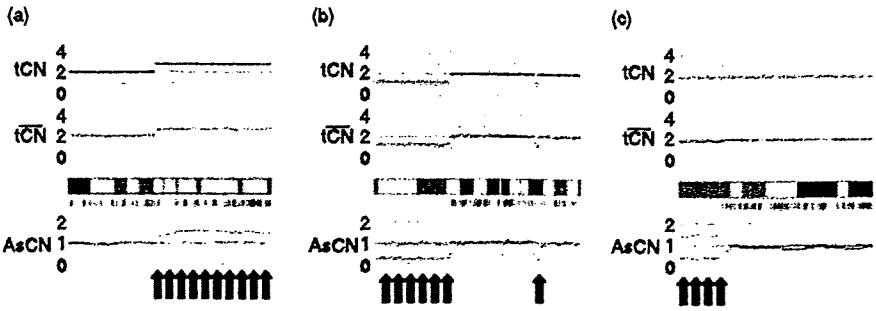

Fig. 1. Representative results of the allele-specific copy number analysis using anonymous references (AsCNAR) program with regard to copy number (CN) alterations detected in our series at particular loci, such as (a) gain (b) chromosomal loss, and (c) uniparental disomy (UPD), which have not been detected using conventional algorithms. The red dots indicate the raw CN plot for each single-nucleotide polymorphism (SNP), and the blue lines indicate the local mean CN of five SNP. The vertical green bar indicates the heterozygous SNP calls.

In the present study, to identify the novel genomic alterations in sporadic HBL cases, we performed high-resolution analyses of genome-wide CN alterations such as gains, losses, allelic imbalances, and amplifications of small chromosomal regions. Due to the high resolution of the SNP arrays and the new algorithm AsCNAR, we could systematically identify several amplifications, deletions, and allelic imbalances, including the UPD.

Materials and Methods

Patients and samples. We obtained 17 primary HBL samples at the time of diagnosis from five patients treated at the Gunma Children's Medical Center and 12 patients treated at different institutes in Japan, including Saitama Children's Medical Center. No patient had received chemo- and/or radiotherapy before the biopsy of the primary tumors. After obtaining informed consent from the parents and approval for the study from the institutional review board of each institute, all the HBL samples were subjected to genomic DNA extraction using the QIAamp DNA Mini Kit (Qiagen, Chatsworth, CA, USA) according to the manufacturer's instructions. Total RNA was extracted from the frozen stocked tumors using the Isogen reagent (Nippon Gene, Osaka, Japan), according to the manufacturer's instructions. The total RNA was reverse transcribed to synthesize cDNA using the Ready-To-Go T-Primed First-Strand Kit (GE Healthcare Bio-Sciences, Piscataway, NJ, USA).

SNP array analysis. The array experiments were performed according to the standard protocol of Affymetrix® GeneChip® Mapping 50K *Xba*I Array (Affymetrix, Inc., Santa Clara, CA, USA). In brief, the total genomic DNA (250 ng from each sample) was first digested with a restriction enzyme (*Xba*I). The digested DNA was then ligated to an appropriate adapter that recognized the four cohesive base pair (bp) overhangs, and polymerase chain reaction (PCR) amplification was performed using a single primer that recognized the adapter sequence. After fragmentation with DNase I, the PCR products were labeled with a biotinylated nucleotide analog using terminal deoxynucleotidyl transferase, and the labeled products were hybridized to the GeneChip® Human Mapping 50K Array for 17 h. Subsequently, the arrays were washed, stained, and scanned.

The genotype calls and the intensity of the SNP probes were determined using GeneChip Operation software (GCOS; Affymetrix, Inc.). The SNP CN and chromosomal regions with gains or losses were individually evaluated using the CNAG[27] and AsCNAR algorithms,[28] which enabled an accurate determination of allele-specific CN as well as the sensitive detection of LOH even in the presence of normal cell contamination of up to 70–80% without requiring constitutive DNA (Fig. 1; http://www.genome.umin.jp).

Validation of CN alterations using the interphase FISH. We performed FISH to validate the CN status obtained using the SNP array analysis. FISH probes were prepared using the BAC clones RP11-185M22, RP11-80P10, and RP11-86M15. Each BAC DNA was purified, and 100 ng of the clone was labeled with digoxigenin-dUTP using random primers; these labeled clones were used as probes for FISH analysis by following the established protocols.[29,30]

Quantitative real-time PCR and reverse transcription (RT)-PCR. Real-time quantitative PCR (RQ-PCR) and real-time quantitative RT-PCR (RQ-RT-PCR) analyses were carried out to quantify the relative CN of several amplifications in the HBL samples and the expression levels of the defender against cell death 1 (*DAD1*), EPH receptor B6 (*EphB6*), *ErbB4*, insulin-like growth factor II (*IGF2*), and *H19* genes using a Power SYBR Green PCR Master Mix (Applied Biosystems, Foster City, CA, USA) with an ABI prism 7700 real-time PCR detection system (Applied Biosystems). The primer pairs were designed using PrimerExpress software (Applied Biosystems) and synthesized by Invitrogen (Carlsbad, CA, USA).The primer sets used for the RQ-PCR experiments are listed in Table 1. Data were captured using Sequence Detection

Table 1. Primers used for polymerase chain reaction (PCR) analyses

Gene	Primer forward	Primer reverse
(Genomic RQ-PCR)		
EphB6	GGACTGCAACTGAACGTCAA	TCTGGAAAGGAAGCAAAGGA
DAD1	GTTATGTCGGCGTCGGTAGT	GTCCCCACGAGGAGACAGTA
(RQ-RT-PCR)		
ERBB4	AACAGCAGTACCGAGCCTTG	CCAGAGGCAGGTAACGAAAC
DAD1	CGAGCCTTTGCTGATTTTCT	TCCAATAAGCTGCCATCTCC
IGF2	CTCTCCGTGCTGTTCTCTCC	TATCGGGAAATGAGGTCAGC
H19	GAAGGAGGTTTAGGGGATCG	TTGCTCTTTCTGCCTGGAAC
(Bisulfite PCR/RQ-PCR)		
H19DMR (Methylated)	GGTACGGTTTTTTTAGGTTTATGTC	ACCCCTACAACCTCCTTACTACG
H19DMR (Unmethylated)	TATGGTTTTTTTAGGTTTATGTTGG	ACCCCTACAACCTCCTTACTACAC

Primers and probes were designed using Primer Express software and MethPrimer software. RQ-PCR, real-time quantitative PCR; RQ-RT-PCR, real-time quantitative reverse transcription-PCR.

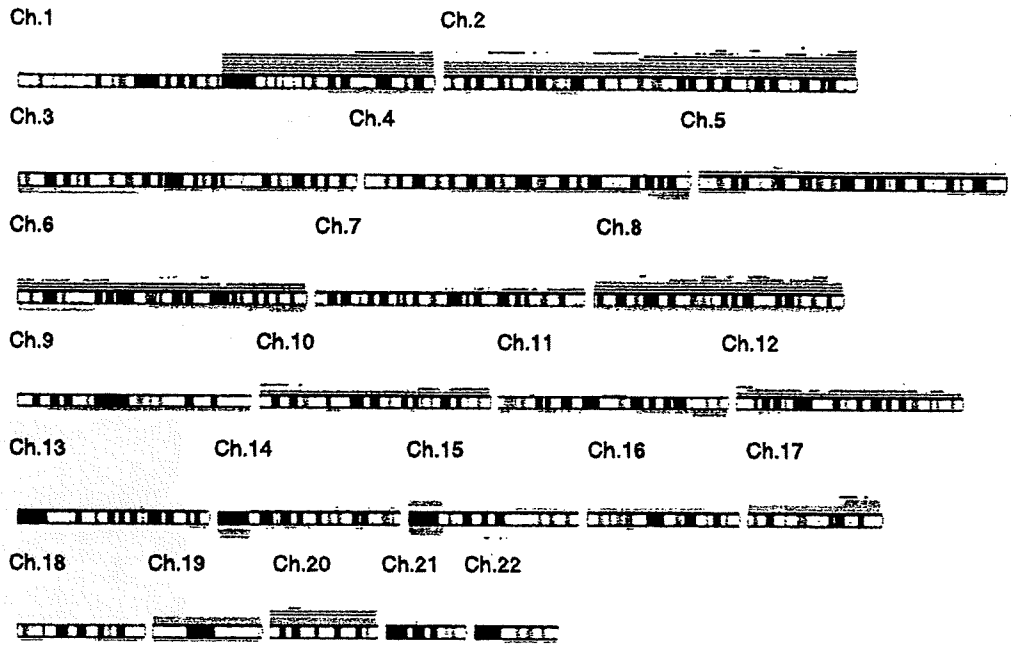

Fig. 2. Overview of the DNA copy number (CN) gains and losses detected in 17 hepatoblastoma (HBL) samples. A gain is indicated by the red bar above the chromosome ideogram, and a loss is indicated by the green bar under the chromosome ideogram. Each horizontal line represents an aberration detected in a single tumor.

software (version 1.7a; Applied Biosystems). For each primer pair, a standard curve was generated from five-fold serial dilution from approximately 50–80 pg of control DNA from a healthy individual. The amounts of genomic DNA and cDNA used in each test and the reference marker for all HBL samples were calculated using the appropriate standard curve. Normalization was performed using the β-actin gene as the internal control.

Sodium bisulfite modification and methylation-specific PCR. The genomic DNA from the tumor samples was treated with sodium bisulfite as described previously.[31] Briefly, 1 μg of DNA was denatured with sodium hydroxide and modified with sodium bisulfite. The modified DNA was then purified with the Wizard® DNA Clean-Up System (Promega, Madison, WI, USA), precipitated with ethanol, resuspended in Tris-EDTA (TE) buffer (pH 8.0), and either used immediately or stored at –20°C until use. The bisulfite-modified DNA was amplified with primer pairs for the methylated and unmethylated complete sequences upstream of the *H19* promoter CpG islands in the HBL samples with UPD in 11p15. The primer pairs were designed using MethPrimer software,[32] and synthesized by Invitrogen. The primers for methylation-specific DNA and unmethylation-specific primers are listed in Table 1. Normal lymphocyte DNA was used as the control. PCR was carried out in a 25 μL reaction volume using Ex Taq Hot Start Version (TaKaRa Bio Inc., Kyoto, Japan). The PCR conditions were as follows: 1 cycle at 95°C for 10 min; followed by 35 cycles of 94°C for 30 s, 60°C for 30 s, 72°C for 2 min; and a final extension at 72°C for 5 min. The PCR products were separated on 3% agarose gels and visualized under UV illumination after ethidium bromide staining. To quantify the ratio of the methylation status, we also carried out the methylation-specific RQ-PCR analysis.

