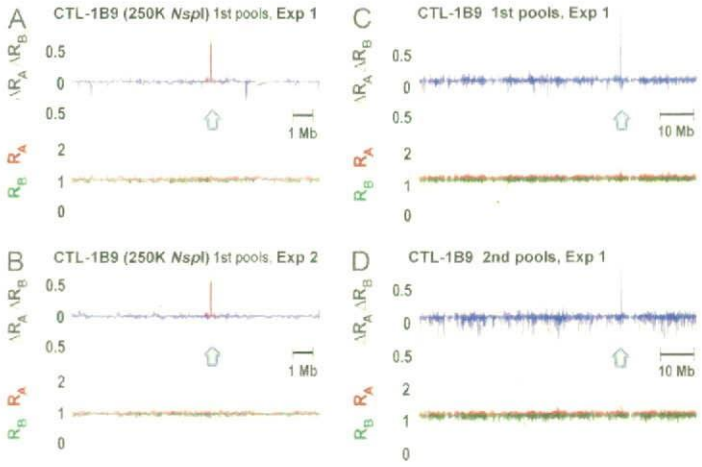

Figure 4. Reproducible detection of association with the immunophenotypes determined by CTL-1B9 at the *BCL2A1* locus. The maximum test statistic value was observed at a single SNP (rs1879894) within 15q25.1 in duplicate experiments for the first pools consisting of 57 CTX$^+$ and 38 CTX$^-$ B-LCLs (A-C). The peak association at the same SNP was reproduced in the experiments with the second pools consisting of 75 CTX$^+$ and 34 CTX$^-$ LCLs (D). Test statistic values ($\Delta R_A \Delta R_B$) are plotted by blue lines together with their R_A (red) and R_B (green) values. The expected $\Delta R_A \Delta R_B$ values multiplied by r^2 correlation coefficients for the adjacent SNPs within 500 kb from the SNP rs1879894 are overlaid by red lines (A,B).

application of WGASs to transplantation immunology, which provides a simple but robust method to fine-map the genetic loci of minor H antigens whose expression is readily determined by standard immunophenotyping with CTL clones established from patients who have undergone transplantation.

The current WGA/CTL method has several desirable features that should contribute to the acceleration of minor H locus mapping. In comparing the method to those of linkage analysis and other nongenetic approaches, including direct peptide sequencing of chemically purified minor H antigens[5,6,10,13] and conventional

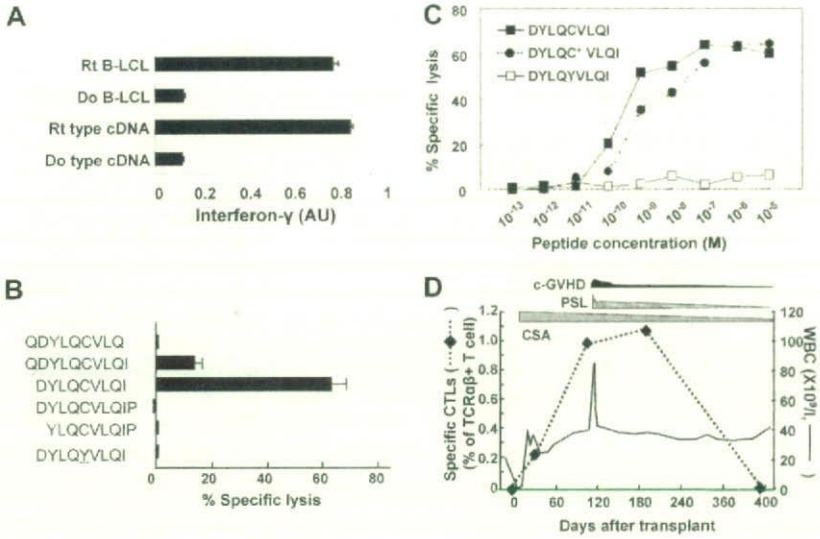

Figure 5. Identification of the CTL-1B9 minimal minor H epitope. (A) Interferon-γ production from CTL-1B9 against HLA-A*2402–transduced 293T cells transfected with plasmid encoding full-length *BCL2A1* cDNA cloned from either the recipient (Rt) from whom CTL-1B9 was isolated or his donor (Do). Rt B-LCL and Do B-LCL were used as positive and negative controls, respectively. Secreted interferon-γ was measured by ELISA and is expressed in arbitrary units (AUs) corresponding to optical density at 630 nm. Results are typical of 2 experiments and data are the mean plus or minus SD of triplicates. (B) A peptide reconstitution assay was conducted to determine the minimal epitope for CTL-1B9. Nonameric peptide (DYLQCVLQI), 2 nonameric peptides shifted by one amino acid to N- or C-terminus, N- and C-terminal extended decameric peptides, and its allelic counterpart (DYLQ*Y*VLQI) were synthesized and tested by adding to antigen-negative donor B-LCL at 10 nM in a standard ^{51}Cr release assay. Results are typical of 2 experiments and data are the mean plus or minus SD of triplicates. (C) Titration of the candidate minor H peptide by epitope reconstitution assay. Chromium-labeled donor B-LCLs were distributed to wells of 96-well round-bottomed plates, pulsed with serial dilutions of the indicated peptides for 30 minutes at room temperature, and then used as targets for CTL-1B9 in a standard ^{51}Cr release assay. A cysteinylated peptide (indicated by an asterisk) was included as an alternative form of the potential epitope. Results are typical of 2 experiments. (D) Tracking of ACC-1^C–specific T cells in the recipient's peripheral blood. In order to longitudinally analyze the kinetics of the ACC-1^C–specific CTLs in peripheral blood from the patient from whom CTL-1B9 was established, a real-time quantitative PCR was conducted. Complementary DNAs of peripheral blood mononuclear cells from the donor and patient before and after HSCT were prepared from the patient. Real-time PCR analysis was performed using a TaqMan assay as described previously.[9] The primers and fluorogenic probe sequences spanning the CTL-1B9 complementarity-determining region 3 (CDR3) were used to detect T cells carrying the CDR3 sequences identical to that of CTL-1B9. The primers and fluorogenic probe sequences spanning constant region of TCR beta chain (TCRBC) mRNA were used as internal control. Samples were quantified with the comparative CT method. The delta CT value was determined by subtracting the average CT value for TCRBC from the average CTL-1B9 CDR3 CT value. The standard curve for the proportion of CTL-1B9 among TCRαβ$^+$ T cells was composed by plotting mean delta CT values for each ratio, and the percentages of T cells carrying the CDR3 sequence identical to CTL-1B9 were calculated by using this standard curve. During this period, quiescent chronic GVHD, which required steroid treatment, developed; however, involvement of immune reaction to ACC-1^C minor H antigen was unlikely since its frequency increased even after resolution of most chronic GVHD symptoms. c-GVHD, chronic GVHD; CSA, cyclosporine A; PSL, prednisolone; WBC, white blood cell count.

Table 2. Correlation of *BCL2A1* sequence polymorphisms with susceptibility to CTL-1B9

	HLA-A*2402–positive B-LCLs								
	Rt	Do	UR1	UR2	UR3	UR4	UR5	UR6	UR7
Cytolysis by CTL-1B9	+	−	+	+	+	+	+	−	−
Detected SNP, position*									
rs1138357, 238	G/A	A	G	G	G/A	G/A	G/A	A	A
rs1138358, 299	T/G	G	T	T	T/G	T/G	T/G	G	G
rs3826007, 427	G	G/A	G	G	G	G	G/A	G/A	G

Rt indicates recipient; Do, donor; UR, unrelated; +, yes; and −, no.

*Nucleotide positions are shown according to the NM_004092.2 mRNA sequence, available at http://www.ncbi.nlm.nih.gov/ as GEO accession GSE10044.

expression cloning,[8,9,11] there are differences in terms of power, sensitivity, and specificity. Direct sequencing of minor H antigen peptide guarantees that the purified peptide is surely present on the cell surface as antigen, but it requires highly specialized equipment and personnel. Expression screening of cDNA libraries is also widely used and has become feasible with commercially available systems. However, it depends highly on the quality of the cDNA library and expression levels of the target genes. In addition, it often suffers from false-positive results due to the forced expression of cDNA clones under a strong promoter. The current method of WGA/CTL genetically determines the relevant minor H antigen locus, not relying on highly technical protein chemistry using specialized equipment, or repetitive cell cloning procedures. It is also not affected by the expression levels of the target antigens.

As a genetic approach, the current method based on genetic association has several advantages over conventional linkage analysis: the mapping resolution has been greatly improved from several Mb in the conventional linkage analysis to the average haplotype block size of less than 100 kb,[17,25-27] usually containing a handful of candidate genes, compared with the dozens as typically found in linkage analysis. This means that the effort needed for the subsequent epitope mapping will be substantially reduced. In fact, the 115 kb region identified for CTL-2A12 contains 4 genes compared with 38 genes as revealed by the previous linkage study (data not shown), and the candidate gene was uniquely identified within the 26 kb region for CTL-1B9, for which linkage analysis had failed due to very rare segregating pedigrees among the CEPH panels with this trait (now ACC-1[C]; data not shown).[15,16] In addition, before moving on to epitope mapping, it would be possible to evaluate the clinical relevance of the minor H antigens by examining the tissue distribution of their expression, based on widely available gene expression databases such as Genomic Institute of the Novartis Research Foundation (GNF, http://symatlas.gnf.org/SymAtlas/).[28]

Second, the required sample size is generally small, and should be typically no more than 100 B-LCLs for common minor H alleles. This is in marked contrast to the association studies for common diseases, in which frequently thousands of samples are required.[17,25-27] In the current approach, sufficiently high test statistic values could be obtained for the relevant loci with a relatively small sample size, since the minor H allele is correctly segregated between the CTX^+ and CTX^- pools by the highly specific immunologic assay. Combined with high accuracy in allelic measurements, this feature allows for the use of pooled DNAs in WGAS, which substantially saves cost and time, compared with the genotyping of individual samples. Unexpectedly, our method allows for a considerable degree of error in the immunophenotyping, indicating the robustness of the current method; in fact, the minor H locus for CTL-2A12 was successfully identified in spite of the presence of 8 (~10%) immunophenotyping errors. When the minor H allele has an extreme allele frequency (eg, $< 5\%$ or $> 95\%$), which could be predicted by preliminary immunophenotyping, WGAS/CTL may not be an efficient method of mapping, due to the impractically large numbers of B-LCLs that would need to be screened to obtain enough CTX^+ or CTX^- B-LCLs. However, such minor H antigens would likely have limited clinical impact or applicability.

Sensitivity of the microarray analysis seems to be very high when the target SNP has good proxy SNPs on the array, because we were able to correctly identify the single SNP correlated with the target of CTL-1B9 from more than 500 000 SNP markers. On the other hand, genome coverage of the microarray is definitely important. In our experiments on CTL-2A12, the association was successfully identified by the marker SNPs showing r^2 values of approximately 0.74 with the target locus of ACC-6. Since the GeneChip 500 K array set captures approximately 65% of all the HapMap phase II SNPs with more than 0.74 of r^2,[29] and higher coverage will be obtained with the SNP 6.0 arrays having more than 1 000 K SNP markers, these arrays can be satisfactorily used as platforms for the WGA/CTL method.

As shown in the current study, the intrinsic sensitivity and specificity of the WGA/CTL method in detecting associated SNPs were excellent. In other words, as long as target SNPs are captured in high r^2 values with one or more marker SNPs within the Affymetrix 500 K SNP set, there is a high likelihood of capturing the SNP with the current approach. To evaluate the probability of a given minor H antigen being captured in high r^2 with marker SNPs, we checked the maximum r^2 values of known minor H antigen SNPs with the Affymetrix 500 K SNPs, according to empirical data from the HapMap project (www.hapmap.org). Among 13 known minor H antigens, 7 have their entries (designated minor H SNP) in the HapMap phase II SNP set (HA-3,[30] HA-8,[31] HB-1,[11] ACC-1 and ACC-2,[7] LB-ADIR-1F,[10] and 7A7-PANE1[13]), and were used for this purpose (note that absence of their entries in the HapMap data set does not necessarily mean that they could not be captured by a particular marker SNP set). As shown in Table S4, all 7 minor H SNPs are captured by at least one flanking SNP that is included in the Affymetrix 500 K SNP set with r^2 values of more than 0.74 in at least one HapMap panel. The situation should be more favorable in the recently available SNP 6.0 array set with 1 000 K SNPs, indicating the genome coverage with currently available SNP arrays would be sufficient to capture typical minor H antigens with our approach.

Most patients who have received allo-HSCT could be a source of minor H antigen-specific CTL clones to be used for this assay, since the donor T cells are in vivo primed and many CTL clones could be established using currently available methods. In fact, substantial numbers of CTL clones have been established worldwide and could serve as the probes to identify novel minor H antigens.[32,33] Once constructed, a panel of B-LCLs, including those transduced with HLA cDNAs, could be commonly applied to immunophenotyping with different CTL clones, especially when

CTLs are obtained from the same ethnic group. In addition, by adopting other immunophenotyping readouts such as production of IL-2 from $CD4^+$ T cells, this method could be applied to identification of MHC class II–restricted minor H antigens which have crucial roles in controlling CTL functions upstream. This may open a new field in the study of allo-HSCT since MHC class II–restricted mHags have been technically difficult to identify by conventional methods.

Finally, the discovery of ACC-1^C as a novel minor H antigen indicates that all the mismatched transplants at this locus could be eligible for allo-immune therapies, since we have previously demonstrated that the counter allele also encodes a minor H antigen, ACC-1^Y, which is preferentially expressed and presented on blood components including leukemic cells and may serve as a target of allo-immunity.[7,34] Indeed, CTLs specific for ACC-2, an HLA-B44–restricted minor H antigen restricted by the third exonic SNP on *BCL2A1*,[7] was independently isolated from the peripheral blood of a patient with recurrent leukemia re-entering complete remission after donor lymphocyte infusion.[32] The number of eligible allo-HSCT recipients has now been effectively doubled, accounting for 50% of transplants with HLA-A24 or 20% of all transplantations performed in the Asian population. In conclusion, we have described a simple but powerful method for minor H mapping to efficiently accelerate the discovery of novel minor H antigens that will be needed to contribute to our understanding of the molecular mechanism of human allo-immunity.

Acknowledgments

The authors thank Dr W. Ho for critically reading the manuscript; Dr Keitaro Matsuo, Dr Hiroo Saji, Dr Etsuko Maruya, Dr Mamoru Ito, Ms Keiko Nishida, Dr Ayako Demachi-Okamura, Ms Yasuko Ogino, Ms Hiromi Tamaki, and the staff members of the transplant center, donor centers, and the Japan Marrow Donor Program office for their generous cooperation and expert technical assistance.

This study was supported in part by Scientific Research on Priority Areas (B01; no.17016089) from the Ministry of Education, Culture, Science, Sports, and Technology, Japan; Research on Human Genome, Tissue Engineering Food Biotechnology and the Second and Third Team Comprehensive 10-year Strategy for Cancer Control (no. 26), from the Ministry of Health, Labor, and Welfare, Japan; and a Grant-in-Aid from Core Research for Evolutional Science and Technology (CREST) of Japan.

Authorship

Contribution: T.K. performed most immunologic experiments and preparation of pooled DNA and quantative PCR, analyzed data, and wrote the manuscript; Y.N. performed the majority of genetic analyses and analyzed the data; H.T. performed T-cell receptor analysis and designed q-PCR primers and probes; G.Y. contributed to the organization of software for linkage analysis and simulation; S.M. prepared the pooled DNA; M.O., K.M., Y.K. and Y.M. collected clinical data and specimens; T.T. and K.K. contributed to data analysis and interpretation, and to the writing of the article; S.O. and Y.A. supervised the entire project, designed and coordinated most of the experiments in this study, contributed to manuscript preparation, and are senior coauthors.

