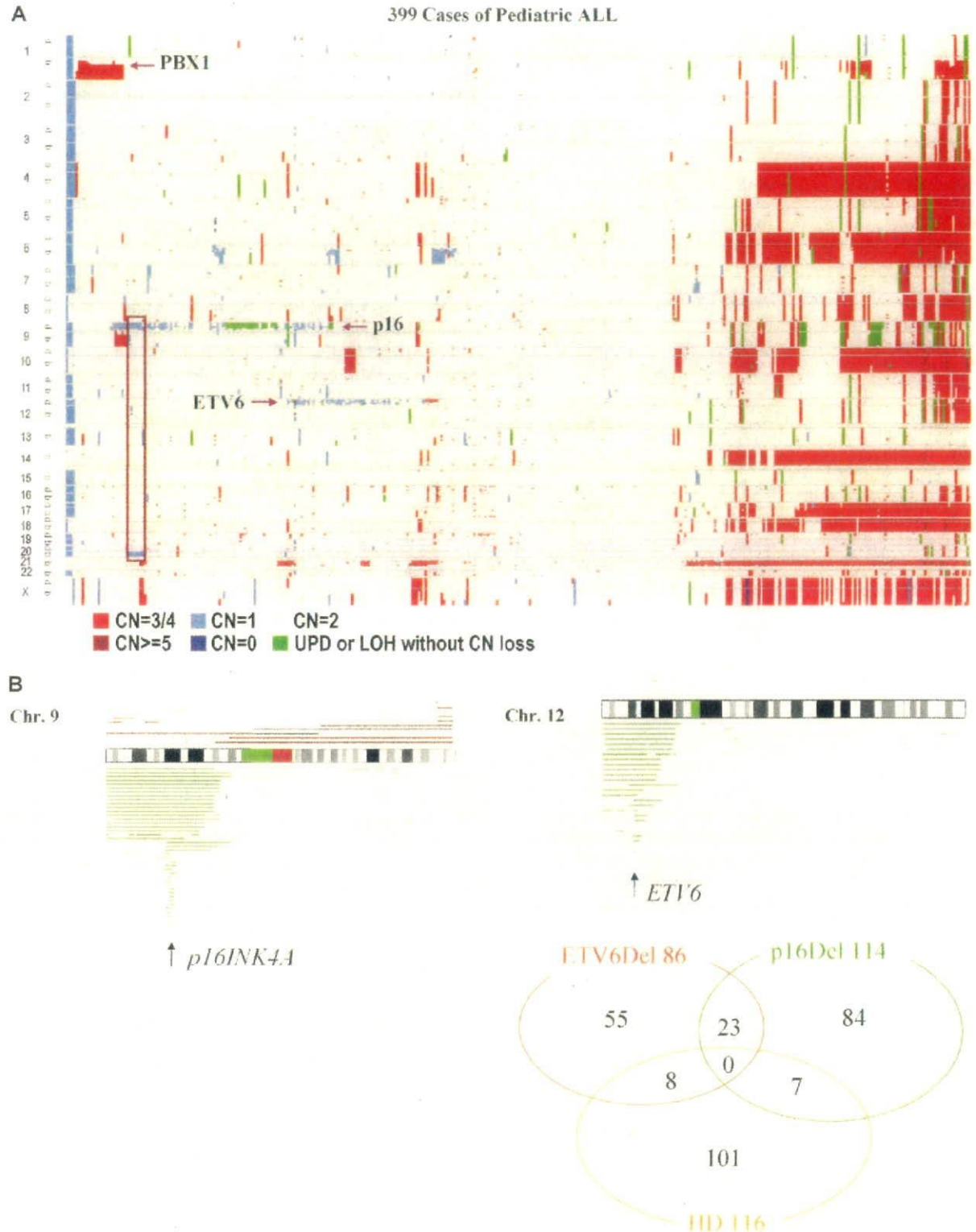

Figure 2. Allelokaryotyping of pediatric ALL. (A) Genetic clustering of 393 ALL samples. Genetic status of each chromosomal region is visualized. Vertical axis: chromosomes, p: short arms, q: long arms. Horizontal axis: individual cases. CN: copy number of alleles. UPD: uniparental disomy. Locations of *PBX1*, *INK4A/ARF*(p16), and *ETV6* genes are indicated. A rectangle indicates cases having 9p and 20q deletions. (B) Three common genomic abnormalities detected in ALL by SNP-chip analysis. Left panel: Deletion of 9p is frequently detected; the arrow indicates the commonly deleted region (CDR) where the p16INK4A gene is located. Right panel: Deletion of 12p often occurs. The arrow indicates that the CDR is where the ETV6 gene is localized. Green lines under the chromosome indicate the deleted regions in individual cases. Brown lines above the chromosome indicate the duplicated regions. Only representative cases are shown. Green and red bands on idiograms indicate centromeres and noncoding regions, respectively. Bottom panel: Venue diagram of 3 common abnormalities detected in this study. Numbers of respective cases in each category are indicated. HD: ALL with hyperdiploid (chromosomes > 50). ETV6Del: ALL with deletion of ETV6 gene. p16Del: ALL with deletion of p16INK4A gene.

Table 3. Clinical features of ALL cases associated with one of three common genetic abnormalities

	Genetic abnormality, no. (%)	Others, no. (%)	*P*
p16Del-ALL			
Age			
1 to 9 y	76 (68)	231 (80)	—
Older than 9 y	35 (32)	57 (20)	.017
WBC			
Below $10^2 \times 10^9$/L	85 (77)	277 (96)	—
Over $10^2 \times 10^9$/L	26 (23)	11 (4)	.001
Non-T lineage	71 (62)	263 (96)	—
T-lineage	83 (38)	11 (4)	.001
ETV6Del-ALL			
Age			
1 to 9 y	75 (87)	232 (74)	—
Older than 9 y	11 (13)	81 (26)	.009
Non-T lineage	89 (95)	261 (85)	—
T-lineage	4 (5)	45 (15)	.015
ETV6/RUNX1			
Positive	53 (66)	43 (15)	—
Negative	27 (34)	243 (85)	.001
HD-ALL			
Age			
1 to 9 y	101 (89)	206 (72)	—
Older than 9	13 (11)	79 (28)	.001
WBC			
Below $10^2 \times 10^9$/L	112 (98)	250 (88)	—
Over $10^2 \times 10^9$/L	2 (2)	35 (12)	.001
Non-T lineage	110 (100)	229 (82)	—
T-lineage	0 (0)	49 (18)	.001
ETV6/RUNX1	HD-ALL	Others	
Positive	8 (8)	88 (34)	—
Negative	97 (92)	173 (66)	.001

p16Del-ALL indicates ALL with deletion of *p16INK4A* gene; ETV6Del-ALL, ALL with deletion of *ETV6* gene; HD-ALL, ALL with hyperdiploidy (chromosomes >50); —, not applicable; WBC, white blood cell count in peripheral blood ($\times 10^9$/L) at diagnosis; *and ETV6/RUNX1, ETV6/RUNX1* fusion was examined by RT-PCR and/or FISH analysis.

tumor content; UPD in samples with as low as 20% of tumor contents are clearly identified (Figure 1Aiv). Whole and partial chromosome UPDs were observed in 95 cases; 6 cases showed both whole and partial chromosome UPD (Figure 3B). Most of the whole chromosome UPD was detected in HD-ALL (Figure 3C). UPD of whole chromosome 9 was the most common whole chromosome UPD (Figure 3B). In contrast, UPD involving part of chromosomes was most often found in non–HD-ALL cases (Figure 3C). Recurrent partial chromosome UPD was detected in many chromosomal regions (Figure 3B). We frequently found whole chromosome 9 UPD (18 cases) and 9p UPD (30 cases) (Figure 3B). *INK4A* deletion was often found in 9p UPD (23 of 30, 77%), while it was rare in whole chromosome 9 UPD (1 of 18, 6%) (Figure 3D).

Relationship between genetic abnormalities

Recurrent abnormalities described above were compared with each other (3 common abnormalities and 26 nonrandom alterations) to detect relationships between these abnormalities (Table S1). Strong correlations between abnormalities of 12p and 21q were detected, duplications of 12p and 21q often occurred simultaneously, and duplications of 21q were accompanied by deletion of the *ETV6* gene that was localized on 12p. ETV6Del ALL frequently had additional changes, including duplications (21q and 1q) and deletions (3p21, 1q, *FHIT*, 15q, and 4q). Abnormalities involving chromosome X, including *DMD* (Xp21.2) deletion, were frequently accompanied by deletions of 8p, 4q, and 6q. Deletion of 20q often occurred with either *p16INK4A* deletion or duplication of 21q (Figure 2A and Table S1).

Impact of nonrandom genetic abnormalities on prognosis

We analyzed prognosis of cases showing nonrandom abnormalities listed in Table 4 and found that the recurrent abnormalities had no impact on the event-free survival (EFS; data not shown) except for amplification/duplication of chromosome 9q. Our initial, early analysis found that EFS of pediatric ALL patients was not impacted by *ETV6* deletion either with or without *ETV6/RUNX1*. Nine cases with 9q amplification/duplication showed a poor prognosis (6 patients relapsed within 3 years; Figure S3A), although the number of cases having this abnormality is too small to reach a significant conclusion. Of these cases, 3 also had duplication of part of chromosome 22 (Figure S3B), suggesting that duplication of 9q is part of an extra copy of the Philadelphia chromosome. These 3 cases showed *BCR/ABL* positively by FISH/RT-PCR analysis (data not shown). Two other cases showed high copy number amplification that encompassed *ABL* and *NUP214* genes (Figure S3C), suggesting these cases had a *NUP214/ABL* fusion.[21] The ALL cells of these 2 cases were steroid-resistant and T-cell lineage phenotype.

Children with HD-ALL without gain of either chromosome 17 or 18 had a worse prognosis (Figure 4). Furthermore, children with HD-ALL and no extra copies of chromosomes 17 and 18 had a significantly worse prognosis ($P < .001$), with 53% EFS at 5 years compared with a 90% 5-year EFS in the other HD-ALL cohort (Figure 4).

Discussion

SNP-chip analysis is a reliable method to detect gene dosage, which was validated by direct sequencing of SNP sites and quantitative PCR in this study. To detect HD-ALL, DNA index is not a sensitive method since contaminated normal cells (DNA index 1.0) decrease the levels of DNA index of hyperdiploid leukemic cells. Although karyotyping is a good method to detect HD-ALL, sufficient number of high-quality chromosomal metaphases is not always obtained from the leukemic cells. SNP-chip analysis may be a more useful and reliable method to detect this subtype of ALL.

Molecular allelokaryotyping of a large series of pediatric ALL samples showed 3 major abnormalities: deletion of *p16INK4A*, deletion of *ETV6*, and hyperdiploidy. Besides these 3 common abnormalities, a number of novel, nonrandom changes were found in ALL. Some of them showed a very narrow commonly deleted region, which was limited to one target gene, including *FHIT* (3p14.2), *TBL1XR1* (3q26.3), and *DMD* (Xp21.2). *DMD* is the causative gene for Duchenne-type muscular dystrophy.[22] While germ-line inactivating mutations of this gene cause the disease, no association has been made between this disease and cancers, including ALL. Since *DMD* is an extremely large gene (2.4 Mb), deletion of it may occur as a result of instability of genomic DNA in ALL cells.

Other investigators and we have found in pediatric ALL a number of deleted genes, including transcriptional factors involved in B-cell differentiation.[19,20] However, since no point mutations of those genes, except PAX5, were detected,[19] it is unclear that these transcriptional factors associated with B-cell development are target genes of these deletions. Mullighan et al[19] and we have found that PAX5 gene is frequently involved in deletions and translocations (N.K., S.O., M.Z., et

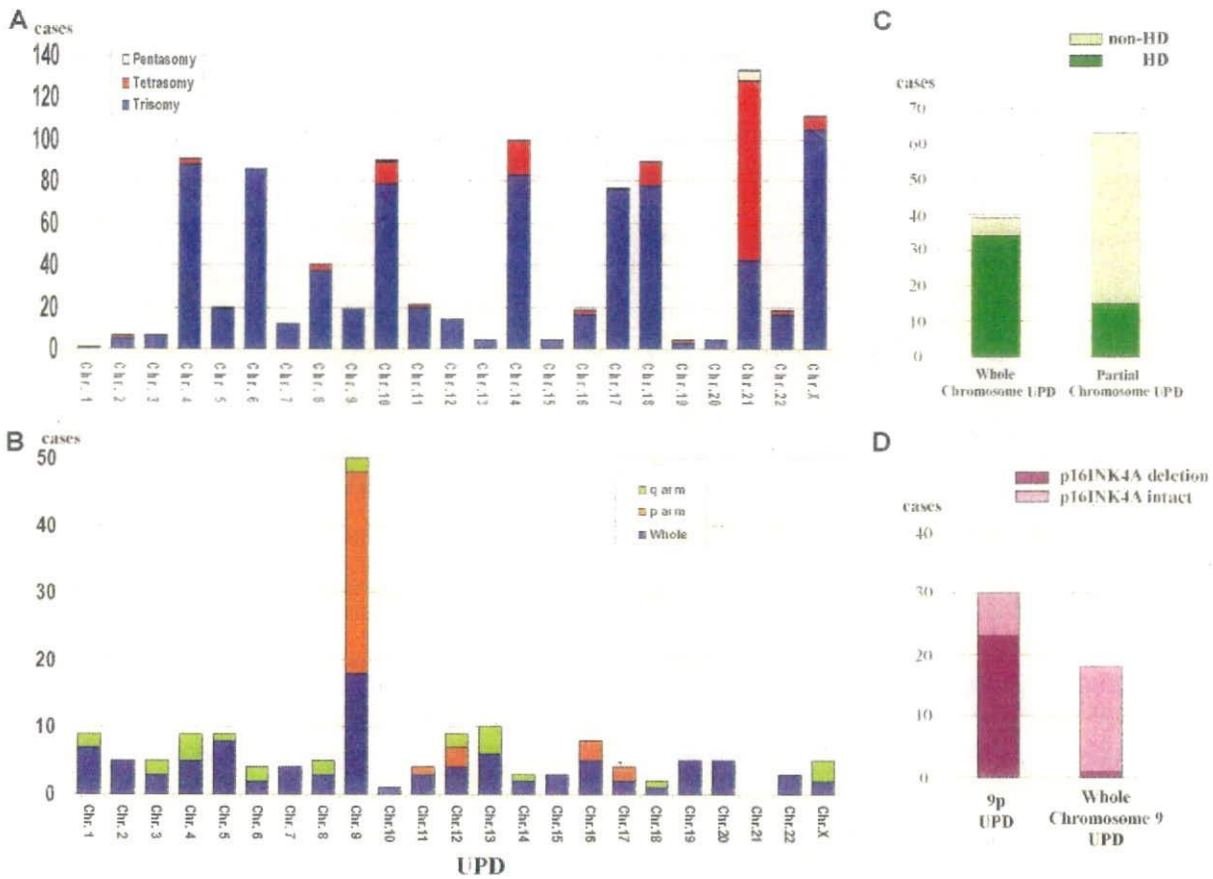

Figure 3. Numerical chromosomal changes and uniparental disomy in pediatric ALL. (A) Frequency of pentasomy/tetrasomy/trisomy affecting each chromosome. For X chromosome, trisomy (105 cases) contains trisosomy X in male pateints (67 cases) and disomy X in female patients (38 cases). All tetrasomy X were female patients. (B) Frequency of uniparental disomy (UPD). Whole: cases with whole chromosome UPD; p arm: cases with UPD of short arm; and q arm: cases with UPD of long arm. Chr: chromosome. UPD involving X chromosome was detected only in female cases. (C) Distribution of whole and partial chromosome UPD in HD and non–HD-ALL. Whole chromosome UPD is frequently detected in HD-ALL. Thirty-four cases with whole chromosome UPD were HD-ALL. Partial UPD is frequently detected in non–HD-ALL. Fifteen of 63 cases with partial UPD were HD-ALL. (D) Frequency of deletion of p16*INK4A* gene in whole chromosome 9 UPD and 9p UPD. Twenty-three cases showed deletion of p*16INK4A*, out of a total of 30 cases with 9p UPD. One case had p16*INK4A* deletion from a total 18 ALL samples with whole chromosome 9 UPD.

al, manuscript submitted), suggesting that impairment of PAX5 is associated with leukomogenesis.