Results

Detection of CN alterations in HBL samples. We investigated 17 HBL samples obtained from the sporadic cases of HBL by using the Affymetrix® GeneChip® 50K *Xba*I Mapping Array. Although these specimens did not contain paired control DNA and had varying degrees of normal tissue contamination, the genomic alterations were accurately determined in most specimens by our CNAG/AsCNAR program (Fig. 1). The real CN and LOH status was inferred from the observed signal ratios of the tumor to the reference, based on the hidden Markov models implemented in the CNAG/AsCNAR program; these are summarized in Fig. 2. The CN data were validated at a number of SNP sites using FISH analysis of the cell nuclei extracted from the HBL samples (Fig. 3). The CN data obtained using the FISH analyses were consistent with those obtained using SNP mapping.

Numerical chromosomal aberrations were observed in 15 HBL samples (88%), excluding two HBL samples (HBL_22 and HBL_250). These 15 cases had variable degrees of CN gains and losses; however, the gains including the amplifications were more frequent than the losses (Table 2 and Fig. 2). Total or partial gains in chromosomes 1q and 2 were the most frequent aberrations detected in eight of the 17 patients (47%). The gain in chromosome 8 was the second most frequent aberration detected in five of the 17 samples (29%). The gains in chromosomes 17q and 20 were observed in 24% of the cases (four of 17 cases). The LOH in chromosomes 4q and 11q was observed in three (18%) and two (12%) of the 17 samples, respectively. However, these regions were usually large, and we could not determine the presence or absence of alterations in specific genes within these regions.

High-grade amplification and common deletion. High-grade amplifications are of particular interest because they may indicate the loci of oncogenes. The regions with high-grade amplification were defined as segments with at least five SNP loci with an inferred CN of >5. High-grade amplifications of 7q34 and 14q11.2 were observed in five (29%) and nine (53%) HBL samples, respectively. For the validation of the amplifications observed using the SNP array, FISH analysis and genomic RQ-PCR were performed. To determine the genes that are potentially affected at 14q11.2, several genes localized at the 14q11.2 chromosomal region with overlap or proximity to the BAC-RP11-85M16 were examined using the UCSC browser (www.genome.ucsc.edu). Genes that map to these regions include *EphB6* and *DAD1*, which are identified as the negative regulators of apoptosis. These two genes were subjected to RQ-PCR. FISH analysis with RP11-85M16 BAC clone probe showed multiple signals, confirming

doi: 10.1111/j.1349-7006.2007.00710.x

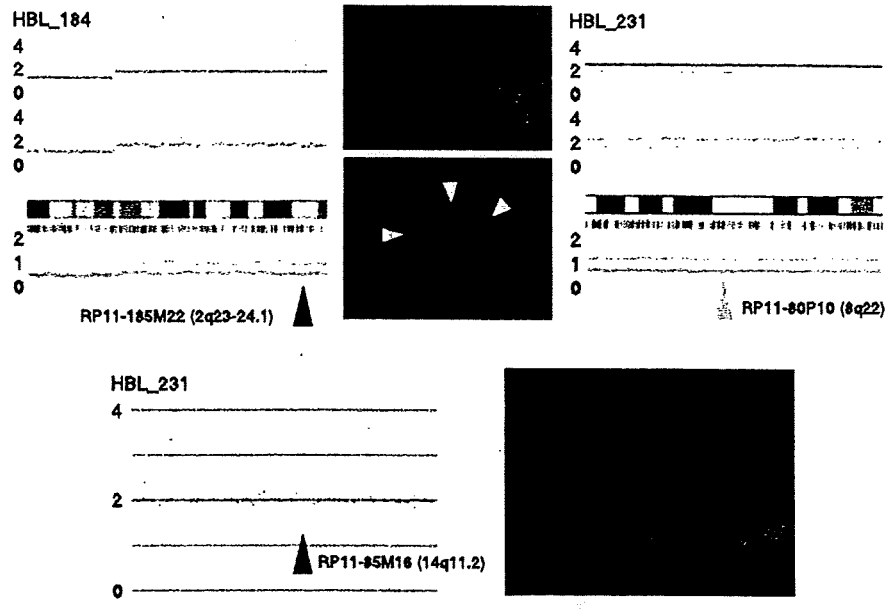

Fig. 3. Representative results of array analysis of hepatoblastoma (HBL) samples (HBL_184 and HBL_231). Fluorescent *in situ* hybridization analysis with BAC probes confirmed the detected changes. We detected three signals from chromosomes 2q and 8q. At the high-amplification region of chromosome 14q, three and more signals were detected.

Table 2. Chromosomal aberrations in 17 primary hepatoblastoma (HBL) samples

Sample	Copy number gain	Copy number loss	Uniparental disomy
HBL_4	1q, 2q34, 5p13.1, 17q23.3-qter	3p13-pter, 3q13.11, 6q14.1-qter, 11q23.1-qter	Not detected
HBL_7	1q, 2, 8, 14q11.2, 20	not detected	11p15.4-pter
HBL_8	7q34	not detected	Not detected
HBL_9	1q, 2q14.1-qter, 6p, 7, 14q11.2	6p12.1, 9p21.1	Not detected
HBL_12	8q11.23, 10q21.3, 10q26.13, 14q11.2, 22q13.31	7q35	Not detected
HBL_14	2p16.3-p22.3, 2p23.1, 2q11.2-q14.1, 2q33.1-q34, 3p21.33-p22.1, 3p24.2, 3p25.1, 3p25.2, 4q32.2-q32.3, 5p13.2, 6q14.3-q16.1, 7q, 11p15.1, 10p13-pter, 11q22.2-q22.3, 12p13.2-pter, 14q23.3-q31.1, 15q22.31-q26.2, 16p12.3, 20p11.23	1q31.1-qter, 2p12-14, 3, 4q, 5p14.1-pter, 5q32-qter, 6p12.13-pter, 6q11.1, 6q25.1-qter, 8, 9, 12p11.1-13.1, 17q24.3, 18p11.21-11.32, 18q21.1-qter, 19, 22	Not detected
HBL_22	Not detected	Not detected	Not detected
HBL_27	1q, 2q24.2-24.3, 7q34, 14q11.2	4q32.3-qter, 16p12.1	Not detected
HBL_28	3p26.1, 7q34, 14q11.2, 20	2p24.1	11p14.3-pter
HBL_34	1q, 2, 7q34, 14q11.2, 17	4q34.1-qter	Not detected
HBL_36	1q32.1-qter	1p13.3-pter, 4q21.22-qter, 5p13.1	Not detected
HBL_37	1q, 2, 5, 6, 7q34, 8, 10, 12, 14, 14q11.2, 15, 16q22.1-pter, 17, 19, 20	Not detected	11p15.2-pter, 16q22.2-qter
HBL_184	2q14.2-qter, 3p24.3, 4q33, 10p14, 11p14.3, 14q11.2	Not detected	Not detected
HBL_185	6p, 21q21.2	Not detected	Not detected
HBL_231	8, 14q11.2, 19, 20	Not detected	11p15.4-pter
HBL_246	1q, 2, 5, 6, 8, 10, 12, 13, 16, 17, 19, 20, 21, 22	Not detected	4, 9
HBL_250	Not detected	Not detected	Not detected

CN gains at 14q11.2 (Fig. 3). Further, in RQ-PCR analysis, the CN gain of *EphB6* and *DAD1* was evident in all samples that showed high-grade amplification in SNP array (data not shown). Other high-grade amplifications are listed in Table 3. The size of these amplicons was typically less than 1 Mb, and the possible genes present in these regions are summarized in the same table. All these candidate genes, except *MMP7*, have not been reported previously with regard to HBL.[33]

Homozygous deletions are also of particular interest because they may indicate a tumor suppressor gene. However, homozygous deletions were not identified in any sample.

CN neutral LOH (UPD). LOH can be more sensitively detected with the CNAG/AsCNAR algorithms by evaluating the allele-specific CN than from the grossly reduced heterozygous SNP calls, particularly when the SNP shows no CN losses. The UPD regions were identified in five of the 17 samples. In four samples (HBL_7, HBL_28, HBL_37, and HBL_231), 11p15 was the common UPD region (Fig. 4a). Other UPD regions were observed within chromosomes 4, 9, and 16q22 (Table 2). The candidate target genes that map to the UPD region located within 11p15 include *IGF2* and *H19*. Methylation-specific PCR analysis was performed for the four HBL samples having UPD within 11p15 to identify the origin of the amplified allele. The methylation status of the differential methylated region (DMR) of *H19* is shown in Fig. 4b. Hypermethylation of the *H19* DMR was detected in all HBL samples having UPD within 11p15; however, normal lymphocyte DNA showed the mosaic methylation pattern. In general, the *H19* DMR is hypermethylated on the paternal allele

Table 3. High-grade amplifications in hepatoblastoma (HBL) samples

Cytoband	Implicated region (base pairs)		Candidate target genes in the region
	Start–end	Size	
2q34	211 193 864–212 239 181	1 045 318	*ErbB4*
3p25.2	11 888 124–12 876 175	988 052	*RAF1*
7q34	141 721 559–142 076 238	354 680	*EphB6*
11q22.2-q22.3	101 394 973–102 830 195	1 435 223	*MMP1, 7, 20*
14q11.2	21 426 631–22 130 392	703 762	*DAD1*