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

Correspondence: Seishi Ogawa, Department of Hematology and Oncology, Department of Regeneration Medicine for Hematopoiesis, The 21st Century COE Program, 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 Yoshiki Akatsuka, Division of Immunology, Aichi Cancer Center Research Institute, 1-1 Kanokoden, Chikusa-ku, Nagoya 464-8681, Japan; e-mail: yakatsuk@aichi-cc.jp.

References

1. Thomas ED Sr. Stem cell transplantation: past, present and future. Stem Cells. 1994;12:539-544.
2. Childs RW, Barrett J. Nonmyeloablative allogeneic immunotherapy for solid tumors. Annu Rev Med. 2004;55:459-475.
3. Goulmy E. Human minor histocompatibility antigens: new concepts for marrow transplantation and adoptive immunotherapy. Immunol Rev. 1997;157:125-140.
4. Bleakley M, Riddell SR. Molecules and mechanisms of the graft-versus-leukaemia effect. Nat Rev Cancer. 2004;4:371-380.
5. den Haan JM, Meadows LM, Wang W, et al. The minor histocompatibility antigen HA-1: a diallelic gene with a single amino acid polymorphism. Science. 1998;279:1054-1057.
6. Pierce RA, Field ED, Mutis T, et al. The HA-2 minor histocompatibility antigen is derived from a diallelic gene encoding a novel human class I myosin protein. J Immunol. 2001;167:3223-3230.
7. Akatsuka Y, Nishida T, Kondo E, et al. Identification of a polymorphic gene, BCL2A1, encoding two novel hematopoietic lineage-specific minor histocompatibility antigens. J Exp Med. 2003;197:1489-1500.
8. Warren EH, Vigneron NJ, Gavin MA, et al. An antigen produced by splicing of noncontiguous peptides in the reverse order. Science. 2006;313:1444-1447.
9. Kawase T, Akatsuka Y, Torikai H, et al. Alternative splicing due to an intronic SNP in HMSD generates a novel minor histocompatibility antigen. Blood. 2007;110:055-1063.
10. van Bergen CA, Kester MG, Jedema I, et al. Multiple myeloma-reactive T cells recognize an activation-induced minor histocompatibility antigen encoded by the ATP-dependent interferon-responsive (ADIR) gene. Blood. 2007;109:4089-4096.
11. Dolstra H, Fredrix H, Maas F, et al. A human minor histocompatibility antigen specific for B cell acute lymphoblastic leukemia. J Exp Med. 1999;189:301-308.
12. de Rijke B, van Horssen-Zoetbrood A, Beekman JM, et al. A frameshift polymorphism in P2X5 elicits an allogeneic cytotoxic T lymphocyte response associated with remission of chronic myeloid leukemia. J Clin Invest. 2005;115:3506-3516.
13. Brickner AG, Evans AM, Mito JK, et al. The PANE1 gene encodes a novel human minor histocompatibility antigen that is selectively expressed in B-lymphoid cells and B-CLL. Blood. 2006;107:3779-3786.
14. Warren EH, Otterud BE, Linterman RW, et al. Feasibility of using genetic linkage analysis to identify the genes encoding T cell-defined minor histocompatibility antigens. Tissue Antigens. 2002;59:293-303.
15. Consortium TIH. The International HapMap Project. Nature. 2003;426:789-796.
16. Consortium TIH. A haplotype map of the human genome. Nature. 2005;437:1299-1320.
17. Risch N, Merikangas K. The future of genetic studies of complex human diseases. Science. 1996;273:1516-1517.
18. Hirschhorn JN, Daly MJ. Genome-wide association studies for common diseases and complex traits. Nat Rev Genet. 2005;6:95-108.
19. Kennedy GC, Matsuzaki H, Dong S, et al. Large-scale genotyping of complex DNA. Nat Biotechnol. 2003;21:1233-1237.
20. 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.
21. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics. 2005;21:263-265.
22. Parker KC, Bednarek MA, Coligan JE. Scheme for ranking potential HLA-A2 binding peptides based on independent binding of individual peptide side-chains. J Immunol. 1994;152:163-175.
23. Kubo RT, Sette A, Grey HM, et al. Definition of specific peptide motifs for four major HLA-A alleles. J Immunol. 1994;152:3913-3924.
24. Dolstra H, de Rijke B, Fredrix H, et al. Bi-directional allelic recognition of the human minor histocompatibility antigen HB-1 by cytotoxic T lymphocytes. Eur J Immunol. 2002;32:2748-2758.
25. Easton DF, Pooley KA, Dunning AM, et al. Genome-wide association study identifies novel breast cancer susceptibility loci. Nature. 2007;447:1087-1093.
26. Gudmundsson J, Sulem P, Manolescu A, et al.

Genome-wide association study identifies a second prostate cancer susceptibility variant at 8q24. Nat Genet. 2007;39:631-637.

27. Zeggini E, Weedon MN, Lindgren CM, et al. Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. Science. 2007;316:1336-1341.
28. Su AI, Cooke MP, Ching KA, et al. Large-scale analysis of the human and mouse transcriptomes. Proc Natl Acad Sci U S A. 2002;99:4465-4470.
29. Nannya Y, Taura K, Kurokawa M, Chiba S, Ogawa S. Evolution of genome-wide power of genetic association studies based on empirical data from the HapMap project. Hum Mol Genet. 2007;16:3494-3505.
30. Spierings E, Brickner AG, Caldwell JA, et al. The minor histocompatibility antigen HA-3 arises from differential proteasome-mediated cleavage of the lymphoid blast crisis (Lbc) oncoprotein. Blood. 2003;102:621-629.
31. Brickner AG, Warren EH, Caldwell JA, et al. The immunogenicity of a new human minor histocompatibility antigen results from differential antigen processing. J Exp Med. 2001;193:195-206.
32. Kloosterboer FM, van Luxemburg-Heijs SA, van Soest RA, et al. Minor histocompatibility antigen-specific T cells with multiple distinct specificities can be isolated by direct cloning of IFNgamma-secreting T cells from patients with relapsed leukemia responding to donor lymphocyte infusion. Leukemia. 2005;19:83-90.
33. Tykodi SS, Warren EH, Thompson JA, et al. Allogeneic hematopoietic cell transplantation for metastatic renal cell carcinoma after nonmyeloablative conditioning: toxicity, clinical response, and immunological response to minor histocompatibility antigens. Clin Cancer Res. 2004;10:7799-7811.
34. Kenny JJ, Knobloch TJ, Augustus M, Carter KC, Rosen CA, Lang JC. GRS, a novel member of the Bcl-2 gene family, is highly expressed in multiple cancer cell lines and in normal leukocytes. Oncogene. 1997;14:997-1001.

Frequent genomic abnormalities in acute myeloid leukemia/myelodysplastic syndrome with normal karyotype

Tadayuki Akagi,[1#] Seishi Ogawa,[2,3,4] Martin Dugas,[5] Norihiko Kawamata,[1] Go Yamamoto,[2] Yasuhito Nannya,[2] Masashi Sanada,[3,4] Carl W. Miller,[1] Amanda Yung,[1] Susanne Schnittger,[6] Torsten Haferlach,[6] Claudia Haferlach,[6] and H. Phillip Koeffler[1]

[1]Division of Hematology and Oncology, Cedars-Sinai Medical Center, UCLA School of Medicine, Los Angeles, CA, USA; [2]Department of Hematology and Oncology, and [3]Department of Cell Therapy and Transplantation Medicine and the 21st Century COE Program, Graduate School of Medicine, University of Tokyo, Tokyo, Japan; [4]Core Research for Evolutional Science and Technology, Japan Science and Technology Agency, Tokyo, Japan; [5]Department of Medical Informatics and Biomathematics, University of Munster, Munster, Germany; [6]MLL Munich Leukemia Laboratory, Munich, Germany

CH and HPK are co-last authors.

[#]Current address: Department of Stem Cell Biology, Graduate School of Medical Science, Kanazawa University. 13-1 Takara-machi, Kanazawa, Ishikawa 920-8640, Japan

Acknowledgments: we thank members of our laboratory for helpful discussions.

Funding: this work was supported by NIH grants as well as the Parker Hughes Fund. HPK is the holder of the Mark Goodson endowed Chair in Oncology Research and is a member of the Jonsson Cancer Center and the Molecular Biology Institute, UCLA. MD and TH are supported by the European Leukemia Network (funded by the 6th Framework Program of the European Community). The study is dedicated to the memory of David Golde, a mentor and friend.

Manuscript received March 5, 2008. Revised version arrived September 17, 2008. Manuscript accepted October 6, 2008.

Correspondence:
Tadayuki Akagi, Ph.D, Division of Hematology and Oncology, Cedars-Sinai Medical Center, UCLA School of Medicine 8700 Beverly Blvd, Los Angeles, CA90048, USA.
E-mail: tadayuki@staff.kanazawa-u.ac.jp

The online version of this article contains a supplementary appendix.

ABSTRACT

Background

Acute myeloid leukemia is a clonal hematopoietic malignant disease; about 45-50% of cases do not have detectable chromosomal abnormalities. Here, we identified hidden genomic alterations and novel disease-related regions in normal karyotype acute myeloid leukemia/myelodysplastic syndrome samples.

Design and Methods

Thirty-eight normal karyotype acute myeloid leukemia/myelodysplastic syndrome samples were analyzed with high-density single-nucleotide polymorphism microarray using a new algorithm: allele-specific copy-number analysis using anonymous references (AsCNAR). Expression of mRNA in these samples was determined by mRNA microarray analysis.

Results

Eighteen samples (49%) showed either one or more genomic abnormalities including duplication, deletion and copy-number neutral loss of heterozygosity. Importantly, 12 patients (32%) had copy-number neutral loss of heterozygosity, causing duplication of either mutant *FLT3* (2 cases), *JAK2* (1 case) or *AML1/RUNX1* (1 case); and each had loss of the normal allele. Nine patients (24%) had small copy-number changes (< 10 Mb) including deletions of *NF1*, *ETV6/TEL*, *CDKN2A* and *CDKN2B*. Interestingly, mRNA microarray analysis showed a relationship between chromosomal changes and mRNA expression levels: loss or gain of chromosomes led, respectively, to either a decrease or increase of mRNA expression of genes in the region.

Conclusions

This study suggests that at least one half of cases of normal karyotype acute myeloid leukemia/myelodysplastic syndrome have readily identifiable genomic abnormalities, as found by our analysis; the high frequency of copy-number neutral loss of heterozygosity is especially notable.

Key words: normal karyotype acute myeloid leukemia/myelodysplastic syndrome, SNP-chip, CNN-LOH.

Citation: Akagi T, Ogawa S, Dugas M, Kawamata N, Yamamoto G, Nannya Y, Sanada M, Miller CW, Yung A, Schnittger S, Haferlach T, Haferlach C, and Koeffler HP. Frequent genomic abnormalities in acute myeloid leukemia/myelodysplastic syndrome with normal karyotype. Haematologica 2009; Haematologica 2009; 94:213-223. doi:10.3324/haematol.13024

Introduction

Acute myeloid leukemia (AML) is a clonal malignant hematopoietic disease characterized by a block in differentiation, resulting in accumulation of immature myeloid cells.[1,2] Karyotypic analyses have revealed several frequent chromosomal translocations producing fusion genes associated with AML. The t(8;21)(q22;q22) translocation is one of these abnormal karyotypes, and this translocation produces *AML1-ETO* fusion products.[3,4] The *AML1-ETO* blocks hematopoietic differentiation and enhances self-renewal of human and murine hematopoietic stem cells.[5,6] The fusion product apparently binds to *AML1* target genes and represses their transcription.[5,6] The inv(16)(p13q22) or t(16;16)(p13;q22) produces the leukemogenic *CBFB-MYH11* fusion gene which blocks differentiation of hematopoietic stem cells by inhibiting the function of *AML1*.[7,8] Acute promyelocytic leukemia cells usually have t(15;17)(q22;q11-21) producing *PML-RARA* fusion products which also behave as a transcriptional repressor.[9,10] Other frequent translocations include t(9;11), t(6;11), inv(3)/t(3;3) and t(6;9).[11] Trisomy 8, 11, 13, 21 and 22, and deletion of chromosome 5/5q, 7/7q, 17/17p and 20/20q also occur moderately frequently.[2,11,12] About 45-50% of AML patients have no detectable chromosomal abnormalities.[13,14] In general, these individuals with a normal karyotype in their leukemic cells show an intermediate prognosis.[13,14]

Besides chromosomal abnormalities, the leukemic cells can have a variety of mutations involving individual genes. Activating mutations of the receptor tyrosine kinase, FMS-like tyrosine kinase 3 (*FLT3*) occur in about 30% AML patients; two major mutant forms occur: an internal tandem duplication (ITD) or a point mutation in the tyrosine kinase domain (TKD).[15] Activating mutations at codon 12, 13 or 61 of either the *NRAS* or *KRAS* occur in 25% and 15% of AML patients, respectively.[1,16] About 10-15% of AML samples have inactivating mutations of C/EBPα whose wild-type function is to enhance differentiation.[17,18] Nucleophosmin1 (NPM1) is mutated in 50-60% of AML samples with normal karyotype.[13,19] This protein has an important role in ribosome biogenesis, including nuclear export of ribosomal proteins. Mutant NPM1 has an aberrant nuclear export signal and remains localized in the cytoplasm.[20]

Single-nucleotide polymorphism microarray (SNP-chip) analysis is a new technique to examine the genome including any copy-number changes and loss of heterozygosity (LOH).[21-25] Importantly, SNP-chip analysis can reveal cryptic abnormalities such as a small copy-number changes (< 10 Mb) or copy-number neutral loss of heterozygosity [CNN-LOH, also called uniparental disomy (UPD)] that cannot be detected by karyotype analysis. In addition, comparative genomic hybridization cannot detect CNN-LOH. SNP-chip analysis has been used in chronic lymphocytic leukemia,[24,25] childhood acute lymphoblastic leukemia,[26,27] juvenile myelomonocytic leukemia,[28] follicular lymphoma,[29] multiple myeloma,[30] and AML.[31,32,50,51]

In the present study, we identified hidden abnormalities and novel disease-related genomic regions using 250 K SNP-chip analysis in samples from patients with normal karyotype AML/myelodysplastic syndrome (MDS). The use of CNAG (copy-number analysis for Affymetrix GeneChips) program[21] and a new algorithm AsCNAR (allele-specific copy-number analysis using anonymous references)[25] provided a highly sensitive technique to detect CNN-LOH, as well as, copy-number changes in AML/MDS genomes.

Design and Methods

Patients' samples

Samples from 30 anonymized patients with normal karyotype AML and 8 anonymized patients with normal karyotype MDS (age, 33-88 years; median, 62 years) were examined. These samples were isolated from bone marrow at diagnosis. The patients' age, gender, diagnosis, white blood cell count (WBC), karyotype and additional mutations of *FLT3* and *NPM1* are summarized in Table 1. This study was approved by Cedars-Sinai Medical Center (IRB number 4485).

