parental alleles is statistically different between panels A (0.38) and B (0.55) ($P < .0001$, by t test), which is explained by the residual tumor component within the bone marrow sample in complete remission (1% blast) used as a paired reference (W150673CR) (*C*). AI in the 9p arm was also sensitively detected in *JAK2* mutation–positive MPD cases. UPD may be carried only by a very small population (~20% estimated from the mean deviation of AsCNs in 9p) (IMF_10) (*D*), or by two discrete populations within the same case (PV_06), as indicated by two-phased dissociation of AsCN graphs (*pink and green arrows*) (*F*). AI in 9p is mainly caused by UPD but may be caused by gains of one parental allele without loss of the other allele (*E*), both of which are not discriminated by conventional allele measurements. Blue and pink bars are UPD and AI calls, respectively, from the HMM-based LOH detection algorithm. Other features are identical to those indicated in figure 1.

In fact, this algorithm precisely identifies known LOH regions, as well as regions with AI, in intentionally mixed tumor samples containing as little as 20% (for LOH without CN loss) to 25% (LOH with CN loss) tumor contents (fig. 2*A*–2*C*). Note that this large gain in sensitivity is obtained without the expense of specificity, which is very close to 100%, as observed with other algorithms (fig. 2*D*). In AsCNAR, small regions of AI (<1 million bases in length) are difficult to detect in samples contaminated with normal cells. However, such regions are also difficult to detect using other algorithms (data not shown).

Identification of UPD in Primary Tumor Samples

To examine further the strength of the newly developed algorithms for AsCN and LOH detection, we explored UPD regions in 85 primary acute leukemia samples, including 39 AML and 46 ALL samples, on GeneChip 50K Xba SNP arrays, since recent reports identified frequent (~20%) occurrence of this abnormality in AML.[23,24] In the SNP call–based LOH inference algorithm, 16 UPD regions were identified in 14 cases, 8 (20.5%) AML and 6 (13.0%) ALL. However, the frequencies were almost doubled with the AsCNAR algorithm; a total of 28 UPD loci were identified in 25 cases, including 14 (35.9%) AML and 11 (23.9%) ALL (fig. 3*A* and table 1). In 5 of the 25 UPD-positive cases, a matched remission sample was available for AsCN analysis, which provided essentially the same results as AsCNAR, except for one relapsed AML case (W150673). In the latter case, a discrepancy in AsCN shifts in 17p UPD occurred between AsCN analysis with and without a constitutive reference, with more CN shift detected with anonymous references (fig. 4*A* and 4*B*). The discrepancy was, however, explained by the unexpected detection of a subtle UPD change in 17p in the reference sample by

Table 2. AI of 9p in *JAK2* Mutation-Positive MPDs

Case	9p Status by AsCNAR			Detection by SNP Call–Based Method[a]	% *JAK2* Mutation[b]	Allele-Specific PCR[c]		
	Type	Break Point[d]	%UPD[e]			SNP	%UPD[f]	*P*[g]
PV_02	Gain	42.9	99	NA	63	*rs2009991*	84	.004
PV_03	Gain	Whole	60	NA	39	*rs10511431*	63	.008
PV_04	UPD	37.0	93	D	95	5Homo	5Homo	5Homo
PV_08	UPD	34.2	91	D	93	5Homo	5Homo	5Homo
PV_07	UPD	23.8	88	D	90	5Homo	5Homo	5Homo
PV_06	UPD[h]	7.1/35.3	83	D	93	5Homo	5Homo	5Homo
PV_11	UPD	31.2	68	D	76	5Homo	5Homo	5Homo
PV_13	UPD	28.1	66	ND	48	*rs1416582*	64	.001
PV_01	UPD	20.9	56	ND	62	*rs10511431*	49	.007
PV_09	UPD	30.8	38	ND	30	*rs10491558*	32	.020
PV_05	UPD	23.5	32	ND	33	*rs1374172*	31	.010
IMF_04	UPD	33.8	79	D	90	5Homo	5Homo	5Homo
IMF_05	UPD	37.0	58	ND	57	*rs1416582*	49	.004
IMF_07	UPD	20.3	52	ND	50	*rs1416582*	57	.005
IMF_12	UPD[h]	26.8/42.9	52	ND	66	5Homo	5Homo	5Homo
IMF_14	UPD[h]	22.8/33.8	45	ND	56	*rs1374172*	35	.015
IMF_19	UPD	34.4	26	ND	43	*rs10511431*	33	.017
IMF_10	UPD	34.6	21	ND	36	*rs1374172*	21	.049
IMF_15	UPD	33.8	21	ND	17	*rs10511431*	20	.084
IMF_06	UPD	35.3	17	ND	28	*rs1374172*	20	.048
IMF_16	(−)	NA	NA	NA	37	NA	NA	NA
ET_12	Gain	Whole	42	NA	27	*rs2009991*	36	.046
ET_14	UPD	42.9	63	ND	45	*rs1374172*	54	.006
ET_01	UPD	35.4	19	ND	59	*rs10511431*	33	.017
ET_05	(−)	NA	NA	NA	23	NA	NA	NA
ET_08	(−)	NA	NA	NA	42	NA	NA	NA
ET_09	(−)	NA	NA	NA	34	NA	NA	NA
ET_10	(−)	NA	NA	NA	16	NA	NA	NA
ET_15	(−)	NA	NA	NA	27	NA	NA	NA
ET_18	(−)	NA	NA	NA	17	NA	NA	NA
ET_19	(−)	NA	NA	NA	27	NA	NA	NA
ET_21	(−)	NA	NA	NA	55	NA	NA	NA

NOTE.—NA = not applied; (−) = neither UPD nor gain of 9p was detected by AsCNAR analysis.

[a] D = UPD was detected by SNP call–based method; ND = not detected.

[b] Percentage of *JAK2* mutant alleles, as measured by allele-specific PCR.

[c] 5Homo = all five tested SNPs were homozygous.

[d] Position of the break point from the p-telomeric end (values are in Mb). The location of *JAK2* corresponds to 5 Mb.

[e] Percentage of tumor cell populations with either UPD or gain of 9p, as determined by AsCNAR analysis.

[f] Percentage of tumor cell populations with either UPD or gain of 9p, as determined by the allele-specific PCR.

[g] *P* values were derived from one-tailed *t* tests comparing triplicate analyses of the target sample and triplicate analyses of five normal samples.

[h] Two UPD-positive populations exist.

AsCNAR ($P < .0001$, by *t* test) (fig. 4*C*), which offset the CN shift in the relapsed sample, although it was morphologically and cytogenetically diagnosed as in complete remission.

Analysis of 9p UPD in MPDs

Another interesting application of the AsCNAR is the analysis of allelic status in the 9p arm among patients with MPD, which includes PV, ET, and IMF. According to past reports, ~10% (in ET) to ~40% (in PV) of MPD cases with the activating *JAK2* mutation (V617F) show evidence of clonal evolution of dominant progeny that carry the homozygous *JAK2* mutation caused by 9p UPD.[5,7,8] In our series that included 53 MPD cases, the *JAK2* mutation was detected in 32 (60%), of which 13 (41%) showed >50% mutant allele by allele measurement with the use of allele-specific PCR, and thus were judged to have one or more populations carrying homozygous *JAK2* mutations (table 2). This frequency is comparable to that reported elsewhere.[8] However, when the same specimens were analyzed with 50K Xba SNP arrays by use of the AsCNAR algorithm, 20 of the 32 *JAK2* mutation–positive cases were demonstrated to have minor UPD subpopulations (table 2 and fig. 3*B*), in which as little as 17% of UPD-positive populations were sensitively detected (fig. 4*D*). In fact, these minor (<50%) UPD-positive populations in these

cases were also confirmed by allele-specific PCR of SNPs on 9p (table 2). The proportion of 9p UPD–positive components estimated both from allele-specific PCR and from AsCNAR (see the "Material and Methods" section) shows a good concordance (table 2). In some cases, 9p UPD–positive cells account for almost all the *JAK2* mutation–positive population, whereas, in others, they represent only a small subpopulation of the entire *JAK2* mutation–positive population (fig. 5). AsCNAR analysis also disclosed the additional three cases that have 9p gain (9p trisomy) (fig. 4*E*). The 9p trisomy is among the most-frequent cytogenetic abnormalities in MPDs[25] and is implicated in duplication of the mutated *JAK2* allele[6] but could not have been discriminated from UPD or "LOH with CN loss" by use of conventional techniques—for example, allele-specific PCR to measure relative allele dose. Since the proportions of the mutated *JAK2* allele coincide with two-thirds of the observed trisomy components in all three cases, the data suggest that the mutated *JAK2* allele is duplicated in the 9p trisomy cases (table 2). Of particular interest is the unexpected finding of the presence of two discrete populations carrying 9p UPD in three cases, in which the AsCN graph showed a two-phased dissociation along the 9p arm (fig. 4*F*). In the previous observations, homozygous *JAK2* mutations have been reported to be more common in PV cases (~40%) than in ET cases (<~10%). With AsCNAR analysis, the difference in the frequency of 9p UPD becomes more conspicuous; nearly all PV cases (11/11) and IMF cases (9/10) with a *JAK2* mutation had one or more UPD components or other gains of 9p material, whereas only 3 of the 11 *JAK2* mutation-positive ET cases carried a 9p UPD component or gain of 9p ($P = 1.3 \times 10^{-4}$, by Fisher's exact test).

Discussion

The robustness of the AsCNAR method lies in its capacity to measure accurately allele dosage and thereby to detect LOH even in the presence of significant normal cell components, which often occurs in primary tumor samples. In principle, an accurate LOH determination is accomplished only by demonstrating an absolute loss of one parental allele, not simply by detecting AI with conventional allele-measurement techniques. This is especially the case for contaminated samples, where it is essentially impossible to discriminate the origin of the remaining minor-allele component (i.e., differentiating normal cells and tumor cells).[1,3] Nevertheless, and paradoxically, it is these normal cells within the tumor samples that enable determination of AsCNs in AsCNAR. It computes AsCNs on the basis of the strength of heterozygous SNP calls produced from the "contaminated" normal component, which effectively works as "an internal reference," precluding the need for preparing a paired germline reference.

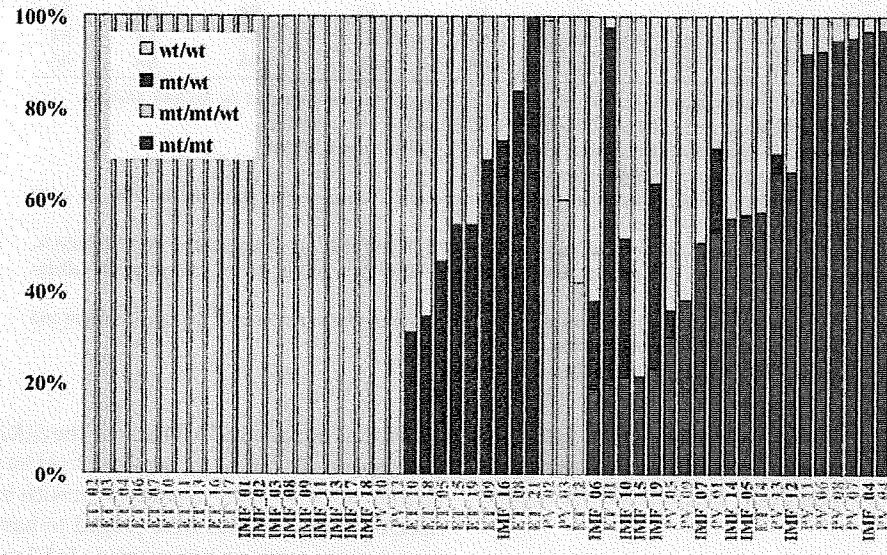

Figure 5. Estimation of tumor populations carrying 9p UPD and the *JAK2* mutation in MPD samples. The populations of 9p UPD–positive components in the 53 MPD cases were estimated by calculation of the mean difference of AsCNs within the UPD regions. Heterozygous (*blue bars*) or homozygous (*red bars*) *JAK2* mutations in MPD samples were also estimated by measurement of *JAK2* mutated alleles and UPD alleles, under the assumption that all the UPD alleles have a *JAK2* mutation. Measurement of *JAK2* mutated alleles was performed by allele-specific PCR. For three cases having trisomy components (*orange bars*), the duplicated allele was assumed to have a *JAK2* mutation, which is the consistent interpretation of the observed fraction of trisomy and mutated *JAK2* alleles for case PV_02 (table 2). mt = *JAK2* mutated allele; wt = wild-type allele.