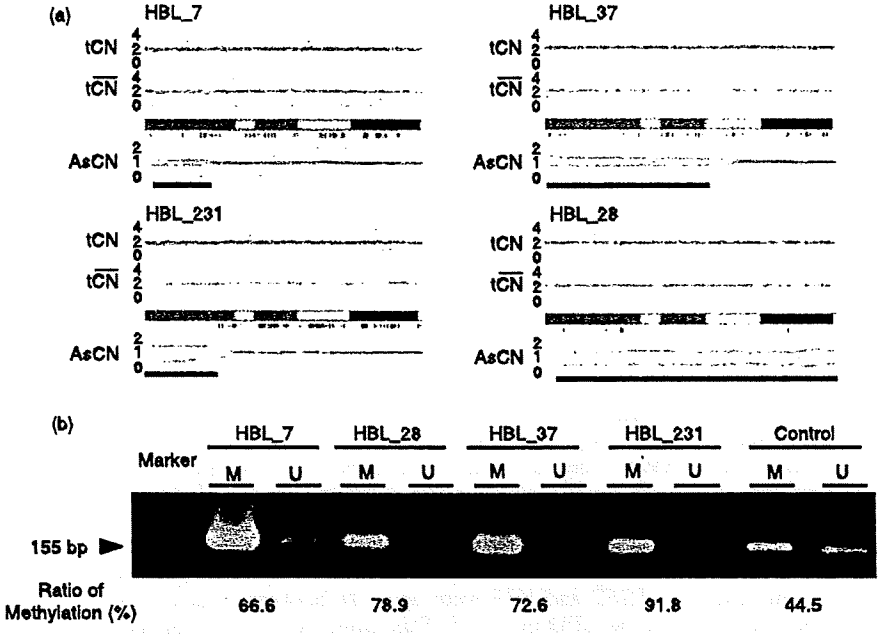

Fig. 4. (a) Copy numbers (CN) of chromosome 11p in four hepatoblastoma (HBL) samples with uniparental disomy (UPD). Although complete CN alterations are not observed, UPD is clearly predicted based on the allele-specific CN alterations (green lines). (b) Methylation-specific polymerase chain reaction (PCR) analysis of the *H19* differential methylated region (DMR). Modified DNA was amplified with primer pairs for methylated and unmethylated complete sequences of the *H19* DMR. *H19* DMR hypermethylation was detected in all HBL samples; however, normal lymphocyte DNA exhibited the mosaic methylation pattern. The results of quantitative real-time methylation-specific PCR analysis are shown below the image depicting the results of electrophoresis.

and hypomethylated on the maternally expressed allele in humans. This indicates that the UPD within this region is considered to be derived from the paternal allele. Furthermore, a low expression level of the non-methylated allele was also observed; methylation-specific RQ-PCR analysis revealed that the ratio of the methylation status ranged from 66.6% to 91.8%.

Expression analyses using RQ-RT-PCR. In order to examine the impact of the abovementioned amplifications and UPD on gene expression, we measured the expression levels of four genes (*DAD1*, *ErbB4*, *IGF2*, and *H19*) through RQ-RT-PCR (Fig. 5). Normal liver total RNA served as the non-neoplastic reference and control. HBL_184 and HBL_231 for which RNA were available showed a high expression of the *ErbB4* gene. However, the expression of *DAD1* was down-regulated in both these samples. The *IGF2* and *H19* genes were oppositely expressed between HBL_184 and HBL_231, having UPD within 11p15.

Discussion

The present study represents the application of the SNP array technology for the genome-wide analysis of CN aberrations in HBL. Several recent studies and our previous research have demonstrated that this technology provided a unique opportunity to assess the DNA CN alterations and LOH simultaneously throughout the entire genome.[24–27,29] As shown in the present analysis, the use of high-resolution SNP arrays improved the ability to identify structural chromosomal aberrations in cancer cells and detect genes affected by these aberrations. Additionally, high-density SNP array analysis with the CN analyzer software can also

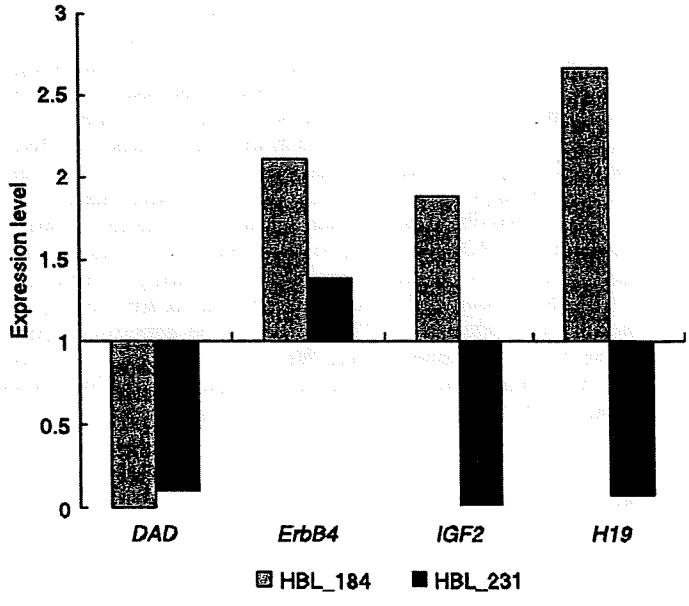

Fig. 5. The results of the expression levels of four genes (defender against cell death 1 [*DAD1*], EPH receptor B6 [*EphB6*], *ErbB4*, insulin-like growth factor II [*IGF2*], and *H19* genes) through real-time quantitative reverse transcription-polymerase chain reaction (RQ-RT-PCR) analyses.

doi: 10.1111/j.1349-7006.2007.00710.x

facilitate the identification of allelic imbalances such as copy-neutral LOH in the absence of a paired normal DNA reference.

The aberrations in chromosomes 1q, 2, 8, and 20 have been noted as the most commonly occurring aberrations in all previous reports,[21,22] as well as in the present study. In the present study, the most frequently detected aberrations were gains in chromosomes 1q and 2 (or 2q), observed in approximately 50% of the cases.

Trisomy in chromosome 1q is a well-known alteration in HBL.[34] Similar 1q imbalances have also been described in other pediatric neoplastic disorders such as lymphoma,[35] Wilms' tumor,[36] and sarcoma,[37] indicating that these aberrations are related to tumor progression. The candidate genes in 1q included the *NTRK1*, *ABL2*, *CD34*, *DAP3* (death receptor protein-3), and caspase-3 genes.[38] The anomalies in chromosome 2, which almost always result in gains in 2q, are also common in HBL. These imbalances are also commonly found in embryonal rhabdomyosarcoma and other pediatric tumors related to BWS. Translocation involving the *PAX3* gene located in 2q35 has been suggested to play a crucial role in the pathogenesis of alveolar rhabdomyosarcoma.[39] Based on this, a genetic link has been suggested between HBL and alveolar rhabdomyosarcoma. The role of the *PAX3* gene in the pathogenesis of HBL is yet to be determined. Additionally, the 2q24–32 region contains several genes that may also have an oncogenic potential. These include a serine/threonine kinase receptor, *ITRAF*, *FRZB*, a secreted antagonist of WNT signaling, and BRCA1-associated RING domain 1 (*BARD1*) genes. However, no specific gene has been identified in the previous,[21,22] and present studies.

The losses in chromosomes 4q and 11q were comprehensively observed. In hepatocellular carcinoma (HCC) cells, Wong *et al.* demonstrated a growth advantage following the loss in the 4q arm.[40] In HCC, 4q21–q22 and 4q35 have been identified as commonly deleted regions, and allelic losses in 4q35 have been associated with a larger tumor size and an aggressive histological tumor type.[41] Previous studies have not reported a significant correlation between HBL with loss in the distal 4q arm and prognosis, but the underlying oncogenic event might be due to the loss of a gene on the distal 4q arm.

Many minimal regions of amplification and deletion were detected using high-density SNP arrays, although homozygous deletion was not identified in any sample. The SNP loci located in 7q34 and 14q11.2 were found to be highly amplified in sporadic HBL samples. The candidate genes at these loci are *EphB6*, *DAD1*, and BCL-like 2 (*BCL2L2*) genes that encode the proteins associated with the execution of cell apoptosis. Gains as well as high amplifications in this region have not been reported previously; however, such an observation will be of particular interest for the discovery of oncogenes involved in the pathogenesis of HBL.

The UPD regions were identified in five of the 17 samples. This is chiefly important because UPD is being particularly considered as a possible mechanism of tumor initiation. During tumorigenesis, UPD is believed to arise due to a mitotic recombination caused by a rare crossover event during mitotic cell division. The products of mitotic recombination are the regions of the genome exhibiting UPD, and both the genomic regions originate from the same parent. We could identify a common UPD on chromosome 11p that is reminiscent of BWS with paternal UPD; in this case, the loss of function of the 11p15 maternal alleles through various mechanisms may be the critical event associated with tumorigenesis and BWS.[42] BWS is a neonatal overgrowth syndrome that predisposes an individual to cancer,[6] and the importance of the maternally active locus in chromosome 11p15 in tumorigenesis is supported by the finding that the loss of imprinted allele and paternal duplication leads to tissue overgrowth and subsequent tumor development. Methylation analysis was performed for the four HBL samples having UPD within 11p15, and hypermethylation of *H19* DMR was detected in all four HBL samples. Because *H19* DMR was hypermethylated on the paternal allele and hypomethylated on the maternally expressed allele in humans, we consider that the UPD within 11p15 was of paternal origin.

Two candidate genes, namely, *IGF2* and *H19*, are located within the telomeric region of chromosome 11p15.5 and have opposite imprinting patterns.[43] In the majority of human tissues, *IGF2* is expressed only from the paternal allele, whereas *H19* is transcribed only from the maternal allele. *H19* is an untranslated gene but has been suggested to function as a tumor suppressor.[44] In fetal and adult organs, the transcriptionally silent *H19* allele was extensively hypermethylated throughout the entire gene and its promoter. On the maternally expressed *H19* allele, *H19* DMR is unmethylated and can bind to the CTCF protein. On the paternal *H19* allele, *H19* DMR is highly methylated. This not only prevents the expression of the imprinted paternal *H19* alleles but also blocks the binding of the CTCF protein.[43] In general, the outcome of UPD with losses of the 11p15 maternal alleles in HBL is the same as that of the loss of imprinting on the inactivated, imprinted, and maternally expressed genes in BWS. Weksberg *et al.* proposed a dual pathway model for tumor development in BWS, wherein methylation defects at *H19* and/or *IGF2* in 11p15 were found to play a role in Wilms' and HBL tumorigenesis.[45] The combined loss of expressions in various 11p15-imprinted genes may contribute to tumorigenesis.