High-density SNP-chip analysis

Genomic DNA was isolated from AML/MDS cells, and the DNA was subjected to GeneChip Human mapping 250 K array NspI microarray (SNP-chip, Affymetrix, Santa Clara, CA, USA) as described previously.[21,25] Hybridization, washing and signal detection were performed on GeneChip Fluidics Station 400 and GeneChip scanner 3000 according to the manufacturer's protocols (Affymetrix). Microarray data were analyzed for determination of both total and allelic-specific copy-number using the CNAG program as previously described[21,25] with minor modifications; the status of copy-numbers as well as CNN-LOH at each SNP was inferred using the algorithms based on hidden Markov models.[21,25] GNAGraph software was used for clustering of AML samples with regards to their copy-number changes, as well as CNN-LOH.[27] Size, position and location of genes were identified with UCSC Genome Browser *http://genome.ucsc.edu*. Copy-number changes, including duplication and deletion, were identified by allele-specific CNAG software.[25,27] These copy-number changes include copy-number variant and physiological deletion at the immunoglobulin and T-cell receptor genes. Copy-number variants as described previously at *http://projects.tcag.ca/variation* and physiological deletions were eliminated manually, and other regions detected by allele-specific CNAG software are listed on Table 4.

Fluorescence in situ hybridization analysis

Bone marrow samples from AML patients were used for interphase fluorescence *in situ* hybridization (FISH) analysis. The FISH studies were performed using the following probes: D5S721 (5p15.2), D5S23 (5p15.2), D7Z1 (centromere of chromosome 7), ABL (9q34.12), EGR1 (5q31.2), D7S486 (7q31), TP53 (17p13.1), D8Z2 (centromere of chromosome 8), AML1 (21q22.12) and BCR (22q11.23) (ABBOTT/VYSIS, Des Plaines, IL, USA). Probes for the 12p13 region [fluorescein-labeled ETV6-

downstream region (483 kb-length) and Texas-red-labeled ETV6-upstream region (264 kb-length)] were used for FISH analysis in case #5. The ETV6 probes were obtained from ABBOTT/VYSIS.

Determination of SNP sequences, JAK2, FLT3, NPM1, and AML1/RUNX1 mutations, and other target genes in cases of CNN-LOH

To determine the SNP sequences, (SNP identities are rs7747259, rs1122637, rs9505293, rs6934027, rs280153 and rs191986) in case #38 chromosome 6p region, the genomic region of each SNP site was amplified by genomic polymerase chain reaction (PCR) using specific primers. For determination of *JAK2* V617F mutation in case #20, genomic PCR was performed with specific primers. PCR products were purified and sequenced. The sequences of the primers are shown in *Online Supplementary Tables S1 and S2*. To determine the *FLT3*-ITD mutation, the PCR reaction was performed with specific primers, and the PCR products were separated on a 2.0% agarose gel stained with ethidium bromide as described previously[34,35] Mutations at exon 12 of the *NPM1* gene were determined using a melting curve-based LightCycler assay (Roche Diagnostics, Mannheim, Germany).[36] Denaturing high-performance liquid chromatography analysis was performed to determine the *AML1/RUNX1* mutation in case #17 as described previously.[37] Alterations of several tyrosine kinase genes including *FGR* (case #3 and #23), *DDR1* (case #2 and 38), *TYK2* (case #2), *MATK* (case #2), *FER* (case #8) and *FGFR4* (case #8) were determined by either nucleotide sequencing of their exons and/or band-shifts of PCR products of exons after their electrophoresis and visualization on a gel (single strand conformation polymorphism), as described previously[38] with minor modifications. The PCR reaction contained genomic DNA, 500 nM of each of the primers, 200 nM of each of the dNTP, 0.5 units of Taq DNA polymerase and 3 μCi [α-32P] dCTP in 20 μL PCR products were diluted 10-fold in the loading buffer (10 mM NaOH, 95% formamide, and 0.05% of both bromophenol blue and xylene cyanol). After denaturation at 94°C for 5 min, 2 mL of the samples were loaded onto a 6% non-denaturating polyacrylamide mutation detection enhancement gel (Bioproducts, Rockland, ME, USA) with 10% (v/v) glycerol and separated at 300 V for 20 h. The gel was dried and subjected to autoradiography.

Table 1. **Baseline clinical characteristics of 38 cases of normal karyotype acute myeloid leukemia/myelodysplastic syndrome.**

Group	Case #	Gender	Age	Type	WBC x10^9/L	*FLT3*	*NPM1*	Karyotype
A	29	M	49	AML M0	13.8	–	–	46,XY
	1	F	33	AML M1	3.5	–	–	46,XX
	14	M	43	AML M1	27.9	–	–	46,XY
	15	F	67	AML M1	365	–	+	46,XX
	6	M	66	AML M2	2.5	–	–	46,XY
	24	M	88	AML M2	6.3	–	–	46,XY
	25	F	61	AML M2	11.9	–	+	46,XX
	33	F	60	AML M2	24.5	+	+	46,XX
	44	M	65	AML M2	62.8	–	–	46,XY
	18	F	43	AML M4	43.1	–	+	46,XX
	19	M	37	AML M4	209	–	+	46,XY
	32	M	45	AML M4	9.5	–	+	46,XY
	34	F	80	AML M4	71.1	+	–	46,XX
	31	F	45	AML M5b	12.7	+	+	46,XX
	16	F	77	MDS RA	5.3	–	–	46,XX
	35	F	68	MDS CMML-1	5.6	–	–	46,XX
	36	F	69	MDS RAEB-1	4.6	–	–	46,XX
	39	F	74	MDS RAEB-1[1]	5.4	–	–	46,XX
	42	F	79	MDS	3.2	–	–	46,XX
B	10	F	49	AML M1	17.0	+	–	46,XX
	4	M	76	AML M2	2.3	–	–	46,XY
	5	F	67	AML M2	1.1	–	–	46,XX
	8	F	75	AML M2	1.8	–	–	46,XX
	9	M	65	AML M2	48.3	–	–	46,XY
	17	M	78	AML M2	1.1	–	–	46,XY
	20	F	65	AML M2	34.3	–	–	46,XX
	23	M	69	AML M2	2.0	–	+	46,XY
	26	M	36	AML M2	14.2	–	+	46,XY
	2	F	71	AML M4	1.9	–	+	46,XX
	7	F	85	AML M4	2.4	–	–	46,XX
	21	F	38	AML M4	37.6	+	+	46,XX
	38	M	53	AML M4	40	+	+	46,XY
	3	F	67	AML M5b	56	+	+	46,XX
	37	F	65	AML M5b	72.8	+	+	46,XX
	12	M	69	MDS RA	11.5	–	–	46,XY
	41	M	77	MDS RA[1]	7.1	–	–	46,XY
	13	F	54	MDS RAEB-2	–	–	–	46,XX
	11	F	60	t-AML M2	1.7	–	–	46,XX

t-AML: therapy-related AML; RA: refractory anemia; RAEB-1 or -2: refractory anemia with excess blasts subtype-1 or -2; CMML: chronic myelomonocytic leukemia; [1]new WHO classification.

Quantitative real-time polymerase chain reaction

Gene-dosages of chromosome 6p24.3 in case #38, and the *MYC* and *CDKN2A* genes in case #20 were determined by quantitative real-time PCR (iCycler, Bio-Rad, Hercules, CA, USA) using Sybr Green. To determine the relative gene dosage of each sample, the chromosome 2p21 region was measured as a control.[27] The copy-number of the 2p21 region was normal, as determined by SNP-chip analysis, in these samples. The delta threshold cycle value (ΔCt) was calculated from the given Ct value by the formula ΔCt = (Ct sample - Ct control). The fold change was calculated as $2^{-\Delta Ct}$. Primer sequences are shown in *Online Supplementary Table S2*.

Gene expression microarray analysis

Total RNA was isolated from AML/MDS cells and processed according to Affymetrix guidelines for analysis with HGU133 Plus 2.0 microarrays. Data were analyzed with R version 2.5.0 using Bioconductor version 2.0.[39] Data were normalized using the robust multi-array average procedure.[39] Since most regions that showed chromosomal abnormalities were not recurring, we were not able to compare individual genes across samples with statistical tests. To assess plausibility of large deletions and amplifications, we subtracted

from each gene (in the respective region) mean expression of this gene in other cases: case #11 was compared with 37 normal karyotype AML/MDS cases; and cases #20, #4 and #5 were compared with other normal karyotype AML/MDS samples. We then calculated a mean expression difference for each region and considered a value below zero to be consistent with deletion and a value above zero to be consistent with amplification.

Results

Proof of principal

To identify hidden abnormalities in AML/MDS with a normal karyotype, 37 samples were analyzed by 250K SNP-chip microarray. One additional case (case #11) had only 13 metaphases and chromosomal abnormalities were not detected on karyotypic analysis; this sample did, however, have numerous genetic abnormalities identified by SNP-chip including hemizygous deletions of 3p25.1-p24.3 (2.29 Mb), 3p24.2-p24.1 (3.96 Mb), 3p23-q12.1 (66.55 Mb), 5q11.2-q-terminal (124.89 Mb), 7q11.23-q36.1 (76.04 Mb), 7q36.2 (0.78 Mb), 11q23.3-q-terminal (18.24 Mb), 17p-terminal-q11.1 (22.48 Mb), and 17q11.2-q12 (4.42 Mb); duplications of 3p24.3 (2.14 Mb), 5p15.31 (1.83 Mb), and 5p14.3-q11.2 (35.53 Mb); and trisomy of chromosomes 8, 21 and 22 (Table 2). To confirm these SNP-chip results, we performed extensive FISH analysis. The number of signals for probes D5S721 (5p15.2), D5S23 (5p15.2), D7Z1 (centromere of chromosome 7) and ABL (9q34.12) was two, and SNP-chip analysis also showed normal copy number (2n) consistent with the SNP-chip data. The EGR1 (5q31.2), D7S486 (7q31) and TP53 (17p13.1) probes revealed one signal; and these regions also showed hemizygous deletion (1n) by SNP-chip analysis. D8Z2 (centromere of chromosome 8), AML1 (21q22.12) and BCR (22q11.23) probes showed three or four signals, and SNP-chip analysis also indicated trisomy (3n) of these chromosomes. Chromosome 9 was normal by both SNP-chip and FISH analyses. As summarized in *Online Supplementary Table S3*, the results of SNP-chip and FISH analyses were completely congruent. Taken together, these results suggest that SNP-chip analysis reflected the genomic changes.

SNP-chip analysis in 37 normal karyotype acute myeloid leukemia/myelodysplastic syndrome samples

SNP-chip analysis of samples from 37 patients with normal karyotype AML/MDS revealed several genomic copy-number changes, as well as CNN-LOH. Nineteen patients (51%) had a normal genome by SNP-chip analysis (group A). In contrast, 18 patients (49%) had one or more genomic abnormalities (group B) (Figure 1). Deletions and/or duplications were found in nine patients (24%). Twelve patients (32%) had CNN-LOH. In group B, 14 cases (78% of the 18 samples) had only one genomic change; one case (6%) had two genomic abnormalities (case #5); two cases (11%) had three changes (case #2 and #4) and one case (5%) had four genomic alterations (case #20).

We also compared the relationship between the genomic changes and the French-American British classification of the 15 AML and 3 MDS samples in group B. In the AML samples, 11 cases had CNN-LOH, three cases had a duplication and seven cases had a deletion. The one AML M1 sample (case #10) had CNN-LOH; and the two AML M5b samples (cases #3 and #37) had CNN-LOH in their chromosomes. In the four AML M4 samples, cases #38, #21 and #2 had CNN-LOH, and cases #2 and #7 had a small deletion. In eight AML M2 samples, five (cases #4, #8, #17, #20, and #23) had CNN-LOH, three (cases #4, #5 and #20) had a duplication, and five (cases # 4, #5, #9, #20 and #26) had a deletion. In the three MDS samples, one sample (case #12) had CNN-LOH, and two samples (cases #13 and #41) had a deletion (Figure 1, Tables 3 and 4). Taken together, these results show that the patients who were categorized as having normal karyotype AML/MDS had easily recognizable deletions, duplications and/or CNN-LOH of their genome.

Table 2. Copy-number changes in case #11 detected by SNP-chip analysis.

Chromosome	Location	Physical localization		Size (Mb)	Status
		Proximal	Distal		
3	3p25.1-p24.3	16,389,202	18,675,075	2.29	Del
	3p24.3	19,589,378	21,731,557	2.14	Dup
	3p24.2-p24.1	24,881,910	28,844,599	3.96	Del
	3p23-q12.1	33,278,003	99,828,897	66.55	Del
5	5p15.31	7,616,335	9,443,217	1.83	Dup
	5p14.3-q11.2	18,603,838	54,129,781	35.53	Dup
	5q11.2-q-ter.	55,738,905	180,629,495	124.89	Del
7	7q11.23-q36.1	71,659,926	147,695,696	76.04	Del
	7q36.2	152,027,450	152,806,031	0.78	Del
8	Trisomy				
11	11q23.3-q-ter.	116,202,097	134,439,182	18.24	Del
17	17p-ter.-q11.1	18,901	22,494,871	22.48	Del
	17q11.2-q12	25,499,505	29,918,396	4.42	Del
21	Trisomy				
22	Trisomy				

AML case #11 had numerous genetic abnormalities. Location and size (Mb) were obtained from UCSC Genome Browser. Copy-number changes previously described as copy-number variant were excluded. Del: deletion; Dup: duplication; ter: terminal.

Chromosomal region and candidate genes in CNN-LOH detected by SNP-chip analysis

Previous studies demonstrated CNN-LOH in AML samples at a frequency of 15-20%.[30,32,50,51,53,54] Our analysis with AsCNAR (allele-specific copy-number analysis using anonymous references) revealed CNN-LOH in 32% of the AML/MDS samples with a normal karyotype; the median size of the CNN-LOH was 30.91 Mb (range, 11.76 Mb–103.77 Mb). We found some cases with a recurrent region of CNN-LOH. Cases #3 and #23 had CNN-LOH on 1p, and the common region of CNN-LOH (30.85 Mb) included the tyrosine kinase genes (*FGR*, *EPHA2* and *EPHB2*) and an imprinted tumor suppressor gene *TP73* (Table 3). Cases #2 and #38 had CNN-LOH on 6p and the common region of CNN-LOH (30.97 Mb) contained the tyrosine kinase gene *DDR1* (Table 3). Cases #4 and #37 had CNN-LOH on 8q and the common region of CNN-LOH (11.76 Mb) contained the tyrosine kinase gene *PTK* (Table 3). CNN-LOH of the

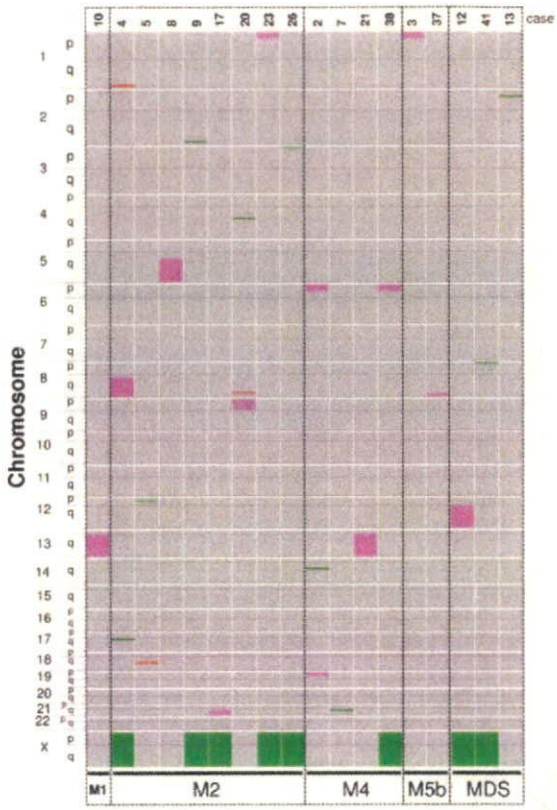

Figure 1. Genomic DNA of 37 acute myeloid leukemia samples with normal karyotype were subjected to SNP-chip analysis; genomic abnormalities are summarized. Pink, green and red bars/boxes indicate CNN-LOH, deletion and duplication, respectively. Nineteen patients (51%) showed no detectable genomic abnormalities (*data not shown*), whereas 18 patients (49%) had one or more genomic abnormalities. Deletion or duplication was found in nine patients (24%), and CNN-LOH occurred in 12 patients (32%). Chromosomal location, size and genes are shown in Tables 3 and 4.

whole region of 13q was found in cases #10 and #21; this region contains the *FLT3*, *FLT1*, *BRCA2* and *RB1* genes (Table 3).