The figure is available in its entirety in the online edition of *The American Journal of Human Genetics.*

Figure 6. Effects of the use of the different reference sets on signal-to-noise (S/N) ratios in CN analysis. The legend is available in its entirety in the online edition of *The American Journal of Human Genetics.*

It far outperforms the SNP call–based LOH-inference algorithms and other methods and definitively determines the state of LOH by sensing CN loss of one parental allele.

In the previously published algorithms, AsCN analysis was enabled by fitting observed array data to a model constructed from a fixed data set from normal samples.[18,21] However, the model that explicitly assumes integer CNs fails to cope with primary tumor samples that contain varying degrees of normal cell components (PLASQ)[18] (fig. 2). Another algorithm (CARAT) requires a large number of references to construct a model by which AsCNs are predicted, but such a model may not necessarily be properly applied to predict AsCNs for the newly processed samples, if the experimental condition for those samples is significantly different from that for the reference samples, which were used to construct the model (fig. 6 and data not shown).[21] Signal ratios between array data from very different experiments could be strongly biased, to the extent that they can no more be properly compensated by conventional regressions. In contrast, AsCNAR uses just a small number of references simultaneously processed with tumor specimens, to minimize difference in experimental conditions between tumor and references, which act as excellent controls in calculating AsCNs, although references analyzed in short intervals also work satisfactorily (data not shown).

The CN analysis software for the Illumina array provides allele frequencies, as well as CNs, by use of a model-based approach, and, as such, it enables AsCN analysis but seems to be less sensitive for detection of AIs.[26] AsCNAR can be easily adapted to other Affymetrix arrays, including 10K and 500K arrays, and may be potentially applied to Illumina arrays.

The probability of finding at least one concordant SNP between a tumor sample and a set of anonymous references is enough with five references, but use of just one reference provides almost an equivalent AsCN profile to that obtained with its paired reference (fig. 7). The sensitivity and specificity of LOH detection with this algorithm are excellent, even in the presence of significant degrees of normal cell components (~70%–80%), which circumvent the need for purifying the tumor components for analysis—for example, by time-consuming microdissection.

Because the AsCNAR algorithm is quite simple, it requires much less computing power and time (several seconds per sample on average laptop computers) than do model-based algorithms. For example, with PLASQ, it takes overnight for model construction and an additional hour for processing each sample.

The high sensitivity of LOH detection by AsCNAR has been validated not only by the analysis of tumor DNA intentionally mixed with normal DNA but also by the analysis of primary leukemia samples. It unveiled otherwise undetected, minor UPD-positive populations within leukemia samples. Especially, the extremely high frequency of 9p UPD or gains of 9p in particular types of *JAK2* mutation–positive MPDs, as well as multiple UPD-positive subclones in some cases, demonstrated how strongly and efficiently a genetic change (point mutation) works to fix the next alteration (mitotic recombination) in the tumor population during clonal evolution in human cancer. Finally, the conspicuous difference in UPD frequency among different MPD subtypes (PV and IMF vs. ET) is noteworthy. This is supported by a recent report that demonstrated the presence of minor subclones carrying exclusively the mutated *JAK2* allele in all PV samples, but in none of the ET samples, by examining a large number of erythroid burst-forming units and Epo-independent erythroid colonies for *JAK2* mutation.[27] Our observation also supports their hypothesis that the biological behavior of these prototypic stem-cell disorders with a continuous disease spectrum could be determined by the components with either homozygous or duplicated *JAK2* mutations.

In conclusion, the AsCNAR with use of high-density oligonucleotide microarrays is a robust method of genomewide analysis of allelic changes in cancer genomes and provides an invaluable clue to the understanding of the genetic basis of human cancers. The AsCNAR algorithm is freely available on our CNAG Web site for academic users.

The figure is available in its entirety in the online edition of *The American Journal of Human Genetics.*

Figure 7. CN profile obtained with the use of a varying number of anonymous references. The legend is available in its entirety in the online edition of *The American Journal of Human Genetics.*

Acknowledgments

This work was supported by Research on Measures for Intractable Diseases, Health and Labor Sciences Research Grants, Ministry of Health, Labor and Welfare, by Research on Health Sciences focusing on Drug Innovation, by the Japan Health Sciences Foundation, by Core Research for Evolutional Science and Technology, Japan Science and Technology Agency, and by Japan Leukemia Research Fund.

Appendix A

AsCNAR

Quadratic Regression

The $\log_2$ signal-ratio, $\log_2 R^{\mathrm{ref}I}_{AB,i}$ is regressed by the quadratic terms (the length $[L_i]$ and the GC content $[M_i]$ of the PCR fragment of the ith SNP) as

$$\log_2 R^{\mathrm{ref}I}_{AB,i} = \alpha L_i^2 + \beta L_i + \chi M_i^2 + \delta M_i + \gamma + \varepsilon_i \ ,$$

where ε_i is the error term and the coefficients of regressions α, β, χ, δ, and γ are dependent on the reference used and are determined to minimize the residual sum of squares (i.e., $\Sigma_i \varepsilon_i^2$). Note that the sum is taken for those SNPs that have concordant SNP calls between the tumor and the reference samples.

We suppose that both allele A DNA and allele B DNA follow the same PCR kinetics, and allele-specific ratios $R^{\mathrm{ref}I}_{A,i}$ and $R^{\mathrm{ref}I}_{B,i}$, respectively, can be regressed by the same parameters, as

$$\log_2 \hat{R}^{\mathrm{ref}I}_{A,i} = \log_2 R^{\mathrm{ref}I}_{A,i} - (\alpha L_i^2 + \beta L_i) - (\chi M_i^2 + \delta M_i) - \gamma$$

and

$$\log_2 \hat{R}^{\mathrm{ref}I}_{B,i} = \log_2 R^{\mathrm{ref}I}_{B,i} - (\alpha L_i^2 + \beta L_i) - (\chi M_i^2 + \delta M_i) - \gamma \ ,$$

and the corrected total CN ratio is

$$\hat{R}^{\mathrm{ref}I}_{AB,i} = \begin{cases} \hat{R}^{\mathrm{ref}I}_{A,i} & \text{for } O_i^{\mathrm{tum}} = O_i^{\mathrm{ref}I} = AA \\ \hat{R}^{\mathrm{ref}I}_{B,i} & \text{for } O_i^{\mathrm{tum}} = O_i^{\mathrm{ref}I} = BB \\ \frac{1}{2}\left(\hat{R}^{\mathrm{ref}I}_{A,i} + \hat{R}^{\mathrm{ref}I}_{B,i}\right) & \text{for } O_i^{\mathrm{tum}} = O_i^{\mathrm{ref}I} = AB \end{cases} .$$

Averaging over the References of Concordance SNPs

Concordant reference sets C_i^K and $C_i^{K,\mathrm{hetero}}$ for each SNP S_i for a given set of references, K, are defined as follows:

$$C_i^K = \{\mathrm{ref}I \,|\, O_i^{\mathrm{tum}} = O_i^{\mathrm{ref}I}, \mathrm{ref}I \in K\}$$

$$C_i^{K,\mathrm{hetero}} = \{\mathrm{ref}I \,|\, O_i^{\mathrm{tum}} = O_i^{\mathrm{ref}I} = AB, \mathrm{ref}I \in K\} \ ,$$

and the averaged CN ratio, $\tilde{R}^K_{AB,i}$, is provided by

$$\tilde{R}^K_{AB,i} = \frac{1}{\#C_i^K} \sum_{\mathrm{ref}I \in C_i^K} \hat{R}^{\mathrm{ref}I}_{AB,i} \ , \ C_i^K \neq \phi$$

where "#" denotes the number of the elements of the set. Similarly, AsCN ratios are obtained by

$$\tilde{R}^K_{A,i} = \frac{1}{\#C_i^{K,\mathrm{hetero}}} \sum_{\mathrm{ref}I \in C_i^{K,\mathrm{hetero}}} \hat{R}^{\mathrm{ref}I}_{A,i} \qquad (C_i^{K,\mathrm{hetero}} \neq \phi) \ .$$

$$\tilde{R}^K_{B,i} = \frac{1}{\#C_i^{K,\mathrm{hetero}}} \sum_{\mathrm{ref}I \in C_i^{K,\mathrm{hetero}}} \hat{R}^{\mathrm{ref}I}_{B,i}$$

Exceptional Handling with Regions of Homozygous Deletion, High Amplification, and LOH

To prevent SNPs within the regions that show homozygous deletion or high-grade amplification from being analyzed as "homozygous SNPs," a homozygous SNP S_i in the tumor sample is redefined as a heterozygous SNP with $O_i^{\mathrm{tum}} = AB$, if $\max(\log_2 \tilde{R}^K_{A,i}, \log_2 \tilde{R}^K_{B,i}) \leq 0.1$ or $\min(\log_2 \tilde{R}^K_{A,i}, \log_2 \tilde{R}^K_{B,i}) \geq -0.1$, where $\tilde{R}^K_{A,i}$ and $\tilde{R}^K_{B,i}$ are calculated supposing SNP S_i is heterozygous. These cutoff values (0.1 and -0.1) are determined by receiver operating characteristic (ROC) curve for detection of gain of the larger allele and loss of the smaller allele in a sample containing 20% tumor cells (data not shown). In addition, SNPs within inferred LOH regions are also analyzed as "heterozygous" SNPs.

Reference Selection

The optimized set of references is selected that minimizes the SD of total CN at the diploid region D,

$$\mathrm{SD}_K(D) = \sqrt{\frac{\sum_{i \in D, C_i^K \neq \phi} (\log_2 \tilde{R}^K_{AB,i})^2}{\#\{i \,|\, i \in D, C_i^K \neq \phi\} - 1}} \ .$$

To do this, instead of testing all possible 2^N combinations of N references, we calculate $\mathrm{SD}_K(D)$ for individual references $K = \{\mathrm{ref1}\}, \{\mathrm{ref2}\}, \{\mathrm{ref3}\}, \ldots, \{\mathrm{ref}N\}$, to order the references such that $\mathrm{SD}_1(D) \leq \ldots \leq \mathrm{SD}_s(D) \leq \mathrm{SD}_{s+1}(D) \leq \ldots \leq \mathrm{SD}_N(D)$, where 1, 2, 3,...,$s$, $s+1$,..., N denotes the ordered references. The optimal set $K(N_0) = \{1,2,3,\ldots,N_0\}$ is determined by choosing N_0 that satisfies $\mathrm{SD}_{K(1)}(D) \geq \ldots \geq \mathrm{SD}_{K(N_0)}(D) < \mathrm{SD}_{K(N_0+1)}(D)$.