This study showed that numerical chromosomal changes and UPD were common genomic abnormalities in pediatric ALL. Interestingly, whereas trisomy 21 was the most common numerical chromosomal abnormality, UPD of chromosome 21 was not detected in our study. Even in 8 cases with Down syndrome who had trisomy 21 in their constitutional DNA, UPD of chromosome 21 was not detected. Although UPD of chromosome 21 in leukemic cells of patients with Down syndrome has been reported,[8] it may be a rare event in pediatric ALL.

Chromosomal mis-segregation occurs when duplicated chromosomes separate improperly during cell division,[23] leading to numerical chromosomal changes in leukemic cells, including HD-ALL. Most of the whole chromosome UPD was detected in HD-ALL, suggesting that whole chromosome UPD is due also to chromosomal mis-segregation. In contrast, UPD involving part of chromosomes was most often found in non–HD-ALL cases. This may suggest that these partial UPDs are not caused by mis-segregation, but by mitotic recombination, which can usually cause chromosomal exchange.[24]

UPD involving chromosome 9 or 9p is a very common abnormality in pediatric ALL. *INK4A* gene may be a target gene of 9p UPD since most of the cases with 9p UPD had deletion of *INK4A* (23 of 30 cases with 9p UPD). For UPD involving whole chromosome 9, INK4A is not a target gene, since only one case with whole chromosome 9 UPD had deletion of this gene (1 of 18, 6%). Which gene is the target of whole chromosome 9 UPD is unclear. Furthermore, 7 cases with 9p UPD had intact *INK4A*, and the target gene of 9p UPD in these cases is also unknown. This is the first report showing that whole chromosome 9 UPD and 9p UPD are common abnormalities in pediatric ALL. Although Mullighan et al also analyzed a large number of pediatric ALL patients by SNP-chip,[19] they did not report this abnormality.

UPD on 9p is associated with activating point mutations of *JAK2* in myeloproliferative disorders (MPDs).[26-28] However, point mutation of *JAK2* in ALL is very rare.[29,30] We examined for *JAK2* mutations at "hot-spots" (exons 12 and 14)[18,26-28] in these 7 cases of ALL with 9p UPD in which deletion of *INK4A* was not detected. We found no mutations of *JAK2* in these cases. Another unidentified gene(s) in the region may be mutated in these cases.

In this study, we found that absence of gain of chromosomes 17 and 18 had an adverse impact on the prognosis of children with HD-ALL. Another large-scale study found that gain of chromosome 17 was associated with a better prognosis in HD-ALL.[31] Although this previous study reported that gain of chromosomes 4

Table 4. Recurrent genetic abnormalities detected by SNP-chip

Chromosomal sites	Type of abnormality	No. of cases (%)	Candidate genes
1q	Duplication	44 (11)	—
1q	Deletion	11 (3)	—
3p21	Deletion	6 (2)	—
3p14.2	Deletion	6 (2)	*FHIT*
3q26.3	Deletion	10 (3)	*TBL1XR1*
4q31	Deletion	7 (2)	—
6q	Deletion	46 (11)	—
7p	Deletion	10 (3)	—
7q34	Deletion	7 (2)	—
8p	Deletion	13 (3)	—
8q	Duplication	9 (2)	—
9q	Dup/amp	9 (2)	*ABL*
10p	Duplication	7 (2)	—
10q24	Deletion	12 (3)	—
11q	Deletion	24 (6)	—
12p	Duplication	13 (3)	—
13q14.2	Deletion	14 (4)	*RB1*
15q	Deletion	7 (2)	—
17p	Deletion	8 (2)	*TP53*
17q	Duplication	10 (3)	—
17q11.2	Deletion	7 (2)	*NF1*
20p12.2	Deletion	6 (2)	—
20q	Deletion	13 (3)	—
Xp21.2	Deletion	8 (2)	*DMD*
Gain of Chr. 21 or 21q in non-HD ALL cases	—	37/283 (13)	—
Gain of Chr. X or Xq in non-HD ALL cases	—	23/283 (8)	—

Nonrandom chromosomal abnormalities (frequency >1.5% of all cases) are listed. 9p deletion and 12p deletions are separately shown in Figure 2B.
HD indicates hyperdiploid (>50 chromosomes); dup/amp, duplication and amplification of the region; and —, not applicable.

and 10 also was associated with a better prognosis, our study showed that change in number of these chromosomes did not have an impact on prognosis. Even though the size of our study is relatively large, it might not be able to detect some important factors associated with survival because the number of cases of this ALL subtype enrolled in our study was too small or advances in treatment of pediatric ALL may have eliminated several factors that previously influenced prognosis.

One of the limitations of SNP-chip analysis is that it cannot detect balanced translocations, which are common abnormalities in ALL, since this technique can detect only allelic dosage. In our correlation study of genomic abnormalities, a strong correlation was found between abnormalities involving 12p and 21q as described above. This, in part, reflected translocations of chromosome 12p and 21q (*ETV6/RUNX1* fusion). Another correlation that we found between p16Del (on 9p) and deletion of 20q, may reflect dic(p13;q11).[9,20,32,33] These strong correlations of gains or loss of genetic materials may, therefore, sometimes suggest unbalanced translocations in ALL. Combination of SNP-chip and karyotyping will be a very strong technique to examine all genomic abnormalities in malignant cells.

Molecular allelokaryotyping of a series of 399 pediatric ALL samples has defined the range of genetic changes that occur in childhood ALL, including those associated with a poor prognosis. SNP-chip analysis is a novel technique that allows a thorough

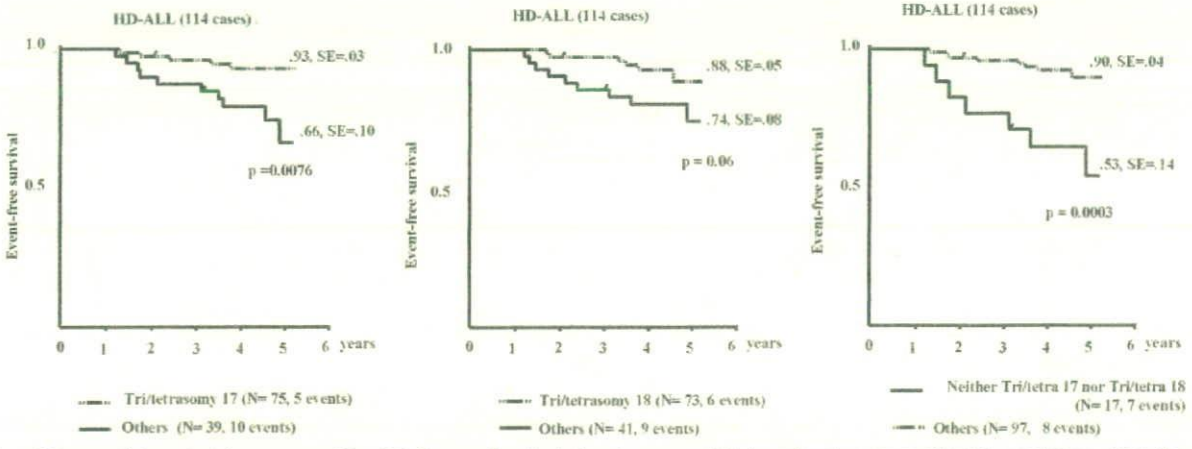

Figure 4. Prognostic impact of chromosomes 17 and 18. Absence of a gain of either chromosome 17 (left panel) or chromosome 18 (middle panel) is associated with poor prognosis in patients with HD-ALL; concurrent absence of a gain of both chromosomes 17 and 18 (right panel) is associated with very poor prognosis. Tri/tetra 18 and 17: HD-ALL with trisomy or tetrasomy 18 or 17. Others: HD-ALL without gain of chromosomes 17 and/or 18.

interrogation of the genome in ALL and identification of clinically significant subgroups of patients.

Acknowledgments

We thank the Parker Hughes Fund and National Institues of Health grants for supporting this study. N.K. is supported by the fellowship from The Tower Cancer Research Foundation. H.P.K. holds the Mark Goodson Chair in Oncology Research at Cedars Sinai and is a member of the Jonsson Cancer Center and the Molecular Biology Institute of UCLA. This work was also supported by grant-in-aid from Department of Health, Welfare, and Labor and from MEXT of the Japanese government, by EU grant FOOD-CT-2005-016320, and a grant from the Deutsche Krebshilfe to C.R.B. The ALL-BFM 2000 trial is supported by 50-2698-Schr 1 of the Deutsche Krebshilfe.

Authorship

Contribution: N.K. and S.O. designed this study, performed experiments, analyzed the data, and wrote the paper. M.Z., K.H., R.K., and W.-D.L. analyzed the data and wrote the paper. M.K., T.F., C.W.M., J.H., and M.S. analyzed the data. M.S., G.Y. and Y.N. performed experiments and analyzed the data. M.S. designed this study and wrote the paper. C.R.B. and H.P.K. designed this study, analyzed the data, and wrote the paper.

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

Correspondence: Seishi Ogawa, 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; or Norihiko Kawamata, Hematology/Oncology, Cedars-Sinai Medical Institute, 8700 Beverly Blvd, Los Angeles, CA 90048; e-mail: kawamatan@cshs.org.

References

1. Armstrong SA, Look AT. Molecular genetics of acute lymphoblastic leukemia. J Clin Oncol. 2005;23:6306-6315.
2. Pui CH, Evans WE. Treatment of acute lymphoblastic leukemia. N Engl J Med. 2006;354:166-178.
3. Pui CH, Relling MV, Downing JR. Acute lymphoblastic leukemia. N Engl J Med. 2004;350:1535-1548.
4. Gunderson KL, Steemers FJ, Lee G, Mendoza LG, Chee MS. A genome-wide scalable SNP genotyping assay using microarray technology. Nat Genet. 2005;37:549-554.
5. Garraway LA, Widlund HR, Rubin MA, et al. Integrative genomic analyses identify MITF as a lineage survival oncogene amplified in malignant melanoma. Nature. 2005;436:117-122.
6. Raghavan M, Lillington DM, Skoulakis S, et al. Genome-wide single nucleotide polymorphism analysis reveals frequent partial uniparental disomy due to somatic recombination in acute myeloid leukemias. Cancer Res. 2005;65:375-378.
7. Teh MT, Blaydon D, Chaplin T, et al. Genomewide single nucleotide polymorphism microarray mapping in basal cell carcinomas unveils uniparental disomy as a key somatic event. Cancer Res. 2005;65:8597-8603.
8. Rogan PK, Close P, Blouin JL, et al. Duplication and loss of chromosome 21 in two children with Down syndrome and acute leukemia. Am J Med Genet. 1995;59:174-181.
9. International HapMap Consortium. A haplotype map of the human genome. Nature. 2005;437:1299-1320.
10. Matsuzaki H, Dong S, Loi H, et al. Genotyping over 100,000 SNPs on a pair of oligonucleotide arrays. Nat Methods. 2004;1:109-111.
11. Nannya Y, Sanada M, Nakazaki K, et al. A robust algorithm for copy number detection using high-density oligonucleotide single nucleotide polymorphism genotyping arrays. Cancer Res. 2005;65:6071-6079.
12. Yamamoto G, Nannya Y, Kato M, et al. Highly sensitive method for genomewide detection of allelic composition in nonpaired, primary tumor specimens by use of affymetrix single-nucleotide-polymorphism genotyping microarrays. Am J Hum Genet. 2007;81:114-126.
13. Bene MC, Castoldi G, Knapp W, et al. Proposals for the immunological classification of acute leukemias: European Group for the Immunological Characterization of Leukemias (EGIL). Leukemia. 1995;9:1783-1786.
14. Ludwig WD, Rieder H, Bartram CR, et al. Immunophenotypic and genotypic features, clinical characteristics, and treatment outcome of adult pro-B acute lymphoblastic leukemia: results of the German multicenter trials GMALL 03/87 and 04/89. Blood. 1998;92:898-909.
15. Schrappe M, Reiter A, Ludwig WD, et al. Improved outcome in childhood acute lymphoblastic leukemia despite reduced use of anthracyclines and cranial radiotherapy: results of trial ALL-BFM 90. German-Austrian-Swiss ALL-BFM Study Group. Blood. 2000;95:3310-3322.
16. Look AT, Melvin SL, Williams DL, et al. Aneuploidy and percentage of S-phase cells determined by flow cytometry correlate with cell phenotype in childhood acute leukemia. Blood. 1982;60:959-967.
17. Whitehead VM, Vuchich MJ, Lauer SJ, et al. Accumulation of high levels of methotrexate polyglutamates in lymphoblasts from children with hyperdiploid (greater than 50 chromosomes) B-lineage acute lymphoblastic leukemia: a Pediatric Oncology Group study. Blood. 1992;80:1316-1323.
18. Scott LM, Tong W, Levine RL, et al. JAK2 exon 12 mutations in polycythemia vera and idiopathic erythrocytosis. N Engl J Med. 2007;356:459-468.
19. Mullighan CG, Goorha S, Radtke I, et al. Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. Nature. 2007;446:758-764.
20. Kuiper RP, Schoenmakers EF, van Reijmersdal SV, et al. High-resolution genomic profiling of childhood ALL reveals novel recurrent genetic lesions affecting pathways involved in lymphocyte differentiation and cell cycle progression. Leukemia. 2007;21:1258–1266.
21. Graux C, Cools J, Melotte C, et al. Fusion of NUP214 to ABL1 on amplified episomes in T-cell acute lymphoblastic leukemia. Nat Genet. 2004;36:1084-1089.
22. Koenig M, Hoffman EP, Bertelson CJ, Monaco AP, Feener C, Kunkel LM. Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary genomic organization of the DMD gene in normal and affected individuals. Cell. 1987;50:509-517.
23. Kops GJ, Weaver BA, Cleveland DW. On the road to cancer: aneuploidy and the mitotic checkpoint. Nat Rev Cancer. 2005;5:773-785.
24. Kotzot D. Complex and segmental uniparental disomy (UPD): review and lessons from rare chromosomal complements. J Med Genet. 2001;38:497-507.
25. Baxter EJ, Scott LM, Campbell PJ, et al. Acquired mutation of the tyrosine kinase JAK2 in human myeloproliferative disorders. Lancet. 2005;365:1054-1061.
26. James C, Ugo V, Le Couedic JP, et al. A unique clonal JAK2 mutation leading to constitutive signalling causes polycythaemia vera. Nature. 2005;434:1144-1148.
27. Kralovics R, Passamonti F, Buser AS, et al. A gain-of-function mutation of JAK2 in myeloproliferative disorders. N Engl J Med. 2005;352:1779-1790.
28. Levine RL, Wadleigh M, Cools J, et al. Activating mutation in the tyrosine kinase JAK2 in polycythemia vera, essential thrombocythemia, and myeloid metaplasia with myelofibrosis. Cancer Cell. 2005;7:387-397.
29. Levine RL, Loriaux M, Huntly BJ, et al. The JAK2V617F activating mutation occurs in chronic myelomonocytic leukemia and acute myeloid leukemia, but not in acute lymphoblastic leukemia or chronic lymphocytic leukemia. Blood. 2005;106:3377-3379.
30. Sulong S, Case M, Minto L, Wilkins B, Hall A, Irving J. The V617F mutation in Jak2 is not found in childhood acute lymphoblastic leukaemia. Br J Haematol. 2005;130:964-965.
31. Sutcliffe MJ, Shuster JJ, Sather HN, et al. High concordance from independent studies by the Children's Cancer Group (CCG) and Pediatric Oncology Group (POG) associating favorable prognosis with combined trisomies 4, 10, and 17 in children with NCI Standard-Risk B-precursor Acute Lymphoblastic Leukemia: a Children's Oncology Group (COG) initiative. Leukemia. 2005;19:734-740.
32. Rieder H, Schnittger S, Bodenstein H, et al. dic(9;20): a new recurrent chromosome abnormality in adult acute lymphoblastic leukemia. Genes Chromosomes Cancer. 1995;13:54-61.
33. Heerema NA, Maben KD, Bernstein J, Breitfeld PP, Neiman RS, Vance GH. Dicentric (9;20)(p11;q11) identified by fluorescence in situ hybridization in four pediatric acute lymphoblastic leukemia patients. Cancer Genet Cytogenet 1996;92:111-115.