Cases #2, #8, #12, #17 and #20 had CNN-LOH on 19p (13.41 Mb), 5q (103.77 Mb), 12q (96.23 Mb), 21q (29.54 Mb) and 9p (43.96 Mb), respectively. Although these regions of CNN-LOH occurred in only one case each, several interesting genes were found in the region, including *INSR*, *TYK2*, and *MATK* (case #2); *APC*, *FER*, *FMS/FLT4*, *PDGFRB*, *ITK* and *FGFR4* (case #8), *AML1/RUNX1* (case #17), and *JAK2* and *TEK* (case #20) (Table 3).

Interestingly, cases #10 and #21 had a *FLT3*-ITD gene mutation (Table 3); case #17 had an *AML1/RUNX1* frameshift caused by a deletion of cytosine at nucleotide 211 (Table 3). Sequencing of *JAK2* in case #20 showed a homozygous canonical *JAK2* mutation [V617F (GTC → TTC)] (Table 3). Each of these mutations occurred at a CNN-LOH. The data suggest that removal of a normal allele and duplication of the mutated allele is favored by the cancer cells.

Validation of copy number-neutral loss of heterozygosity

To validate CNN-LOH, we determined SNP sequences and gene-dosage in a CNN-LOH region using case #38 (Figure 2). If a chromosome has LOH, the nucleotide at the SNP site should not be heterozygous, but should be homozygous. We, therefore, examined six independent SNP sites in case #38 on the chromosome 6p region of CNN-LOH including rs7747259, rs1122637, rs9505293, rs6934027, rs280153 and rs191986. All six SNP sites showed only a single nucleotide; no SNP sites showed heterozygosity (Figure 2B). Each one of these sites is heterozygous in the general population at a frequency varying between 25% and 42% (*Entrez SNP database. http://www.ncbi.nlm.nih.gov/sites/entrez?db=snp*). These results strongly suggest that this region has LOH.

Next, we determined gene-dosage of the region to exclude the possibility of hemizygous deletion. The gene-dosage of 6p24.3 in case #38 was compared to that of normal genomic DNA using quantitative genomic real-time PCR by comparing the ratio between 6p24.3 and the reference genomic DNA, 2p21. As shown in Figure 2C, the amount of DNA at this site for case #38 was almost the same as that for normal genomic DNA, indicating that this region is not deleted. Taken together, our sequence data and gene dosage study validated the results of our SNP-chip analysis, clearly showing CNN-LOH at 6p24.3.

Chromosomal regions of copy-number change detected by SNP-chip analysis

Nine patients (24%) had small copy-number changes including deletions and/or duplications; the median size of the duplications and deletions was 0.3 Mb (range, 0.09–4.33 Mb) and 0.625 Mb (range, 0.11–5.87 Mb), respectively. As shown in Table 4, hemizygous deletions were found at 14q21.2 (0.3 Mb, case #2), 17q11.2 (2.7 Mb, case #4), 12p13.31 - p13.2 (2.91 Mb, case #5), 21q21.2 (0.44 Mb, case #7), 2q36.2 (0.41 Mb, case #9), 2p23.1 (0.56 Mb, case #13), 4q24 (1.08 Mb, case #20), 9p21.3 - p21.2 (5.87 Mb, case #20), 3p26.3 (0.69 Mb, case #26), and 8p23.2 (0.11 Mb, case #41). Cases #4, #5 and #20 had duplication at 1q43 (0.09 Mb), 18q21.2 (0.3 Mb), and 8q24.13 - q24.21 (4.33 Mb), respectively. These regions contain well-known oncogenes and tumor suppressor genes (Table 4). The tumor suppressor genes, *NF1* and *CDKN2A/CDKN2B*, and the transcription factor, *ETV6/TEL* were deleted in cases #4, #20 and #5, respectively; and the oncogene *MYC* was duplicated in case #20.

Validation of copy-number changes

Next, we validated copy-number changes in cases #20 and #5 using different techniques. Case #20 had duplication at 8q24.13 - q24.21 (Figure 3A) and hemizygous deletion at 9p21.3 - p21.2 (Figure 3B); these regions contain the oncogene *MYC* and the tumor suppressor genes *CDKN2A* and *CDKN2B*, respectively. Relative gene-dosages of the *MYC* and *CDKN2A* genes were examined by quantitative genomic real-time PCR with the chromosome 2p21 region as a control. The

level of the *MYC* gene was about 2-fold higher while the level of the *CDKN2A* gene was approximately 10-fold lower compared with normal genomic DNA (Figures 3C and D).

Chromosome 12p13.31 - p13.2 was deleted in case #5; this region contains the transcription factor *ETV6/TEL* (Figure 3E). FISH analysis with a probe of fluorescein-labeled ETV6-downstream (normal copy-number region) revealed two signals and a probe of Texas-red-labeled ETV6-upstream (hemizygous deleted region) revealed one signal (Figure 3F), validating the observations from SNP-chip analysis.

Table 3. Chromosomal regions identified as CNN-LOH.

Case*	FAB	Location	Physical localization Proximal	Distal	Size (Mb)	Genes
3	AML M5b	1p-ter.-p35.2	825,852	31,679,683	30.85	*FGR*
23	AML M2	1p-ter.-p35.1	825,852	33,526,200	32.7	*EPHA2, EPHB2 EPHA8, TP73 LCK* (only #23)
2	AML M4	6p-ter. - p21.3	3119,769	31,094,483	30.97	*DDR1*
38	AML M4	6p-ter. - p21.3	1119,769	33,781,344	33.66	
4	AML M2	8q12.3 - q-ter.	64,069,382	146,106,670	82.04	*PTK2*
37	AML M5b	8q24.22 - q-ter.	134,507,898	146,263,538	11.76	*NBS1* (only #4)
10	AML M1	Whole 13q				*FLT3* (ITD)
21	AML M4	Whole 13q				*FLT1 BRCA2 RB1*
2	AML M4	19p-ter. - p13.13	212,033	13,625,099	13.41	*INSR, TYK2, MATK*
8	AML M2	5q13.3 - q-ter	76,761,338	180,536,297	103.77	*APC, FER, FMS/FLT4 PDGFRB, ITK, FGFR4, NPM1*
12	MDS RA	12q11 - q-ter.	36,144,018	132,377,151	96.23	*HER3*
17	AML M2	21q21.1 - q-ter.	17,346,621	46,885,639	29.54	*AML1/RUNX1* (delC211, frameshift)
20	AML M2	9p-ter. - p21.3 9p21.2 - q21.11	140,524 27,142,682	21,047,062 44,108,554	20.91 16.97	*JAK2* (V617F) *TEK*

*Twelve patients (32%) had CNN-LOH. Physical localization, size (Mb), and genes were obtained from UCSC Genome Browser. Note: Cases #10 and #21 had a mutant form of FLT3-internal tandem repeat [FLT3 (ITD)]. Case #20 had a mutant JAK2 (JAK2 V617F) which is constitutively active, and case #17 had a deletion of cytosine at nucleotide 211 of AML1/RUNX1, resulting in a frameshift. *. Known tyrosine kinase and tumor suppressor genes are shown.*

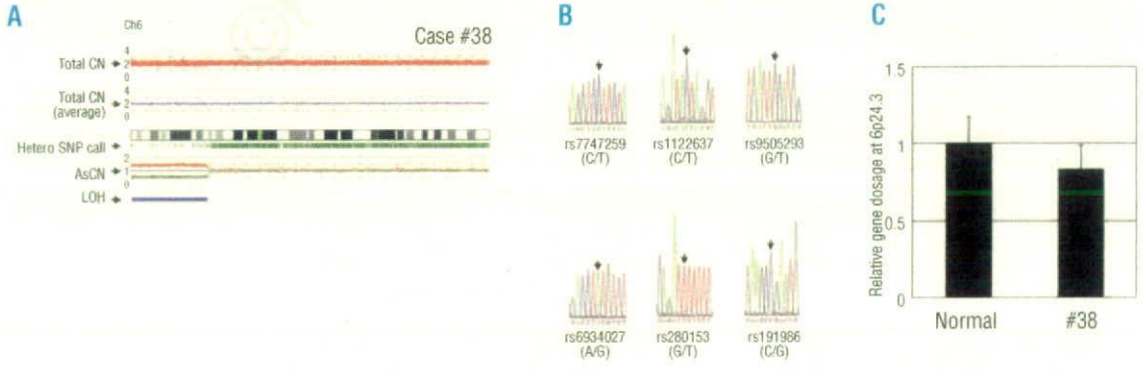

Figure 2. Validation of CNN-LOH (A) Region of CNN-LOH in chromosome 6 of case #38. Red dots represent SNP sites as probes and indicate total copy-number. The blue line represents an average of copy-number and shows gene dosage. Green bars represents heterozygous (hetero) SNP calls. Red and green lines show allele-specific copy-number (AsCN). Blue bars indicate LOH detected by heterozygous SNP calls. (B) Determination of SNP sequences in the 6p region. Six independent SNP sites were sequenced. All six SNP sites contained only a single nucleotide; no SNP site displayed heterozygosity. Results are consistent with CNN-LOH. (C) Determination of gene-dosage in the 6p region. Gene-dosage of 6p24.3 (CNN-LOH region) in case #38 is compared to that in normal genomic DNA using quantitative genomic real-time PCR. Levels of the gene-dosage were determined as a ratio between 6p24.3 and the reference genomic DNA, 2p21.

Table 4. Chromosomal location of small copy-number changes.

Case#	Type	Location	Physical localization Proximal	Physical localization Distal	Size (Mb)	Status	Gene*
4	AML M2	1q43	235,009,590	235,101,866	0.09	Dup	
		17q11.2	25,002,820	27,705,467	2.7	Del	*NF1*
5	AML M2	12p13.31-p13.2	9,312,096	12,218,922	2.91	Del	*ETV6/TEL*
		18q21.2	49,053,520	49,357,887	0.3	Dup	
9	AML M2	2q36.2	225,014,233	225,424,075	0.41	Del	
20	AML M2	4q24	105,640,274	106,723,813	1.08	Del	
		8q24.13 - q24.21	126,445,881	130,777,342	4.33	Dup	*MYC*
		9p21.3 - p21.2	21,063,692	26,935,976	5.87	Del	*CDKN2A, CDKN2B*
26	AML M2	3p26.3	1,221,075	1,911,873	0.69	Del	
2	AML M4	14q21.2	45,915,366	46,216,073	0.3	Del	
7	AML M4	21q21.2	23,126,095	23,566,855	0.44	Del	
41	MDS RA*	8p23.2	3,483,631	3,589,278	0.11	Del	
13	MDS RAEB-2	2p23.1	30,659,972	31,220,245	0.56	Del	

*Nine patients (24%) had deletion and/or duplication. Location, size (Mb), and genes were obtained from UCSC Genome Browser. Copy-number changes previously described as copy-number variant were excluded. Del, deletion. Dup, duplication. *Known oncogenes and tumor suppressor genes are shown.*

Relationship between genomic abnormalities and mutant genes within the region

In our normal karyotype AML/MDS samples, eight cases (21%) had *FLT3*-ITD and 14 cases (37%) had a *NPM1* mutation (Table 1). We compared genomic abnormalities, and *FLT3*-ITD and *NPM1* mutations (*Online Supplementary Table S4*). Both *FLT3*-ITD and *NPM1* were mutated in two samples in group A (11%) and four cases in group B (22%). A single mutation of *FLT3*-ITD was found in one sample in group A (5%) and one case in group B (6%); a single mutation of *NPM1* occurred in five samples in group A (26%) and three samples in group B (17%). These mutations were, therefore, dispersed between both groups A and B

Relationship between genomic abnormalities and gene expression

We compared genomic abnormalities and gene expression. mRNA microarray analysis was done on all samples.[30] First, the level of mRNA expression in case #11 was compared with that in 37 normal karyotype AML samples. Affymetrix microarray analysis showed decreased average gene expression in the deleted regions and increased gene expression for regions with trisomy: the difference of average expression of genes located on deleted regions of chromosomes 5, 7, 17, as well as, trisomy 8, 21 and 22 were -0.21±0.01, -0.16 ±0.013, -0.27±0.018, +0.21±0.012, +0.22±0.022 and +0.15±0.013 (mean difference ± standard error), respectively (Figure 4A and *data not shown*).

Next, we examined the relationship between small copy-number changes and mRNA expression levels in the region. For this analysis, we chose deleted regions on chromosome 9 in case #20 (Figure 3B), chromosome 17 in case #4 (Table 4) and chromosome 12 in case #5 (Figure 3E). The differences in mean expression of genes located in deleted regions of chromosomes 9 (case #20), 17 (case #4), and 12 (case #5) were -0.15±0.07, -0.37±0.07, and -0.23±0.051 (mean difference ± standard error), respectively (Figure 4B). These results showed that large and small copy-number changes led to alterations of mRNA expression. In addition, the difference in mean expression of genes located in the CNN-LOH regions of each sample was comparable to that in normal copy-number samples, suggesting that CNN-LOH does not contribute to aberrant levels of gene expression (*data not shown*).

Discussion

Our genome-wide SNP-chip analysis of normal karyotype AML/MDS showed that 49% of these samples had one or more genomic abnormalities including deletions, duplications and CNN-LOH. Previous studies demonstrated that CNN-LOH occurs in AML samples at a frequency of 15-20%.[31,32,50,51,53,54] Of interest, about 40% of cases of relapsed of AML had CNN-LOH.[52] In our analysis, 32% of samples had CNN-LOH, and these regions of CNN-LOH contain several tyrosine kinase and tumor suppressor genes that may be candidate target genes in normal karyotype AML/MDS. In fact, the *FLT3*-ITD (13q12.2), *JAK2* V617F (9p24.1) and deletion of a cytosine at nucleotide 211 of *AML1/RUNX1* (21q22.12) occurred in areas of CNN-LOH resulting in duplication of these mutant genes and loss of the normal allele. A prior paradigm was that CNN-LOH marked the location of a mutated tumor suppressor gene, but it is clear that CNN-LOH can also be a signpost of an activated (mutated) oncogene. Of note, several CNN-LOH, including a region on chromosome 1p (cases #3 and #23), 6p (cases #2 and #38), 8q (cases #4 and 37) and 13q (cases #10 and #21), occurred in more than one sample. In addition, CNN-LOH of these regions, as well as several other unique CNN-LOH regions in our cohort, were also found in other stud-

ies.[36,37,38] Although these alterations are not frequent, shared regions of CNN-LOH clearly highlight their importance. These findings prompted us to screen genes located in CNN-LOH regions. We focused on tyrosine kinase genes including *FGR* (cases #3 and #23), *DDR1* (cases #2 and #38), *TYK2* (case #2), *MATK* (case #2), *FER* (case #8) and *FGFR4* (case #8), and either determined their exon nucleotide sequences or looked for single strand conformation polymorphism band-shifts of PCR products of the exons. However, these genes did not have detectable mutations (*data not shown*). Nevertheless, we believe that these CNN-LOH, as well as deletions and duplications, are acquired somatic mutations. We examined these regions for known copy-number polymorphisms (web site, *http://projects.tcag.ca/variation*) and found none. Also previously, we compared SNP-chip data between matched samples of acute promyelocytic leukemia and normal genomic DNA from the same individual (Akagi *et al.*, unpublished data) and found that CNN-LOH occurred only in the leukemia samples but not in the corresponding germline DNA. Furthermore, SNP-chip analysis easily detected a deletion on chromosome 3 (0.69 Mb) in case #28 in the AML sample which was not present in the remission bone marrow sample from the same individual (*Online Supplementary Figure S1*). Taken together, these findings suggest that the alterations detected by SNP-chip analysis are somatic mutations.