Note that, in principle, a diploid region cannot be unequivocally determined without doing single-cell–based analysis—for example, FISH or cytogenetics. Otherwise, a diploid region is empirically determined by setting the CN-minimal regions with no AI as diploid, which provides correct estimation of the ploidy in most cases (data not shown).

The figure is available in its entirety in the online edition of *The American Journal of Human Genetics.*

Figure C1. Inference of LOH on the basis of heterozygous SNP calls. The legend is available in its entirety in the online edition of *The American Journal of Human Genetics.*

Appendix C
Inference of LOH Based on Heterozygous SNP Calls

For a given contiguous region $\Omega_{i,j}$ between the ith and jth SNPs ($i \leq j$) and for the complete set of observed SNP calls therein, $O(\Omega_{i,j})$, consider the log likelihood ratio

$$Z(\Omega_{i,j}) \equiv \ln \frac{P(O(\Omega_{i,j}) | \Omega_{i,j} \in \mathrm{LOH})}{P(O(\Omega_{i,j}) | \Omega_{i,j} \notin \mathrm{LOH})},$$

where the ratio is taken between the conditional probabilities that the current observation, $O(\Omega_{i,j})$, is obtained under the assumption that $O(\Omega_{i,j})$ belongs to LOH or not. We assume a constant miscall rate ($q = 0.001$) for all SNP and use the conditional probability that the kth SNP is heterozygous (h_k), depending on the observed $k-1$th SNP call, for partially taking the effect of linkage disequilibrium into account:

$$Z(\Omega_{i,j}) = \ln \frac{\prod_{i \leq k \leq j} \{(1-q)O_k + q(1-O_k)\}}{\prod_{i \leq k \leq j} [\{(1-h_k)(1-q) + h_k q\}O_k + \{(1-h_k)q + h_k(1-q)\}(1-O_k)]}$$

where h_k is calculated using the data from the 96 normal Japanese individuals, whereas O_k takes either 1 or 0, depending on the kth SNP call, with 1 for a homozygous call and 0 for a heterozygous call. For each chromosome, a set of regions, $\Omega_{I_n,J_n}(J_{n-1} < I_n \leq J_n, J_0 = 0)$ ($n = 1,2,3,\ldots$), can be uniquely determined as follows.

Beginning with the SNP at the short arm end (S_0), find the SNP S_{I_n} that satisfies $Z(\Omega_{I_n,J_n}) > 0$ and $Z(\Omega_{i,i}) \leq 0$ for $J_{n-1} < \forall i < I_n$ (fig. C1). Identify the SNP S_{J^+}, such that $Z(\Omega_{I_n,j}) > 0$ for $I_n \leq \forall j \leq J^+$ and $Z(\Omega_{I_n,J^+1}) \leq 0$, or that S_{J^+} is the end of the chromosome (fig. C1). Then, put J_n as $\arg\max_j Z(\Omega_{I_n,j})(I_n \leq j \leq J^+)$ (fig. C1). This procedure is iteratively performed, beginning the next iteration with the SNP S_{I_n+1}, until it reaches to the end of the long arm, generating a set of nonoverlapping regions, $\Omega_{I_1,J_1}, \Omega_{I_2,J_2}, \Omega_{I_3,J_3}, \ldots \Omega_{I_n,J_n}, \ldots$. LOH inference is now enabled by testing each $Z(\Omega_{I_n,J_n})$ against a threshold (25), which is arbitrarily determined from the ROC curve for LOH determination on a DNA sample from a lung cancer cell line, NCI-H2171 (fig. C1). This algorithm is implemented in our CNAG program, which is available at our Web site.

Appendix E
Algorithm for Detection of AI With or Without LOH

The regions with AI are inferred from the AsCN data by use of an HMM, where the real state of AI (a hidden state) is inferred from the observed states of difference in AsCNs of the two parental alleles, which are expressed as dichotomous values ("preset" or "absent") according to a threshold (μ). The emission probabilities at the ith SNP locus (Si) are

$$P(|\log_2 \tilde{R}^K_{A,i} - \log_2 \tilde{R}^K_{B,i}| \leq \mu | Si \in \mathrm{AI}) = \beta$$

$$P(|\log_2 \tilde{R}^K_{A,i} - \log_2 \tilde{R}^K_{B,i}| > \mu | Si \in \mathrm{AI}) = 1 - \beta$$

and

$$P(|\log_2 \tilde{R}^K_{A,i} - \log_2 \tilde{R}^K_{B,i}| > \mu | Si \in \overline{\mathrm{AI}}) = \alpha$$

$$P(|\log_2 \tilde{R}^K_{A,i} - \log_2 \tilde{R}^K_{B,i}| \leq \mu | Si \in \overline{\mathrm{AI}}) = 1 - \alpha$$

(see also the "Material and Methods" section and appendix A for calculation of $\tilde{R}^K_{A,i}$ and $\tilde{R}^K_{B,i}$).

The parameters (μ, α, and β) are determined by the results of 10%, 20%, and 30% tumor samples. Sensitivity and specificity are calculated with varying threshold (μ), where sensitivity is defined as the ratio of detected SNPs of UPD region detected in the 100% tumor sample, specificity is defined as the ratio of nondetected SNPs in normal samples, and α and β parameters are determined from mixed tumor-sample data for each threshold value. Sensitivity and specificity are relatively stable and are within the acceptable range when the threshold is between 0.05 and 0.15 in 20% and 30% tumor samples (fig. E1). We used 0.12, 0.17, and 0.06 for μ, α, and β, respectively, on the basis of 20% tumor-sample data.

Considering that UPD is caused by a process similar to recombination, the Kosambi's map function $(1/2)\tanh(2\theta)$ is used for transition probability, where θ is the distance between two SNPs, expressed in cM units; for simplicity, 1 cM should be 1 Mbp. Thus, the most likely underlying, hidden, real states of AI are calculated for each SNP according to Vitervi's method, by which AI-positive regions are defined by contiguous SNPs with "present" AI calls flanked by either chromosomal end or an "absent" AI call. Next, to determine the LOH status for each AI-positive region (Γ), AsCN states at each SNP locus within Γ are

The figure is available in its entirety in the online edition of *The American Journal of Human Genetics.*

Figure E1. Sensitivity and specificity for determination of AI, LOH, and UPD. The legend is available in its entirety in the online edition of *The American Journal of Human Genetics.*

inferred as "reduced (R)" and "not reduced ($\bar{R}$)" for the smaller AsCNs, and "increased (I)" and "not increased ($\bar{I}$)" for the larger AsCNs, using similar HMMs from the "observed CN states" of the smaller and the larger AsCNs, which are expressed as dichotomous values according to thresholds μ_S and μ_L, respectively. The emission probabilities of these models are

$$P[\min(\log_2 \tilde{R}^K_{A,i}, \log_2 \tilde{R}^K_{B,i}) < \mu_S | Si \in \mathrm{R}] = 1 - \beta_S$$

$$P[\min(\log_2 \tilde{R}^K_{A,i}, \log_2 \tilde{R}^K_{B,i}) \geqslant \mu_S | Si \in \mathrm{R}] = \beta_S$$

$$P[\min(\log_2 \tilde{R}^K_{A,i}, \log_2 \tilde{R}^K_{B,i}) < \mu_S | Si \in \bar{\mathrm{R}}] = \alpha_S$$

$$P[\min(\log_2 \tilde{R}^K_{A,i}, \log_2 \tilde{R}^K_{B,i}) \geqslant \mu_S | Si \in \bar{\mathrm{R}}] = 1 - \alpha_S$$

and

$$P[\max(\log_2 \tilde{R}^K_{A,i}, \log_2 \tilde{R}^K_{B,i}) > \mu_L | Si \in \mathrm{I}] = 1 - \beta_L$$

$$P[\max(\log_2 \tilde{R}^K_{A,i}, \log_2 \tilde{R}^K_{B,i}) \leqslant \mu_L | Si \in \mathrm{I}] = \beta_L$$

$$P[\max(\log_2 \tilde{R}^K_{A,i}, \log_2 \tilde{R}^K_{B,i}) > \mu_L | Si \in \bar{\mathrm{I}}] = \alpha_L$$

$$P[\max(\log_2 \tilde{R}^K_{A,i}, \log_2 \tilde{R}^K_{B,i}) \leqslant \mu_L | Si \in \bar{\mathrm{I}}] = 1 - \alpha_L \ .$$

These parameters (μ_S, α_S, β_S, μ_L, α_L, and β_L) are determined by evaluating sensitivities and specificities of the results for 10%, 20%, and 30% tumor samples, where sensitivities and specificities are calculated the same way as was AI. Sensitivity and specificity are relatively stable for μ_S between -0.03 and -0.13 and are relatively stable for μ_L between 0.04 and 0.09 in 20% and 30% tumor samples (fig. E1). We employed $\mu_S = -0.1$, $\alpha_S = 0.3$, $\beta_S = 0.26$, $\mu_L = 0.08$, $\alpha_L = 0.27$, and $\beta_L = 0.31$ on the basis of the data for 20% tumor content.

Web Resources

The URLs for data presented herein are as follows:

ATCC, http://www.atcc.org/common/cultures/NavByApp.cfm
BACPAC Resources Center, http://bacpac.chori.org/
CNAG, http://www.genome.umin.jp/
dChip, http://www.dchip.org/
Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for *JAK2*, AML, PV, ET, and IMF)
PLASQ, http://genome.dfci.harvard.edu/~tlaframb/PLASQ/