In the present study, we identified that the expression patterns of *IGF2* and *H19* were opposite between genes with and without the UPD in 11p15. This difference in the expression patterns might influence the clinical features of HBL. Further prospective studies are required to reveal any potential correlations between specific LOH and clinical outcomes.

In summary, the analysis of LOH and CN alterations using the SNP microarray in HBL samples revealed significant areas of allelic imbalance. We hypothesize that UPD, in addition to allelic imbalance, constitutes a novel genetic mechanism involved in tumorigenesis. Therefore, detailed characterizations such as functional studies should be conducted to elucidate the significance of the regions detected in this study, many of which may contain the candidate tumor suppressor genes and oncogenes involved in the pathogenesis of HBL.

Acknowledgments

This work was supported in part by a Grant-in-Aid for Cancer Research from the Ministry of Health, Labour and Welfare of Japan, a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan, and a grant from the 21st Century COE Program from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

1 Roebuck DJ, Perilongo G. Hepatoblastoma: an oncological review. *Pediatr Radiol* 2006; **36**: 183–6.
2 Tiao GM, Bobey N, Allen S *et al.* The current management of hepatoblastoma: a combination of chemotherapy, conventional resection, and liver transplantation. *J Pediatr* 2005; **146**: 204–11.
3 Schnater JM, Kohler SE, Lamers WH, von Schweinitz D, Aronson DC. Where do we stand with hepatoblastoma? A review. *Cancer* 2003; **98**: 668–78.
4 Ikeda H, Matsuyama S, Tanimura M. Association between hepatoblastoma and very low birth weight: a trend or a chance? *J Pediatr* 1997; **130**: 557–60.
5 Hughes LJ, Michels VV. Risk of hepatoblastoma in familial adenomatous polyposis. *Am J Med Genet* 1992; **43**: 1023–5.
6 DeBaun MR, Tucker MA. Risk of cancer during the first four years of life in children from The Beckwith–Wiedemann Syndrome Registry. *J Pediatr* 1998; **132**: 398–400.
7 Fukuzawa R, Hata J, Hayashi Y, Ikeda H, Reeve AE. Beckwith–Wiedemann syndrome-associated hepatoblastoma: wnt signal activation occurs later in

tumorigenesis in patients with 11p15.5 uniparental disomy. *Pediatr Dev Pathol* 2003; **6**: 299–306.
8 Little MH, Thomson DB, Hayward NK, Smith PJ. Loss of alleles on the short arm of chromosome 11 in a hepatoblastoma from a child with Beckwith–Wiedemann syndrome. *Hum Genet* 1988; **79**: 186–9.
9 Albrecht S, von Schweinitz D, Waha A, Kraus JA, von Deimling A, Pietsch T. Loss of maternal alleles on chromosome arm 11p in hepatoblastoma. *Cancer Res* 1994; **54**: 5041–4.
10 Oda H, Imai Y, Nakatsuru Y, Hata J, Ishikawa T. Somatic mutations of the APC gene in sporadic hepatoblastomas. *Cancer Res* 1996; **56**: 3320–3.
11 Lengauer C, Kinzler KW, Vogelstein B. Genetic instabilities in human cancers. *Nature* 1998; **396**: 643–9.
12 Yeh YA, Rao PH, Cigna CT, Middlesworth W, Lefkowitch JH, Murty VV. Trisomy 1q, 2, and 20 in a case of hepatoblastoma: possible significance of 2q35-q37 and 1q12-q21 rearrangements. *Cancer Genet Cytogenet* 2000; **123**: 140–3.
13 Nagata T, Mugishima H, Shichino H *et al.* Karyotypic analyses of hepatoblastoma. Report of two cases and review of the literature suggesting chromosomal loci responsible for the pathogenesis of this disease. *Cancer Genet Cytogenet* 1999; **114**: 42–50.
14 Sainati L, Leszl A, Stella M *et al.* Cytogenetic analysis of hepatoblastoma: hypothesis of cytogenetic evolution in such tumors and results of a multicentric study. *Cancer Genet Cytogenet* 1998; **104**: 39–44.
15 Tonk VS, Wilson KS, Timmons CF, Schneider NR. Trisomy 2, trisomy 20, and del (17p) as sole chromosomal abnormalities in three cases of hepatoblastoma. *Genes Chromosomes Cancer* 1994; **11**: 199–202.
16 Park JP, Omvold KT, Brown AM, Mohandas TK. Trisomy 2 and 19, and tetrasomy 1q and 14 in hepatoblastoma. *Cancer Genet Cytogenet* 1999; **115**: 86–7.
17 Balogh E, Swanton S, Kiss C, Jakab ZS, Secker-Walker LM, Olah E. Fluorescence in situ hybridization reveals trisomy 2q by insertion into 9p in hepatoblastoma. *Cancer Genet Cytogenet* 1998; **102**: 148–50.
18 Sainati L, Leszl A, Surace C, Perilongo G, Rocchi M, Basso G. Fluorescence in situ hybridization improves cytogenetic results in the analysis of hepatoblastoma. *Cancer Genet Cytogenet* 2002; **134**: 18–20.
19 Surace C, Leszl A, Perilongo G, Rocchi M, Basso G, Sainati L. Fluorescent in situ hybridization (FISH) reveals frequent and recurrent numerical and structural abnormalities in hepatoblastoma with no informative karyotype. *Med Pediatr Oncol* 2002; **39**: 536–9.
20 Parada LA, Limon J, Iliszko M *et al.* Cytogenetics of hepatoblastoma: further characterization of 1q rearrangements by fluorescence in situ hybridization: an international collaborative study. *Med Pediatr Oncol* 2000; **34**: 165–70.
21 Hu J, Wills M, Baker BA, Perlman EJ. Comparative genomic hybridization analysis of hepatoblastomas. *Genes Chromosomes Cancer* 2000; **27**: 196–201.
22 Weber RG, Pietsch T, von Schweinitz D, Lichter P. Characterization of genomic alterations in hepatoblastomas. A role for gains on chromosomes 8q and 20 as predictors of poor outcome. *Am J Pathol* 2000; **157**: 571–8.
23 Janne PA, Li C, Zhao X *et al.* High-resolution single-nucleotide polymorphism array and clustering analysis of loss of heterozygosity in human lung cancer cell lines. *Oncogene* 2004; **23**: 2716–26.
24 Huang J, Wei W, Zhang J *et al.* Whole genome DNA copy number changes identified by high density oligonucleotide arrays. *Hum Genomics* 2004; **1**: 287–99.
25 Peiffer DA, Le JM, Steemers FJ *et al.* High-resolution genomic profiling of chromosomal aberrations using Infinium whole-genome genotyping. *Genome Res* 2006; **16**: 1136–48.
26 Zhao X, Li C, Paez JG *et al.* An integrated view of copy number and allelic alterations in the cancer genome using single nucleotide polymorphism arrays. *Cancer Res* 2004; **64**: 3060–71.
27 Nannya Y, Sanada M, Nakazaki K *et al.* A robust algorithm for copy number detection using high-density oligonucleotide single nucleotide polymorphism genotyping arrays. *Cancer Res* 2005; **65**: 6071–9.
28 Yamamoto G, Nannya Y, Kato M *et al.* Highly sensitive method for genomewide detection of allelic composition in nonpaired, primary tumor specimens by use of affymetrix single-nucleotide-polymorphism genotyping microarrays. *Am J Hum Genet* 2007; **81**: 114–26.
29 Wong KK, Tsang YT, Shen J *et al.* Allelic imbalance analysis by high-density single-nucleotide polymorphic allele (SNP) array with whole genome amplified DNA. *Nucleic Acids Res* 2004; **32**: e69.
30 Trask BJ. Fluorescence in situ hybridization: applications in cytogenetics and gene mapping. *Trends Genet* 1991; **7**: 149–54.
31 Herman JG, Graff JR, Myohanen S, Nelkin BD, Baylin SB. Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. *Proc Natl Acad Sci USA* 1996; **93**: 9821–6.
32 Li LC, Dahiya R. MethPrimer: designing primers for methylation PCRs. *Bioinformatics* 2002; **18**: 1427–31.
33 Koch A, Waha A, Hartmann W *et al.* Elevated expression of Wnt antagonists is a common event in hepatoblastomas. *Clin Cancer Res* 2005; **11**: 4295–304.
34 Douglass EC, Green AA, Hayes FA, Etcubanas E, Horowitz M, Wilimas JA. Chromosome 1 abnormalities: a common feature of pediatric solid tumors. *J Natl Cancer Inst* 1985; **75**: 51–4.
35 Kaneko Y, Variakojis D, Kluskens L, Rowley JD. Lymphoblastic lymphoma: cytogenetic, pathologic, and immunologic studies. *Int J Cancer* 1982; **30**: 273–9.
36 Kaneko Y, Kondo K, Rowley JD, Moohr JW, Maurer HS. Further chromosome studies on Wilms' tumor cells of patients without aniridia. *Cancer Genet Cytogenet* 1983; **10**: 191–7.
37 Nilsson M, Meza-Zepeda LA, Mertens F, Forus A, Myklebost O, Mandahl N. Amplification of chromosome 1 sequences in lipomatous tumors and other sarcomas. *Int J Cancer* 2004; **109**: 363–9.
38 Kissil JL, Kimchi A. Assignment of death associated protein 3 (DAP3) to human chromosome 1q21 by in situ hybridization. *Cytogenet Cell Genet* 1997; **77**: 252.
39 Turc-Carel C, Lizard-Nacol S, Justrabo E, Favrot M, Philip T, Tabone E. Consistent chromosomal translocation in alveolar rhabdomyosarcoma. *Cancer Genet Cytogenet* 1986; **19**: 361–2.
40 Wong N, Lai P, Lee SW *et al.* Assessment of genetic changes in hepatocellular carcinoma by comparative genomic hybridization analysis: relationship to disease stage, tumor size, and cirrhosis. *Am J Pathol* 1999; **154**: 37–43.
41 Bando K, Nagai H, Matsumoto S *et al.* Identification of a 1-cM region of common deletion on 4q35 associated with progression of hepatocellular carcinoma. *Genes Chromosomes Cancer* 1999; **25**: 284–9.
42 Koufos A, Hansen MF, Copeland NG, Jenkins NA, Lampkin BC, Cavenee WK. Loss of heterozygosity in three embryonal tumours suggests a common pathogenetic mechanism. *Nature* 1985; **316**: 330–4.
43 Hark AT, Schoenherr CJ, Katz DJ, Ingram RS, Levorse JM, Tilghman SM. CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature* 2000; **405**: 486–9.
44 Zhang Y, Shields T, Crenshaw T, Hao Y, Moulton T, Tycko B. Imprinting of human H19: allele-specific CpG methylation, loss of the active allele in Wilms tumor, and potential for somatic allele switching. *Am J Hum Genet* 1993; **53**: 113–24.
45 Weksberg R, Nishikawa J, Caluseriu O *et al.* Tumor development in the Beckwith–Wiedemann syndrome is associated with a variety of constitutional molecular 11p15 alterations including imprinting defects of KCNQ1OT1. *Hum Mol Genet* 2001; **10**: 2989–3000.

doi: 10.1111/j.1349-7006.2007.00710.x

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.