We also found small copy-number changes in some cases. Several features of case #20 are worthy of comment. The *MYC* gene was duplicated, and the *CDKN2A* (*p16/INK4A* and *p14/ARF*) and *CDKN2B* (*p15/INK4B*) genes were hemizygously deleted. Prominent expression of C-MYC protein is associated with stimulation of p14/ARF which inactivates MDM2, producing greater levels of p53 resulting in either apoptosis or slowing of cell growth which allows for DNA repair.[41,42] However, when the *p14/ARF* gene is deleted, C-MYC has an unfettered ability to stimulate growth of the cells. Case #20 had this constellation of changes. Furthermore, this individual had a homozygous *JAK2* mutation. *JAK2* is mutated (codon 617, valine changed to phenylalanine) and constitutively active in nearly 100%, 50% and 30% of samples from patients with polycythemia vera, agnogenic myeloid metaplasia and essential thrombocythemia, respectively, as well as in 1-3% of AML cases.[43-45] We do not know the prior history of this individual.

Some of the deleted genes are of particular interest; first, the tumor suppressor gene *NF1* was deleted in case #4. Children with neurofibromatosis type-1 have inactivating mutations of the *NF1* and an increased risk of developing juvenile myelomonocytic leukemia,[46] and LOH at the *NF1* gene locus occurs in this form of leukemia and other cancers. A recent study showed that three of 103 T-ALL (3%) samples and two of 71 AML samples with *MLL* rearrangements (3%) had deletion of the *NF1* gene region; a mutation in the remaining *NF1* allele was found in three samples, suggesting that *NF1* inactivation might be involved in the development of leukemia. Second, concerning case #5 (deletion of *ETV6/TEL*), *ETV6/TEL* is a transcriptional repressor and is involved in various translocations associated with

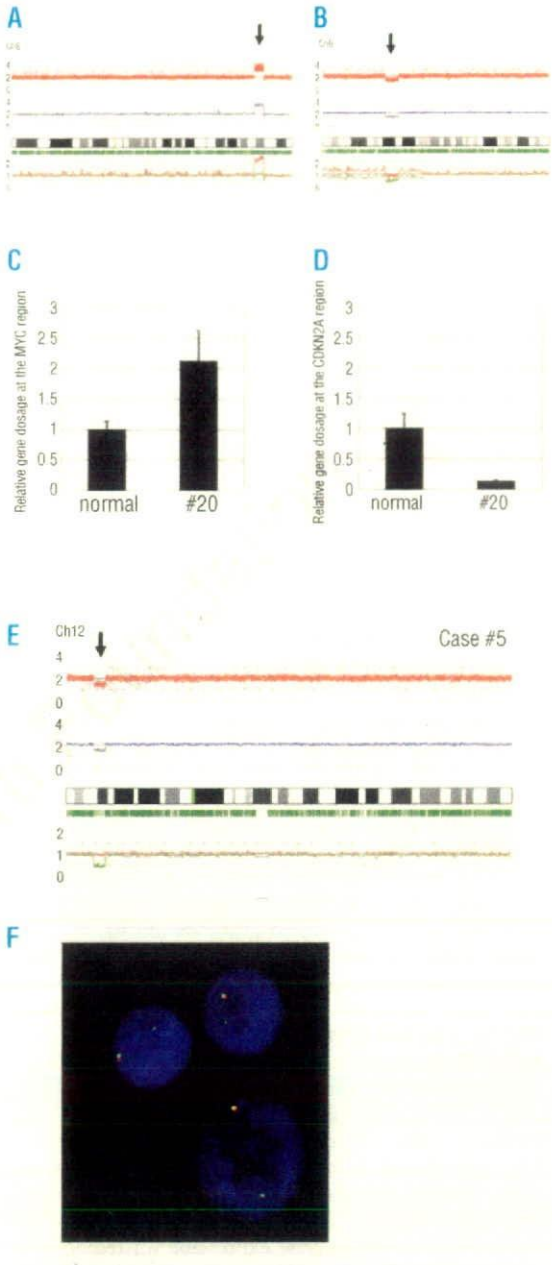

Figure 3. Validation of duplication and deletion: (A) Chromosome 8q24.13-q24.21 is duplicated. This region contains the oncogene *MYC*. (B) Chromosome 9p21.3 - p21.2 shows a deletion. The deleted region contains the tumor suppressor genes *CDKN2A* (*p16/INK4A* and *p14/ARF*) and *CDKN2B* (*p15/INK4B*). (C, D) Gene-dosages of the *MYC* gene (C) and the *CDKN2A* gene (D) region in case #20 are compared to normal genomic DNA by quantitative genomic real-time PCR. Levels of the gene-dosage are determined as a ratio between target gene and the reference genomic DNA, 2p21. (E) Case #5 had hemizygous deletion in chromosome 12p13.31-p13.2; this region contains the transcription factor *ETV6/TEL* gene. Physical localization and size are presented in Table 4. (F) FISH analysis of case #5 with probes for the *ETV6/TEL* region. Probes of fluorescein-labeled ETV6-downstream (normal region by SNP-chip analysis) and Texas-red-labeled ETV6-upstream (hemizygous deleted region by SNP-chip analysis) revealed one and two signals, respectively.

leukemia. About 30% of AML patients have loss of expression of the ETV6/TEL protein;[47,48] mutations of *ETV6/TEL* were found in 2% of AML samples, and these mutants behaved in a dominant-negative fashion.[48] Interestingly, previous array-comparative genome hybridization analysis of normal karyotype AML showed duplication of 8q24.13-q24.21 (including the *MYC* gene) and deletion of 12p12.3 (including the *ETV6* gene);[25] this constellation of alterations was also observed in our study.

Our microarray analysis showed that regions with copy-number loss or gain of chromosomal material

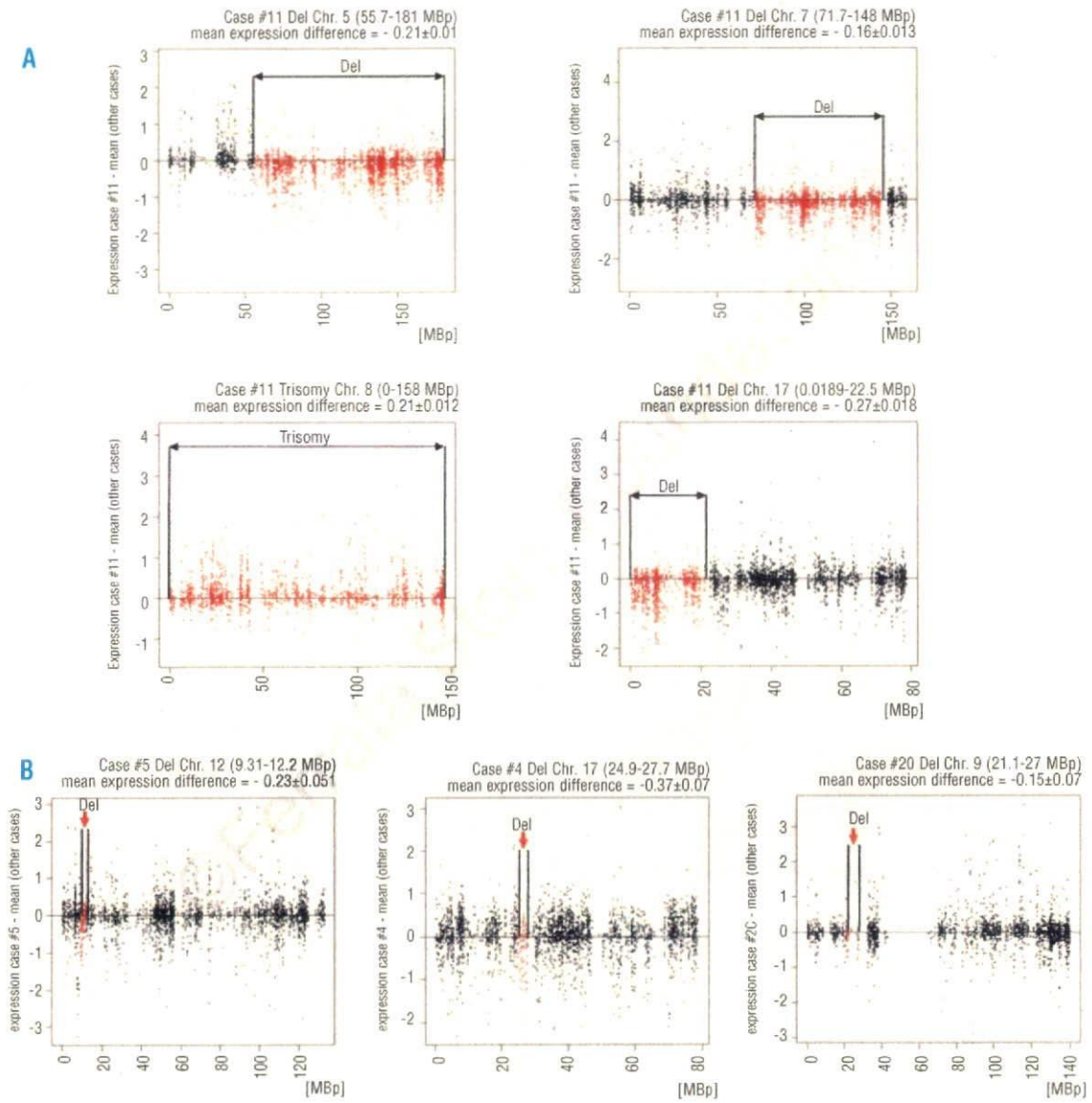

Figure 4. mRNA microarray analyses of acute myeloid leukemia samples. (A) Relationship between genomic abnormality and gene expression in acute myeloid leukemia case #11. mRNA microarray analysis was performed on all samples, and expression levels of acute myeloid leukemia cells from case #11 were compared to those of 37 normal karyotype acute myeloid leukemia samples. Affymetrix microarray analysis showed decreased average gene expression in the deleted regions, and increased average gene expression for trisomy 8: the difference of mean expression of genes located in the deleted region of chromosomes 5 (upper, left), 7 (upper, right), 17 (lower, right) and trisomy 8 (lower, left) were -0.21±0.01, -0.16±0.013, -0.27±0.018, and +0.21±0.012 (mean difference ± standard error), respectively. (B) Expression levels in acute myeloid leukemia cells from cases #20, #4 and #5 were compared with those in 36 normal karyotype AML samples. The differences in mean expression of genes located on the deleted region of chromosome 9 in case #20 (right), chromosome 17 in case #4 (middle), and chromosome 12 in case #5 (left) were -0.15±0.07, -0.37±0.07, and -0.23±0.051 (mean difference ± standard error), respectively. Each spot (black and red, Y-axis) indicates one gene and reflects the difference between each case and the mean of the other cases. Red spots represent genes located on an aberrant chromosome. The X-axis shows the chromosomal location.

were associated with either decreased or increased mRNA expression of genes in that same region, respectively, demonstrating the relationship between chromosomal status and gene expression. From an analysis perspective, we applied a descriptive approach and intended to assess plausibility of data. Some genes do indeed have higher expression values in deleted regions (Figure 4A, red points above zero) than in other cases, and some genes have lower values in trisomy (Figure 4A, red points below zero) than in other cases. However on average, expression in deleted regions is clearly lower than in non-deleted cases.

Because most regions are not recurring, we compared only one sample versus the rest (i.e. case #11 was compared with 37 normal karyotype AML/MDS cases; and cases #20, #4 and #5 were compared with other normal karyotype AML/MDS samples.) Various technical and biological sources of noise can confound the analysis. Overall, expression data appear to be consistent with chromosomal deletions and amplifications of the investigated regions. Further studies in larger cohorts of patients should enable prognostic stratification of patients in relation to their genomic changes and reveal new therapeutic targets.

Authorship and Disclosures

TA performed research, analyzed the data and wrote the paper; SO and MS performed SNP-chip analyses; GY and YN developed the CNAG; NK, AY, CWM and MD assisted in data analyses; SS, CH and TH provided AML samples, performed FISH analysis and aided in data analyses; HPK directed the overall study.

The authors declare no competing financial interests.

References

1. Kelly LM, Gilliland DG. Genetics of myeloid leukemias. Annu Rev Genomics Hum Genet 2002;3:179-98.
2. Estey E, Dohner H. Acute myeloid leukaemia. Lancet 2006;368:1894-907.
3. Erickson P, Gao J, Chang KS, Look T, Whisenant E, Raimondi S, et al. Identification of breakpoints in t(8;21) acute myelogenous leukemia and isolation of a fusion transcript, AML1/ETO, with similarity to Drosophila segmentation gene, runt. Blood 1992;80:1825-31.
4. Miyoshi H, Kozu T, Shimizu K, Enomoto K, Maseki N, Kaneko Y, et al. The t(8;21) translocation in acute myeloid leukemia results in production of an AML1-MTG8 fusion transcript. EMBO J 1993;12:2715-21.
5. Licht JD. AML1 and the AML1-ETO fusion protein in the pathogenesis of t(8;21) AML. Oncogene 2001;20:5660-79.
6. Peterson LF, Zhang DE. The 8;21 translocation in leukemogenesis. Oncogene 2004;23:4255-62.
7. Liu P, Tarlé SA, Hajra A, Claxton DF, Marlton P, Freedman M, et al. Fusion between transcription factor CBF beta/PEBP2 β and a myosin heavy chain in acute myeloid leukemia. Science 1993;261:1041-4.
8. Shigesada K, van de Sluis B, Liu PP. Mechanism of leukemogenesis by the inv(16) chimeric gene CBFB/PEBP2B-MHY11. Oncogene 2004;23:4297-307.
9. de Thé H, Chomienne C, Lanotte M, Degos L, Dejean A. The t(15;17) translocation of acute promyelocytic leukaemia fuses the retinoic acid receptor alpha gene to a novel transcribed locus. Nature 1990;347:558-61.
10. de Thé H, Lavau C, Marchio A, Chomienne C, Degos L, Dejean A. The PML-RAR-α fusion mRNA generated by the t(15;17) translocation in acute promyelocytic leukemia encodes a functionally altered RAR. Cell 1991;66:675-84.
11. Byrd JC, Mrózek K, Dodge RK, Carroll AJ, Edwards CG, Arthur DC, et al. Pretreatment cytogenetic abnormalities are predictive of induction success, cumulative incidence of relapse, and overall survival in adult patients with de novo acute myeloid leukemia: results from Cancer and Leukemia Group B (CALGB 8461). Blood 2002;100:4325-36.
12. Grimwade D, Walker H, Oliver F, Wheatley K, Harrison C, Harrison G, et al. The importance of diagnostic cytogenetics on outcome in AML: analysis of 1,612 patients entered into the MRC AML 10 trial. The Medical Research Council Adult and Children's Leukaemia Working Parties. Blood 1998;92:2322-33.
13. Mrozek K, Marcucci G, Paschka P, Whitman SP, Bloomfield CD. Clinical relevance of mutations and gene-expression changes in adult acute myeloid leukemia with normal cytogenetics: are we ready for a prognostically prioritized molecular classification? Blood 2007;109:431-48.
14. Caligiuri MA, Schichman SA, Strout MP, Mrózek K, Baer MR, Frankel SR, et al. Molecular rearrangement of the ALL-1 gene in acute myeloid leukemia without cytogenetic evidence of 11q23 chromosomal translocations. Cancer Res 1994;54:370-3.
15. Gilliland DG, Griffin JD. The roles of FLT3 in hematopoiesis and leukemia. Blood 2002;100:1532-42.
16. Reuter CW, Morgan MA, Bergmann L. Targeting the Ras signaling pathway: a rational, mechanism-based treatment for hematologic malignancies? Blood 2000;96:1655-69.
17. Pabst T, Mueller BU, Zhang P, Radomska HS, Narravula S, Schnittger S, et al. Dominant-negative mutations of CEBPA, encoding CCAAT/enhancer binding protein-alpha (C/EBPα), in acute myeloid leukemia. Nat Genet 2001;27:263-70.
18. Gombart AF, Hofmann WK, Kawano S, Takeuchi S, Krug U, Kwok SH, et al. Mutations in the gene encoding the transcription factor CCAAT/enhancer binding protein alpha in myelodysplastic syndromes and acute myeloid leukemias. Blood 2002;99:1332-40.
19. Falini B, Mecucci C, Tiacci E, Alcalay M, Rosati R, Pasqualucci L, et al. Cytoplasmic nucleophosmin in acute myelogenous leukemia with a normal karyotype. N Engl J Med 2005;352:254-66.
20. Falini B, Nicoletti I, Martelli MF, Mecucci C. Acute myeloid leukemia carrying cytoplasmic/mutated nucleophosmin (NPMc+ AML): biologic and clinical features. Blood 2007;109:874-85.
21. Nannya Y, Sanada M, Nakazaki K, Hosoya N, Wang L, Hangaishi A, et al. A robust algorithm for copy number detection using high-density oligonucleotide single nucleotide polymorphism genotyping arrays. Cancer Res 2005;65:6071-9.
22. Engle LJ, Simpson CL, Landers JE. Using high-throughput SNP technologies to study cancer. Oncogene 2006;25:1594-601.
23. Yamamoto G, Nannya Y, Kato M, Sanada M, Levine RL, Kawamata N, 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;81:114-26.
24. Pfeifer D, Pantic M, Skatulla I, Rawluk J, Kreutz C, Martens UM, et al. Genome-wide analysis of DNA copy number changes and LOH in