References

1. Mei R, Galipeau PC, Prass C, Berno A, Ghandour G, Patil N, Wolff RK, Chee MS, Reid BJ, Lockhart DJ (2000) Genome-wide detection of allelic imbalance using human SNPs and high-density DNA arrays. Genome Res 10:1126–1137
2. Horvath A, Boikos S, Giatzakis C, Robinson-White A, Groussin L, Griffin KJ, Stein E, Levine E, Delimpasi G, Hsiao HP, et al (2006) A genome-wide scan identifies mutations in the gene encoding phosphodiesterase 11A4 (*PDE11A*) in individuals with adrenocortical hyperplasia. Nat Genet 38:794–800
3. Lindblad-Toh K, Tanenbaum DM, Daly MJ, Winchester E, Lui WO, Villapakkam A, Stanton SE, Larsson C, Hudson TJ, Johnson BE, et al (2000) Loss-of-heterozygosity analysis of small-cell lung carcinomas using single-nucleotide polymorphism arrays. Nat Biotechnol 18:1001–1005
4. Knudson AG (2001) Two genetic hits (more or less) to cancer. Nat Rev Cancer 1:157–162
5. Baxter EJ, Scott LM, Campbell PJ, East C, Fourouclas N, Swanton S, Vassiliou GS, Bench AJ, Boyd EM, Curtin N, et al (2005) Acquired mutation of the tyrosine kinase JAK2 in human myeloproliferative disorders. Lancet 365:1054–1061
6. James C, Ugo V, Le Couedic JP, Staerk J, Delhommeau F, Lacout C, Garcon L, Raslova H, Berger R, Bennaceur-Griscelli A, et al (2005) A unique clonal JAK2 mutation leading to constitutive signalling causes polycythaemia vera. Nature 434: 1144–1148
7. Kralovics R, Passamonti F, Buser AS, Teo SS, Tiedt R, Passweg JR, Tichelli A, Cazzola M, Skoda RC (2005) A gain-of-function mutation of JAK2 in myeloproliferative disorders. N Engl J Med 352:1779–1790
8. Levine RL, Wadleigh M, Cools J, Ebert BL, Wernig G, Huntly BJ, Boggon TJ, Wlodarska I, Clark JJ, Moore S, et al (2005) Activating mutation in the tyrosine kinase JAK2 in polycythemia vera, essential thrombocythemia, and myeloid metaplasia with myelofibrosis. Cancer Cell 7:387–397
9. Kennedy GC, Matsuzaki H, Dong S, Liu WM, Huang J, Liu G, Su X, Cao M, Chen W, Zhang J, et al (2003) Large-scale genotyping of complex DNA. Nat Biotechnol 21:1233–1237
10. Zhao X, Li C, Paez JG, Chin K, Janne PA, Chen TH, Girard L, Minna J, Christiani D, Leo C, et al (2004) An integrated view of copy number and allelic alterations in the cancer genome using single nucleotide polymorphism arrays. Cancer Res 64:3060–3071
11. Huang J, Wei W, Zhang J, Liu G, Bignell GR, Stratton MR, Futreal PA, Wooster R, Jones KW, Shapero MH (2004) Whole genome DNA copy number changes identified by high density oligonucleotide arrays. Hum Genomics 1:287–299
12. Bignell GR, Huang J, Greshock J, Watt S, Butler A, West S, Grigorova M, Jones KW, Wei W, Stratton MR, et al (2004) High-resolution analysis of DNA copy number using oligonucleotide microarrays. Genome Res 14:287–295
13. Wang ZC, Buraimoh A, Iglehart JD,Richardson AL (2006) Genome-wide analysis for loss of heterozygosity in primary and recurrent phyllodes tumor and fibroadenoma of breast using single nucleotide polymorphism arrays. Breast Cancer Res Treat 97:301–309
14. Zhou X, Mok SC, Chen Z, Li Y, Wong DT (2004) Concurrent analysis of loss of heterozygosity (LOH) and copy number abnormality (CNA) for oral premalignancy progression using the Affymetrix 10K SNP mapping array. Hum Genet 115:327–330
15. Matsuzaki H, Dong S, Loi H, Di X, Liu G, Hubbell E, Law J, Berntsen T, Chadha M, Hui H, et al (2004) Genotyping over 100,000 SNPs on a pair of oligonucleotide arrays. Nat Methods 1:109–111
16. Nannya Y, Sanada M, Nakazaki K, Hosoya N, Wang L, Hangaishi A, Kurokawa M, Chiba S, Bailey DK, Kennedy GC, et al (2005) A robust algorithm for copy number detection using high-density oligonucleotide single nucleotide polymorphism genotyping arrays. Cancer Res 65:6071–6079
17. Beroukhim R, Lin M, Park Y, Hao K, Zhao X, Garraway LA, Fox EA, Hochberg EP, Mellinghoff IK, Hofer MD, et al (2006) Inferring loss-of-heterozygosity from unpaired tumors using

high-density oligonucleotide SNP arrays. PLoS Comput Biol 2:e41
18. Laframboise T, Harrington D, Weir BA (2007) PLASQ: a generalized linear model-based procedure to determine allelic dosage in cancer cells from SNP array data. Biostatistics 8: 323–336
19. Kralovics R, Teo SS, Li S, Theocharides A, Buser AS, Tichelli A, Skoda RC (2006) Acquisition of the V617F mutation of JAK2 is a late genetic event in a subset of patients with myeloproliferative disorders. Blood 108:1377–1380
20. Wang L, Ogawa S, Hangaishi A, Qiao Y, Hosoya N, Nanya Y, Ohyashiki K, Mizoguchi H, Hirai H (2003) Molecular characterization of the recurrent unbalanced translocation der(1;7)(q10;p10). Blood 102:2597–2604
21. Huang J, Wei W, Chen J, Zhang J, Liu G, Di X, Mei R, Ishikawa S, Aburatani H, Jones KW, et al (2006) CARAT: a novel method for allelic detection of DNA copy number changes using high density oligonucleotide arrays. BMC Bioinformatics 7:83
22. Dugad R, Desai U (1996) A tutorial on hidden Markov models. Technical report SPANN-96.1. Signal Processing and Artificial Neural Networks Laboratory, Bombay, India
23. Raghavan M, Lillington DM, Skoulakis S, Debernardi S, Chaplin T, Foot NJ, Lister TA, Young BD (2005) Genome-wide single nucleotide polymorphism analysis reveals frequent partial uniparental disomy due to somatic recombination in acute myeloid leukemias. Cancer Res 65:375–378
24. Fitzgibbon J, Smith LL, Raghavan M, Smith ML, Debernardi S, Skoulakis S, Lillington D, Lister TA, Young BD (2005) Association between acquired uniparental disomy and homozygous gene mutation in acute myeloid leukemias. Cancer Res 65:9152–9154
25. Najfeld V, Montella L, Scalise A, Fruchtman S (2002) Exploring polycythaemia vera with fluorescence in situ hybridization: additional cryptic 9p is the most frequent abnormality detected. Br J Haematol 119:558–566
26. Peiffer DA, Le JM, Steemers FJ, Chang W, Jenniges T, Garcia F, Haden K, Li J, Shaw CA, Belmont J, et al (2006) High-resolution genomic profiling of chromosomal aberrations using Infinium whole-genome genotyping. Genome Res 16: 1136–1148
27. Scott LM, Scott MA, Campbell PJ, Green AR (2006) Progenitors homozygous for the V617F mutation occur in most patients with polycythemia vera, but not essential thrombocythemia. Blood 108:2435–2437

Leukemia (2007) 21, 992–997

www.nature.com/leu

ORIGINAL ARTICLE

Unbalanced translocation der(1;7)(q10;p10) defines a unique clinicopathological subgroup of myeloid neoplasms

M Sanada[1,7], N Uike[2], K Ohyashiki[3], K Ozawa[4], W Lili[1], A Hangaishi[5], Y Kanda[5], S Chiba[6], M Kurokawa[5], M Omine[7], K Mitani[8] and S Ogawa[1,9]

[1]Department of Regeneration Medicine for Hematopoiesis, Graduate School of Medicine, University of Tokyo, Tokyo, Japan; [2]Department of Hematology/Oncology, National Kyushu Cancer Center, Fukuoka, Japan; [3]Department of Hematology, Tokyo Medical University, Tokyo, Japan; [4]Department of Hematology, Jichi Medical School, Tochigi, Japan; [5]Department of Hematology/Oncology, Graduate School of Medicine, University of Tokyo, Tokyo, Japan; [6]Cell Therapy and Transplantation Medicine, Graduate School of Medicine, University of Tokyo, Tokyo, Japan; [7]Division of Hematology, Internal Medicine, Showa University Fujigaoka Hospital, Yokohama, Japan; [8]Department of Hematology, Dokkyo University School of Medicine, Tochigi, Japan and [9]Core Research for Evolutional Science and Technology, Japan Science and Technology Agency, Saitama, Japan

The unbalanced translocation, der(1;7)(q10;p10), is one of the characteristic cytogenetic abnormalities found in myelodysplastic syndromes (MDS) and other myeloid neoplasms. Although described frequently with very poor clinical outcome and possible relationship with monosomy 7 or 7q- (−7/7q-), this recurrent cytogenetic abnormality has not been explored fully. Here we analyzed retrospectively 77 cases with der(1;7)(q10;p10) in terms of their clinical and cytogenetic features, comparing with other 46 adult −7/7q- cases without der(1;7)(q10;p10). In contrast with other −7/7q- cases, where the abnormality tends to be found in one or more partial karyotypes, der(1;7)(q10;p10) represents the abnormality common to all the abnormal clones and usually appears as a sole chromosomal abnormality during the entire clinical courses, or if not, is accompanied only by a limited number and variety of additional abnormalities, mostly trisomy 8 and/or loss of 20q. der(1;7)(q10;p10)-positive MDS cases showed lower blast counts ($P<0.0001$) and higher hemoglobin concentrations ($P<0.0075$) at diagnosis and slower progression to acute myeloid leukemia ($P=0.0043$) than other −7/7q- cases. der(1;7)(q10;p10) cases showed significantly better clinical outcome than other −7/7q cases ($P<0.0001$). In conclusion, der(1;7)(q10;p10) defines a discrete entity among myeloid neoplasms, showing unique clinical and cytogenetic characteristics.
Leukemia (2007) **21**, 992–997. doi:10.1038/sj.leu.2404619; published online 22 February 2007
Keywords: MDS; AML; MPD; t(1;7); cytogenetics

Introduction

The unbalanced translocation, der(1;7)(q10;p10), is a nonrandom chromosomal abnormality found in myelodysplastic syndromes (MDS) as well as acute myeloid leukemia (AML), and less frequently in myeloproliferative disorders (MPD).[1] It is presented typically with an International System for Chromosome Nomenclature (2005) description,[2] 46,XY(or XX), +1,der(1;7)(q10;p10) and characterized by the presence of one of the two possible derivative chromosomes with two copies of apparently normal chromosome 1 and a single copy of the intact chromosome 7, leading to allelic imbalance of trisomy 1q and monosomy 7q. It was found consistently in 1.5–6% in MDS, 0.2–2.1% in AML and rarely in MPD depending on the literature.[3–9] Detailed molecular studies using quantitative fluorescent *in situ* hybridization analysis disclosed that the translocation is likely to occur through a mitotic recombination within the large clusters of homologous centromere alphoid sequences, D1Z7 on chromosome 1 and D7Z1 on chromosome 7. The breakpoints are widely distributed within these alphoids clusters but typically spare the short arm third of both alphoids, suggesting that it may not be a dicentric chromosome but that der(1;7)(q10,q10) is a more appropriate karyotypic description for this translocation.[10]

On the other hand, however, the clinical aspect of this translocation is less clearly defined compared with its genomic structure. While the common clinical features thus far reported include a previous history of chemo- and/or radiotherapies in more than half cases,[11–13] presence of eosinophilia,[13] trilineage dysplasia, high rates of progression to AML in MDS cases and poor prognosis with less than 1 year of median survival,[3,14,15] these features have not been fully investigated because of the rarity of this translocation compared with other well-characterized cytogenetic groups found in AML/MDS. Also it is unclear that its pathogenetic link to monosomy 7 or partial deletion of 7q (−7/7q-) other than der(1;7)(q10;p10), although some authors speculated that der(1;7)(q10;p10) represents a mere variant of the former.

In this paper, to clarify these points, we analyzed retrospectively 77 cases with der(1;7)(q10;p10) for their clinical and genetic/cytogenetic characteristics in comparison with other 46 cases with −7/7q-. A literary review of 125 der(1;7)(q10;p10) cases was also provided.