Molecular Allelokaryotyping of Early-stage, Untreated Chronic Lymphocytic Leukemia

Sören Lehmann, MD[1,2]
Seishi Ogawa, MD[3]
Sophie D. Raynaud, MD[4]
Masashi Sanada, MD[3]
Yasuhito Nannya, MD[3]
Michel Ticchioni, MD[5]
Christian Bastard, MD[6]
Norihiko Kawamata, MD[1]
H. Phillip Koeffler, MD[1]

[1] Department of Hematology/Oncology, Cedars-Sinai Medical Center, University of California at Los Angeles School of Medicine, Los Angeles, California.

[2] Department of Hematology, Karolinska University Hospital, Huddinge, Sweden.

[3] Regeneration Medicine for Hematopoiesis, University of Tokyo, Tokyo, Japan.

[4] Genetics Laboratory, Archet Hospital, Nice, France.

[5] Immunology Laboratory, Archet Hospital, Nice, France.

[6] Genetic Oncology Laboratory, University of Rouen, Rouen, France.

The first 3 authors contributed equally to this work.

The last 2 authors contributed equally to this work.

Address for reprints: Norihiko Kawamata, MD, Department of Hematology/Oncology, Cedars-Sinai Medical Center/UCLA School of Medicine, 8700 Beverly Boulevard, Los Angeles, CA 90048; Fax: (310) 423-0443; E-mail: kawamatan@cshs.org

Received June 7, 2007; revision received September 10, 2007; accepted September 18, 2007.

BACKGROUND. To the authors' knowledge, genetic abnormalities in early-stage chronic lymphocytic leukemia (CLL) have not been examined fully. Single nucleotide polymorphism (SNP) genomic array (SNP-chip) is a new tool that can detect copy number changes and uniparental disomy (UPD) over the entire genome with very high resolution.

METHODS. The authors performed SNP-chip analysis on 56 samples from patients with early-stage, untreated CLL. To validate the SNP-chip data, fluorescence in situ hybridization (FISH) analysis was performed at selected sites. Expression levels of ZAP-70 and the mutational status of immunoglobulin heavy-chain gene also were examined.

RESULTS. SNP-chip analysis easily detected nearly all changes that were identified by FISH, including trisomy 12, deletion of *TP53* (17p13), deletion of *ATM* (11q22), and deletion of 13q14. Only 10 of 56 CLL samples (18%) had no genomic abnormalities. Excluding the 4 common abnormalities mentioned above, 25 CLL samples (45%) had a total of 45 copy number changes detected by SNP-chip analysis. Four samples had 6q deletion at 6q21 that involved the *AIM1* gene. UPD was detected in 4 samples; 2 samples involved whole chromosome 13 resulting in homozygous deletion of micro-RNA-15a (miR-15a)/miR-16-1. CLL samples with deletion of 13q14 and trisomy 12 were mutually exclusive.

CONCLUSIONS. Genetic abnormalities, including whole chromosome 13 UPD, are very common events in early-stage CLL. SNP-chip analysis can detect small genetic abnormalities in CLL and may be able to support or even supplant FISH and cytogenetics. ***Cancer* 2008;112:1296–305.** © *2008 American Cancer Society.*

KEYWORDS: uniparental disomy, micro-RNA, *ATM*, *TP53*, cyclin D2.

Chronic lymphocytic leukemia (CLL) constitutes the most common type of leukemia in the western world.[1] Its clinical course spans from asymptomatic disease to a rapidly progressing course necessitating urgent and intensive treatment.[2] Several prognostic factors have been identified that predict the time to onset of therapy and the survival of patients, including clinical stage as defined by Rai and Binet, lymphocyte doubling time, β2 microglobulin, and expression of CD38 and ZAP-70.[3–5] However, a more robust manner to subclassify CLL is by identifying the genomic changes in the malignant clone. The relative heterogeneity of the disease is reflected by the occurrence of different genetic abnormalities in distinct subclasses of patients. In addition, a strong correlation exists between specific genetic aberrations and the clinical course of the disease.[3–5] CLL can be divided into 2 categories based on the mutational status of the variable region of the immunoglobulin heavy-chain (IgVH). Somatic hypermutation of the *IgVH* gene is observed in >50% of patients, and its presence is associated with a more benign clinical

DOI 10.1002/cncr.23270
Published online 1 February 2008 in Wiley InterScience (www.interscience.wiley.com).

course.[3–5] By using a sensitive tool such as fluorescence in situ hybridization (FISH), genetic abnormalities can be identified in 80% of patients with CLL.[6] Common genetic changes include deletions of the short arm of chromosome 17 and the long arm of chromosome 11 or 13 as well as trisomy of chromosome 12.[6] Deletion of 13q14 is the most common abnormality in CLL[3–5] and, along with hypermutation of *IgVH*, is associated with a good prognosis.[3–5]

A crucial question is how these genetic events participate in the pathogenesis of CLL. Although major uncertainties exist regarding the causative genes, the common deleted regions at 11q and 17p harbor the *ATM* and *TP53* genes, respectively; and these genes also can be mutated in CLL.[3–5] The 13q14 deletion appears more complex. Despite thorough searches by several groups, no mutated or causative gene has been identified within the common deleted region (CDR).[7,8] However, Calin et al recently demonstrated decreased expression and mutations in 2 mirco-RNAs located between exons 4 and 5 of the *DLEU2* gene, in the CDR at 13q14.[9,10] Nevertheless, the role of these micro-RNAs for the pathogenesis of CLL remains unclear. Beside these common abnormalities, genetic alterations in early-stage CLL have not been fully examined.

Recently, high-density, single nucleotide polymorphism (SNP) genomic arrays (SNP-chips) have become available as a tool to screen for new genetic aberrations in malignant cells.[11,12] SNP-chip analysis allows the detection of copy number changes and allelic loss without copy number change, so-called uniparental disomy (UPD), with a resolution that was not obtained with previous methods.[11,12] We named this novel analysis *molecular allelokaryotyping*, because this technique allows us to examine genetic status at the molecular level and to evaluate allele-specific gene dosage in cancer cells.[13] In the current study, we performed molecular allelokaryotyping on 56 samples of early-stage (Binet stage A) CLL using the 50-k Xba GeneChip from Affymetrix (50,000 SNP probes).

MATERIALS AND METHODS

Patients and Samples

We studied 56 untreated patients with B-cell CLL at Binet stage A who were diagnosed using standard immunophenotypic criteria.[4] All patients had typical morphologic features of CLL and had immunophenotypic scores of 4 or 5 according to the scoring system proposed by Matutes et al.[14] The percentage of CLL cells in peripheral blood ranged between 38% and 98% at diagnosis (mean, 77%). Lymphocyte doubling times in all patients before treatment were >12 months. Thirty patients were men, and 26 were women; and the median age at diagnosis was 69 years (range, 51-86 years). Peripheral blood mononuclear cells were purified by Ficoll gradient and were used as sample material. Chromosomal analysis of CLL cells was performed according to the standard technique. Results of chromosomal analysis were obtained from 49 patients. This protocol was approved by the local ethical committee Comites de Protection des Personnes dans la Recherche Biomedicale (CPPRB) of the Central University Hospital in Nice.

SNP-Chip Analysis

SNP-chips (GeneChip Human mapping 50-k array *Xba*I 240; Affymetrix Japan, Tokyo, Japan) were used for this study. Fragmentation and labeling of DNAs were performed by using a GeneChip resequencing kit (Affymetrix Japan) according to the manufacturer's protocols. Hybridization, washing and signal detection were performed on a GeneChip Fluidics Station 400 and a GeneChip scanner 3000 according to the manufacturer's protocols (Affymetrix, Japan). The data were analyzed by a newly developed program, the copy number analysis for Affymetrix GeneChips (CNAG) program, as described previously.[12,13] By using this program, we were able to obtained allele-specific gene dosage levels of the samples (without matched control samples) by comparing the data with pooled data of normal DNA from volunteers.[13] The results were observed and all genomic abnormalities were examined manually.

Allelelic Dosage Analysis by Real-time Polymerase Chain Reaction

Allelic dosage was measured by real-time polymerase chain reaction (PCR) using genomic DNA extracted from the patients or healthy volunteers. The primers used for allelic dosage analysis are listed in the Supplement Table (available at URL: http://research.csmc.edu/paper/kawamata/CLL.pdf. Accessed on December 25, 2007).

FISH

Peripheral blood smear samples from patients were used for interphase FISH analysis. The FISH studies were performed using the LSI-p53 (17p13)/LSI-*ATM* (11q22) and LSI-13S319 (13q14) 13q34/CEP-12 (chromosome 12) Multicolor Probe sets purchased from Abbott Molecular Inc. (Rungis Cedex, France). Three hundred nuclei were analyzed for each probe set. The cutoff for positive values was 20% for deletion of *ATM* and *TP53*, 9% for 13q14 deletions, and 15% for trisomy 12.

Mutation Status Analysis of the *IgVH* Gene

Genomic DNA was isolated by using standard proteinase K/phenol/chloroform extraction methods.[15] Amplification of the VDJ rearrangements was performed by PCR using framework region 1 and JH family-specific consensus primers, as described previously.[16] The VDJ nucleotide sequences were determined directly and were compared with reported human genome sequences.[16] When a sample demonstrated that <2% of base pairs differed from the reported consensus sequence, it was considered ummutated.[16]

Determination of Genomic Sequences of the *TP53* and *AIM1* Genes

To determine nucleotide sequences, genomic DNA from CLL cells was used as a template. Genomic DNA encoding exons 4, 5, 6, 7, and 8 of the *TP53* gene and all 20 exons of the *AIM1* gene were amplified by PCR with the specific primers. The PCR products were separated in 2.5% agarose gel and purified. Sequences were determined directly by using the purified PCR products. The PCR product of the *TP53* gene was cloned further into pCR 2.1 TOPO vector (Invitrogen, Carlsbad, Calif), and the product was sequenced.

Analysis of Methylation Status of the AIM1 Gene

Genomic DNA was modified by sodium bisulfate using the EZ DNA Methylation Kit (Zymo Research, Orange, Calif). The CpG island of the *AIM1* gene (from −26 to +898; the transcription initiation site was considered as +1) in CLL cells was amplified from the modified genomic DNA as a template by using bisulfate-modified, DNA-specific primers (sense primer: 5′-TTG GAG GTA GAG GGA GAG TTT TT-3′; antisense primer: 5′-CAC CCT CTT AAA TAA AAA CTC C-3′). The PCR products were digested with *BstUI* restriction endonucleases to analyze the methylation status of the CpG island, as reported previously.[17]

ZAP Expression

ZAP-70 expression was determined by flow cytometry on lymphoid cells as described previously with minor modifications.[16] Briefly, cells were incubated with phycoerythrin (PE)-CD56, PE-CD3, peridin-chlorophyll protein-indodicarbocyanine 5-CD19 (Becton-Dickinson, San Jose, Calif), and APC-CD5 antibodies (Pharmingen, San Jose, Calif), then fixed and permeabilized by using the Fix & Perm kit according to manufacturer's instructions (Caltag Laboratories, Burlingame, Calif). Cells were incubated with either an anti-ZAP-70 (Upstate, Charlottesville, Va) or an isotype control antibody followed by fluorescein isothiocyanate-antimouse immunoglobulin antibody (SouthernBiotech, Birmingham, Ala). Cells were washed extensively between each step. The percentage of CLL cells that expressed ZAP-70 was calculated by using the gating strategy as described previously.[16] A CLL sample was considered ZAP-70 positive when >20% the CD19-positive/CD5-positive cells expressed ZAP-70.