CLL using high-density SNP arrays. Blood 2007;109:1202-10.
25. Lehmann S, Ogawa S, Raynaud SD, Sanada M, Nannya Y, Ticchioni M, et al. Molecular allelokaryotyping of early stage untreated chronic lymphocytic leukemia. Cancer 2008; 112:1296-305.
26. Mullighan CG, Goorha S, Radtke I, Miller CB, Coustan-Smith E, Dalton JD, et al. Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. Nature 2007; 446:758-64.
27. Kawamata N, Ogawa S, Zimmermann M, Kato M, Sanada M, Hemminki K, et al. Molecular allelokaryotyping of pediatric acute lymphoblastic leukemias by high resolution single nucleotide polymorphism oligonucleotide genomic microarray. Blood 2008;111:776-84.
28. Flotho C, Steinemann D, Mullighan CG, Neale G, Mayer K, Kratz CP, et al. Genome-wide single-nucleotide polymorphism analysis in juvenile myelomonocytic leukemia identifies uniparental disomy surrounding the NF1 locus in cases associated with neurofibromatosis but not in cases with mutant RAS or PTPN11. Oncogene 2007;26:5816-21.
29. Fitzgibbon J, Iqbal S, Davies A, O'shea D, Carlotti E, Chaplin T, et al. Genome-wide detection of recurring sites of uniparental disomy in follicular and transformed follicular lymphoma. Leukemia 2007;21:1514-20.
30. Walker BA, Leone PE, Jenner MW, Li C, Gonzalez D, Johnson DC, et al. Integration of global SNP-based mapping and expression arrays reveals key regions, mechanisms, and genes important in the pathogenesis of multiple myeloma. Blood 2006;108:1733-43.
31. Raghavan M, Lillington DM, Skoulakis S, Debernardi S, Chaplin T, Foot NJ, et al. Genome-wide single nucleotide polymorphism analysis reveals frequent partial uniparental disomy due to somatic recombination in acute myeloid leukemias. Cancer Res 2005;65:375-8.
32. Fitzgibbon J, Smith LL, Raghavan M, Smith ML, Debernardi S, Skoulakis S, et al. Association between acquired uniparental disomy and homozygous gene mutation in acute myeloid leukemias. Cancer Res 2005;65:9152-4.
33. Tyybäkinoja A, Elonen E, Piippo K, Porkka K, Knuutila S. Oligonucleotide array-CGH reveals cryptic gene copy number alterations in karyotypically normal acute myeloid leukemia. Leukemia 2007;21:571-4.
34. Kiyoi H, Naoe T, Yokota S, Nakao M, Minami S, Kuriyama K, et al. Internal tandem duplication of FLT3 associated with leukocytosis in acute promyelocytic leukemia: Leukemia Study Group of the Ministry of Health and Welfare (Kohseisho). Leukemia 1997;11: 1447-52.
35. Schnittger S, Schoch C, Dugas M, Kern W, Staib P, Wuchter C, et al. Analysis of FLT3 length mutations in 1003 patients with acute myeloid leukemia: correlation to cytogenetics, FAB subtype, and prognosis in the AMLCG study and usefulness as a marker for the detection of minimal residual disease. Blood 2002; 100:59-66.
36. Schnittger S, Schoch C, Kern W, Mecucci C, Tschulik C, Martelli MF, et al. Nucleophosmin gene mutations are predictors of favorable prognosis in acute myelogenous leukemia with a normal karyotype. Blood 2005;106:3733-9.
37. Dicker F, Haferlach C, Kern W, Haferlach T, Schnittger S. Trisomy 13 is strongly associated with AML1/RUNX1 mutations and increased FLT3 expression in acute myeloid leukemia. Blood 2007;110: 1308-16.
38. Yin D, Xie D, Hofmann WK, Miller CW, Black KL, Koeffler HP. Methylation, expression, and mutation analysis of the cell cycle control genes in human brain tumors. Oncogene 2002;21:8372-8.
39. Bolstad BM, Irizarry RA, Astrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics 2003;19:185-93.
40. Haferlach T, Kohlmann A, Schnittger S, Dugas M, Hiddemann W, Kern W, et al. A global approach to the diagnosis of leukemia using gene expression profiling. Blood 2005;106:1189-98.
41. Hemann MT, Bric A, Teruya-Feldstein J, Herbst A, Nilsson JA, Cordon-Cardo C, et al. Evasion of the p53 tumour surveillance network by tumour-derived MYC mutants. Nature 2005;436:807-11.
42. Dang CV, O'Donnell KA, Juopperi T. The great MYC escape in tumorigenesis. Cancer Cell 2005;8:177-8.
43. Baxter EJ, Scott LM, Campbell PJ, East C, Fourouclas N, Swanton S, et al. Acquired mutation of the tyrosine kinase JAK2 in human myeloproliferative disorders. Lancet 2005; 365:1054-61.
44. Lee JW, Kim YG, Soung YH, Han KJ, Kim SY, Rhim HS, et al. The JAK2 V617F mutation in de novo acute myelogenous leukemias. Oncogene 2006;25:1434-6.
45. Illmer T, Schaich M, Ehninger G, Thiede C; DSIL2003 AML study group. Tyrosine kinase mutations of JAK2 are rare events in AML but influence prognosis of patients with CBF-leukemias. Haematologica 2007; 92:137-8.
46. Side LE, Emanuel PD, Taylor B, Franklin J, Thompson P, Castleberry RP, et al. Mutations of the NF1 gene in children with juvenile myelomonocytic leukemia without clinical evidence of neurofibromatosis, type 1. Blood 1998;92:267-72.
47. Hernández JM, González MB, García JL, Ferro MT, Gutiérrez NC, Marynen P, et al. Two cases of myeloid disorders and a t(8;12)(q12;p13). Haematologica 2000;85: 31-4.
48. Barjesteh van Waalwijk van Doorn-Khosrovani S, Spensberger D, de Knegt Y, Tang M, Löwenberg B, Delwel R. Somatic heterozygous mutations in ETV6 (TEL) and frequent absence of ETV6 protein in acute myeloid leukemia. Oncogene 2005;24:4129-37.
49. R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2007; ISBN 3-900051-07-0.
50. Serrano E, Carnicer MJ, Orantes V, Estivill C, Lasa A, Brunet S, et al. Uniparental disomy may be associated with microsatellite instability in acute myeloid leukemia (AML) with a normal karyotype. Leuk Lymphoma 2008;49:1178-83.
51. Gupta M, Raghavan M, Gale RE, Chelala C, Allen C, Molloy G, et al. Novel regions of acquired uniparental disomy discovered in acute myeloid leukemia. Genes Chromosomes Cancer 2008;47:729-39.
52. Raghavan M, Smith LL, Lillington DM, Chaplin T, Kakkas I, Molloy G, et al. Segmental uniparental disomy is a commonly acquired genetic event in relapsed acute myeloid leukemia. Blood 2008;112:814-21.
53. Gorletta TA, Gasparini P, D'Elios MM, Trubia M, Pelicci PG, Di Fiore PP. Frequent loss of heterozygosity without loss of genetic material in acute myeloid leukemia with a normal karyotype. Genes Chromosomes Cancer 2005;44: 334-7.
54. Tyybäkinoja A, Elonen E, Vauhkonen H, Saarela J, Knuutila S. Single nucleotide polymorphism microarray analysis of karyotypically normal acute myeloid leukemia reveals frequent copy number neutral loss of heterozygosity. Haematologica 2008; 93:631-2.

LETTERS

Oncogenic mutations of ALK kinase in neuroblastoma

Yuyan Chen[1,2,3]*, Junko Takita[1,2,3]*, Young Lim Choi[4]*, Motohiro Kato[1,3], Miki Ohira[5], Masashi Sanada[2,3,6], Lili Wang[2,3,6], Manabu Soda[4], Akira Kikuchi[7], Takashi Igarashi[1], Akira Nakagawara[5], Yasuhide Hayashi[8], Hiroyuki Mano[4,6] & Seishi Ogawa[2,3,6]

Neuroblastoma in advanced stages is one of the most intractable paediatric cancers, even with recent therapeutic advances[1]. Neuroblastoma harbours a variety of genetic changes, including a high frequency of *MYCN* amplification, loss of heterozygosity at 1p36 and 11q, and gain of genetic material from 17q, all of which have been implicated in the pathogenesis of neuroblastoma[2–5]. However, the scarcity of reliable molecular targets has hampered the development of effective therapeutic agents targeting neuroblastoma. Here we show that the anaplastic lymphoma kinase (ALK), originally identified as a fusion kinase in a subtype of non-Hodgkin's lymphoma (NPM–ALK)[6–8] and more recently in adenocarcinoma of lung (EML4–ALK)[9,10], is also a frequent target of genetic alteration in advanced neuroblastoma. According to our genome-wide scans of genetic lesions in 215 primary neuroblastoma samples using high-density single-nucleotide polymorphism genotyping microarrays[11–14], the *ALK* locus, centromeric to the *MYCN* locus, was identified as a recurrent target of copy number gain and gene amplification. Furthermore, DNA sequencing of *ALK* revealed eight novel missense mutations in 13 out of 215 (6.1%) fresh tumours and 8 out of 24 (33%) neuroblastoma-derived cell lines. All but one mutation in the primary samples (12 out of 13) were found in stages 3–4 of the disease and were harboured in the kinase domain. The mutated kinases were autophosphorylated and displayed increased kinase activity compared with the wild-type kinase. They were able to transform NIH3T3 fibroblasts as shown by their colony formation ability in soft agar and their capacity to form tumours in nude mice. Furthermore, we demonstrate that downregulation of *ALK* through RNA interference suppresses proliferation of neuroblastoma cells harbouring mutated *ALK*. We anticipate that our findings will provide new insights into the pathogenesis of advanced neuroblastoma and that *ALK*-specific kinase inhibitors might improve its clinical outcome.

To identify oncogenic lesions in neuroblastoma, we performed a genome-wide analysis of primary tumour samples obtained from 215 neuroblastoma patients using high-density single-nucleotide polymorphism (SNP) arrays (Affymetrix GeneChip 250K *Nsp*I) (Supplementary Table 1). Twenty-four neuroblastoma-derived cell lines were also analysed (Supplementary Table 2). Interrogating over 250,000 SNP sites, this platform permits the identification of copy number changes at an average resolution of less than 12 kilobases (kb)[13,14].

Analysis of this large number of samples, consisting of varying disease stages, permitted us to obtain a comprehensive registry of genomic lesions in neuroblastoma (Supplementary Figs 1 and 2). A gain of chromosomes, often triploid or hyperploid (defined by mean copy number of >2.5), was a predominant feature of neuroblastoma genomes in the lower stages. Ploidy generally correlated with the clinical stage, where non-hyperploid cases were significantly associated with stage 4 disease ($P = 4.13 \times 10^{-5}$, trend test) (Supplementary Fig. 3 and Supplementary Table 3). 17q gains, frequently in multiple copies (3≤ copy number <5), were a hallmark of the neuroblastoma genome[4] and were found in most neuroblastoma cases. Copy number gains tended to spare chromosomes 3, 4, 10, 14 and 19 (Supplementary Figs 2 and 3). Notably, these chromosomes often had copy number losses including 1p (22.8%), 3p (8.8%), 4p (5.1%), 6q (7.0%), 10q (9.8%), 11q (19.5%), 14q (3.7%), 19p (7.4%) and 19q (5.1%), implicating the pathogenic role of 'relative' gene dosages.

After excluding known copy number variations, we identified a total of 28 loci undergoing high-grade amplifications (copy number ≥5) (Supplementary Table 4). These lesions fell into relatively small genomic segments, having a mean size of 361 kb, which accelerated the identification of gene targets in these regions (Supplementary Table 4 and Supplementary Fig. 4). The candidate gene targets included *TERT* (5p15.33), *HDAC3* (5q31.3), *IGF2* (11p15.1), *MYEOV* (11q13.3), *FGF7* (15q21.1) and *CDH13* (16q23.3). However, many of them were not recurrent but found only in a single case. Although the recurrent lesions were mostly explained by the amplification of *MYCN* at 2p24, as found in 50 out of 215 (23%) of the primary cases, we identified another peak of recurrent amplification at 2p23 (Fig. 1a), which consisted of amplicons in five primary cases and in one neuroblastoma-derived cell line, NB-1 (Supplementary Fig. 5). This peak was located at the centromeric margin of the common copy number gains in chromosome 2p, which was created by copy number gains in 109 samples mostly from non-hyperploid stage 4 cases. The minimum overlapping amplification was defined by the amplicons found in the NB-1 cell line (Supplementary Fig. 5) and contained a single gene, the anaplastic lymphoma kinase (*ALK*), which has previously been reported to be overexpressed in neuroblastoma cases[15]. Although five of the six samples showing *ALK* amplification also had *MYCN* amplification, one primary case (NT056) lacked a *MYCN* peak and the amplification was confined to the *ALK*-containing locus. In interphase fluorescent *in situ* hybridization (FISH) analysis of NB-1, *MYCN* and *ALK* loci were amplified in separate amplicons (Fig. 1b), indicating that the 2p23 amplicons containing *ALK* were unlikely to represent merely 'passenger' events of *MYCN* amplification but actively contributed to the pathogenesis of neuroblastoma.