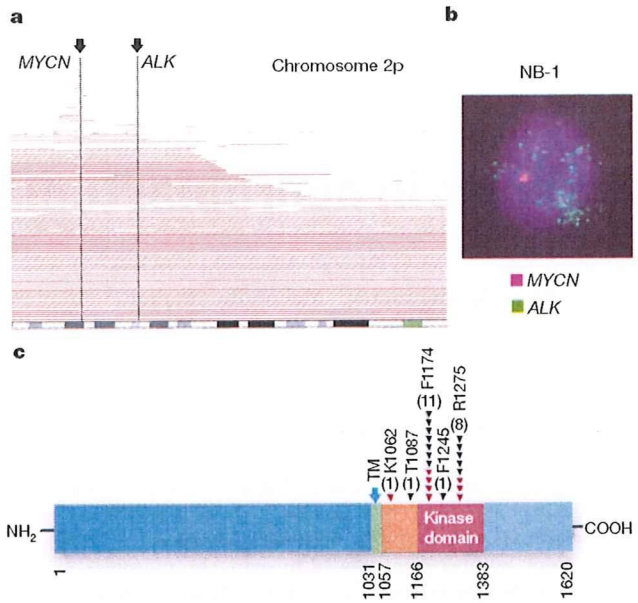

Figure 1 | Common 2p gains/amplifications and *ALK* mutations in neuroblastoma samples. a, Recurrent copy number gains on the 2p arm. High-grade amplifications are shown by light-red horizontal lines, whereas simple gains are shown by dark-red lines. Two common peaks of copy number gains and amplifications 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).

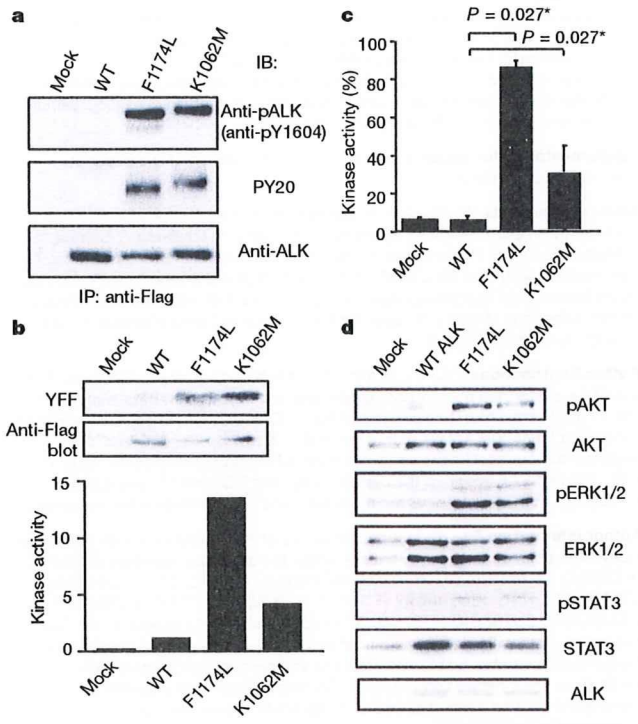

Figure 2 | Kinase activity of ALK mutants and their downstream signalling. **a**, Stably expressed ALK and its mutants (F1174L and K1062M) were immunoprecipitated with an anti-Flag antibody and subjected to western blot analysis with anti-pY1604 (upper panel) or PY20 (middle panel). An anti-ALK blot 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).

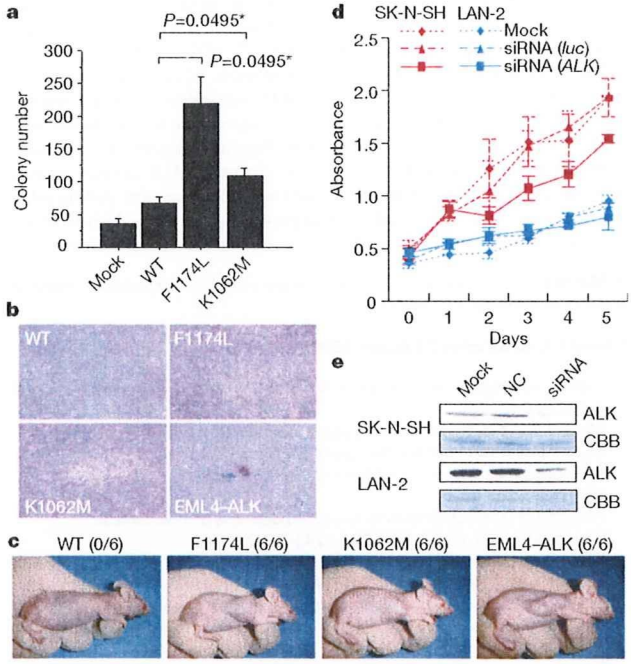

Figure 3 | Oncogenic role of ALK mutations. **a**, Colony assays for NIH3T3 cells stably expressing wild-type as well as mutant ALK (F1174L and K1062M). The average numbers of colonies in triplicate experiments are plotted and standard deviation is indicated. Results showing statistically significant 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)[15], 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′-AATTGCGGCCGCTACTTGTCATCGTCGTCCTTGTAGTCGGGGCCCAGGCTG 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).

ELSEVIER

Cancer Genetics and Cytogenetics 186 (2008) 115–119

CANCER GENETICS AND CYTOGENETICS

Short communication

MNX1–ETV6 fusion gene in an acute megakaryoblastic leukemia and expression of the *MNX1* gene in leukemia and normal B cell lines

Takeshi Taketani[a,b], Tomohiko Taki[c], Masahiro Sako[d], Takefumi Ishii[d], Seiji Yamaguchi[a], Yasuhide Hayashi[e,*]

[a]*Department of Pediatrics, Shimane University Faculty of Medicine, Izumo, Shimane, Japan*
[b]*Division of Blood Transfusion, Shimane University Hospital, Matsue, Shimane, Japan*
[c]*Department of Molecular Laboratory Medicine, Kyoto Prefectural University of Medicine Graduate School of Medical Science, Kyoto, Japan*
[d]*Department of Pediatric Hematology/Oncology, Osaka City General Hospital, Osaka, Japan*
[e]*Department of Hematology/Oncology, Gunma Children's Medical Center, 779 Shimohakoda, Hokkitsu, Shibukawa, Gunma 377-8577, Japan*

Received 5 March 2008; received in revised form 11 June 2008; accepted 27 June 2008

Abstract Patients with infant acute myeloid leukemia (AML) who carry a t(7;12)(q36;p13) translocation have been reported to have a poor clinical outcome. *MNX1–ETV6* fusion transcripts (previously *HLXB9–ETV6*) were rarely detected in AML patients having t(7;12)(q36;p13). A 23-month-old girl with acute megakaryoblastic leukemia (AMKL) exhibited chromosome abnormalities, including add(7)(q22), and del(12)(p12p13). Southern blot analysis of bone marrow cells showed an *ETV6* gene rearrangement. Reverse transcriptase-polymerase chain reaction (RT-PCR) followed by sequence analysis revealed the presence of an *MNX1–ETV6* fusion gene. The patient responded well to chemotherapy, achieved complete remission, and at writing had been in complete remission for 60 months. The *MNX1* expression by RT-PCR was significantly more frequent in Epstein–Barr virus–transformed B-cell lines derived from normal adult lymphocytes than in leukemic cell lines. This represents a novel case of an AMKL patient with *MNX1–ETV6* fusion transcripts who had a good prognosis.

1. Introduction

Many recurrent chromosomal translocations are involved in acute myeloid leukemia (AML) [1]. AML with 12p13 translocations have been reported to involve the ETS variant gene 6 (*TEL* oncogene) (*ETV6*) [2]. In cases of AML carrying 12p13 abnormalities, a recurrent translocation t(7;12)(q36;p13) is found in 20%–30% of infant cases [3–5]. Fluorescence in situ hybridization assay is needed to evaluate this translocation, because it is difficult to detect by conventional karyotyping [3–5]. AML patients with this translocation are characterized by age under 20 months at diagnosis, thrombocytosis, high percentage of CD34-positive cells, presence of additional chromosomal abnormalities, including trisomy 19 or trisomy 8 (or both), and a poor prognosis [3–5]. An *MNX1–ETV6* fusion gene (previously *HLXB9–ETV6*) was identified in two pediatric AML patients having t(7;12)(q36;p13) [6]; however, heterogeneity of the 7q36 and 12p13 translocations was reported [5,7,8]. Thus, *MNX1–ETV6* fusion gene in AML patients having t(7;12) is infrequently reported [7,8].

We describe the case of a 23-month-old AML patient with add(7)(q22), del(12)(p12p13), and *MNX1–ETV6* fusion transcript; the child has remained alive for 5 years. We also report the expression of the *MNX1* gene in several leukemic and normal Epstein–Barr virus-transformed cell lines.

2. Case report

A 23-month-old girl was admitted to Osaka City General Hospital because of appetite loss and pallor. Blood examination showed a white blood cell count of 10,520/μL with 55.5% blasts, a hemoglobin level of 7.0 g/dL, and a platelet count of 164,000/μL. She had a mediastinal mass, but no hepatosplenomegaly. Bone marrow examination revealed a nuclear cell count of 30,000/μL with 71.2% blasts. The

* Corresponding author. Tel.: +81-279-52-3551, ext. 2200; fax: +81-279-52-2045.
E-mail address: hayashiy-tky@umin.ac.jp (Y. Hayashi).