Materials and methods

Patients

The patients included in this study are those who were diagnosed and treated as having MDS, AML and MPD at one of the collaborating hospitals during 1 January 1990 through 30 September 2005, for whom at least one cytogenetic report showed either der(1;7)(q10;p10), monosomy 7, or other unbalanced abnormalities showing loss of genetic materials of the long arm of chromosome 7. −7/7q- was required to be

Correspondence: Dr S Ogawa, Department of Hematology and Oncology, Department of Regeneration Medicine for Hematopoiesis, Graduate School of Medicine, University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo 113-8655, Japan.
E-mail: sogawa-tky@umin.ac.jp
Received 24 August 2006; revised 25 January 2007; accepted 25 January 2007; published online 22 February 2007

confirmed in at least three independent karyotypes, whereas other abnormalities were found in at least two karyotypes.[2] Complex karyotypes were defined by the presence of more than two independent karyotypes with at least three unrelated aberrations according to the previous studies.[16,17] Clinical records were reviewed retrospectively at each hospital according to the above criteria and after fully anonymized, the cases who met the criteria were reported according to an indicated form with regard to patients' age, sex, diagnosis (French–American–British (FAB)), date of diagnosis, cytogenetics (G-banding), histories of chemo-radiotherapies, cellularity and blast counts of bone marrow, complete peripheral blood counts at diagnosis, presence or absence of leukemic transformation, and date and cause of death. The list of the collaborating hospitals is the Hematology unit of Akita University Hospital and Dokkyo University Hospital, Honma Hospital, Jichi Medical School Hospital, National Kyushu Cancer Center, Nishio Municipal Hospital, Saitama Medical School Hospital, Showa University Fujigaoka Hospital, Tokyo Medical School Hospital, the Hospital of Tokyo Women's Medical University and the University of Tokyo Hospital. The diagnosis and sub-classification of MDS and AML at each collaborating hospital were made according to the FAB criteria.[18] Mutation study was performed for 20 cases each with der(1;7)(q10;p10) and −7/7q-, for which genomic DNA from bone marrow was available. Another 20 MDS cases showing normal karyotypes were also subjected to the mutation analysis, according to the approval from the ethical committee, University of Tokyo (Approval No. 948-1).

Statistical analysis

The statistical difference in each clinical feature between der(1;7)(q10;p10) and other −7/7q cases was tested by 2×2 contingency tables using the Fischer's exact test or the Student's *t*-test. Overall survival was estimated using the Kaplan–Meier method and the statistic significance was calculated using log-rank tests. After possible association with overall survival was individually tested for a number of variables, including age, bone marrow blasts, peripheral blood counts, chromosomal abnormalities and history of anticancer therapies, proportional hazard modeling was used for identifying the independent risk factors that influence overall survival, where those factors that showed potential significance in the univariate tests with $P<0.10$ were subjected to the multivariate analysis using backward stepwise selection of covariates. All *P*-values were two-sided and *P*-values of 0.05 or <0.05 were considered statistically significant.

Mutation analysis

Mutation status of the *runx1/AML1 and N-ras* genes was tested on aforementioned 60 cases. Exons 3, 4, 5, 6, 7b and 8 of the *runx1* gene and exons 1 and 2 of the *N-ras* gene were amplified from genomic DNA by polymerase chain reaction (PCR) as described previously.[19,20] Sequencing was performed using an ABI Prism 3100 Genetic Analyzer with the same primers as used in PCR amplification.

Literary review of der(1;7)(q10;p10) cases

Reported cases of der(1;7)(q10;p10) were retrieved from the Mitelman Database of Chromosome Aberrations in Cancer 2006[21] according to the following karyotypic descriptions: der(1;7)(q10;p10); der(1)t(1;7)(p11;p11); +t(1;7)(p11;p11),−7; der(1;7)(p10;q10); and dic(1;7)(p11;q11).

Results

Patients' characteristics

In total, 123 cases, including 22 AML, 98 MDS and 3 MPD, were analyzed. The demographic features of these 123 cases are summarized in Table 1. Seventy-seven cases were der(1;7)(q10;p10), which contain 20 cases reported previously.[22,23] The other 46 cases had −7/7q- other than der(1;7)(q10;p10). Strong male predominance was evident in both groups, especially in der(1;7)(q10;p10). The cases having der(1;7)(q10;p10) accounted for 2.3% (10 cases) of 427 cases who were diagnosed as having AML or MDS at the University of Tokyo Hospital. −7/7q- was found in 38 cases among the 427 cases, of which 22 and 16 cases were diagnosed as having MDS and AML, respectively. Both showed a high median age of the disease onset but it was higher in der(1;7)(q10;p10) than for −7/7q- ($P=0.0027$). In our series, 32.9% of der(1;7)(q10;p10) and 25.0% of −7/7q- cases had one or more prior histories of chemotherapies and/or radiotherapies for some malignancies. In both groups, more cases are diagnosed as MDS than as AML. Increased eosinophil counts in der(1;7)(q10;p10) have been underscored in some literatures,[13,24,25] but in our series it was not so conspicuous and only six der(1;7)(q10;p10) and one −7/7q- cases had prominent eosinophilia (>450/μl).

Cytogenetic features

Karyotypic analysis revealed a number of cytogenetic features which contrast der(1;7)(q10;p10) to other 7q- abnormalities (Supplementary Table 1 and 2). der(1;7)(q10;p10) was always present at the time of clinical diagnosis and even when multiple subclones exist, it involves all the abnormal karyotypes. In 45 (58.4%) cases, it appeared as a sole chromosomal abnormality during the observation periods. The remaining 32 cases had additional chromosomal abnormalities, but they were limited in number and mostly consisted of trisomy 8 (18 cases) and/or loss of 20q (10 cases). In contrast, −7/7q- appeared as the solitary abnormality was less common (28.3%) and it was more typically accompanied by other frequently complex abnormalities, where 24 of the 46 −7/7q- cases showed more than four additional unrelated chromosomal abnormalities. Indeed, the mean number of additional abnormalities was significantly higher in the −7/7q cases than in der(1;7)(q10;p10) cases ($P<0.0001$) (Table 1). Moreover, in some cases, it appeared only in partial karyotypes (−7/7q- cases 5, 6, 22, 23, 28, 31 and 35) and evolved during the course of the diseases. 5q- was the most common abnormality that was found in association with −7/7q- (23/46), whereas trisomy 8 and 20q- were very rare in this group.

der(1;7)(q10;p10) in MDS

Since the majority of cases in this study were diagnosed as MDS (FAB classification[18]), it is of interest to focus the analysis on those cases having MDS (Table 2). In fact, clinical features were considerably different between both cytogenetic groups in MDS. According to the FAB classification, der(1;7)(q10;p10) cases were more likely to be classified into RA ($P=0.0002$), whereas RAEB/RAEB-t were the leading diagnosis in the −7/7q- group ($P=0.0003$). Similarly, der(1;7)(q10;p10) cases were more frequently classified into favorable risk groups than −7/7q- cases ($P<0.0001$) in the International Prognostic Scoring System (IPSS),[16] where more Int-1-risk cases ($P<0.0001$) and less high-risk cases ($P<0.0001$) were diagnosed in der(1;7)(q10;p10) than in −7/7q- cases. In accordance with this was that bone marrow

Table 1 Clinical characteristics of patients with der(1;7) and −7/7q-

	der(1;7)	*−7/7q-*	P-*value***	*der(1;7) Mittelman**	P-*value****
Number of patients	77	46		125	
Male/female	68/9	36/10	0.1963	74/51	<0.0001
Median age	67 (17–88)	58 (21–78)	0.0027	58 (7–86)	<0.0001
Positive prior history of chemoradiotherapy	25/76 (32.9%)	10/40 (25.0%)	0.4045		
Time to diagnosis (month)	105 (8–224)	76 (15–135)	0.1057		
Diagnosis					
MDS	64	34	0.2512	77	0.0015
AML	10	12	0.0888	33	0.0329
M0	2	1		2	
M1	1	2		3	
M2	3	4		4	
M4	2	1		8	
M5	0	0		0	
M6	0	3		0	
M7				3	
AML, NS	1	0		13	
MLL(Ph+)	1	1		0	
MPD	3	0		13	
Additional chromosomal abnormalities					
Total	32 (41.6%)	33 (71.7%)	0.0015	45 (36.0%)	0.4581
Trisomy 8	18	1	0.0014	25	0.5983
Trisomy 21	2	1		8	0.3235
Trisomy 9	1	0		4	0.6513
del(20q)	10	4	0.5666	2	0.0014
−5/5q-	1	23	<0.0001	3	
Number of additional chromosomal abnormalities					
1	23	7	0.0837	29	0.3223
2	7	2	0.4935	7	0.3975
3	1	4	0.0645	4	0.6513
⩾4	2	24	<0.0001	5	0.7108
PB					
WBC ($\times 10^3/\mu l$)	3.0 (0.8–39)	3.35 (0.3–56.9)	0.1072		
Hb (g/dl)	9.4 (2.8–13.8)	7.9 (3.8–15.0)	0.0150		
PLT ($\times 10^4/\mu l$)	8.1 (1.3–87)	5.75 (0.6–70.8)	0.2542		
Eosino ($/\mu l$)	45 (0–16932)	0 (0–740)	0.2720		
$>450/\mu l$	6	1			
BM					
Hypercellular	16/69	17/42	0.0585		
Normocellular	31/69	16/42	0.5544		
Hypocellular	22/69	9/42	0.2795		

Abbreviations: AML, acute myeloid leukemia; BM, bone marrow; Eosino, Median eosinophil count; Hb, median hemoglobin level; MDS, myelodysplastic syndrome; MLL, mixed lineage leukemia; MPD, myeloproliferative disorders; NS, not specified; PB, peripheral blood; PLT, median platelet count; WBC, median white blood cell count.
*Non-Asian cases from the Mitelman database of chromosome aberrations in cancer.
***P*-values are calculated between der(1;7)(q10;p10) and −7/7q- cases in the current series or ****P*-values between ours and Mittelman's der(1;7)(q10;p10) cases.

blast counts were significantly lower in der(1;7)(q10;p10) cases than in −7/7q- cases ($P<0.0001$). Also anemia tended to be less severe at the time of diagnosis in der(1;7)(q10;p10) than in −7/7q- cases ($P=0.0075$). More than half cases in both groups were transformed into AML but der(1;7)(q10;p10) cases showed significantly slower progression to AML than −7/7q- cases ($P=0.0043$). Accordingly more patients tend to have been treated by chemotherapy in −7/7q- group than in der(1;7)(q10;p10) group ($P=0.049$). In total, 57 deaths had occurred during the observation period. Infection was the most common cause of deaths in the der(1;7)(q10;p10) group without leukemic transformation (11 of 16 informative cases), whereas MDS patients were more frequently transformed to AML before death in the −7/7q- group (14/22 in −7/7q- vs 14/35 in der(1;7)(q10;p10)).

Allogeneic stem cell transplantation was performed in two der(1;7)(q10;p10) and three −7/7q- cases. The two der(1;7)(q10;p10) cases survived 1184 and 1508 days after transplantation, whereas two of the three −7/7q- cases succumbed to death within a year due to relapse or complication of the transplantation. Among non-transplanted cases, the der(1;7)(q10;p10) group showed significantly better clinical outcome than other −7/7q- cases, although there was significant heterogeneity with regard to therapies. The median overall survivals were 710 and 272 days in der(1;7)(q10;p10) and other −7/7q- cases, respectively ($P<0.0001$) (Figure 1). In univariate analyses of all MDS cases, >10% blast counts ($P=0.0006$) and cytopenia in two or more lineages ($P=0.0189$) were also extracted as significant risk factors. However, after the backward step-wise liner regressions, only −7/7q- karyotypes ($P<0.0001$)

and 60 years or more ages ($P=0.0148$) were extracted as independently significant risk factors, indicating der(1;7)(q10;p10) and −7/7q- karyotypes define separate risk groups (Table 3).