Correlation Analysis of Genetic Abnormalities and Clinical Status

The clinical status was available for 51 of 56 patients. Correlation of each genetic abnormality and prognosis were examined by chi-square test.

RESULTS

Validation of SNP-Chip Analysis

Initially, we investigated the specificity and sensitivity of SNP-chip analysis to detect known genomic abnormalities as defined by FISH. These studies validated the efficacy of SNP-chip analysis. Five of our 56 CLL samples (9%) had loss of 1 *ATM* (11q22) allele detected by FISH. These samples had 27%, 58%, 58%, 59%, and 59% of cells with heterozygous loss of *ATM* detected by FISH. SNP-chip analysis detected all 5 of these samples with 11q deletion (Figs. 1, 2). SNP-chip also identified 1 sample that had deletion of the 11q22 region that was not detected by FISH (Fig. 2, Patient 44). FISH analysis in this sample indicated that 14% of cells had 11q22 deletion, and the sample was diagnosed as negative for this deletion. However, SNP-chip analysis clearly demonstrated a deletion of 11q14.2 through q24.2 (Fig. 2). Because data on allele-specific gene dosage indicate levels of paternal and maternal alleles, difference in gene dosage between 2 parental alleles can be observed clearly. Hemizygous deletion of 11q14.2 through q24.2 in this specimen is observed clearly by allele-specific gene dosage analysis (Fig. 2).

Three samples (5%) had loss of 1 allele of *TP53* (17p13) in 30%, 39%, and 55% of cells, as detected by FISH. Each of these deletions also was identified by SNP-chip analysis (Fig. 1). Four of our patients (7%) had trisomy 12 in 21%, 25%, 29%, and 54% of cells detected by FISH. SNP-chip analysis identified gain of whole chromosome 12 in all of these patients (Fig. 1).

Deletion of 13q14, which contains micro-RNA-15a (miR-15a)/miR-16-1, is the most common genetic abnormality in CLL.[9] Of the 56 samples of CLL, 33 samples (59%) had a heterozygous and/or homozygous loss of the critical region on 13q14 identified by FISH (Table 1). SNP-chip analysis per-

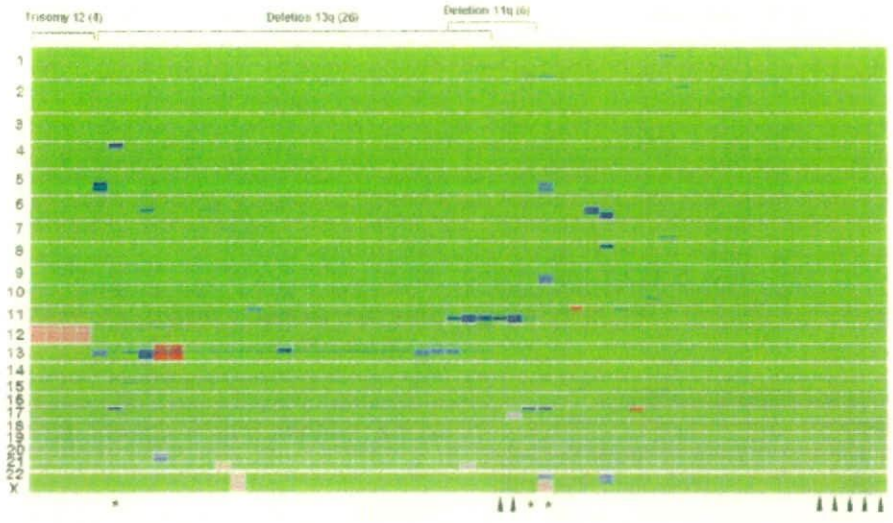

FIGURE 1. Genomic alterations detected by single nucleotide polymorphism (SNP)-chip analysis. Displayed are deletions (blue), amplifications (pink), and uniparental disomy (red) that were identified by SNP-chip analysis. Vertical axis shows chromosomes 1 through X; horizontal axis shows data from 56 patients with chronic lymphocytic leukemia. Upper and lower boxes for each chromosome represent the short and long arms, respectively, of each chromosome. Samples with deletion of *TP53* (17p13) are indicted by asterisks. Samples with 13q14 deletion, which was detected only by fluorescence in situ hybridization analysis, are indicated by triangles. Numbers in parenthesis indicate the number of patients who had each common genetic abnormality identified by SNP-chip analysis.

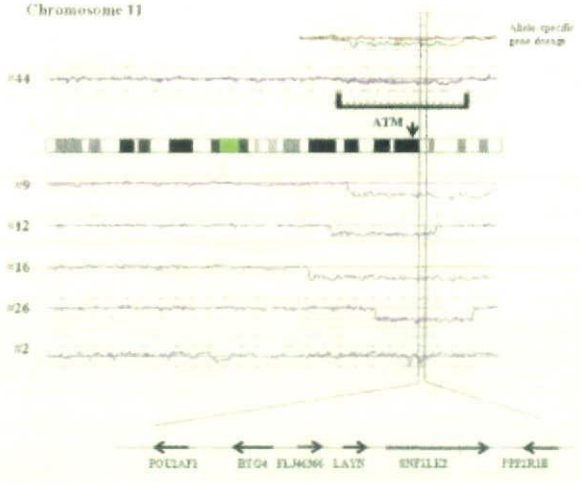

FIGURE 2. Common deleted region of 11q. Single nucleotide polymorphism-chip data from 6 samples of chronic lymphocytic leukemia with deletion of 11q are shown. Each blue line indicates the gene dosage level of each patient. Top: Allele-specific gene dosage levels. Note that the gene dosage level of 1 parental allele (green line) is lower than that of the other parental allele (red line), indicating hemizygous deletion of this region. A vertical arrow above the chromosomal panel indicates the location of the *ATM* gene, one of the commonly deleted regions. The first patient had a slight deletion of 11q that also was detected as hemizygous deletion of *ATM1* in 14% cells by fluorescence in situ hybridization (see text). The bottom line schematically displays genes in a second commonly deleted region on 11q. Each horizontal arrow indicates the direction and size of each of the 6 genes involved in this second commonly deleted region of 11q.

formed on the same samples identified 26 of those 33 samples (78%) (Fig. 1) (Table 1). Seven samples that had 13q14 deletion detected by FISH, but not by SNP-chip analysis, had an average of 15% cells (range, 10%-25%) with deletion of 13q14.

Furthermore, to validate the SNP-chip data, we measured allelic dosage levels by real-time PCR (Fig. 3). We used the 2p12 region as an internal control, because it was not involved in any of the abnormalities in our study. The data demonstrated that the region involved in 13q14 deletions detected by SNP-chip analysis had low allelic dosage, but the regions that were not involved in 13q14 deletion possessed intact levels of allelic dosage (Fig. 3, bottom). Only a representative example is shown in Figure 3.

Novel Copy Number Changes

Four common genetic abnormalities, including trisomy 12 (4 patients), deletion of *TP53* (17p13) (3 patients), deletion of *ATM* (11q22) (6 cases), and deletion of 13q14 (26 patients) were identified, as discussed above. All 6 samples that had *ATM* deletion also had other genetic abnormalities, including deletions of 13q14 and 17p13, detected by SNP-chip/FISH analysis (Fig. 1). In 5 of these samples with *ATM* deletion, the deleted regions were very large; the sixth sample (Patient 2) had 2 small deletions very close together (Fig. 2). The first deletion included the *ATM* gene, and the second deletion con-

TABLE 1
Comparison Between Fluorescence in Situ Hybridization and Single Nucleotide Polymorphism Chip Analysis for Detecting 13q Deletions in Chronic Lymphocytic Leukemia

Patient	Percent of cells with Monoallelic/Biallelic loss at 13q by FISH*	Result by SNP Chip†
5	33/21	Loss (H)
6	27/15	Loss
7 (P)	13/—	No loss
8 (P)	43/11	Loss
9 (P)	43/28	Loss (H)
12	14/—	No loss
14	23/3	Loss
16	11/—	No loss
18	10/—	No loss
19	15/7	No loss
20 (P)	65/—	Loss
21	17/—	Loss
22 (P)	52/—	Loss
23	7/53	Loss (H)
24	3/30	Loss
25	45/—	Loss
26 (P)	47/—	Loss
28 (P)	27/49	Loss (H)
30	27/35	Loss (H)
31	66/5	Loss
32	12/—	Loss
38	19/—	Loss
40	38/—	Loss
41 (P)	25/—	No loss
42	20/7	Loss
43	18/—	No loss
44	15/—	Loss
46 (P)	66/8	Loss
47	32/—	Loss
49	24/3	Loss
51	18/—	Loss
52	24/—	Loss
54	25/64	Loss (H)

FISH indicates fluorescence in situ hybridization; SNP, single nucleotide polymorphism; (H), homozygous deletion was detected by SNP-chip analysis; (P), patients with rapidly progressive disease.* The middle column lists the number of cells with either monoallelic or biallelic loss of D13S319 signal on 13q by FISH (— indicates no cells with biallelic loss).
† The right column shows loss of 13q identified by SNP-chip analysis.

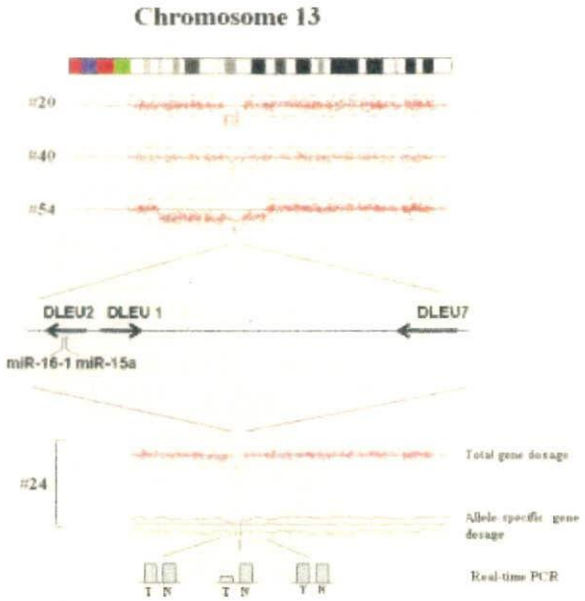

FIGURE 3. Commonly deleted region of 13q14 and whole chromosome 13 uniparental disomy (UPD) with homozygous deletion of 13q14. Top: Representative patients with chronic lymphocytic leukemia (CLL) who had 13q14 deletion detected by single nucleotide polymorphism (SNP)-chip analysis. Individual orange dots represent calculated SNP-chip probe signals, indicating gene dosage levels. The first and second samples show hemizygous deletion of 13q14, and the third sample shows a large hemizygous deletion of 13q and a small homozygous deletion inside of the hemizygous deletion. Middle: Genes in the commonly deleted region of 13q14 are illustrated schematically. Horizontal arrows indicate the direction of transcription and the size of each gene. Two micro-RNAs, miR-15a and miR-16-1, are in intron 4 of the *DLEU2* gene. Bottom: Whole chromosome 13 UPD is shown. Red/green lines represent allele-specific gene dosage. Allele-specific gene dosage lines show loss of 1 parental chromosome (green line) and duplication of the other (red line). Duplicated alleles (red line) have deletion of 13q14 (arrow) leading to homozygous deletion of this region. Each graph shows the allelic dosage levels of 13q14.3 (commonly deleted region), 13q14.1 (centromeric), and 13q21.1 (telomeric) in CLL and normal cells measured by real-time polymerase chain reaction (PCR). T indicates tumor cells; N, reference DNA from a normal, healthy volunteer.

tained 6 other genes (*POU2AF1*, *BTG4*, *FLJ46266*, *LAYN*, *SNFLK2*, and *PPP2R1B*) (Fig. 2). This second commonly deleted region on chromosome 11q was deleted in all 6 samples.

Similar to our findings with *ATM* deletions, all 3 samples that had a 17p deletion had additional genetic abnormalities (Fig. 1). Trisomy 12 was the sole abnormality in all 4 samples as demonstrated by SNP-chip analysis (Fig. 1).

Except for the above-described recurrent abnormalities, the vast majority of genomic changes that have not been described previously were identified by our SNP-chip analysis. Twenty-five samples (45%) had a total of 45 copy number abnormalities other than 13q14 deletion, 11q deletion, 17p deletion, and trisomy 12 identified in SNP-chip analysis (Tables 2, 3).

Of the additional changes that were identified in the SNP-chip analysis, 25 changes were either large deletions (14 samples), duplications (4 samples), UPD (4 samples), trisomy (2 samples), or monosomy (1 sample) that contained >50 genes (Table 2). In addition, 20 changes were either small deletions (13 samples) or duplications (7 samples) and often contained as few as 1 gene or as many as 44 genes

TABLE 2
Single Nucleotide Polymorphism Chip Analysis: Large Copy Number Changes in Chronic Lymphocytic Leukemia*

Chromosomal region	Patient no.	Type of abnormality	Base pair localization
1q32.3-q42	29	Del	208,481,195-227,717,345
4p14-pter	20	Del	pter-34,979,810
4p13-p14	20	Del	40,684,216-44,307,737
5q14.2-q23.2	8	Del	81,990,043-125,733,062
5q14.3-q31.3	29	Del	90,287,988-130,387,114
5q31.3-q33.2	29	Del	143,517,709-154,316,746
5q33.3-q34	29	Del	159,203,291-165,777,302
6q14-q23.2	4	Del	80,810,590-133,976,522
6q16.1-q22.1	30	Del	95,629,083-117,659,576
6q22.33-qter	15	Del	127,493,358-qter
8p12-pter	15	Del	pter-32632781
9q21.11-qter	29	Del	69,208,391-qter
11p14.3-p12	31	Del	25,735,615-35,446,467
11p15.1-pter	35	UPD	pter-19,040,053
Chromosome 13	23, 24	UPD	
17q21.32-qter	12	Dup	42,719,806-qter
17p11.2-pter	48	UPD	pter-21,950,207
Monosomy 21	24		
Trisomy 22	5		
22q11.23-qter	9	Dup	32,159,222-qter
Trisomy X	22		
Xq21.31-pter	15	Del	pter-88,615,115
Xp11.21-pter	29	Del	pter-55,661,160
Xq12-qter	29	Dup	65,546,625-qter

* Large copy number changes other than 11q−, +12, 13q−, and 17p− detected by the single nucleotide polymorphism chip. The table lists deletions (Del), duplications (Dup), uniparental disomy (UPD), trisomy, monosomy, and base pair locations.