0165-4608/08/$ – see front matter
doi:10.1016/j.cancergencyto.2008.06.009

blasts were negative for myeloperoxidase staining and platelet peroxidase staining electron-microscopically. Flow cytometric analysis showed that the blasts expressed CD41, CD36, CD13, CD33, CD15, and CD7 antigens, suggesting megakaryoblastic origin. Conventional G-banding chromosomal analysis revealed a karyotype of 46,XX,add (7)(q22),del(12)(p12p13) in all 20 bone marrow cells examined (Fig. 1).

The patient was diagnosed as having AMKL (M7 subtype, based on the French–American–British classification), and was treated on the Japanese Childhood AML Cooperative Study Group Protocol, AML99 [9]. She obtained complete remission with induction chemotherapy (cytarabine, etoposide, and mitoxantrone). Thereafter, she was treated with five additional courses of intensive chemotherapy (high-dose cytarabine, etoposide, idarubicin, and mitoxantrone). As of writing, she had been in complete remission for 60 months after diagnosis.

3. Materials and methods

3.1. Southern blot analysis

High molecular weight DNA was extracted from bone marrow cells of the patient by proteinase K digestion and phenol–chloroform extraction [10]. Ten micrograms of DNA was digested with *Eco*RI, subjected to electrophoresis on 0.7% agarose gels, and transferred to nylon membrane, and hybridized to cDNA probes ^{32}P-labeled by the random hexamer method [10]. The probes used were a 516-bp *MNX1* cDNA fragment (nucleotide nt598 to nt1114; GenBank accession no. NM_005515; previously *HLXB9*).

3.2. Expression of WT1 mRNA and mutation of FLT3

WT1 mRNA was examined for detection of minimal residual disease as previously reported [11]. Internal tandem duplication and mutation of *FLT3* were examined as previously reported [10].

3.3. Reverse transcriptase–polymerase chain reaction and nucleotide sequencing

MNX1–ETV6 chimeric mRNA was detected by reverse transcriptase–polymerase chain reaction (RT-PCR) as described previously [12]. Total RNA was extracted from the leukemic cells of the patient using the acid guanidine thiocyanate–phenol chloroform method [12]. Total RNA (4 μg) was reverse-transcribed to cDNA, using a cDNA synthesis kit (GE Healthcare Bio-Science, Piscataway, NJ) [12]. PCR was performed with AmpliTaq Gold DNA polymerase (Applied Biosystems, Foster City, CA; Tokyo, Japan), using the reagents recommended by the manufacturer. The primers used and PCR conditions were as described previously [6]. The PCR products were subcloned into pCR2.1 vector (Invitrogen, Carlsbad, CA) and sequenced by the fluorometric method using the BigDye Terminator cycle sequencing kit (Applied Biosystems).

3.4. Expression of the MNX1 gene by RT-PCR in leukemic cell lines

To analyze the expression pattern of the *MNX1* gene in leukemic cell lines, RT-PCR was performed. Fifty-nine cell lines were examined, as follows [12]; 10 B-precursor ALL cell lines (LC4-1, NALM-6, NALM-24, NALM-26, UTP-2, RS4;11, SCMC-L10, KOCL-33, KOCL-45, KOCL-69), 9

Fig. 1. G-banding karyotype of the leukemic cells in a pediatric patient with acute megakaryoblastic leukemia: 46,XX,add(7)(q22),del(12)(p12p13).

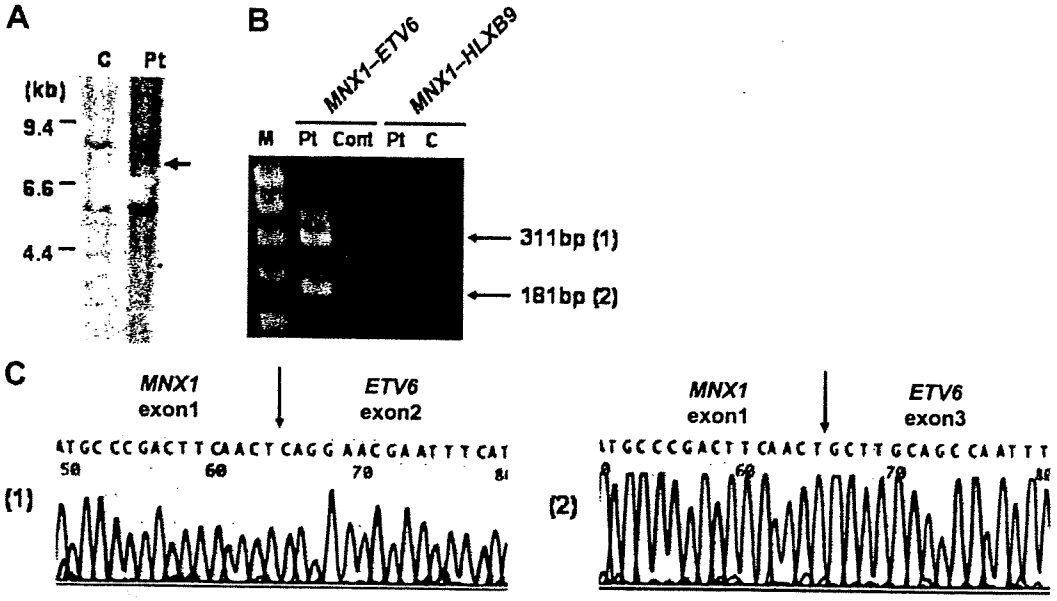

Fig. 2. Detection of the *MNX1–ETV6* fusion gene (previously *HLXB9–ETV6*). (A) Rearrangement of the *MNX1* gene by Southern blotting with *Eco*RI digestion. The arrow indicates a rearranged band of the *MNX1* gene; C, control; Pt, patient. (B) The *MNX1–ETV6* fusion transcript identified by reverse transcriptase–polymerase chain reaction (RT-PCR). Lanes 2 and 3, *MNX1–ETV6* fusion transcript; lanes 4 and 5, *ETV6–MNX1* fusion transcript. C, control; M, size marker; Pt, patient; (C) Nucleotide and amino acid sequencing of two *MNX1–ETV6* fusion transcripts.

B-ALL cell lines (BALM-1, BALM-13, BALM-14, BJAB, DAUDI, RAJI, RAMOS, BAL-KH, NAMALLA), 9 T-ALL cell lines (RPMI-8402, MOLT-14, THP-6, PEER, H-SB2, HPB-ALL, L-SAK, L-SMY, KCMC-T), 8 AML cell lines (YNH-1, ML-1, KASUMI-3, KG-1, inv-3, SN-1, NB4, HEL), 6 acute monocytic leukemia cell lines (THP-1, IMS/M1, CTS, P31/FUJ, MOLM-13, KOCL-48), 5 chronic myelogenous leukemia cell lines (MOLM-1, MOLM-7, TS9;22, SS9;22, K-562), 2 acute megakaryoblastic leukemia cell lines (CMS, CMY), and 10 Epstein-Barr virus transformed B lymphocyte (EBV-B) cell lines derived from normal adult peripheral lymphocytes. Five normal BM samples were also examined. RT-PCR mixtures and conditions were the same as described [10]. The primers used for RT-PCR were HLXB9-658F (5′-GGCATGATCCTGCCTAAGAT-3′) (sense primer) and HLXB9-1092R (TGCT GTAGGGGAAATGGTCGTCG) (antisense primer) [6].

4. Results and discussion

The karyotype of the patient's leukemic cells was 46,XX,add(7)(q22),del(12)(p12p13), suggesting that both *ETV6* and *MNX1* were involved in this chromosomal abnormality. With informed consent from the patient's parents, DNA and total RNA were extracted from bone marrow cells of the patient. Southern blot analysis of DNA from leukemic cells of the patient using the *MNX1* probe showed a rearranged band (Fig. 2A). We performed RT-PCR for *MNX1–ETV6* chimeric mRNA and obtained two RT-PCR products, of 311 bp and 181 bp (Fig. 2B). Sequence analysis of these PCR products showed that one product was an in-frame fusion transcript of exon 1 of *MNX1* to exon 3 of *ETV6*, and the other was an out-of-frame fusion transcript of exon 1 of *MNX1* to exon 2 of *ETV6* (Fig. 2C). These transcripts were the same as previously reported [6]. The reciprocal *ETV6–MNX1* transcript was not detected (Fig. 2B).

The *WT1* mRNA level was 3,400 copies/μg RNA at diagnosis, but decreased to <50 copies/μg RNA after remission. Neither internal tandem duplication nor mutation of *FLT3* were found in this patient, suggesting that the prognosis is not poor [1].

Table 1
Expression of the *MNX1* gene in leukemia and EBV-B cell lines by reverse transcriptase–polymerase chain reaction

Cell line	Cells examined, no.	Cells expressing *MNX1*, no. (%)
ALL	28	5 (17.9)
B precursor	10	0 (0)
B	9	2 (22.2)
T	9	3 (33.3)
AML	16	3 (18.8)
AML	8	1 (12.5)
AMoL	6	2 (33.3)
AMKL	2	0 (0)
CML	5	1 (20)
EBV-B	10	7 (70)
normal BM	5	0 (0)

Abbreviations: ALL, acute lymphoblastic leukemia; AMKL, acute megakaryoblastic leukemia; AML, acute myeloid leukemia; AMoL, acute monocytic leukemia; B, B-cell; BM, bone marrow; CML, chronic myelogenous leukemia; EBV-B, Epstein–Barr virus-transformed human B lymphocytes; T, T-cell.