Table 2 Clinical characteristics of MDS patients with der(1;7) and −7/7q-

	der(1;7)	*−7/7q-*	P-*value*
MDS	64	34	
Male/female	57/7	27/7	0.2310
Median age	67 (17–88)	60 (22–78)	0.0974
De novo MDS	44	20	0.9999
Positive prior history of chemoradiotherapy	20	8	
WBC ($\times 10^3/\mu l$)	3.0 (0.8–34.0)	2.4 (0.3–24.0)	0.7043
Hb (g/dl)	9.0 (2.8–13.7)	7.8 (3.8–12.1)	0.0075
PLT ($\times 10^4/\mu l$)	8.0 (1.3–87)	8.1 (0.8–71)	0.9293
BM blast (%)	3.0 (0–14.0)	7.2 (0–29)	<0.0001
Trisomy 8	17	1	0.0048
del(20q)	10	4	0.7651
−5/5q-	0	20	<0.0001
Diagnosis			
RA	39	7	0.0002
RAEB	9	12	0.0203
RAEBt	6	9	0.0379
CMMoL	4	3	0.6908
MDS, NS	6	3	0.9999
IPSS			
High	5	14	0.0002
Int-2	29	17	0.6766
Int-1	31	2	<0.0001
Therapy			
Chemotherapy	11/59	10/24	0.0491
Allo HSCT	2	3	
Transformation to AML (%)	33 (51.6%)	22 (64.7%)	0.2853
Median time to transform (month)	12.0	4.23	0.0043
Median overall survival (month)	23.7	9.07	<0.0001
Death (infection)	39 (18)	22 (8)	

Abbreviations: AML, acute myeloid leukemia; BM, bone marrow; CMMoL, chronic myelomonocytic leukemia; Hb, median hemoglobin level; HSCT, hematopoietic stem cell transplantation; IPSS, international prognostic scoring system; MDS, myelodysplastic syndrome; NS, not specified; PLT, median platelet count; RA, refractory anemia; RAEB, refractory anemia with excess of blasts; RAEBt, RAEB in transformation; WBC, median white blood cell count.

Mutation of runx1 *and* N-ras *genes*

Since high rates of *runx1* mutations have recently been reported in association with −7/7q- cases,[19,26,27] we examined mutation status of *runx1* in 20 der(1;7)(q10;p10) as well as 20 −7/7q- cases together with additional 20 MDS cases without cytogenetic abnormalities. More *runx1* mutations were found in der(1;7)(q10;p10), but it was not significant (7 der(1;7)(q10;p10), 2 −7/7q- and two other cases with normal karyotypes). There were no correlations between mutations and specific FAB subtypes, blast counts or over all survivals, although the sample numbers being too small. *N-ras* mutaions were found in one der(1;7)(q10;p10) cases (at codon 12), two −7/7q- cases (at codon 12) and two normal karyotypes (at codons 12 and 13), which are comparable to the rates reported previously in MDS.[20,28,29]

Literary review of 125 der(1;7)(q10;p10) cases reported previously in the literatures

In total, 164 entries were retrieved from the Mitelman database as having der(1;7)(q10;p10), of which 39 Asian cases were excluded from the further analyses to prevent duplicated retrievals and to explore the ethnic difference in the clinical and cytogenetic pictures of this translocation (Table 1). They had almost similar demographic and cytogenetic features to the current series, but still showed marked difference in several respects. Compared with Asian cases, the male preponderance

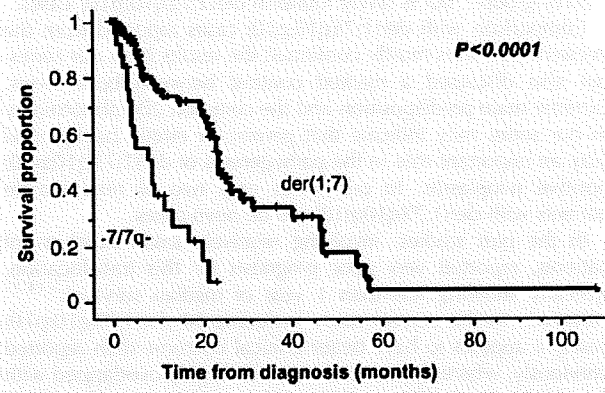

Figure 1 Comparison of survival between der(1;7)(q10;p10) and other −7/7q- cases in MDS. Kaplan–Meier curves of overall survival in MDS cases are compared between der(1;7)(q10;p10) and other −7/7q-. *P*-values in log rank tests are shown.

Table 3 Factors on overall survival

	Univariate		*Multivariate*	
	P-*value*	*HR (95% CI)*	P-*value*	*HR (95% CI)*
MDS cases				
der(1;7) vs −7/7q-	<0.0001	4.240 (2.202–8.164)[a]	<0.0001	4.787 (2.455–9.333)[a]
Age >60	0.0655	1.724 (0.959–3.096)	0.0148	2.088 (1.155–3.774)
Cycytopenia[b] 2/3	0.0189	1.984 (1.107–3.546)		
BM blast >10%	0.0006	2.810 (1.518–5.205)		
Additional cytogenetic changes	0.1598	1.470 (0.856–2.526)		
Secondary MDS	0.1872	1.473 (0.836–2.470)		

Abbreviations: BM, bone marrow; CI, cumulative interval; HR, hazard ratio; MDS, myelodysplastic syndrome.
[a]Hazard ratio of −7/7q- to der(1;7) group.
[b]Found in more than two lineages as defined by neutrophil count <1800/μl, platelets <10,000/μl, Hb <10 g/dl.

is not prominent and diagnosis of MDS and del(20q) abnormalities seem to be less common in the Western cases.

Discussion

We retrospectively analyzed 77 der(1;7)(q10;p10) cases in order to disclose clinical and cytogenetic features of this translocation, focusing on comparison to other −7/7q- cases, since loss of 7q material is a cardinal consequence of this unbalanced translocation as well as gain of 1q material. In spite of the common 7q loss, we found significant differences between both cytogenetic groups in terms of hematological pictures, clinical outcome and cytogenetic characteristics.

der(1;7)(q10;p10) cases tend to be presented with milder anemia, lower blast counts, thus more RA diagnosis, and to show slower progression to AML and have significantly better clinical outcome than other −7/7q- cases. The cytogenetic profiles exhibit a more striking contrast between both groups. der(1;7)(q10;p10) appears as the sole chromosomal abnormality in more than half cases and if not, the additional abnormalities are limited in number and variation, consisting mostly of trisomy 8 and loss of 20q, although the latter has not been described in the Western literatures.[14] In contrast, −7/7q- as the sole abnormality is less common in adult cases but more likely to coexist with other frequently complex abnormalities of partial karyotypes.[4] Common additional abnormalities in −7/7q- cases include −5/5q-, which predicts grave clinical outcomes among −7/7q- group[30] but is rarely found in der(1;7)(q10;p10) cases.

Comparison with der(1;7)(q10;p10) cases reported from the Western countries mostly confirmed the results from our series, but also disclosed a marked contrast between both series. Extreme male predisposition and the common 20q abnormality in our series may indicate that genetic or ethnic background play an important role in the pathogenesis of der(1;7)(q10;p10) positive neoplasms. In our series only two of nine female patients with der(1;7)(q10;p10) are *de novo* cases.

In the past studies, involving relatively small numbers of patients, reported very poor prognosis for this translocation, typically showing less than 1 year of median survival.[3,14,15] However, in the current study including 64 der(1;7)(q10;p10) cases, it appears to have better clinical outcome than reported previously, where the survival curve is roughly overlapped with that of the Int-2 IPSS category with a median survival of 23 months for der(1;7)(q10;p10) cases, which was significantly longer than 9 months of −7/7q-positive MDS cases. This may be explained by the fact that −7/7q- cases showed higher bone marrow blast counts and lower hemoglobin concentration than der(1;7)(q10;p10) cases, which were extracted as significant risk factors in univariate analyses (Table 3). Unexpectedly, however, in multivariate analysis using backward stepwise selection of covariates, only the −7/7q- karyotypes and 60 years or more ages were selected as independent risk factors. Indeed, when the effects of blast counts and cytopenia were adjusted by a proportional hazard model, difference in karyotypes still remains to be significant ($P=0.0034$). These observations suggest that der(1;7)(q10;p10) defines a separate prognostic group that shows significantly better clinical outcome than −7/7q- cases and also indicate a possibility that the der(1;7)(q10;p10) may be more appropriately assigned to an intermediate rather than high-risk karyotype in the IPSS[16] or GCEGCH[17] scoring system for better prediction of prognosis. Unfortunately, however, we were not able to test the latter possibility using our series due to a small number of the cases.

With regard to the pathogenetic role of der(1;7)(q10;p10), no specific molecular targets have been identified for this translocation. Since their breakpoints are distributed widely within the large (~0.5–3 Mb) alphoid cluster regions on the centromeres of chromosome 1 and chromosome 7,[10] no specific gene target at the breakpoints is likely to be involved in the recombination event, but loss of 7q and/or gain of 1q should play a role in the pathogenesis of this translocation. In this point of views, it may be worth mentioning that gains of 1q material are also among recurrent chromosomal abnormalities in MDS.[6,31] Of particular note is that similar 'dicentric' translocations are found in MDS and most frequently involve chromosome 1, resulting in trisomy 1q.[15]

On the other hand, it is still open to questions whether additional genetic hits are required to develop der(1;7)(q10;p10)-positive MDS. According to the cytogenetic profiles of der(1;7)(q10;p10) and −7/7q- cases, different pathogenetic models might be postulated in both cytogenetic groups. der(1;7)(q10;p10) represents a relatively early genetic event and a few additional genetic insults, including +8 and/or 20q-, could be involved in the neoplastic evolution. On the other hand, −7/7q- is likely to be shared by more heterogeneous subgroups and could be a later genetic event during neoplastic process which shows discrete cytogenetic profiles from those involved in der(1;7)(q10;p10). *Runx1* and *N-ras* genes seem to be among common targets of both cytogenetic groups. Especially, a relatively high incidence of *runx*1 mutations (7/20) in our der(1;7)(q10;p10) series should be further confirmed, considering a possible link between runx1 mutations and 7q loss.[19,26]

Conclusion

der(1;7)(q10;p10) defines a unique clinicopathological entity of myeloid neoplasm having a distinct cytogenetic profile and clinical picture. In our study on heterogeneous MDS patients, their clinical outcome is not as bad as reported previously, but still poor with 23 months of median survival, which argues for importance of stem cell transplantation as a potentially curative treatment, and also for development of novel therapeutic approaches.

Acknowledgements

We are grateful to the late Professor Hisamaru Hirai, who initially promoted and encouraged this work. We also thank Dr Masaaki Takatoku of Jichi Medical School and Dr Hiroshi Harada of Showa University Fujigaoka Hospital, Dr Yukihiro Arai of Dokkyo University, Dr Akira Matsuda and Dr Motohiro Misumi of Saitama Medical School, Dr Yuta Koyama of Honma Hospital, Dr Ikuo Miura of Akita University, Dr Hideaki Mizoguchi of Tokyo Women's Medical University and Dr Takuhei Murase of Nishio Municipal Hospital for providing case reports. This work was supported by Research on Measures for Intractable Diseases, Health and Labor Sciences Research Grants, Ministry of Health, Labor and Welfare and by Research on Health Sciences focusing on Drug Innovation, The Japan Health Sciences Foundation.