(Table 3). Overlapping deletions were detected on 5q, 6q, and Xp (Fig. 4). Two samples had deletion of 5q, and 2 samples had deletion of Xp, as shown in Figure 3A. Four samples had 6q deletions at 6q21 (Fig. 3B). Small alterations that were detected by the SNP-chip analysis involved genes that could be associated with the development of CLL (Table 3). Because only 1 gene, *AIM1*, was involved in the deletion of 6q21 (Fig. 3B), we examined all coding exons and the methylation status of this gene in the 4 samples with deletion of 6q; we observed neither point mutations nor methylation of this gene in these samples (data not shown).

CDR on 13q and UPD

The CDR of 13q14 was localized previously to a 300-kilobase (kb)-long interval on 13q14.2 through 14.3.[18] The CDR in our samples, as identified by SNP-chip analysis, spanned a region of approximately 350 kb that contained this CDR (Figs. 1, 3). Figure 3 shows representative samples of 13q14 deletions. Recently, frequent deletion or mutation of 2 micro-RNAs, miR-15a and miR-16-1, in this region have been reported in CLL.[9,10] All of our samples with 13q− contained these 2 micro-RNAs as a CDR. In our SNP-chip analysis, 6 samples demonstrated homozygous deletion of the 13q14 region (Fig. 3 and data not shown). FISH analysis revealed that these 6 samples had >20% of cells with biallelic loss of 13q14 (Table 1). The presence of cells with biallelic loss of this region

TABLE 3
Single Nucleotide Polymorphism Chip Analysis: Small Copy Number Changes in Chronic Lymphocytic Leukemia*

Chromosomal region	Patient no.	Del/Dup	BP localization	No. of genes	Potential target genes
1q23.2-q23.3	26	Del	156,937, 948-157,642,778	16	
1q32.1	9	Del	199,921,145-200,239,831	7	*BTG2*
1p36.11	10	Del	26,042,344-23,056,134	41	*CD52*
1q42.2-q42.3	2	Del	230,852,980-231,417,785	2	*IRF2BP2*
2p22.2-p22.1	55	Del	37,169,037-39,532,146	14	*SOS1*
4q34.1	40	Dup	175,786,683-176,935,240	2	*ADAM29*
4q35.2	45	Dup	188,796,058-190,012,617	3	*ZFP42*
6q21	15, 28	Del	107,003,343-107,132,119	1	*AIM1*
7q32.1-q32.3	10	Del	128,719,303-129,992,732	13	
9p22.3-p22.2	6	Dup	16,279,143-17,104,199	1	*BNC2*
10q22.1-q22.2	3	Del	72,697,898-72,223,192	44	
10q23.31	3	Del	90,526,429-91,254,269	10	*FAS*
10q23.3	20	Dup	92,020,599-92,812,916	3	*ANKARD1*
12p12.3	24	Del	14,675,168-14,981,176	7	*H2AFJ*
15p13.2-p13.3	8	Dup	28,069,811-30,756,749	5	*KLF13*
15q15.3	6	Del	41,676,268-42,010,479	11	*ELL3*
17q21.31	35	Dup	44,640,555-45,267,454	8	*WNT3*
17q11.2	2	Del	25,744,867-27,046,466	13	*NF1*
20q13.13	53	Dup	45,666,937-46,358,010	1	*PREX1*

Del indicates deletion; Dup, duplication; BP, base pair.
* Small copy number changes that involved <50 genes are listed. Candidate target genes are shown in the far right column.

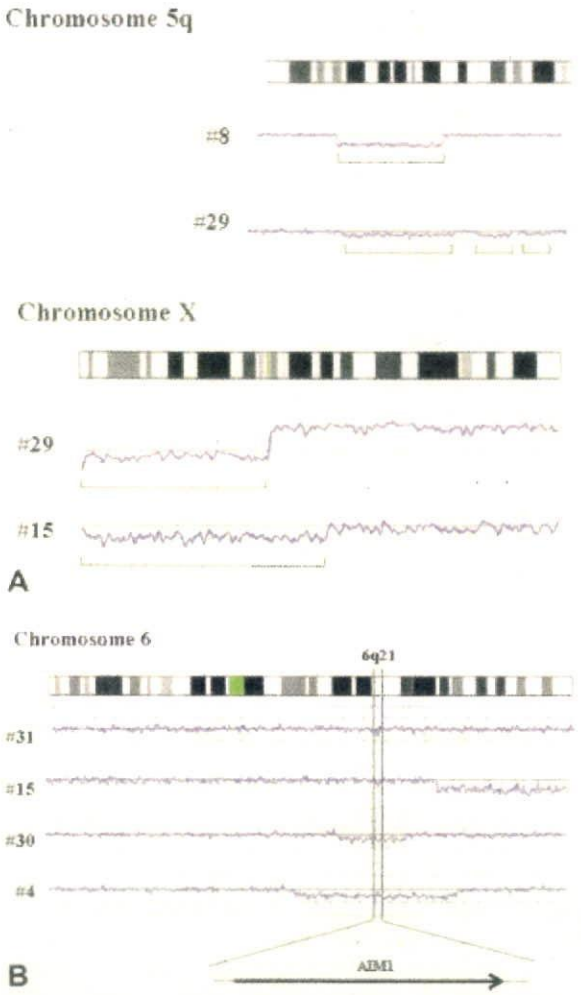

FIGURE 4. Recurrent deletions of 5q, Xp, and 6q detected by single nucleotide polymorphism (SNP)-chip analysis. (A) Deletions of 5q and Xp were detected in 2 patients, respectively, as indicated by brackets. (B) SNP-chip data from 4 samples of chronic lymphocytic leukemia with deletion of 6q. The commonly deleted region is at 6q21 and involves only 1 gene, *AIM1*. A horizontal line at the bottom indicates the direction of transcription of the *AIM1* gene. Each line indicates the gene dosage level of each patient.

was detected in 13 samples, including these 6 samples, by FISH analysis (Table 1).

Unlike comparative genomic hybridization, SNP-chip analysis can detect regions of loss of heterozygosity without copy number changes, so-called UPD.[13] Four of 56 CLL samples had UPD. It is noteworthy that 2 samples with homozygous deletion of 13q14 demonstrated whole chromosome 13 UPD (Table 2) (Fig. 3).

Of the other 2 UPD regions, 1 was located at 11p (11p15.1-pter), and the other was located at 17p (17p11.2-pter), spanning the *TP53* gene (Table 2). Because UPD often is associated with mutations, we sequenced the *TP53* gene at exons 4, 5, 6, 7, and 8 (hotspots of mutations). Codon 72 of *TP53* was arginine (72R) rather than proline (72P) in this instance; this is a polymorphic site (data not shown).

Correlation Analysis of Genetic Abnormalities and Clinical Course

We also examined ZAP-70 expression levels and hypermutation of the *IgVH* gene. Among our 51 patients with early-stage CLL, 19 patients were positive for ZAP-70, and 17 patients had no hypermutation in *IgVH*. Data on the clinical course were available for 51 patients. Nine patients progressed to stage B, 2 patients progressed to stage C, and 3 additional patients began chemotherapy as they progressed clinically (lymphocyte doubling time was <12 months), although they did not fulfill the criteria for stage B. The characteristics of these patients are shown in Table 4. Five of 14 patients had either *ATM* and/or *TP53* deletions, 5 patients had somatic hypermutation of *IgVH*, 8 patients were positive for ZAP-70, and 6 patients were negative for ZAP-70. Even in this small cohort of rapidly progressing patients, deletions of *TP53*, nonmutational status of *IgVH*, and ZAP-70 expression were associated significantly with rapid disease progression (Table 5). Additional genomic abnormalities other than the 4 common ones were not found to be associated significantly with worsening of the disease in our patients (Table 5).

We also compared the genomic abnormalities identified by SNP-chip and/or FISH analysis (Table 6) (Supplement Table). Trisomy 12 and deletion of 13q14 were significantly exclusive ($P = .01$). Deletion of *TP53* (17p13) was associated significantly with the presence of additional genetic abnormalities other than the 4 common genetic abnormalities ($P = .034$). The presence of 13q14 deletion was associated with the absence of additional abnormality ($P = .046$). We compared ZAP-70 expression and *IgVH* mutational status with genomic abnormalities (Supplement Table). ZAP-70 positivity was correlated with nonmutated status of *IgVH* ($P < .001$). Trisomy 12, deletion of 11q22, and deletion of 13q14 were correlated with nonhypermutation of *IgVH* ($P = .033$, $P = .027$, and $P = .027$, respectively) (see Supplement Table). Deletion of 11q22 was correlated with positive ZAP-70 expression ($P = .05$) (Supplement Table).

DISCUSSION

In our 56 samples of early-stage CLL samples, only 10 sample showed no genetic abnormalities by either

TABLE 4
Characteristics of Patients With Progressive Disease

Patient no.	Status[a]	Age, y	Sex	FISH[b]	Additional abnormality	IgVH	ZAP-70	Karyotype[c]
2	At	86	M	17p−, 11q−	Del 17q11.2	−	−	N
9	At	73	M	11q−	Del 1q32.1 Dup 22q11.23−qter	−	+	N
13	At	66	M	13q−	—	+	+	−Y
7	B	53	M	13q−	—	−	+	N
8	B	79	M	13q−	—	+	−	N
10	B	82	M	13q−	—	−	−	N
17	B	58	M	12+	—	−	+	+12
20	B	66	W	13q−,17p−	Del 4p14-pter Del 4p13-p14 Dup 10q23.3	+	−	ND
22	B	65	W	13q−	Trisomy X	+	−	N
26	B	51	W	11q−	Del 1q23.2-q23.3	−	−	Add(11)(q13)
28	C	81	M	13q−	—	−	+	Del (11)(q22)
29	C	64	W	17p−	Del 1q32.3-q42 Del 5q14.3-q31.3 Del 5q31.3-q33.2 Del 5q33.3-q34 Del 9q21.11-qter Del Xp11.21-pter Dup Xq12-qter	+	+	Iso(X)(q)
41	B	70	W	13q−	—	−	+	N
46	B	82	M	13q−	—	−	+	Complex

FISH indicates fluorescence in situ hybridization; IgVH, immunoglobulin heavy-chain; M, man; Del, deletion; −, absence; +, presence; Dup, duplicate; W, woman; Add, addition.

[a] At, therapy initiated but still in stage A; B, progression to stage B; C, progression to stage C.

[b] FISH analysis examined deletions of 11q22 (*ATM*), 17p13 (*TP53*), 13q14, and trisomy 12.

[c] N, normal karyotype, ND, not determined; Complex, complex abnormalities of karyotype. This column describes chromosomal abnormalities (details of karyotypes are shown in Supplement Table 1).

SNP-chip or FISH analysis. The data suggest that genetic abnormalities, including gain, loss, and UPD of genetic materials, frequently occur at an early stage of CLL. In addition to well-documented common genetic abnormalities, including trisomy 12, deletions of *ATM* (11q22), *TP53* (17p13), and 13q14 in CLL, we observed that deletions of 5q, 6q, and Xp were frequent in early-stage CLL. We examined the *AIM1* gene as a target of this deletion. However, neither genetic nor epigenetic abnormalities of this gene were observed in the remaining intact allele. Haploinsufficiency of this gene may be sufficient to contribute to leukemogenesis. Haploinsufficiency of some tumor suppressor genes reportedly suffices to contribute to caricinogenesis.[19] Deletion of 11q is very large, and several genes are deleted in many patients. However, it is believed that the *ATM* gene is the target gene of this deletion.[20] We observed a second commonly deleted region telomeric to the *ATM* gene. This region may contain another tumor suppressor gene associated with the development of CLL. Concurrent deletion of *ATM* and a tumor suppressor gene in this second region may contribute to a poor prognosis for patients with CLL similar to our Patients 2, 9, and 26, who had progressive disease.

In this study, we also analyzed expression levels of ZAP-70 and the mutational status of *IgVH*. We demonstrated for the first time that ZAP-70 expression levels were correlated with 11q22 deletion in early-stage CLL. Although the number of patients was small, we observed that nonhypermutation of *IgVH* was correlated with trisomy 12, deletion of 11q22, and 13q14 in early-stage CLL.

In the current study, we compared data from SNP-chip analysis with data from FISH analysis on several well documented hotspots of deletion, including *TP53*, *ATM*, and 13q14. The sensitivity of FISH analysis was better than that of our SNP-chip analysis at the 13q14 site only when a very minor subclone had the small deletion. In contrast, when a deleted region was large, SNP-chip clearly was able to detect the deletion of 11q22 (*ATM*), even if the clone with this abnormality was not predominant (Fig. 2). We used a 50-k SNP-chip, and this chip had 6 SNP probes encompassing the CDR on 13q14. Now, higher resolution SNP-chips (250-500 k chips) are available and have >30 probes in this region.

TABLE 5
Correlation of Disease Progression and Genomic Status*

TP53 (17p13) deletion	+	−	
Disease progression			
+	3	11	
−	0	38	$P = .002$
ATM (11q22) deletion	+	−	
Disease progression			
+	3	11	
−	3	35	$P = .122$
Additional genomic abnormalities[†]	+	−	
Disease progression			
+	6	8	
−	17	21	$P = .872$
ZAP-70 expression	+	−	
Disease progression			
+	7	7	
−	12	26	$P < .001$
IgVH status	Nonmutated	Mutated	
Disease progression			
+	9	5	
-	8	30	$P < .001$

+ Indicates presence; −, absence; IgVH, immunoglobulin heavy-chain.

* Correlations of disease progression and genomic abnormalities were examined by using the chi-square test. The clinical status of 51 patients was examined, and *P* values are shown.

† Additional abnormalities included genomic abnormalities (see Tables 2 and 3) other than the 4 common genetic abnormalities (trisomy 12, deletion of *TP53*, deletion of *ATM*, and deletion of 13q14).

TABLE 6
Correlation of Genomic Abnormalities

Trisomy 12	+	−	
13q14 Deletion			
+	0	33	
−	4	19	$P = .010$
TP53 (17q13) deletion	+	−	
Additional abnormality*			
+	3	22	
−	0	31	$P = .034$
13q14 Deletion	+	−	
Additional abnormality*			
+	12	14	
−	21	10	$P = .046$

+ Indicates presence; −, absence.

* Additional abnormalities included genomic abnormalities (see Tables 2 and 3) other than the 4 common genetic abnormalities (trisomy 12, deletion of *TP53*, deletion of *ATM*, and deletion of 13q14). *P* values are shown.