Table 3. High-grade amplifications in hepatoblastoma (HBL) samples

Cytoband	Implicated region (base pairs)		Candidate target genes in the region
	Start–end	Size	
2q34	211 193 864–212 239 181	1 045 318	*ErbB4*
3p25.2	11 888 124–12 876 175	988 052	*RAF1*
7q34	141 721 559–142 076 238	354 680	*EphB6*
11q22.2-q22.3	101 394 973–102 830 195	1 435 223	*MMP1, 7, 20*
14q11.2	21 426 631–22 130 392	703 762	*DAD1*

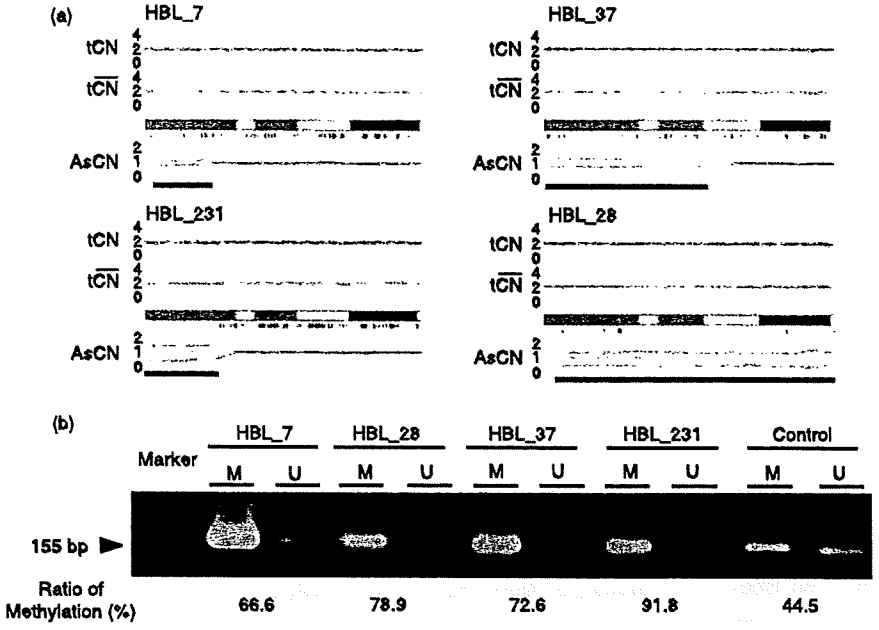

Fig. 4. (a) Copy numbers (CN) of chromosome 11p in four hepatoblastoma (HBL) samples with uniparental disomy (UPD). Although complete CN alterations are not observed, UPD is clearly predicted based on the allele-specific CN alterations (green lines). (b) Methylation-specific polymerase chain reaction (PCR) analysis of the *H19* differential methylated region (DMR). Modified DNA was amplified with primer pairs for methylated and unmethylated complete sequences of the *H19* DMR. *H19* DMR hypermethylation was detected in all HBL samples; however, normal lymphocyte DNA exhibited the mosaic methylation pattern. The results of quantitative real-time methylation-specific PCR analysis are shown below the image depicting the results of electrophoresis.

and hypomethylated on the maternally expressed allele in humans. This indicates that the UPD within this region is considered to be derived from the paternal allele. Furthermore, a low expression level of the non-methylated allele was also observed; methylation-specific RQ-PCR analysis revealed that the ratio of the methylation status ranged from 66.6% to 91.8%.

Expression analyses using RQ-RT-PCR. In order to examine the impact of the abovementioned amplifications and UPD on gene expression, we measured the expression levels of four genes (*DAD1*, *ErbB4*, *IGF2*, and *H19*) through RQ-RT-PCR (Fig. 5). Normal liver total RNA served as the non-neoplastic reference and control. HBL_184 and HBL_231 for which RNA were available showed a high expression of the *ErbB4* gene. However, the expression of *DAD1* was down-regulated in both these samples. The *IGF2* and *H19* genes were oppositely expressed between HBL_184 and HBL_231, having UPD within 11p15.

Discussion

The present study represents the application of the SNP array technology for the genome-wide analysis of CN aberrations in HBL. Several recent studies and our previous research have demonstrated that this technology provided a unique opportunity to assess the DNA CN alterations and LOH simultaneously throughout the entire genome.[24–27,29] As shown in the present analysis, the use of high-resolution SNP arrays improved the ability to identify structural chromosomal aberrations in cancer cells and detect genes affected by these aberrations. Additionally, high-density SNP array analysis with the CN analyzer software can also

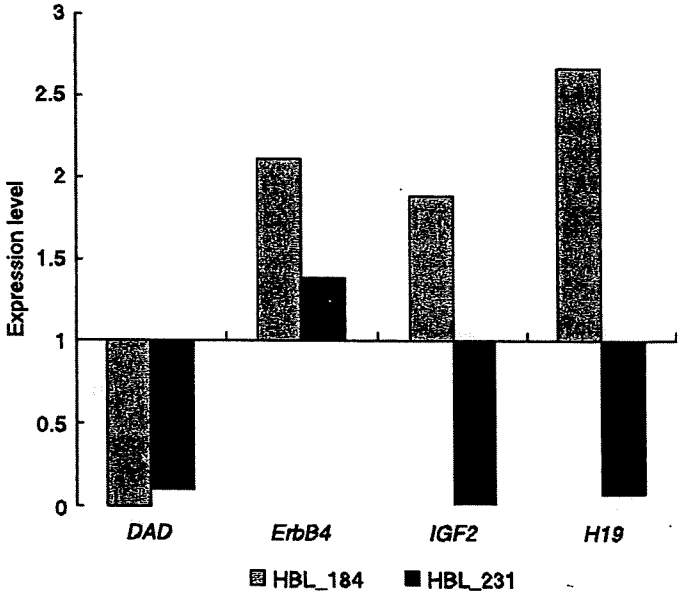

Fig. 5. The results of the expression levels of four genes (defender against cell death 1 [*DAD1*], EPH receptor B6 [*EphB6*], *ErbB4*, insulin-like growth factor II [*IGF2*], and *H19* genes) through real-time quantitative reverse transcription-polymerase chain reaction (RQ-RT-PCR) analyses.

doi: 10.1111/j.1349-7006.2007.00710.x

High incidence of t(7;12)(q36;p13) in infant AML but not in infant ALL, with a dismal outcome and ectopic expression of *HLXB9*. Genes Chromosomes Cancer 2006;45:731–9.

[6] Beverloo HB, Panagopoulos I, Isaksson M, van Wering E, van Drunen E, de Klein A, Johansson B, Slater R. Fusion of the homeobox gene *HLXB9* and the *ETV6* gene in infant acute myeloid leukemias with the t(7;12)(q36;p13). Cancer Res 2001;61:5374–7.

[7] Simmons HM, Oseth L, Nguyen P, O'Leary M, Conklin KF, Hirsch B. Cytogenetic and molecular heterogeneity of 7q36/12p13 rearrangements in childhood AML. Leukemia 2002;16:2408–16.

[8] Tosi S, Hughes J, Scherer SW, Nakabayashi K, Harbott J, Haas OA, Cazzaniga G, Biondi A, Kempski H, Kearney L. Heterogeneity of the 7q36 breakpoints in the t(7;12) involving *ETV6* in infant leukemia. Genes Chromosomes Cancer 2003;38:191–200.

[9] Shimada A, Taki T, Tabuchi K, Tawa A, Horibe K, Tsuchida M, Hanada R, Tsukimoto I, Hayashi Y. *KIT* mutations, and not *FLT3* internal tandem duplication, are strongly associated with a poor prognosis in pediatric acute myeloid leukemia with t(8;21): a study of the Japanese Childhood AML Cooperative Study Group. Blood 2006;107:1806–9.

[10] Taketani T, Taki T, Sugita K, Furuichi Y, Ishii E, Hanada R, Tsuchida M, Sugita K, Ida K, Hayashi Y. *FLT3* mutations in the activation loop of tyrosine kinase domain are frequently found in infant ALL with *MLL* rearrangements and pediatric ALL with hyperdiploidy. Blood 2004;103:1085–8.

[11] Inoue K, Ogawa H, Yamagami T, Soma T, Tani Y, Tatekawa T, Oji Y, Tamaki H, Kyo T, Dohy H, Hiraoka A, Masaoka T, Kishimoto T, Sugiyama H. Long-term follow-up of minimal residual disease in leukemia patients by monitoring *WT1* (Wilms tumor gene) expression levels. Blood 1996;88:2267–78.

[12] Taketani T, Taki T, Shibuya N, Kikuchi A, Hanada R, Hayashi Y. Novel *NUP98–HOXC11* fusion gene resulted from a chromosomal break within exon 1 of *HOXC11* in acute myeloid leukemia with t(11;12)(p15;q13). Cancer Res 2002;62:4571–4.

[13] Deguchi Y, Kehrl JH. Selective expression of two homeobox genes in CD34-positive cells from human bone marrow. Blood 1991;78: 323–8.

[14] Deguchi Y, Yamanaka Y, Theodossiou C, Najfeld V, Kehrl JH. High expression of two diverged homeobox genes, HB24 and HB9, in acute leukemias: molecular markers of hematopoietic cell immaturity. Leukemia 1993;7:446–51.

[15] Harrison KA, Druey KM, Deguchi Y, Tuscano JM, Kehrl JH. A novel human homeobox gene distantly related to proboscipedia is expressed in lymphoid and pancreatic tissues. J Biol Chem 1994;269: 19968–75.