Because an oncogene can be activated by gene amplification and/or mutation, to search for possible mutations we performed DNA heteroduplex formation analysis[16] and genomic DNA sequencing for the exons 20 to 28 of *ALK*, which encompass the juxtamembrane and kinase domains (Supplementary Table 5). In total, we identified eight nucleotide changes in 21 neuroblastoma samples, 13 out of 215

[1]Department of Pediatrics, [2]Cell Therapy and Transplantation Medicine, [3]Cancer Genomics Project, Graduate School of Medicine, The University of Tokyo, Tokyo 113-8655, Japan. [4]Division of Functional Genomics, Jichi Medical University, Toguchi 329-0498, Japan. [5]Division of Biochemistry, Chiba Cancer Center Research Institute, Chiba 260-8717, Japan. [6]Core Research for Evolutional Science and Technology, Japan Science and Technology Agency, Saitama, 332-0012, Japan. [7]Division of Hematology/Oncology, Saitama Children's Medical Center, Saitama 339-8551, Japan. [8]Gunma Children's Medical Center, Shibukawa 377-8577, Japan.
*These authors contributed equally to this work.

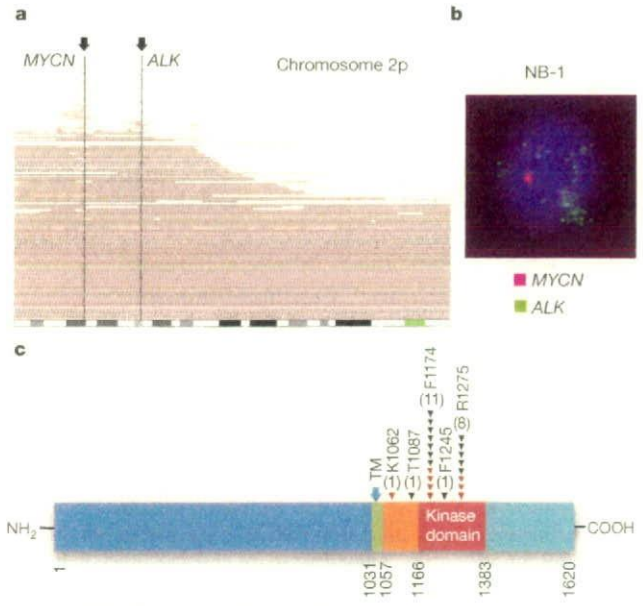

Figure 1 | Common 2p gains/amplifications and *ALK* mutations in neuroblastoma samples. **a**, Recurrent copy number gains on the 2p arm. High-grade amplifications are shown by light-red horizontal lines, whereas simple gains are shown by dark-red lines. Two common peaks of copy number gains and amplifications in the *MYCN* and *ALK* loci are indicated by arrows. The cytobands in 2p are shown at the bottom. **b**, Interphase FISH analysis of NB-1 showing high-grade amplification of *MYCN* (red) and *ALK* loci (green). The amplified *MYCN* locus appears as a single large signal. **c**, Distribution of the eight *ALK* mutations found in 21 neuroblastoma samples. The positions of the mutated amino acids are indicated by black (primary samples) and red (cell lines) arrowheads. The number of mutations at each site is shown at the top of the arrowheads. TM, transmembrane.

(6.1%) primary samples and 8 out of 24 (33%) cell lines, which resulted in seven types of amino acid substitutions at five different positions (Table 1 and Supplementary Fig. 6). They were not found in either the genomic DNA collected from 50 healthy volunteers or in the SNP databases at the time of preparing this manuscript. In fact, somatic origins of missense changes were confirmed in 9 out of 13 primary cases, for which DNA was obtained from the peripheral blood or the tumour-free bone marrow specimens (Supplementary Fig. 6). On the other hand, T1087I (ACC>ATC), found in case NT126, had a germline origin and thus it could not be determined whether the T1087I change was a rare non-functional polymorphism or represented a pathogenic germline mutation. For other changes found in three primary cases (NT128, NT217 and NT218) and cell lines, normal DNA was not available but they were likely to represent oncogenic mutations because they were identical to common somatic changes (F1174L or R1275Q) or shown to have oncogenic potential in functional assays (K1062M).

Most mutations occurred within the kinase domain (20 out of 22 or 91%), which clearly showed two mutation hotspots at F1174 and R1275 (Fig. 1c). A neuroblastoma-derived cell line, SJNB-2, had a homozygous *ALK* mutation of R1275Q, which was probably due to uniparental disomy of chromosome 2 (Supplementary Fig. 7a). Another case (NT074) harboured two different mutations, F1174L and R1275Q, but it remains to be determined whether both are on the same allele. ALK mutations within the kinase domain occurred at amino acid positions that are highly conserved across species and during molecular evolution (Supplementary Figs 8 and 9). According to the conserved structure of other insulin receptor kinases we predicted that F1174 is located at the end of the Cα1 helix, whereas the other two are on the two β-sheets: before the catalytic loop (β6, F1245) and within the activation loop (β9, R1275) (Supplementary Fig. 7b, c)[17]. Thus, conformational changes due to amino acid substitutions at these positions might be responsible for the aberrant activity of the mutant kinases.

Table 1 | *ALK* mutations/amplifications in neuroblastoma samples

Sample	Age (months)	Stage	*MYCN**	Clinical outcome	Mutations/ amplifications	Nucleotide substitution	Origin of mutations
NT126	99	4	−	Dead	T1087I	ACC>ATC	Germ line
NT218	8	1	−	Alive	F1174L	TTC>TTG	ND
NT074	34	3	+	Dead	F1174L	TTC>TTA	Somatic
					R1275Q	CGA>CAA	
NT160	12	4	+	Dead	F1174L	TTC>TTA	Somatic
NT217	24	4	+	Dead	F1174L	TTC>TTA	ND
NT190	48	4	+	Alive	F1174L	TTC>TTA	Somatic
NT060	163	3	−	Alive	F1174C	TTC>TGC	Somatic
NT162	28	4	+	Dead	F1174V	TTC>GTC	Somatic
NT195	24	4	+	Alive	F1245L	TTC>TTG	Somatic
NT055	6	3	−	Alive	R1275Q	CGA>CAA	Somatic
NT128	8	4	−	Dead	R1275Q	CGA>CAA	ND
NT164	54	4	+	Dead	R1275Q	CGA>CAA	Somatic
NT200	133	4	−	Dead	R1275Q	CGA>CAA	Somatic
SCMC-N5†	-	-	+	-	K1062M	AAG>ATG	ND
SJNB-4†	-	-	+	-	F1174L	TTC>TTA	ND
LAN-1†	-	-	+	-	F1174L	TTC>TTA	ND
SCMC-N2†	-	-	+	-	F1174L	TTC>TTA	ND
SK-N-SH†	-	-	−	-	F1174L	TTC>TTA	ND
SJNB-2†‡	-	-	+	-	R1275Q	CGA>CAA	ND
LAN-5†	-	-	+	-	R1275Q	CGA>CAA	ND
TGW†	-	-	+	-	R1275Q	CGA>CAA	ND
NT204	12	1	+	Alive	Amplification	-	-
NT056	11	3	−	Dead	Amplification	-	-
NT071	36	3	+	Alive	Amplification	-	-
NT165	19	4	+	Dead	Amplification	-	-
NT169	7	4	+	Dead	Amplification	-	-
NB-1†	-	-	+	-	Amplification	-	-

ND, not determined.
* Presence (+) or absence (−) of *MYCN* amplification in FISH analysis. All cases where there was an absence of *MYCN* amplification (−) were also checked for possible *MYCN* mutations by sequencing of all *MYCN* exons, but no *MYCN* mutations were identified.
† Cell lines.
‡ Homozygous mutation.

ALK mutation highly correlated with *MYCN* amplification ($P = 1.55 \times 10^{-4}$, Fisher's exact test; Supplementary Table 6) where 14 out of 21 mutations coexisted with *MYCN* amplification. Regardless of the status of *MYCN* amplification, 12 of the 13 mutations were found in patients with advanced stage neuroblastoma (Table 1). However, whereas *MYCN* amplification and stage 4 were significant risk factors for poor survival, the mutation/amplification status of *ALK* was not likely to have a major impact on survival (Supplementary Fig. 10 and Supplementary Table 7), although the statistical power of the current analysis was largely limited in order to detect a marginal hazard.

To evaluate the impact of ALK mutations on kinase activity, we generated Flag-tagged constructs of ALK and its mutants, F1174L and K1062M, which were stably expressed in NIH3T3 cells, and examined their phosphorylation status and *in vitro* kinase activity. The ALK mutants stably expressed in NIH3T3 cells were phosphorylated according to western blot analysis using an antibody specific for phosphorylated ALK (anti-pY1604) and a PY20 blot after anti-Flag immunoprecipitation of the mutant kinases (Fig. 2a), whereas the wild-type kinase was not phosphorylated. The immunoprecipitated ALK mutants also showed increased tyrosine kinase activity *in vitro* when compared with wild-type ALK. This was shown using both a universal substrate for tyrosine kinase (poly-GluTyr) and the synthetic YFF peptide[18], which was derived from a sequence of the activation loop of *ALK* (Fig. 2b, c). In accordance with these findings, downstream molecules of ALK signalling including AKT, STAT3 and ERK[15] were activated in cells expressing mutant ALK, as shown by their increased phosphorylation (Fig. 2d).

Next, we investigated the oncogenic potential of these mutants. NIH3T3 cells stably expressing mutant kinases showed increased colony formation in soft agar compared with the wild-type protein (Fig. 3a and Supplementary Fig. 11). The tumorigenicity of these ALK mutants was further assayed by injecting 1.0×10^7 NIH3T3 cells into nude mice. The NIH3T3 cells transfected with the ALK mutants showed focus-forming capacity and developed subcutaneous tumours (6 out of 6 inoculations) 21 days after inoculation, whereas the mock and wild-type ALK-transfected cells did not (0 out of 6 inoculations) (Fig. 3b, c). Finally, we examined the effect of ALK inhibition on the proliferation of neuroblastoma-derived cell lines. RNA interference (RNAi)-mediated *ALK* knockdown resulted in reduced cell proliferation of SK-N-SH cells harbouring the F1174L mutation, but the effects were less clear in wild-type ALK-expressing LAN-2 cells (Fig. 3d, e). Of particular interest is a recent report that 5 out of 17 neuroblastoma-derived cell lines, including SK-N-SH and NB-1, frequently showed high sensitivity to the specific ALK inhibitor TAE684 (ref. 19).

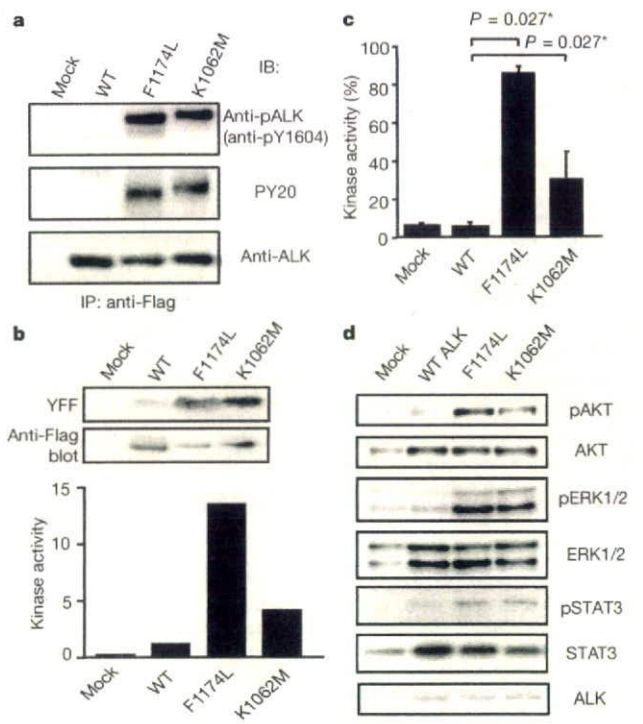

Figure 2 | Kinase activity of ALK mutants and their downstream signalling. **a**, Stably expressed ALK and its mutants (F1174L and K1062M) were immunoprecipitated with an anti-Flag antibody and subjected to western blot analysis with anti-pY1604 (upper panel) or PY20 (middle panel). An anti-ALK blot of precipitated kinases is also displayed (bottom panel). **b**, *In vitro* kinase assay for wild-type ALK kinase and its mutants using the synthetic YFF peptide as a substrate, where kinase activity is expressed as relative values to that for wild-type kinase based on the densities in the autoradiogram. **c**, Kinase activity was also assayed for the poly-GluTyr peptide. Significantly different measurements are indicated by asterisks with *P* values. Bars show mean (±s.d.) in three independent experiments. **d**, Western blot analyses of NIH3T3 cells expressing wild-type and mutant ALK for phosphorylated forms of AKT (pAKT), ERK (pERK1/2) and STAT3 (pSTAT3). The total amount of each molecule is also displayed (AKT, ERK1/2, and STAT3) together with an anti-ALK blot (ALK).

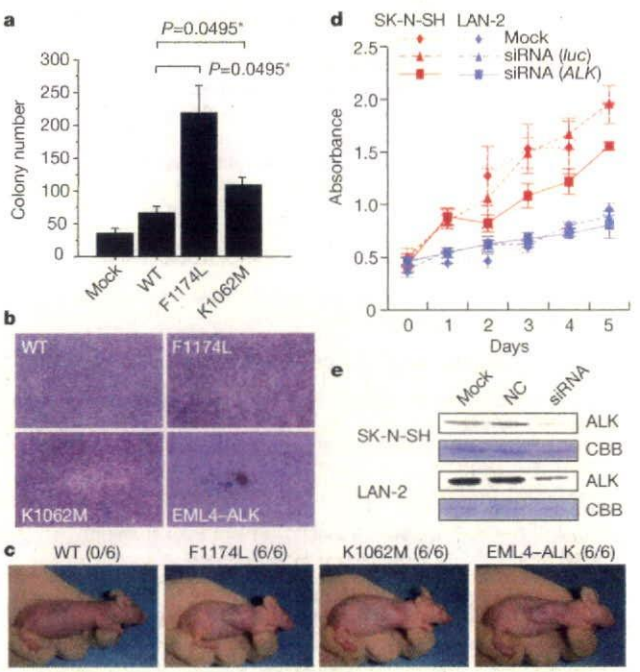

Figure 3 | Oncogenic role of ALK mutations. **a**, Colony assays for NIH3T3 cells stably expressing wild-type as well as mutant ALK (F1174L and K1062M). The average numbers of colonies in triplicate experiments are plotted and standard deviation is indicated. Results showing statistically significant differences as compared with experiments using wild-type ALK are indicated by asterisks with *P* values. **b**, **c**, NIH3T3 cells were transfected with wild-type and mutant ALK (F1174L, K1062M and EML4–ALK) and subjected to a focus forming assay (**b**) as well as an *in vivo* tumorigenicity assay in nude mice (**c**). **d**, Effect of RNAi-mediated ALK knockdown on cell proliferation in neuroblastoma cell lines expressing either the F1174L mutant (SK-N-SH) or wild-type ALK (LAN-2). Cell growth was measured using the Cell Counting Kit-8 after knockdown experiments using *ALK*-specific siRNAs (siRNA *ALK*), control siRNAs (siRNA *luc*), or mock experiments, where absorbance was measured in triplicate and averaged for each assay. To draw growth curves, the mean ± s.d. of the averaged absorbance in three independent knockdown experiments is plotted. **e**, Successful knockdown of ALK protein was confirmed by anti-ALK blots (ALK) using Coomassie brilliant blue G-250 (CBB) staining as loading controls. NC, control siRNA; siRNA, *ALK* siRNA.