References

1 Willem P, Pinto M, Bernstein R. Translocation t(1;7) revisited. Report of three further cases and review. *Cancer Genet Cytogenet* 1988; **36**: 45–54.

2 Shaffer L, Tommerup NE. *An International System for Human Cytogenetic Nomenclature (2005).* Karger: Basel, Switzerland, 2005.
3 Horiike S, Taniwaki M, Misawa S, Nishigaki H, Okuda T, Yokota S *et al.* The unbalanced 1;7 translocation in *de novo* myelodysplastic syndrome and its clinical implication. *Cancer* 1990; **65**: 1350–1354.
4 Mauritzson N, Albin M, Rylander L, Billstrom R, Ahlgren T, Mikoczy Z *et al.* Pooled analysis of clinical and cytogenetic features in treatment-related and *de novo* adult acute myeloid leukemia and myelodysplastic syndromes based on a consecutive series of 761 patients analyzed 1976–1993 and on 5098 unselected cases reported in the literature 1974–2001. *Leukemia* 2002; **16**: 2366–2378.
5 Sandberg AA, Morgan R, Hecht BK, Hecht F. Translocation (1;7)(p11;p11): a new myeloproliferative hematologic entity. *Cancer Genet Cytogenet* 1985; **18**: 199–206.
6 Lee DS, Kim SH, Seo EJ, Park CJ, Chi HS, Ko EK *et al.* Predominance of trisomy 1q in myelodysplastic syndromes in Korea: is there an ethnic difference? A 3-year multi-center study. *Cancer Genet Cytogenet* 2002; **132**: 97–101.
7 Toyama K, Ohyashiki K, Yoshida Y, Abe T, Asano S, Hirai H *et al.* Clinical implications of chromosomal abnormalities in 401 patients with myelodysplastic syndromes: a multicentric study in Japan. *Leukemia* 1993; **7**: 499–508.
8 Mertens F, Johansson B, Heim S, Kristoffersson U, Mitelman F. Karyotypic patterns in chronic myeloproliferative disorders: report on 74 cases and review of the literature. *Leukemia* 1991; **5**: 214–220.
9 Reilly JT, Snowden JA, Spearing RL, Fitzgerald PM, Jones N, Watmore A *et al.* Cytogenetic abnormalities and their prognostic significance in idiopathic myelofibrosis: a study of 106 cases. *Br J Haematol* 1997; **98**: 96–102.
10 Wang L, Ogawa S, Hangaishi A, Qiao Y, Hosoya N, Nanya Y *et al.* Molecular characterization of the recurrent unbalanced translocation der(1;7)(q10;p10). *Blood* 2003; **102**: 2597–2604.
11 Scheres JM, Hustinx TW, Geraedts JP, Leeksma CH, Meltzer PS. Translocation 1;7 in hematologic disorders: a brief review of 22 cases. *Cancer Genet Cytogenet* 1985; **18**: 207–213.
12 Morrison-DeLap SJ, Kuffel DG, Dewald GW, Letendre L. Unbalanced 1;7 translocation and therapy-induced hematologic disorders: a possible relationship. *Am J Hematol* 1986; **21**: 39–47.
13 Imai Y, Yasuhara S, Hanafusa N, Ohsaka A, Enokihara H, Tomizuka H *et al.* Clonal involvement of eosinophils in therapy-related myelodysplastic syndrome with eosinophilia, translocation t(1;7) and lung cancer. *Br J Haematol* 1996; **95**: 710–714.
14 Pedersen B. Survival of patients with t(1;7)(p11;p11). Report of two cases and review of the literature. *Cancer Genet Cytogenet* 1992; **60**: 53–59.
15 Pedersen B, Norgaard JM, Pedersen BB, Clausen N, Rasmussen IH, Thorling K. Many unbalanced translocations show duplication of a translocation participant. Clinical and cytogenetic implications in myeloid hematologic malignancies. *Am J Hematol* 2000; **64**: 161–169.
16 Greenberg P, Cox C, LeBeau MM, Fenaux P, Morel P, Sanz G *et al.* International scoring system for evaluating prognosis in myelodysplastic syndromes. *Blood* 1997; **89**: 2079–2088.
17 Sole F, Luno E, Sanzo C, Espinet B, Sanz GF, Cervera J *et al.* Identification of novel cytogenetic markers with prognostic significance in a series of 968 patients with primary myelodysplastic syndromes. *Haematologica* 2005; **90**: 1168–1178.
18 Bennett JM, Catovsky D, Daniel MT, Flandrin G, Galton DA, Gralnick HR *et al.* Proposals for the classification of the myelodysplastic syndromes. *Br J Haematol* 1982; **51**: 189–199.
19 Christiansen DH, Andersen MK, Pedersen-Bjergaard J. Mutations of AML1 are common in therapy-related myelodysplasia following therapy with alkylating agents and are significantly associated with deletion or loss of chromosome arm 7q and with subsequent leukemic transformation. *Blood* 2004; **104**: 1474–1481.
20 Plata E, Viniou N, Abazis D, Konstantopoulos K, Troungos C, Vaiopoulos G *et al.* Cytogenetic analysis and RAS mutations in primary myelodysplastic syndromes. *Cancer Genet Cytogenet* 1999; **111**: 124–129.
21 Mitelman F, Johansson B, Martens FE. Mitelman database of chromosome aberrations in cancer. In: http://cgap.nci.nih.gov/Chromosomes/Mitelman.
22 Hsiao HH, Ito Y, Sashida G, Ohyashiki JH, Ohyashiki K. *De novo* appearance of der(1;7)(q10;p10) is associated with leukemic transformation and unfavorable prognosis in essential thrombocythemia. *Leuk Res* 2005; **29**: 1247–1252.
23 Hsiao HH, Sashida G, Ito Y, Kodama A, Fukutake K, Ohyashiki JH *et al.* Additional cytogenetic changes and previous genotoxic exposure predict unfavorable prognosis in myelodysplastic syndromes and acute myeloid leukemia with der(1;7)(q10;p10). *Cancer Genet Cytogenet* 2006; **165**: 161–166.
24 Forrest DL, Horsman DE, Jensen CL, Berry BR, Dalal BI, Barnett MJ *et al.* Myelodysplastic syndrome with hypereosinophilia and a nonrandom chromosomal abnormality dic(1;7): confirmation of eosinophil clonal involvement by fluorescence *in situ* hybridization. *Cancer Genet Cytogenet* 1998; **107**: 65–68.
25 Kim SH, Suh C, Choi SJ, Kim JG, Lee JH, Kim SB *et al.* Myelodysplastic syndrome that progressed to acute myelomonocytic leukemia with eosinophilia showing peculiar chromosomal abnormality: a case report. *J Korean Med Sci* 1999; **14**: 448–450.
26 Niimi H, Harada H, Harada Y, Ding Y, Imagawa J, Inaba T *et al.* Hyperactivation of the RAS signaling pathway in myelodysplastic syndrome with AML1/RUNX1 point mutations. *Leukemia* 2006; **20**: 635–644.
27 Pedersen-Bjergaard J, Christiansen DH, Desta F, Andersen MK. Alternative genetic pathways and cooperating genetic abnormalities in the pathogenesis of therapy-related myelodysplasia and acute myeloid leukemia. *Leukemia* 2006; **20**: 1943–1949.
28 Padua RA, West RR. Oncogene mutation and prognosis in the myelodysplastic syndromes. *Br J Haematol* 2000; **111**: 873–874.
29 Christiansen DH, Andersen MK, Desta F, Pedersen-Bjergaard J. Mutations of genes in the receptor tyrosine kinase (RTK)/RAS-BRAF signal transduction pathway in therapy-related myelodysplasia and acute myeloid leukemia. *Leukemia* 2005; **19**: 2232–2240.
30 Smith SM, Le Beau MM, Huo D, Karrison T, Sobecks RM, Anastasi J *et al.* Clinical-cytogenetic associations in 306 patients with therapy-related myelodysplasia and myeloid leukemia: the University of Chicago series. *Blood* 2003; **102**: 43–52.
31 Sole F, Espinet B, Sanz GF, Cervera J, Calasanz MJ, Luno E *et al.* Incidence, characterization and prognostic significance of chromosomal abnormalities in 640 patients with primary myelodysplastic syndromes. Grupo Cooperativo Espanol de Citogenetica Hematologica. *Br J Haematol* 2000; **108**: 346–356.

Supplementary Information accompanies the paper on the Leukemia website (http://www.nature.com/leu)

Histone deacetylase inhibitors trichostatin A and valproic acid circumvent apoptosis in human leukemic cells expressing the RUNX1 chimera

Ko Sasaki, Tetsuya Yamagata and Kinuko Mitani[1]

Department of Hematology, Dokkyo Medical University School of Medicine, 880 Kitakobayashi, Mibu-machi, Shimotsuga-gun, Tochigi 321-0293, Japan

(Received July 12, 2007/Revised October 22, 2007/Accepted October 24, 2007/Online publication February 4, 2008)

Disturbance of the normal functions of wild-type RUNX1 resulting from chromosomal translocations or gene mutations is one of the major molecular mechanisms in human leukemogenesis. RUNX1-related chimeras generated by the chromosomal translocations repress transcriptional activity of wild-type RUNX1 by recruiting the co-repressor/histone deacetylase complex. Thus, histone deacetylase inhibitors are expected to restore normal functions of wild-type RUNX1 and thereby affect the growth and differentiation ability of leukemic cells expressing the chimera. We investigated the *in vitro* effects of histone deacetylase inhibitors, trichostatin A and valproic acid, on human leukemic cell lines such as SKNO-1 and Kasumi-1 expressing RUNX1/ETO, Reh expressing TEL/RUNX1 and SKH-1 co-expressing RUNX1/EVI1 and BCR/ABL. We also employed K562 cells expressing BCR/ABL without such a chimera as a control. Treatment with each inhibitor increased acetylated histone 4 in all of these cell lines. Interestingly, proliferation of SKNO-1, Kasumi-1, SKH-1 and Reh cells was significantly suppressed after 3-day culture with trichostatin A or valproic acid, when compared to that of K562 cells. We observed cell cycle arrest and apoptotic induction in the RUNX1 chimera-expressing cells by the propidium iodide staining. Up- and downregulation of cell cycle regulator genes appeared to be the molecular basis for the former, and activation of both extrinsic and intrinsic apoptotic caspases for the latter. We propose histone deacetylase inhibitors to be an attractive choice in the molecular targeting therapy of RUNX1-related leukemia. (*Cancer Sci* 2008; 99: 414–422)

Histone acetylation is maintained by the net balance between the opposing actions of histone acetyltransferase and histone deacetylase (HDAC), and represents one of the fundamental mechanisms by which chromatin remodeling and transcriptional activity are regulated.[1] When the balance is shifted towards histone acetylation, the nucleosome structure opens and associating DNA unfolds. This makes the transcription machinery consisting of transcription factors and RNA polymerase II accessible to target genome and thus results in transcriptional activation. On the other hand, when shifted towards histone deacetylation, compact nucleosome forms condensed chromatin, to which transcriptional machinery becomes inaccessible and thus transcription is silent. This fine regulation of transcription controls cellular processes of proliferation, differentiation and apoptotic induction.[2]