These higher resolution SNP-chips will have greater sensitivity to detect this small deletion in CLL comparable to the sensitivity of FISH analysis. We believe that the higher resolution (250-500 k) SNP-chip will be able to supplant FISH and cytogenetics.

One of the advantages of our method is that we can detect allele-specific gene dosage levels without matched normal controls. We used pooled data from several hundred normal DNA samples, and our computerized software selected optimal reference data for individual CLL samples.[13] This new technique allowed us precisely to detect UPD, one of the common genetic abnormalities observed in cancers, without matched control DNA.[13] It is noteworthy that we identified 2 patients with UPD of the whole chromosome 13 among patients with homozygous deletion of 13q14 (Patients 23 and 24). This finding pinpoints the usefulness of SNP-chip analysis to detect copy number neutral loss of heterozygosity. It also suggests that homozygosity of the whole chromosome 13 is a step in the clonal evolution of CLL, resulting in the loss of both alleles at the CDR on 13q14. We also observed UPD and deletions that involved 17p13, which contains the tumor suppressor gene *TP53*. Sequencing the *TP53* gene in the sample with UPD showed a previously reported polymorphic nucleotide change at codon 72 in both alleles. It is unclear whether *TP53* is the target gene of this UPD.

Recently, Pfeifer et al reported an SNP-chip analysis of 70 CLL samples using 10- to 50-k Affymetrix GeneChip genomic microarrays.[21] The difference between our study and theirs is that we examined patients with early-stage CLL who were not treated at the time the samples were collected, whereas Pfeifer et al analyzed patients in all stages of CLL, and approximately 50% of their patients previously had received chemotherapy. Chemotherapeutic reagents, especially alkylating reagents (which are used frequently to treat CLL[22]), damage DNA and lead to therapy-related genetic abnormalities.[23] Furthermore, treatments with chemotherapeutic reagents select the resistant clones in CLL, and drug-resistant clones become dominant after prolonged treatment.[24] Those clones may be different from initial CLL malignant cells. Our patients were not treated with any reagents and most likey had genetic events that occurred during the early stage of development of CLL. Although Pfeifer et al observed duplication/amplification of 2p involving *Rel* and *BCL11A*,[21] we did not observe these genetic abnormalities. Duplication/amplification of *Rel* and *BCL11A* may be associated with either advanced disease or secondary events caused by chemotherapeutic treatment.

Pfeifer et al also reported that patients who had CLL with biallelic loss of 13q14 had a worse prognosis.[21] In our study, 15 of 33 patients who had 13q14 deletion had clones with biallelic loss of this region (percentages of clones with biallelic loss, 3%-64%) detected by FISH. Only 4 of those 15 patients with biallelic loss of 13q14 (percentage of cells with biallelic loss: 8%, 11%, 28%, and 35%) developed progressive disease.

Therefore, biallelic loss of 13q14 did not appear have an adverse effect on the prognosis of patients CLL in our study ($P = .953$) (Supplementary Table).

We demonstrated that 13q14 deletion and trisomy 12 were exclusive of each other. MiR15a/16-1 commonly are deleted in patients with 13q14 deletions. MiR15a/16-1 are micro-RNAs that block the function of a variety of target genes.[25] Cyclin D2 is one of the target genes of these micro-RNAs (available at URL: http://www.targetscan.org/ Accessed on August 1, 2007) and is localized on 12p13. Overexpression of cyclin D2 is observed frequently in CLL.[26] Deletion of miR15a/16-1 may lead to restoration of cyclin D2 expression in CLL; likewise, trisomy 12 may be an alternative means to enhance expression of this gene. Cyclin D2 may be a common downstream target of both deletion of 13q14 and trisomy 12. Unfortunately, we were not able to obtain expression levels of cyclin D2 or miR15a/16-1 in our samples. Future analysis will elucidate the correlations of these important genes in the development of CLL.

In summary, the results from the current study suggest that SNP arrays can accurately detect the common abnormalities in CLL. The technique also revealed a large number of new abnormalities, especially smaller regions that were undetected by prior methods. Several of these abnormalities involve genes of potential interest, because they may be causative of CLL and will require further study. The SNP-chip is a rapid, robust assay that can survey the entire genome, provide diagnostic and prognostic clarity, and help identify new therapeutic targets. The cost of the assay is decreasing quickly, and it has the potential to replace conventional techniques.

REFERENCES

1. Zwiebel JA, Cheson BD. Chronic lymphocytic leukemia: staging and prognostic factors. *Semin Oncol.* 1998;25:42–59.
2. Seiler T, Dohner H, Stilgenbauer S. Risk stratification in chronic lymphocytic leukemia. *Semin Oncol.* 2006;33:186–194.
3. Binet JL, Caligaris-Cappio F, Catovsky D, et al. Perspectives on the use of new diagnostic tools in the treatment of chronic lymphocytic leukemia. *Blood.* 2006;107:859–861.
4. Oscier D, Fegan C, Hillmen P, et al. Guidelines on the diagnosis and management of chronic lymphocytic leukaemia. *Br J Haematol.* 2004;125:294–317.
5. Shanafelt TD, Geyer SM, Kay NE. Prognosis at diagnosis: integrating molecular biologic insights into clinical practice for patients with CLL. *Blood.* 2004;103:1202–1210.
6. Dohner H, Stilgenbauer S, Benner A, et al. Genomic aberrations and survival in chronic lymphocytic leukemia. *N Engl J Med.* 2000;343:1910–1916.
7. Migliazza A, Bosch F, Komatsu H, et al. Nucleotide sequence, transcription map, and mutation analysis of the 13q14 chromosomal region deleted in B-cell chronic lymphocytic leukemia. *Blood.* 2001;97:2098–2104.
8. Bullrich F, Fujii H, Calin G, et al. Characterization of the 13q14 tumor suppressor locus in CLL: identification of ALT1, an alternative splice variant of the LEU2 gene. *Cancer Res.* 2001;61:6640–6648.
9. Calin GA, Dumitru CD, Shimizu M, et al. Frequent deletions and down-regulation of micro-RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proc Natl Acad Sci USA.* 2002;99:15524–15529.
10. Calin GA, Ferracin M, Cimmino A, et al. MicroRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. *N Engl J Med.* 2005;353:1793–1801.
11. Lindblad-Toh K, Tanenbaum DM, Daly MJ, et al. Loss-ofheterozygosity analysis of small-cell lung carcinomas using single-nucleotide polymorphism arrays. *Nat Biotechnol.* 2000;18:1001–1005.
12. Nannya Y, Sanada M, Nakazaki K, et al. A robust algorithm for copy number detection using high-density oligonucleotide single nucleotide polymorphism genotyping arrays. *Cancer Res.* 2005;65:6071–6079.
13. Yamamoto G, Nannya Y, Kato M, et al. Highly sensitive method for genome-wide detection of allelic composition in non-paired, primary tumor specimens using SNP genotyping micorarrays. *Am J Hum Genet.* 2007;81:114–126.
14. Matutes E, Owusu-Ankomah K, Morilla R, et al. The immunological profile of B-cell disorders and proposal of a scoring system for the diagnosis of CLL. *Leukemia.* 1994;8:1640–1645.
15. Sambrook J, Russell DW. Molecular Cloning: A Laboratory Manual. 3rd ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 2001.
16. Crespo M, Bosch F, Villamor N, et al. ZAP-70 expression as a surrogate for immunoglobulin-variable-region mutations in chronic lymphocytic leukemia. *N Engl J Med.* 2003;348:1764–1775.
17. Xiong Z, Laird PW. COBRA: a sensitive and quantitative DNA methylation assay. *Nucleic Acids Res.* 1997;25:2532–2534.
18. Kalachikov S, Migliazza A, Cayanis E, et al. Cloning and gene mapping of the chromosome 13q14 region deleted in chronic lymphocytic leukemia. *Genomics.* 1997;42:369–377.
19. Largaespada DA. Haploinsufficiency for tumor suppression: the hazards of being single and living a long time. *J Exp Med.* 2001;193:F15–F18.
20. Stankovic T, Weber P, Stewart G, et al. Inactivation of ataxia telangiectasia mutated gene in B-cell chronic lymphocytic leukaemia. *Lancet.* 1999;353:26–29.
21. Pfeifer D, Pantic M, Skatulla I, et al. Genome-wide analysis of DNA copy number changes and LOH in CLL using high-density SNP arrays. *Blood.* 2007;109:1202–1210.
22. Robak T, Kasznicki M. Alkylating agents and nucleoside analogues in the treatment of B cell chronic lymphocytic leukemia. *Leukemia.* 2002;16:1015–1027.
23. Karp JE, Smith MA. The molecular pathogenesis of treatment-induced (secondary) leukemias: foundations for treatment and prevention. *Semin Oncol.* 1997;24:103–113.
24. Bosanquet AG, Bell PB. Novel ex vivo analysis of nonclassical, pleiotropic drug resistance and collateral sensitivity induced by therapy provides a rationale for treatment strategies in chronic lymphocytic leukemia. *Blood.* 1996;87:1962–1971.
25. Yoon S, De Micheli G. Prediction of regulatory modules comprising microRNAs and target genes. *Bioinformatics.* 2005;21(suppl 2):ii93–ii100.
26. Delmer A, Ajchenbaum-Cymbalista F, Tang R, et al. Overexpression of cyclin D2 in chronic B-cell malignancies. *Blood.* 1995;85:2870–2876.

Identification of human minor histocompatibility antigens based on genetic association with highly parallel genotyping of pooled DNA

Takakazu Kawase,[1,2] Yasuhito Nannya,[3-5] Hiroki Torikai,[1] Go Yamamoto,[3-5] Makoto Onizuka,[6] Satoko Morishima,[1] Kunio Tsujimura,[7] Koichi Miyamura,[5,8] Yoshihisa Kodera,[5,8] Yasuo Morishima,[5,9] Toshitada Takahashi,[10] Kiyotaka Kuzushima,[1] Seishi Ogawa,[3-5] and Yoshiki Akatsuka[1,5]

[1]Division of Immunology, [2]Division of Epidemiology and Prevention, Aichi Cancer Center Research Institute, Nagoya; [3]Department of Hematology/Oncology and [4]21st Century COE Program, Graduate School of Medicine, University of Tokyo, Tokyo; [5]Core Research for Evolutional Science and Technology, Japan Science and Technology Agency, Saitama; [6]Department of Genetic Information, Division of Molecular Life Science, Tokai University School of Medicine, Isehara; [7]Department of Microbiology and Immunology, Hamamatsu University School of Medicine, Hamamatsu; [8]Department of Hematology, Japanese Red Cross Nagoya First Hospital, Nagoya; [9]Department of Hematology and Cell Therapy, Aichi Cancer Center Hospital, Nagoya; and [10]Aichi Comprehensive Health Science Center, Aichi Health Promotion Foundation, Chita-gun, Japan

Minor histocompatibility (H) antigens are the molecular targets of allo-immunity responsible both for the development of antitumor effects and for graft-versus-host disease (GVHD) in allogeneic hematopoietic stem cell transplantation (allo-HSCT). However, despite their potential clinical use, our knowledge of human minor H antigens is largely limited by the lack of efficient methods of their characterization. Here we report a robust and efficient method of minor H gene discovery that combines whole genome association scans (WGASs) with cytotoxic T-lymphocyte (CTL) assays, in which the genetic loci of minor H genes recognized by the CTL clones are precisely identified using pooled-DNA analysis of immortalized lymphoblastoid cell lines with/without susceptibility to those CTLs. Using this method, we have successfully mapped 2 loci: one previously characterized (*HMSD* encoding ACC-6), and one novel. The novel minor H antigen encoded by *BCL2A1* was identified within a 26 kb linkage disequilibrium block on chromosome 15q25, which had been directly mapped by WGAS. The pool size required to identify these regions was no more than 100 individuals. Thus, once CTL clones are generated, this method should substantially facilitate discovery of minor H antigens applicable to targeted allo-immune therapies and also contribute to our understanding of human allo-immunity. (Blood. 2008;111: 3286-3294)

Introduction

Currently, allogeneic hematopoietic stem cell transplantation (allo-HSCT) has been established as one of the most effective therapeutic options for hematopoietic malignancies[1] and is also implicated as a promising approach for some solid cancers.[2] Its major therapeutic benefits are obtained from allo-immunity directed against patients' tumor cells (graft-versus-tumor [GVT] effects). However, the same kind of allo-immune reactions can also be directed against normal host tissues resulting in graft-versus-host disease (GVHD). In HLA-matched transplants, both GVT and GVHD are initiated by the recognition of HLA-bound polymorphic peptides, or minor histocompatibility (H) antigens, by donor T cells. Minor H antigens are typically encoded by dichotomous single nucleotide polymorphism (SNP) alleles, and may potentially be targeted by allo-immune reactions if the donor and recipient are mismatched at the minor H loci. Identification and characterization of minor H antigens that are specifically expressed in hematopoietic tissues, but not in other normal tissues, could contribute to the development of selective antileukemic therapies while minimizing unfavorable GVHD reactions, one of the most serious complications of allo-HSCT.[3,4] Unfortunately, the total number of such useful minor H antigens that are currently molecularly characterized is still disappointingly small, including HA-1,[5] HA-2,[6] ACC-1[Y] and ACC-2,[7] DRN-7,[8] ACC-6,[9] LB-ADIR-1F,[10] HB-1,[11] LRH-1,[12] and 7A7-PANE1,[13] limiting the number of patients eligible for such GVT-oriented immunotherapy.

Several techniques have been developed to identify novel minor H antigens targeted by CTLs generated from patients who have undergone transplantation. Among these, linkage analysis based on the cytotoxicity of the CTL clones against panels of lymphoblastoid cell lines (B-LCLs) from large pedigrees was proposed as a novel genetic approach,[14] and has been successfully applied to identify novel minor H epitopes encoded by the *BCL2A1* and *P2RX5* genes.[7,12] Nevertheless, the technology is still largely limited by its resolution, especially when large segregating families are not available. Linkage analysis using B-LCL panels from the Centre d'Etude du Polymorphisme Humain (CEPH) could only localize minor H loci within a range of 1.64 Mb to 5.5 Mb, which still contained 11 to 46 genes,[7,12,14] thus requiring additional selection procedures to identify the actual minor H genes.

On the other hand, clinically relevant minor H antigens might be associated with common polymorphisms within the human

Submitted October 22, 2007; accepted December 19, 2007. Prepublished online as *Blood* First Edition paper, January 4, 2008; DOI 10.1182/blood-2007-10-118950.

T.K. and Y.N. contributed equally to this work.

The online version of this article contains a data supplement.