Through the genome-wide analysis of genetic lesions in neuroblastoma, we identified novel oncogenic *ALK* mutations in advanced neuroblastoma. Combined with the cases having a high-grade amplification of the *ALK* gene, aberrant ALK signalling was likely to be involved in 11% (16 out of 151) of the advanced neuroblastoma cases. Because ALK kinase has been shown to be deregulated only in the form of a fusion kinase in human cancers, including lymphoma and lung cancer, the identification of oncogenic mutations in *ALK* not only increases our understanding of the molecular pathogenesis of advanced neuroblastoma, but also adds a new paradigm to the concept of 'ALK-positive human cancers' in that the mutated *ALK* kinases themselves might participate in human cancers. Our results again highlight the power of genome-wide studies to clarify the genetic lesions in human cancers[20–22]. Given that *ALK* mutations are preferentially involved in advanced neuroblastoma cases having a poor prognosis, our findings implicate that *ALK* inhibitors may improve the clinical outcome of children suffering from intractable neuroblastoma.

METHODS SUMMARY

Genomic DNA from 215 patients with primary neuroblastoma and 24 neuroblastoma-derived cell lines was analysed on GeneChip SNP genotyping microarrays (Affymetrix GeneChip 250K *Nsp*I). After appropriate normalization of mean array intensities, signal ratios were calculated between tumours and anonymous normal references in an allele-specific manner, and allele-specific copy numbers were inferred from the observed signal ratios based on the hidden Markov model using CNAG/AsCNAR software[13,14]. *ALK* mutations were examined by DNA heteroduplex analysis and/or genomic DNA sequencing[16]. Full-length cDNAs for mutant *ALK* were isolated by high-fidelity PCR and inserted into pcDNA3 and pMXS. The expression plasmids were transfected into NIH3T3 cells using Effectene Transfection Reagent (Qiagen) or by calcium phosphate methods[9]. Western blot analysis of mutant ALK kinases, *in vitro* kinase assays, and tumour formation assays in nude mice were performed as previously described[9]. This study was approved by the ethics boards of the University of Tokyo and of the Chiba Cancer Center Research Institute.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 3 June; accepted 28 August 2008.

1. Maris, J. M., Hogarty, M. D., Bagatell, R. & Cohn, S. L. Neuroblastoma. *Lancet* **369**, 2106–2120 (2007).
2. Maris, J. M. *et al.* Loss of heterozygosity at 1p36 independently predicts for disease progression but not decreased overall survival probability in neuroblastoma patients: a Children's Cancer Group study. *J. Clin. Oncol.* **18**, 1888–1899 (2000).
3. Attiyeh, E. F. *et al.* Chromosome 1p and 11q deletions and outcome in neuroblastoma. *N. Engl. J. Med.* **353**, 2243–2253 (2005).
4. Bown, N. *et al.* Gain of chromosome arm 17q and adverse outcome in patients with neuroblastoma. *N. Engl. J. Med.* **340**, 1954–1961 (1999).
5. Brodeur, G. M., Seeger, R. C., Schwab, M., Varmus, H. E. & Bishop, J. M. Amplification of N-myc in untreated human neuroblastomas correlates with advanced disease stage. *Science* **224**, 1121–1124 (1984).
6. Shiota, M. *et al.* Anaplastic large cell lymphomas expressing the novel chimeric protein p80NPM/ALK: a distinct clinicopathologic entity. *Blood* **86**, 1954–1960 (1995).
7. Morris, S. W. *et al.* Fusion of a kinase gene, ALK, to a nucleolar protein gene, NPM, in non-Hodgkin's lymphoma. *Science* **263**, 1281–1284 (1994).
8. Fujimoto, J. *et al.* Characterization of the transforming activity of p80, a hyperphosphorylated protein in a Ki-1 lymphoma cell line with chromosomal translocation t(2;5). *Proc. Natl Acad. Sci. USA* **93**, 4181–4186 (1996).
9. Soda, M. *et al.* Identification of the transforming *EML4–ALK* fusion gene in non-small-cell lung cancer. *Nature* **448**, 561–566 (2007).
10. Rikova, K. *et al.* Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer. *Cell* **131**, 1190–1203 (2007).
11. Kennedy, G. C. *et al.* Large-scale genotyping of complex DNA. *Nature Biotechnol.* **21**, 1233–1237 (2003).
12. Matsuzaki, H. *et al.* Genotyping over 100,000 SNPs on a pair of oligonucleotide arrays. *Nature Methods* **1**, 109–111 (2004).
13. Nannya, Y. *et al.* A robust algorithm for copy number detection using high-density oligonucleotide single nucleotide polymorphism genotyping arrays. *Cancer Res.* **65**, 6071–6079 (2005).
14. Yamamoto, G. *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.* **81**, 114–126 (2007).
15. Osajima-Hakomori, Y. *et al.* Biological role of anaplastic lymphoma kinase in neuroblastoma. *Am. J. Pathol.* **167**, 213–222 (2005).
16. Donohoe, E. Denaturing high-performance liquid chromatography using the WAVE DNA fragment analysis system. *Methods Mol. Med.* **108**, 173–187 (2005).
17. Hu, J., Liu, J., Ghirlando, R., Saltiel, A. R. & Hubbard, S. R. Structural basis for recruitment of the adaptor protein APS to the activated insulin receptor. *Mol. Cell* **12**, 1379–1389 (2003).
18. Donella-Deana, A. *et al.* Unique substrate specificity of anaplastic lymphoma kinase (ALK): development of phosphoacceptor peptides for the assay of ALK activity. *Biochemistry* **44**, 8533–8542 (2005).
19. McDermott, U. *et al.* Genomic alterations of anaplastic lymphoma kinase may sensitize tumors to anaplastic lymphoma kinase inhibitors. *Cancer Res.* **68**, 3389–3395 (2008).
20. Garraway, L. A. *et al.* Integrative genomic analyses identify MITF as a lineage survival oncogene amplified in malignant melanoma. *Nature* **436**, 117–122 (2005).
21. Mullighan, C. G. *et al.* Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature* **446**, 758–764 (2007).
22. Kawamata, N. *et al.* Molecular allelokaryotyping of pediatric acute lymphoblastic leukemias by high-resolution single nucleotide polymorphism oligonucleotide genomic microarray. *Blood* **111**, 776–784 (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank H. P. Koeffler for critically reading and editing the manuscript. We also thank M. Matsumura, Y. Ogino, S. Ichimura, S. Sohma, E. Matsui, Y. Yin, N. Hoshino and Y. Nakamura for their technical assistance. This work was supported by the Core Research for Evolutional Science and Technology, Japan Science and Technology Agency and by a Grant-in-Aid from the Ministry of Health, Labor and Welfare of Japan for the third-term Comprehensive 10-year Strategy for Cancer Control.

Author Contributions Y.C., Y.L.C. and J.T. contributed equally to this work. M.K. and M.Sa. performed microarray experiments and subsequent data analyses. Y.C. and J.T. performed mutation analysis of *ALK*. Y.C., Y.L.C., J.T., M.So., L.W. and H.M. conducted functional assays of mutant ALK. A.N., M.O., T.I., A.K. and Y.H. prepared tumour specimens and were involved in statistical analysis. A.N., Y.H., H.M., J.T. and S.O. designed the overall study, and S.O. and J.T. wrote the manuscript. All authors discussed the results and commented on the manuscript.

Author Information The nucleotide sequences of *ALK* mutations detected in this study have been deposited in GenBank under the accession numbers EU788003 (K1062M), EU788004 (T1087I), EU788005 (F1174L; TTC/TTA), EU788006 (F1174L; TTC/TTG), EU788007 (F1174C), EU788008 (F1174V), EU788009 (F1245L) and EU788010 (R1275Q). The copy number data as well as the raw microarray data will be accessible from http://www.ncbi.nlm.nih.gov/geo/ with the accession number GSE12494. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to S.O. (sogawa-tky@umin.net) or Y.H. (hayashiy-tky@umin.ac.jp).

METHODS

Specimens. Primary neuroblastoma specimens were obtained during surgery or biopsy from patients who were diagnosed with neuroblastoma and admitted to a number of hospitals in Japan. In total, 215 primary neuroblastoma specimens were subjected to SNP array analysis after informed consent was obtained from the parents of each patient. The patients were staged according to the International Neuroblastoma Staging System[23]. The clinicopathological findings are summarized in Supplementary Table 1. Twenty-four neuroblastoma-derived cell lines were also analysed by SNP array analysis (Supplementary Table 2). The SCMC-N2, SCMC-N4 and SCMC-N5 cell lines were established in our laboratory[24,25]. The SJNB series of cells and the UTP-N-1[26] cell line were gifts from A. T. Look and A. Inoue, respectively. The other cell lines used were obtained from the Japanese Cancer Resource Cell Bank (http://cellbank.nibio.go.jp/).

Microarray analysis. High molecular mass DNA was isolated from tumour specimens as well as from the peripheral blood or the bone marrow as described previously[24]. The DNA was subjected to SNP array analysis using Affymetrix GeneChip Mapping 50K and/or 250K arrays (Affymetrix) according to the manufacturer's suggested protocol. The scanned array images were processed with Gene Chip Operation software (GCOS)[13], followed by SNP calls using GTYE. Genome-wide copy number measurements and loss of heterozygosity detection were performed using CNAG/AsCNAR algorithms[14], which enabled an accurate determination of allele-specific copy numbers.

Confirmation of SNP array data. FISH and/or genomic PCR analysis confirmed the results of SNP array analyses as described previously[13]. PCR primer sets were designed to amplify several adjacent fragments inside and outside of the homozygously deleted regions in tumour samples.

Mutation analysis. Mutations in the *ALK* gene were examined in 239 neuroblastoma samples, including 24 cell lines, by denaturing high-performance liquid chromatography (DHPLC) using the WAVE system (Model 4500; Transgenomic) according to the manufacturer's suggested protocol[16]. The samples showing abnormal conformations were subjected to direct sequencing analysis using an ABI PRISM 3100 Genetic Analyser (Applied Biosystems). Using direct sequencing, mutation analysis of *MYCN* was also performed in seven cases with *ALK* alterations but not *MYCN* amplification. The primer sets used in this study are listed in Supplementary Table 5.

Transforming potential of *ALK* mutants. Total RNA was extracted from SJNB-1 (wild type), SCMC-N2 (F1174L) and SCMC-N5 (K1062M) cells as described previously[26]. First-strand cDNA was synthesized from RNA using Transcriptor Reverse Transcriptase and an oligo (dT) primer (Roche Applied Science). The resulting cDNA was then amplified by PCR using the KOD-Plus-Ver.2 DNA polymerase (Toyobo) and the primers sense 5′-TCAGAAGCTTTACCAAGGACTGTTCAGAGC-3′ and antisense 5′-AATTGCGGCCGCTACTTGTCATCGTCGTCCTTGTAGTCGGGCCCAGGCTG GTTCATGC-3′, thereby introducing a HindIII site at the 5′ terminus and a NotI site and a Flag sequence at the 3′ terminus. The HindIII–NotI fragments of *ALK* cDNA were subcloned into pcDNA3 to generate expression plasmids. After resequencing to confirm that they had no other mutations, the *ALK* plasmids were used for transfection into NIH3T3 cells using Effectene Transfection Reagent (Qiagen) according to the suggested manufacturer's protocol. The transfected NIH3T3 cells were selected in 800 μg ml^{-1} G418 for 2 weeks to obtain stably expressing clones.

To evaluate the phosphorylation status of ALK mutants, the cell lysates of stable clones were immunoprecipitated with antibodies to Flag (Sigma) and the resulting precipitates were subjected to western blot analysis with the antibody specific to pTyr 1604 (Cell Signaling Technology) of ALK and the generic anti-phosphotyrosine antibody (PY20). The *in vitro* kinase activity of ALK mutants was measured using a non-radioactive isotope solid-phase enzyme-linked immunosorbent assay using the Universal Tyrosine Kinase Assay kit (Takara) according to the manufacturer's suggested protocol. We also performed the *in vitro* kinase assay with the synthetic YFF peptide (Operon Biotechnologies) as described previously[18]. For anchorage-independent growth analysis, 1×10^3 stably transfected NIH3T3 cells were mixed in 0.3% agarose with 10% FBS-DMEM and plated on 0.6% agarose-coated 35-mm dishes. After culture for 28 days, the colonies of >0.1 mm in diameter were counted. The quantification of the colonies was from three independent experiments. To investigate the downstream signalling of ALK, western blot analysis was performed using the anti-ERK1/2, anti-phospho-ERK1/2, anti-AKT, anti-phospho-AKT, anti-STAT3 and anti-phospho-STAT3 antibodies (Cell Signaling Technology)[15].

The cDNA mutant of *ALK* was also inserted into the pMXS plasmid and the constructs were introduced into NIH3T3 cells by the calcium phosphate method as described previously[9]. The cells were then either cultured for 21 days or injected subcutaneously at six sites in three nude mice.

Inhibition of ALK through RNAi-mediated knockdown. To suppress the expression of the ALK protein, two different pairs of *ALK* siRNAs (*ALK* siRNA1 and *ALK* siRNA2) were obtained (Qiagen)[15]. The sequences were 5′-GAGUCUGGCAGUUGACUUCdTdT-3′ for *ALK* siRNA1 and 5′-GCUCCGGCGUGCCAAGCAGdTdT-3′ for *ALK* siRNA2. A siRNA, targeting a sequence in firefly (*Photinus pyralis*) luciferase mRNA (*luc* siRNA), was used as a negative control (Qiagen)[15]. The sequences of *luc* siRNA were as follow: sense 5′-CGUACGCGGAAUACUUCGAdTdT-3′ and antisense 5′-UCGAAGUAUUCCGCGUACGdTdT-3′. Gene knockdown was achieved in SK-N-SH and LAN-2 cells using HiPerFect transfection reagent following the manufacturer's suggested instructions (Qiagen). To assess the effect of ALK knockdown on cell growth, these cells were seeded in 96-well plates at a concentration of 8.0×10^3 cells per well 24 h before transfection and assayed using the Cell Counting Kit-8 (Wako).

Statistical analysis. The significance of the correlation between *MYCN* amplification and *ALK* mutation was tested according to the conventional 2×2 contingency table using Fisher's exact test. The significance of the differences in kinase activity between wild-type and mutant ALK kinases was examined by the Mann–Whitney *U*-test based on the measured percentage activity of kinases in the precipitates of the corresponding samples. The significance of the differences in colony formation between wild-type and mutant ALK kinases was also examined by the Mann–Whitney *U*-test. The size of the hazards from possible risk factors, including International Neuroblastoma Staging System stages, *MYCN* status and *ALK* mutation/amplification were estimated by Cox regression analysis assuming a proportional hazard model using Stata software. Correlation between ploidy and clinical stage was tested by nptrend test.

23. Smith, E. I., Haase, G. M., Seeger, R. C. & Brodeur, G. M. A surgical perspective on the current staging in neuroblastoma—the International Neuroblastoma Staging System proposal. *J. Pediatr. Surg.* **24**, 386–390 (1989).
24. Takita, J. *et al.* Allelotype of neuroblastoma. *Oncogene* **11**, 1829–1834 (1995).
25. Takita, J. *et al.* Absent or reduced expression of the caspase 8 gene occurs frequently in neuroblastoma, but not commonly in Ewing sarcoma or rhabdomyosarcoma. *Med. Pediatr. Oncol.* **35**, 541–543 (2000).
26. Takita, J. *et al.* Allelic imbalance on chromosome 2q and alterations of the caspase 8 gene in neuroblastoma. *Oncogene* **20**, 4424–4432 (2001).