Transcriptional dysregulation, resulting from chimeric gene formation by chromosomal translocation, is one of the major pathogeneses in human leukemia.[3,4] The representative examples include *RUNX1*-targeted chimeric genes such as *RUNX1/ETO* by t(8;21)(q22;q22) in acute myelogenous leukemia,[5] *RUNX1/EVI1* by t(3;21)(q26;q22) in myelodysplastic syndrome or chronic myelocytic leukemia-transformed leukemia,[6] and *TEL/RUNX1* by t(12;21)(p13;q22) in pre-B cell acute lymphoblastic leukemia.[7] Wild-type RUNX1 is a member of the RUNX family of transcription factors that plays a critical role in differentiation of the myeloid progenitor *in vitro* and establishment of fetal liver hematopoiesis.[8] Conditional deletion of the gene causes disturbance in maturation of megakaryocyte and development of both T and B lymphocytes *in vivo*.[9] In the RUNX1 molecule, the N-terminal Runt domain is required for DNA binding to the PEBP2 site, while the C-terminal transactivation domain is necessary for recruiting histone acetyltransferase CREB binding protein and monocytic leukemia zinc finger protein.[8] The RUNX1-related chimeric genes encode an aberrant transcription factor with N-terminal portion of RUNX1 fused to almost the entire of the fusion partner in RUNX1/ETO,[5] and RUNX1/EVI1,[6] or the N-terminal portion of TEL fused to almost the entire RUNX1 in TEL/RUNX1.[7] The chimeric transcription factors dominantly interfere with molecular and biological functions of wild-type RUNX1 through HDAC recruitment to the RUNX1-targeting promoters via co-repressors such as N-CoR and mSin3A in RUNX1/ETO,[10] TEL/RUNX1,[11] and CtBP in RUNX1/EVI,[12,13] which constitutes underpinning molecular mechanisms in the development of RUNX1 chimera-type leukemia.[14] Thus, repression of HDAC activity is presumed to impede oncogenic effect of the chimeric molecules and restore normal functions of RUNX1.

Histone deacetylase inhibitors (HDACi) are members of a new class of chemical agents that modulate gene expression by hyperacetylating histone.[15] Depending on the chemical structure, they are mainly classified into five groups: (i) hydroxamic acid-derived compounds; (ii) short-chain fatty acids; (iii) synthetic benzamide derivatives; (iv) cyclic tetrapeptides; and (v) miscellaneous compounds. HDACi not only suppresses enzymatic activity of HDAC but also releases HDAC from associating transcription factors.[16] Although expression of only 2–5% of genes is significantly activated after the treatment with HDACi, they could be attractive agents for anticancer therapy,[17] if transcription of tumor suppressor genes is recovered. Specifically in RUNX1-related leukemia, HDACi is expected to inactivate HDAC that is bound to the chimeric transcription factors and thus suppress their transforming activity. To establish the possible effects of HDACi in such leukemia, we treated human leukemia cell lines carrying the RUNX1 chimera; namely, SKNO-1[18] and Kasumi-1[19] with RUNX1/ETO; SKH-1[6] with RUNX1/EVI1; and Reh[20] with TEL/RUNX1 with HDACi such as trichostatin A (TSA) and valproic acid (VPA). TSA and VPA belong to a member of hydroxamic acids and short-chain fatty acids, respectively. We employed K562 cells[21] as a control lacking the chimeric transcription factors but possessing constitutively activated non-receptor type tyrosine kinase BCR/ABL. Administration of TSA or VPA exaggerated histone acetylation in all the cell lines examined. Interestingly, the proliferative ability of SKNO-1, Kasumi-1, Reh and SKH1 cells was significantly suppressed by TSA or VPA in

[1]To whom correspondence should be addressed. E-mail: kinukom-tky@umin.ac.jp

doi: 10.1111/j.1349-7006.2007.00699.x

a dose-dependent manner, while that of K562 cells was not. Upon treatment, cell cycle arrest and subsequent apoptotic induction were noticed in SKNO-1, Kasumi-1 and Reh cells. Importantly, polymerase chain reaction (PCR) array analysis demonstrated the increased and decreased expression of cell cycle regulators because of the treatment in these cells. All these data suggest the possibility that HDACi might be a useful modality in the treatment of leukemia expressing RUNX1-related chimera.

Materials and Methods

Cell lines and treatment with HDACi. The human leukemia cell lines SKNO-1 Kasumi-1, Reh, SKH-1 and K562 were maintained in RPMI-1640 medium supplemented with 10% of fetal calf serum. Cells were treated with trichostatin A (TSA; Sigma-Aldrich, St. Louis, MO, USA) or valproic acid (VPA; Sigma-Aldrich) for 24 or 48 h. TSA dissolved in ethanol and VPA in distilled water were added to culture media at concentrations of 10 or 20 ng/mL and 0.5, 1.0 or 2.5 mM, respectively.

Western analysis. Fifty μg of cell lysates obtained from human leukemia cell lines without or with the HDACi treatment at various concentrations was subjected to western analysis using anti-acetylated histone 4 (Upstate, Lake Placid, NY, USA), anticaspase-3 (BD Bioscience, San Diego, CA, USA), anticaspase-8 (BD Bioscience), anticaspase-9 (Cell Signaling Technology, Danvers, MA, USA), anti-PARP (BD Bioscience) and anti-glyceraldehyde 3-phosphate dehydrogenase (GAPDH; EMD Biosciences, San Diego, CA, USA) antibodies as described previously.[22]

Proliferation assay. Exponentially proliferating human leukemia cells at a concentration of 1×10^5/mL were exposed to TSA or VPA at various concentrations for the indicated time intervals. The numbers of viable cells were counted excluding non-viable cells by Trypan blue staining. These experiments were performed in triplicate.

Flow cytometric analysis. At each time point, harvested cells were washed with phosphate-buffered saline (PBS) and stained with propidium iodide (PI) at a concentration of 50 μg/mL at room temperature for 10 min. The DNA content of nuclei was determined with a FACS Calibur flow cytometer (BD Bioscience). The cell cycle data were acquired and analyzed by ModFit LT software (Verity Software House, Topsham, ME, USA). To analyze the proportion of cells under the process of apoptosis, sub-G1 fractions were also evaluated with ModFit LT software. Cells were also analyzed for surface expression of DR5 by direct staining with antihuman DR5-FITC conjugate (Alexis Biochemicals, Lausen, Switzerland). In brief, cells were stained with 100 μL of PBS containing 5 μg/mL of anti-DR5 antibody for 10 min at room temperature, washed twice, and then analyzed with a FACS Calibur flow cytometer (BD Bioscience).

PCR-based array analysis. Exponentially proliferating human leukemia cells at a concentration of 1×10^5/mL were treated with 20 ng/mL of TSA or 2.5 mM of VPA for 12 h. After the treatment, total RNA was extracted from cells using an RNAeasy Kit (Qiagen, Hilden, Germany) and subjected to RT^2 Profiler PCR Array system (APHS-020 A; SuperArray Bioscience, Frederick, MD, USA). The obtained data were analyzed under the manufacturer's program.

Results

TSA and VPA stimulate histone acetylation. At first, we analyzed the effect of TSA and VPA on the acetylation of histone 4 in human leukemia cell lines SKNO-1, Kasumi-1, SKH-1, Reh and K562. SKNO-1 and Kasumi-1 express RUNX1/ETO, and Reh expresses TEL/RUNX1, while K562 cells possess the constitutively activated tyrosine kinase BCR/ABL. SKH-1 was established from a patient with blastic crisis of chronic myelocytic leukemia and expresses both RUNX1/EVI1 and BCR/ABL. We exposed these cells to 10 or 20 ng/mL of TSA or 0.5, 1.0 or 2.5 mM of VPA for 24 h. Western analysis with anti-acetylated histone 4 antibody showed low levels of acetylated protein in all the cell lines before the exposure to TSA or VPA (Fig. 1). The treatment with either HDACi led to a robust concentration-dependent increase of acetylated histone 4. For Reh and SKH-1, we observed higher amounts of the acetylated histone 4 protein after the treatment with TSA than that with VPA. These data confirmed the potent HDACi activity of TSA and VPA in these leukemic cells.

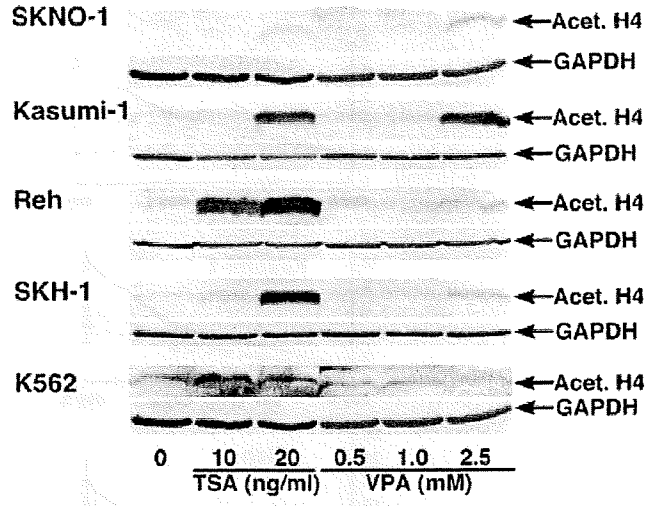

Fig. 1. Treatment with histone deacetylase inhibitors (HDACi) increases the acetylated histone 4. Acetylated histone 4 proteins in SKNO-1, Kasumi-1, Reh, SKH-1 and K562 were detected by anti-acetylated histone 4 antibody after the treatment with trichostatin A (TSA) or valproic acid (VPA) for 24 h at the concentrations indicated. The amount of loaded protein was shown by anti-glyceraldehyde 3-phosphate dehydrogenase (GAPDH) bolt.

TSA and VPA induce concentration-dependent growth repression in SKNO-1, Kasumi-1, SKH-1 and Reh cells, but not in K562 cells. To investigate the biological influence of TSA and VPA on human leukemic cell, we treated these cell lines with the HDACi at the same concentrations as in the above assays for 3 days and counted the number of viable cells. Upon the exposure to TSA, cell growth of SKNO-1, Kasumi-1 and Reh cells was partially or completely inhibited, while that of K562 cells was unchanged (Fig. 2a). The same treatment showed an intermediate effect on SKH-1 cells between the RUNX1 chimera-positive cells and the BCR/ABL-positive cells. On the other hand, incubation with VPA caused more serious changes of cell proliferation and survival in all of these cell lines (Fig. 2b). At the highest concentration, the cell numbers were declined in SKNO-1 and Kasumi-1 cells, suggesting that VPA provoked cell death. SKH-1 and Reh cells also hardly proliferated under the same condition. Even in K562 cells, the growth kinetics became markedly slower. These data indicated that the leukemic cells expressing the RUNX1-related chimeras are affected more seriously with the TSA or VPA treatment than K562 cells and lose their ability to proliferate and survive.

TSA and VPA inhibit cell cycle progression and cause apoptosis in SKNO-1, Kasumi-1 and Reh cells. To clarify the potential effect of TSA and VPA on cell cycle, we performed DNA-content analysis with PI staining after 48-h culture in medium containing 10 or 20 ng/mL of TSA (Fig. 3a) or 1.0 or 2.5 mM of VPA (Fig. 3b). Before the HDACi exposure, the S fraction was relatively higher in SKNO-1, Reh, SKH-1 and K562 cells, and lower in Kasumi-1 cells. When treated with either TSA or VPA, we observed a decrease in the S-phase fraction in a dose-dependent manner in