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

population, and therefore could be ideal targets of genetic association studies, considering recent advances of large-scale genotyping technologies and the assets of the International HapMap Project.[15,16] In this alternative genetic approach using the extensive linkage disequilibrium (LD) found within the human genome, target loci can be more efficiently localized within relatively small haplotype blocks without depending on limited numbers of recombination events, given the large number of genotyped genetic markers.[17] Moreover, since the presence of a target minor H allele in individual target cells can be determined by ordinary immunologic assays using minor H antigen-specific CTLs, the characterization of minor H antigens should be significantly more straightforward than identifying alleles associated with typical common complex diseases, for which typically weak-to-moderate genetic effects have been assumed.[18]

In this report, we describe a high-performance, cost-effective method for the identification of minor H antigens, in which whole genome association scans (WGASs) are performed based on SNP array analysis of pooled DNA samples constructed from cytotoxicity-positive (CTX^+) and cytotoxicity-negative (CTX^-) B-LCLs as determined by their susceptibility to CTL clones. Based on this method, termed WGA/CTL, we were able to map the previously characterized ACC-6 minor H locus to a 115-kb block containing only 4 genes, including *HMSD*.[9] Moreover, using the same approach, a novel minor H antigen encoded by the *BCL2A1* gene was identified within a 26-kb block containing only *BCL2A1* on chromosome 15q25. Surprisingly, the pool size required to identify these regions was no more than 100 individuals. Thus, this WGA/CTL method has significant potential to accelerate the discovery of minor H antigens that could be used in more selective, and thus more effective, allo-immune therapies in the near future.

Methods

Cell isolation and cell cultures

This study was approved by the institutional review board of the Aichi Cancer Center and the University of Tokyo. All blood or tissue samples were collected after written informed consent was obtained in accordance with the Declaration of Helsinki. B-LCLs were derived from allo-HSCT donors, recipients, and healthy volunteers. B-LCLs were maintained in RPMI 1640 medium supplemented with 10% fetal calf serum, 2 mM L-glutamine, 1 mM sodium pyruvate.

Generation of CTL lines and clones

CTL lines were generated from peripheral blood mononuclear cells (PBMCs) obtained after transplantation by stimulation with irradiated (33 Gy) recipient PBMCs harvested before HSCT, thereafter stimulated weekly in RPMI 1640 supplemented with 10% pooled human serum and 2 mM L-glutamine. IL-2 was added on days 1 and 5 after the second and third stimulations. CTL clones were isolated by standard limiting dilution and expanded as previously described.[7] CTL-1B9 was isolated from PBMCs harvested on day 30 after transplantation from a patient receiving a marrow graft from his HLA-identical sibling (HLA A11, A24, B39, B51, Cw7, Cw14), and CTL-2A12 has been described recently.[9]

Chromium release assay

Target cells were labeled with 0.1 mCi (3.7 MBq) of ^{51}Cr for 2 hours, and 10^3 target cells/well were mixed with CTL at the effector-to-target (E/T) ratio indicated in a standard 4-hour cytotoxicity. All assays were performed at least in duplicate. Percent specific lysis was calculated as follows: ((Experimental cpm − Spontaneous cpm) / (Maximum cpm − Spontaneous cpm)) × 100.

Immunophenotyping by enzyme-linked immunosorbent assay

B-LCL cells (20 000 per well, which had been retrovirally transduced with restriction HLA cDNA for individual CTLs, if necessary) were plated in each well of 96-well round-bottomed plates, and corresponding CTL clones (10 000 per well) were added to each well. After overnight incubation at 37°C, 50 μL supernatant was collected and released IFN-γ was measured by standard enzyme-linked immunosorbent assay (ELISA).

Construction of pooled DNA and microarray experiments

Genomic DNA was individually extracted from immunophenotyped B-LCLs. After DNA concentrations were measured and adjusted to 50 μg/mL using the PicoGreen dsDNA Quantitation Reagent (Molecular Probes, Eugene, OR), the DNA specimens from CTX^+ and CTX^- B-LCLs were separately combined to generate individual pools. DNA pools were analyzed in pairs using Affymetrix GeneChip SNP-genotyping microarrays (Affymetrix, Tokyo, Japan) according to the manufacturer's protocol,[19,20] where 2 independent experiments were performed for each array type (for more detailed statistical analysis for generated microarray data, see Document S1, available on the *Blood* website; see the Supplemental Materials link at the top of the online article).

Estimation of LD blocks

LD structures of the candidate loci were evaluated based on empirical data from the International Hap Map Project (http://www.hapmap.org/).[15] LD data for the relevant HapMap panels were downloaded from the HapMap web site and further analyzed using Haploview software (http://www.broad.mit.edu/mpg/haploview/).[21]

Transfection of 293T cells and ELISA

Twenty thousand 293T cells retrovirally transduced with HLA-A*2402 were plated in each well of 96-well flat-bottomed plates, cultured overnight at 37°C, then transfected with 0.12 μg of plasmid containing full-length *BCL2A1* cDNA generated from either the patient or his donor using Trans IT-293 (Mirus, Madison, WI). B-LCLs of the recipient and his donor were used as positive and negative controls, respectively. Ten thousand CTL-1B9 cells were added to each well 20 hours after transfection. After overnight incubation at 37°C, 50 μL of supernatant was collected and IFN-γ was measured by ELISA.

SNP identification by direct sequencing

Complementary DNA prepared from B-LCLs was polymerase chain reaction (PCR) amplified for the coding region of *BCL2A1* using the following primers: sense: 5′-AGAAGATGACAGACTGTGAATTTGG -3′; antisense: 5′-TCAACAGTATTGCTTCAGGAGAG-3′.

PCR products were purified and directly sequenced with the same primer and BigDye Terminator kit (version 3.1) by using ABI PRISM 3100 (Applied Biosystems, Foster City, CA).

Confirmatory SNP genotyping

Genotyping was carried out using fluorogenic 3′-minor groove binding (MGB) probes in a PCR assay. PCR was conducted in 10-μL reactions containing both allelic probes, 500 nM each of the primers, 1 × TaqMan Universal PCR Master Mix (Applied Biosystems), and 1 μL (100 ng) DNA. PCR cycling conditions were as follows: predenature, 50°C for 2 minutes, 95°C for 10 minutes, followed by 35 cycles of 92°C for 15 seconds and 60°C for 1 minute in a GeneAmp PCR System 9700 (Applied Biosystems). The PCR products were analyzed on an ABI 7900HT with the aid of SDS 2.2 software (Applied Biosystems).

Epitope reconstitution assay

The candidate *BCL2A1*-encoded minor H epitope and its allelic counterpart (DYLQYVLQI) peptides were synthesized by standard Fmoc chemistry. ^{51}Cr-labeled CTX^- donor B-LCLs were incubated with graded concentrations of the peptides and then used as targets in standard cytotoxicity assays.

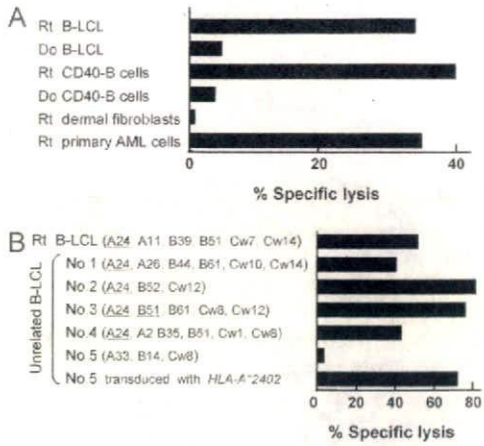

Figure 1. Specificity of CTL-1B9 against hematopoietic cells and its restriction HLA. (A) The cytolytic activity of CTL-1B9 was evaluated in a standard 4-hour ^{51}Cr release assay (E/T ratio, 20:1). Targets used were B-LCL, CD40-activated (CD40-B) B cells, dermal fibroblasts, and primary acute myeloid leukemia cells from the recipient (Rt), and B-LCL and CD40-B cells from his donor (Do). Rt dermal fibroblasts were pretreated with 500 U/mL IFN-γ and 10 ng/mL TNF-α for 48 hours before ^{51}Cr labeling. (B) Cytolytic activity of CTL-1B9 against a panel of B-LCLs derived from unrelated individuals, each of whom shared 1 or 2 class I MHC allele(s) with the recipient from whom the CTL-1B9 was generated. The shared HLA allele(s) with the recipient are underlined. B-LCLs (no. 5) which did not share any HLA alleles with the recipient, were retrovirally transduced with HLA-A*2402 cDNA and included to confirm HLA-A*2402 restriction by CTL-1B9. Results are typical of 2 experiments and data are the mean plus or minus the standard deviation (SD) of triplicates.

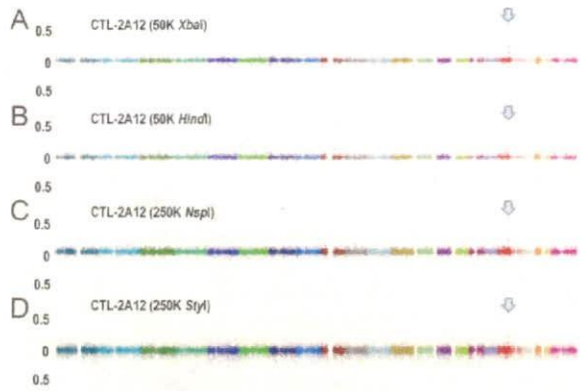

Figure 2. Whole genome association scans performed with pooled DNA generated based on immunophenotyping with CTL-2A12. Pooled DNAs generated from 44 CTX$^+$ and 44 CTX$^-$ B-LCLs were analyzed with 50 K *Xba*I (A), 50 K *Hind*III (B), 250 K *Nsp*I (C), and 250 K *Sty*I (D) arrays. Test statistics were calculated for all SNPs and plotted in the chromosomal order. In all SNP array types, a common association peak is observed at 18q21, to which the minor H antigen for CTL-2A12, encoded by the *HMSD* gene, had been mapped based on expression cloning[9] (arrows).

Results

CTL-based typing and SNP array analysis of pooled DNA

CTL-2A12 and CTL-1B9 are CTL clones established from the peripheral blood of 2 patients with leukemia who had received HLA-identical sibling HSCTs. Each clone demonstrated specific lysis against the B-LCLs of the recipient but not against donor B-LCLs, indicating recognition of minor H antigen (Figure 1A and Kawase et al[9]). The minor H antigen for CTL-2A12 had been previously identified by expression cloning[9]; on the other hand, the target minor H antigen for the HLA-A24–restricted CTL-1B9 clone, which was apparently hematopoietic lineage–specific (Figure 1A) and present in approximately 80% of the Japanese population (data not shown), had not yet been determined. Using these CTL clones, a panel of B-LCLs expressing the restriction HLA (HLA-B44 for CTL-2A12 and HLA-A24 for CTL-1B9) endogenously or retrovirally transduced, were subjected to "immunophenotyping" for the presence or absence of the minor H antigen by ELISA and, if necessary, by standard chromium release assay (CRA). Based on the assay results, for CTL-2A12 we initially collected 44 cytotoxicity-positive (CTX$^+$) and 44 cytotoxicity-negative (CTX$^-$) B-LCLs after screening 132 B-LCLs, while 57 CTX$^+$ and 38 CTX$^-$ B-LCLs were obtained from 121 B-LCLs for CTL-1B9. From these sets of B-LCL panels, pools of DNA were generated and subjected to analysis on Affymetrix GeneChip 100 K and 500 K microarrays in duplicate.[19,20]

Detection of association between minor H phenotypes and marker SNPs

Genetic mapping of the minor H locus was performed by identifying marker SNPs that showed statistically significant deviations in allele-frequencies between CTX$^+$ and CTX$^-$ pools based on the observed allele-specific signals in the microarray experiments. For this purpose, we evaluated the deviations of observed allele ratios between CTX$^+$ and CTX$^-$ pools for each SNP on a given array (Document S1). An SNP was considered as positive for association if its test statistic exceeded an empirically determined threshold that provided a "genome-wide" P value of .05 in duplicate experiments (Document S1, Figures S1,S2, and Table S1). Threshold values for different pool sizes are also provided in Table S2 for further experiments. The positive SNPs eventually obtained for both CTLs are summarized in Table 1, where the 10 SNPs showing the highest test statistics are listed for individual experiments.

Mapping of the minor H loci by WGASs

All the SNPs significantly associated with susceptibility to CTL-2A12 were correctly mapped within a single 115 kb LD block at chromosome 18q21 containing the *HMSD* gene (Figures 2 and 3A), which had been previously shown to encode the ACC-6 minor H antigen recognized by CTL-2A12.[9] According to the above criteria, no false-positive SNPs were reported in any array types (Table 1). Confirmation genotyping of individual B-LCLs from both panels revealed none of the 44 that had been immunophenotyped as CTX$^-$ were misjudged, while 8 of the 44 CTX$^+$ B-LCLs were found to actually carry no minor H-positive allele for ACC-6, which was likely due to the inclusion of individual B-LCLs showing borderline cytotoxicity (data not shown).

On the other hand, positive association of the target minor H antigen with CTL-1B9 was detected in 2 independent loci: SNP rs1879894 at 15q25.1 in 250 K *Nsp*I (Table 1, Figure 4A-B, and Figure S5) and SNP rs1842353 at 8q12.3 in 50 K *Hind*III (Table 1 and Figure S3A). We eventually focused on rs1879894, as it showed a much more significant genome-wide P value than SNP rs1842353 (Table 1). In contrast to the CTL-2A12 case, where many mutually correlated SNPs around the most significant one created a broad peak in the statistic plots (Figure 2 arrows and Figure S3), the adjacent SNPs (rs6495463 and rs2562756; Figure 3B solid arrows) around rs187894 (Figure 3B open arrow) did not show large test statistic values, reflecting the fact that no marker SNPs on 100 K and 500 K arrays exist in high LD (Figure 3B dashed red lines encompassing 26 kb) with this SNP according to the HapMap data. To further confirm the association, we generated additional B-LCL pools consisting of 75 CTX$^+$ and 34 CTX$^-$

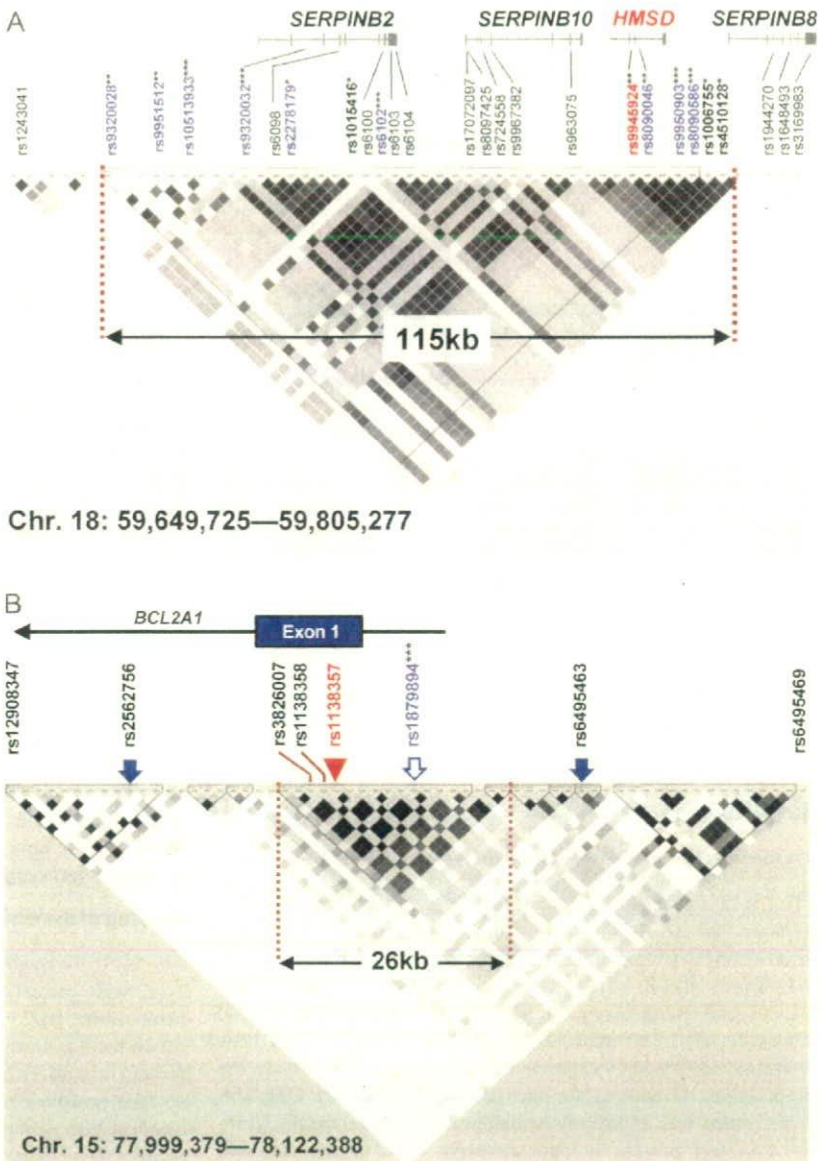

Figure 3. Linkage disequilibrium (LD) block mapped by CTL-2A12 and CTL-1B9. (A) An LD block map identified by pairwise r^2 plot from HapMap CEU data are overlayed with SNPs available from Affymetrix GeneChip SNP-genotyping microarrays (arrows) and 4 genes in the 115 kb block. SNPs that emerged repeatedly in the 2 independent experiments are indicated in blue. The genomewide *P* values for positive SNPs are shown as follows: $*P < .05$; $**P < .01$; $***P < .001$. The intronic SNP (rs9945924) controlling the alternative splicing of *HMSD* transcripts and expression of encoded ACC-6 minor H antigen is indicated in red. (B) LD blocks identified by pairwise r^2 plot from HapMap JPT data are overlayed with SNPs available from Affymetrix GeneChip SNP-genotyping microarrays (arrows) and exon 1 of the *BCL2A1* gene. The only SNP showing a high association with CTL-1B9 immunophenotypes (rs1879894) is shown as an open arrow. The nonsynonymous SNP (rs1138357) controlling the expression of the minor H antigen recognized by CTL-1B9 is indicated by a red arrowhead. ***SNP with genomewide $P < .001$. The 2 SNPs adjacent to the 26 kb LD block (rs2562756 and rs6495463) never gave a significant genomewide *P* value.

B-LCLs from another set of 128 B-LCLs, and performed a WGAS. As expected, the WGAS of the second pools also identified the identical SNP with the highest test statistic value in duplicate experiments, unequivocally indicating that this SNP is truly associated with the minor H locus of interest (Figure 4C,D and Table S3). The association was also detected when the references in the first and second pools were swapped (data not shown).

Identification of the minor H epitope recognized by CTL-1B9

The LD block containing SNP rs1879894 that was singled out from more than 500 000 SNP markers with 2 sets of DNA pools only encodes exon 1 of *BCL2A1* (Figure 3B). To our surprise, this was the region to which we had previously mapped an HLA-A24–restricted minor H antigen, ACC-1[Y].[7] We first confirmed that full-length *BCL2A1* cDNA cloned only from the recipient but not his donor could stimulate interferon-γ secretion from CTL-1B9 when transduced into donor B-LCL (Figure 5A), indicating that *BCL2A1* is a bona fide gene encoding minor H antigen recognized by CTL-1B9. We next genotyped 3 nonsynonymous SNPs in the *BCL2A1* exon 1 sequence (Figure 3B) and comparison was made between the genotypes and the susceptibility to CTL-1B9 of 9 HLA-A*2402$^+$ B-LCLs, including ones generated from the recipient (from whom CTL-1B9 was established) and his donor. Susceptibility to CTL-1B9 correlated completely with the presence of guanine at SNP rs1138357 (nucleotide position 238, according to the mRNA sequence for NM_004049.2) and thymine at SNP rs1138358 (nucleotide position 299) (Table 2), suggesting that the expression of the minor H epitope recognized by CTL-1B9 is controlled by either of these SNPs. We searched for nonameric amino acid sequences spanning the 2 SNPs using BIMAS software,[22] since most reported HLA-A*2402 binding peptides contain 9 amino acid residues.[23] Among these, a nonameric peptide, DYLQCVLQI (the polymorphic residue being underlined), has a predicted binding score of 75 and was considered as a candidate minor H epitope. As shown in Figure 5B, the DYLQCVLQI was strongly recognized by CTL-1B9, whereas its allelic counterpart,

Table 1. Positive SNPs from pooled DNA analysis

CTL-2A12, Exp 1				CTL-2A12, Exp 2				CTL-1B9, Exp 1				CTL-1B9, Exp 2			
rsID	Chr	Position	$\Delta R_{A\Delta R_B}$	rsID	Chr	Position	$\Delta R_{A\Delta R_B}$	rsID	Chr	Position	$\Delta R_{A\Delta R_B}$	rsID	Chr	Position	$\Delta R_{A\Delta R_B}$
50K X bal															
rs10513933	18	59699669	0.366*	rs10513933	18	59699669	0.511†	rs1363258	5	103297593	0.239	rs10499174	6	131209689	0.352*
rs9320028	18	59668150	0.255‡	rs9320028	18	59668150	0.360*	rs726083	3	67093729	0.203	rs30058	5	122325602	0.240
rs6102	18	59721450	0.221	rs10485873	7	3503743	0.157	rs639243	5	31392931	0.198	rs150724	16	61960443	0.213
rs724533	23	116440574	0.137	rs219323	14	59510440	0.150	rs1936461	10	56519024	0.186	rs1993129	8	63618836	0.208
rs1341112	6	104919391	0.136	rs10506892	12	82478539	0.147	rs763876	12	94922502	0.186	rs356946	13	69066751	0.201
rs470490	18	61182216	0.136	rs10492269	12	97786333	0.144	rs958404	7	133054441	0.179	rs2869268	4	86421898	0.184
rs2826718	21	21471423	0.134	rs10483466	14	35986827	0.139	rs10486727	7	41672315	0.178	rs287002	12	40312537	0.183
rs10506697	12	73241741	0.128	rs5910124	23	116408616	0.137	rs2833488	21	32010112	0.176	rs1146808	13	67688608	0.182
rs10506891	12	82393029	0.127	rs10512545	17	66337079	0.134	rs379212	5	60977687	0.172	rs10501287	11	42446011	0.180
rs308995	14	59657919	0.125	rs295678	5	58186928	0.131	rs1954004	14	58627872	0.170	rs564993	5	31393476	0.177
50K *Hind*III															
rs9320032	18	59712191	0.486†	rs9320032	18	59712191	0.506†	rs1842353	8	63617543	0.244*	rs9300692	13	101216476	0.225‡
rs8090046	18	59773066	0.207‡	rs8090046	18	59773066	0.245*	rs10521202	17	12755289	0.201‡	rs1842353	8	63617543	0.210‡
rs1474220	2	108525317	0.193‡	rs10498752	6	41876488	0.210‡	rs7899961	10	59696431	0.198‡	rs10520983	5	31314700	0.195‡
rs10498752	6	41876488	0.178	rs1941538	18	37994337	0.176	rs9320974	6	124421441	0.197‡	rs1334375	13	80897038	0.173
rs2298578	21	21632551	0.167	rs7682770	4	152748018	0.174	rs10520983	5	31314700	0.179	rs10519164	15	75412758	0.163
rs7516032	1	91618962	0.165	rs1445862	5	3675257	0.169	rs1862446	5	147460749	0.170	rs9322063	6	146852196	0.152
rs5030938	10	70645922	0.164	rs4696976	4	21058616	0.167	rs1358778	20	13266796	0.169	rs8067384	17	37926265	0.150
rs1883041	21	44921845	0.158	rs5030938	10	70645922	0.165	rs1873790	4	83422480	0.166	rs10521202	17	12755289	0.147
rs3902916	4	189045176	0.155	rs3902916	4	189045176	0.165	rs1220724	4	70888705	0.162	rs7914904	10	62749969	0.141
rs1000551	20	58709208	0.154	rs1883041	21	44921845	0.164	rs9300692	13	101216476	0.157	rs1220724	4	70888705	0.141
250K *Nsp*I															
rs9950903	18	59781783	0.534†	rs9950903	18	59781783	1.036†	rs1879894	15	78055874	0.846†	rs1879894	15	78055874	1.072†
rs1463835	3	23539615	0.532†	rs8090586	18	59781864	0.518†	rs9646294	16	6110019	0.484†	rs6771859	3	190642054	0.387†
rs16975459	18	37802275	0.383*	rs6473170	8	80664840	0.338*	rs17734332	5	134945240	0.365†	rs10512261	9	98804394	0.299*
rs8090586	18	59781864	0.367*	rs4510128	18	59782312	0.310‡	rs566619	7	41381538	0.345*	rs12122772	1	60384564	0.287*
rs16872621	4	22081055	0.312‡	rs1006755	18	59782026	0.300‡	rs17737566	6	50345280	0.310*	rs2153155	4	26034162	0.248‡
rs870582	6	125097114	0.301‡	rs7039378	9	118735938	0.258	rs3849955	9	28350374	0.285*	rs17126896	14	53320494	0.246‡
rs1015416	18	59720363	0.270‡	rs1860563	16	6418899	0.258	rs4616156	13	86581518	0.273*	rs1328652	13	35607527	0.240
rs2155907	11	97599883	0.227	rs4699126	4	105709109	0.212	rs2484698	1	217474460	0.263*	rs7021551	9	27446645	0.237
rs2112948	5	50994294	0.222	rs10275055	7	156212079	0.204	rs17139603	11	79638632	0.262*	rs252817	5	106752487	0.237
rs2919747	2	129681506	0.217	rs1526411	7	124658309	0.201	rs2156737	4	100642529	0.246‡	rs10772587	12	12681356	0.235
250K *Sty*I															
rs6102	18	59721450	0.597†	rs6102	18	59721450	0.495†	rs9383925	6	151975774	0.819†	rs201204	6	104842863	0.688†
rs9951512	18	59690885	0.374*	rs9945924	18	59771746	0.407*	rs6497397	16	19646258	0.311‡	rs12556155	23	108836419	0.442†
rs6496897	15	90493249	0.320‡	rs9951512	18	59690885	0.317‡	rs917252	7	22219990	0.289‡	rs4791422	17	10605304	0.435†
rs9945924	18	59771746	0.315‡	rs1983205	3	157782892	0.314‡	rs1019403	3	7823997	0.260‡	rs7749012	6	106459559	0.336*
rs12707805	8	107404746	0.303‡	rs950865	5	2720684	0.307‡	rs17053134	5	155373544	0.259‡	rs509951	5	31385483	0.308‡
rs10971778	9	33893184	0.296‡	rs2278179	18	59715512	0.292‡	rs11710880	3	72214965	0.246	rs16879024	8	32225711	0.256‡
rs6565076	16	81487818	0.294‡	rs10427722	22	36417752	0.289‡	rs17167866	7	13919264	0.237	rs2100054	15	75293482	0.252
rs2278179	18	59715512	0.291‡	rs17156659	7	82046820	0.271	rs10867062	9	137935241	0.237	rs11811023	1	143805934	0.240
rs7806238	7	29906442	0.290‡	rs4502324	18	4811261	0.262	rs5925800	23	23278707	0.235	rs17382798	15	75256074	0.231
rs965888	18	38062658	0.283‡	rs1348428	2	225927288	0.260	rs2255831	4	146614313	0.234	rs2030302	17	12526591	0.231

Significant SNPs that appeared on both experiments are underlined.
*Genomewide $P < .01$.
†Genomewide $P < .001$.
‡Genomewide $P < .05$.

DYLQ*Y*VLQI, was not. Decameric peptide, QDYLQ*C*VLQI, on the other hand, appeared to be weakly recognized; however, it is likely that the nonameric form was actually being presented after N-terminal glutamine cleavage by aminopeptidase in the culture medium. Because it was possible that the cystine might be cysteinylated, recognition of synthetic peptides DYLQCVLQI and cysteinylated DYLQC*VLQI were assayed using CTL-1B9. Half-maximal lysis for the former was obtained at a concentration of 200 pM, whereas recognition of the latter was several-fold weaker (Figure 5C). Thus, we concluded that DYLQCVLQI defines the cognate HLA-A*2402–restricted CTL-1B9 epitope, now designated ACC-1^C. This incidentally provides a second example of products from both dichotomous SNP alleles being recognized as HLA-A*2402–restricted minor H antigens, the first example being the HB-1 minor H antigen.[24] Finally, real-time quantitative PCR revealed that T cells carrying the complementarity-determining region 3 sequence identical to CTL-1B9 became detectable in the patient's blood at the frequencies of 0.22%, 0.91%, 1.07% and 0.01% among $TCR\alpha\beta^+$ T cells at days 30, 102, 196, and 395 after transplantation, respectively, suggesting that ACC-1^C minor H antigen is indeed immunogenic (Figure 5D).

Discussion

Recent reports have unequivocally demonstrated that WGASs can be successfully used to identify common variants involved in a wide variety of human diseases.[25-27] Our report represents a novel