Cloning of genes involved in chromosomal translocations by high-resolution single nucleotide polymorphism genomic microarray

Norihiko Kawamata[a,b,c], Seishi Ogawa[b,d], Martin Zimmermann[b,e], Birte Niebuhr[f], Carol Stocking[f], Masashi Sanada[d], Kari Hemminki[g], Go Yamatomo[d], Yasuhito Nannya[d], Rolf Koehler[h], Thomas Flohr[h], Carl W. Miller[a], Jochen Harbott[i], Wolf-Dieter Ludwig[j], Martin Stanulla[e], Martin Schrappe[k], Claus R. Bartram[h,l], and H. Phillip Koeffler[a,l]

[a]Hematology/Oncology, Cedars-Sinai Medical Center/UCLA School of Medicine, Los Angeles, CA 90048; [d]Regeneration Medicine of Hematopoiesis, School of Medicine, University of Tokyo, Tokyo 113-8655, Japan; [e]Department of Pediatric Hematology and Oncology, Children's Hospital, Hannover Medical School, 30625 Hannover, Germany; [f]Molecular Pathology Heinrich-Pette-Institute, 20251 Hamburg, Germany; [g]Division of Molecular Genetic Epidemiology, German Cancer Research Center, 69120 Heidelberg, Germany; [h]Institute of Human Genetics, University of Heidelberg, 69117 Heidelberg, Germany; [i]Department of Hematology and Oncology, Center for Pediatrics, 35390 Giessen, Germany; [j]Department of Hematology, Oncology and Tumor Immunology, Robert-Rössle-Clinic at the HELIOS-Clinic Berlin-Buch, Charité, 13125 Berlin, Germany; and [k]Department of Pediatrics, University of Kiel, 69117 Kiel, Germany

Edited by Joe W. Gray, Lawrence Berkeley National Laboratory, Berkeley, CA and accepted by the Editorial Board June 13, 2008 (received for review November 21, 2007)

High-resolution single nucleotide polymorphism genomic microarray (SNP-chip) is a useful tool to define gene dosage levels over the whole genome, allowing precise detection of deletions and duplications/amplifications of chromosomes in cancer cells. We found that this new technology can also identify breakpoints of chromosomes involved in unbalanced translocations, leading to identification of fusion genes. Using this technique, we found that the *PAX5* gene was rearranged to a variety of partner genes including *ETV6*, *FOXP1*, *AUTS2*, and *C20orf112* in pediatric acute lymphoblastic leukemia (ALL). The 3′ end of the *PAX5* gene was replaced by the partner gene. The PAX5 fusion products bound to PAX5 recognition sequences as strongly as wild-type PAX5 and suppressed its transcriptional activity in a dominant-negative fashion. In human B cell leukemia cells, binding of wild-type PAX5 to a regulatory region of *BLK*, one of the direct downstream target genes of PAX5, was diminished by expression of the PAX5-fusion protein, leading to repression of *BLK*. Expression of PAX5-fusion genes in murine bone marrow cells blocked development of mature B cells. PAX5-fusion proteins may contribute to leukemogenesis by blocking differentiation of hematopoietic cells into mature B cells. SNP-chip is a powerful tool to identify fusion genes in human cancers.

chromatin immunoprecipitation | dominant negative | fusion gene | PAX5 | SNP-chip

Pediatric acute lymphoblastic leukemia (ALL) is the most common malignant disease in children (1–3). It is a genetic abnormality resulting from accumulation of mutations in tumor suppressor genes and oncogenes (1–3). Fusion genes including *ETV6/RUNX1* and *E2A/PBX1* are frequently detected in pediatric ALL (1). Deletion of the *INK4A/ARF* gene (9p21) is also a common abnormality in ALL (1). However, other genetic changes remain to be elucidated in this disease.

Identification of mutated genes in ALL has evolved with improvements in technology. A very recent approach is single nucleotide polymorphism (SNP) analysis using an array based technology (4–6) that allows identification of amplifications, deletions, and allelic imbalances, such as uniparental disomy (represents doubling of the abnormal allele due to somatic recombination or duplication, and loss of the other normal allele) (7, 8). However, SNP-chip analysis is only able to detect changes of gene dosage and is unable to identify balanced translocations, which commonly occur in ALL.

Previously, we analyzed 399 pediatric ALL cases by SNP-chip analysis and found a number of genomic abnormalities, in addition to well known common alterations (9). This technique is sensitive enough to identify genes involved in start sites of deletions/duplications. Indeed, this method allowed us to identify that the *PBX1* gene was involved in start sites of duplication of 1q23 generated by der(19)t(1;19)(q23;p13) (9). Furthermore, correlation analysis of the individual genomic abnormalities suggested the presence of der(12)t(12;21)(p13;q22) and der(21)t(12;21)(p13;q22), as well as dic(9;20)(p13;q11) (9).

In this study, we found that this new technology permitted us to identify genes involved in well known unbalanced translocations including *ETV6/RUNX1*. Further, we found previously undetected fusion genes between *PAX5* and a number of other partner genes by using this technique.

Results

Genes Involved in Unbalanced Translocations Were Identified by SNP-Chip Analysis. Because SNP-chip analysis can only detect changes of gene dosage including deletions, duplications, and amplifications (Fig. 1*A*), this technique is unable to identify balanced translocations (Figs. 1*Aii*). However, when one of a pair of reciprocally translocated chromosomes is lost, SNP-chip analysis can detect this abnormality as partial deletions of involved chromosomes (Fig. 1*Aiii*). Similarly, when one of a pair of reciprocally translocated chromosomes becomes duplicated, SNP-chip can also detect this abnormality as partial duplication of the involved chromosomes (Fig. 1*Aiv*). Furthermore, high resolution SNP-chip analysis allows us to identify the genes involved in these unbalanced translocations.

To prove that SNP-chip analysis can detect unbalanced translocations and the genes involved in these translocations, we

Author contributions: N.K., S.O., B.N., C.S., C.R.B., and H.P.K. designed research; N.K., S.O., B.N., C.S., M. Sanada, G.Y., and Y.N. performed research; N.K., S.O., and M. Sanada contributed new reagents/analytic tools; N.K., S.O., M.Z., M. Sanada, K.H., G.Y., Y.N., R.K., T.F., C.W.M., J.H., W.-D.L., M. Stanulla, M. Schrappe, C.R.B., and H.P.K. analyzed data; and N.K., S.O., M.Z., C.S., M. Stanulla, M. Schrappe, C.R.B., and H.P.K. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. J.W.G. is a guest editor invited by the Editorial Board.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. EU784145, PAX5-FOXP1; EU784146, PAX5-AUTS2; EU784147, PAX5-C20orf112 short isoform; and EU784148, PAX5-C20orf112 long isoform).

[b]N.K., S.O., and M.Z. contributed equally to this work.

[c]To whom correspondence should be addressed at: Hematology/Oncology, Cedars-Sinai Medical Institute/UCLA Geffen School of Medicine, 8700 Beverly Boulevard, Los Angeles, CA 90048. E-mail: kawamatan@cshs.org.

[l]C.R.B. and H.P.K. contributed equally to this work.

This article contains supporting information online at www.pnas.org/cgi/content/full/0711039105/DCSupplemental.

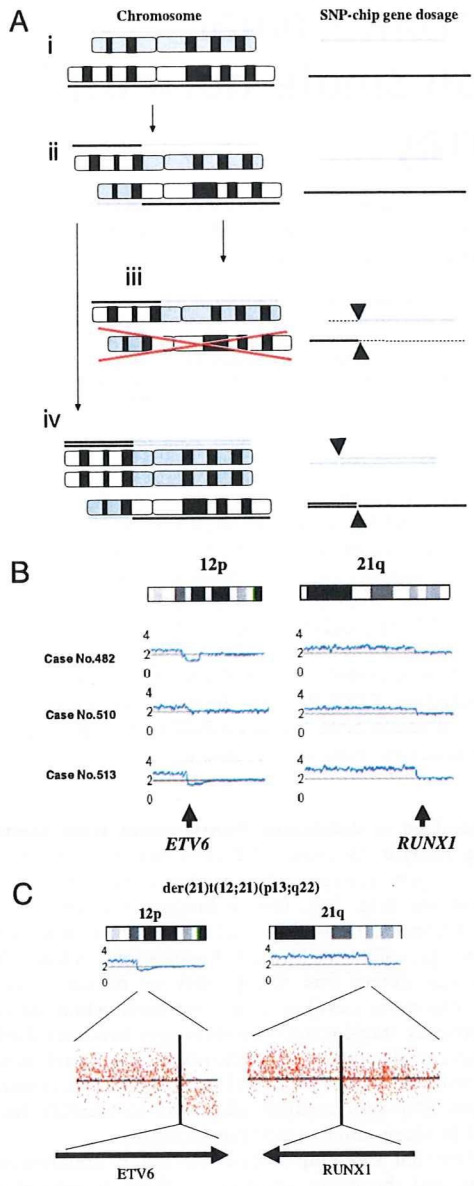

Fig. 1. SNP-chip analysis detected genes involved in unbalanced translocations. (*A*) SNP-chip analysis can identify breakpoints of translocations when one of the paired translocated chromosomes is either lost or duplicated/amplified. (*Left*) Chromosomal status. Gene dosages are indicated either above or beneath the chromosomes. (*Right*) Results of SNP-chip analysis. (*Ai*) Normal chromosomes; gene dosage is normal. (*Aii*) Reciprocal translocation; gene dosage is normal. (*Aiii*) One of the paired translocated chromosomes is lost; gene dosage is lower than normal on the left side of the upper chromosome and the right side of the lower chromosome. Arrow heads indicate the breakpoint of the translocation in each chromosome. (*Aiv*) One of the paired translocated chromosomes is duplicated; gene dosage is higher than normal on the right side of the upper chromosome and the left side of the lower chromosome. Arrow heads indicate the breakpoint of this translocation in each chromosome. (*B*) Representative cases with unbalanced translocation of der(21)t(12;21)(p13;q22). (*Left*) Start sites of duplication at 12p13 involving the *ETV6* gene. (*Right*) Start sites of duplication at 21q22 involving the *RUNX1* gene. SNP-chip data of representative cases with dup(12)(p13) and dup(21)(q22) are shown. These abnormalities were validated by FISH and/or RT-PCR (data not shown). Results of SNP-chip data were visualized by CNAG software. Lines above each chromosome show total gene dosage; level 2 indicates diploid (2N) amount of DNA, which is normal. (*C*) Magnified view of SNP-chip data. (*Upper*) Start sites of duplications at 12p13 and 21q22 are magnified. Signals of individual probe signals are shown. Vertical lines indicated the positions of start sites of duplications. (*Lower*) Genes involved in the start sites of duplications.

analyzed cases having extra copies of *ETV6/RUNX1* fusion genes generated by der(21)t(12;21)(p13;q22) (Fig. 1*B*), which were initially identified by FISH and/or RT-PCR (data not shown). SNP-chip was clearly able to identify this abnormality as duplications involving chromosome 12 and 21 (Fig. 1*B*). Further, the result of high-resolution (250k) SNP-chip clearly identified *ETV6* (12p13) and *RUNX1* (21q22) as the target genes involved in this unbalanced translocation (Fig. 1*C*).

PAX5 Gene Is Frequently Fused to Partner Genes. Our previous data showed the presence of dic(9;20)(p13;q11) in 11 cases of ALL (9), 5 of which had deletion 9p13.2-pter. These 5 cases had start sites of this deletion at 9p13.2 mapping to the *PAX5* gene (Fig. 2*A* and data not shown). This prompted us to reexamine all cases of B-ALL that had deletion of 9p [supporting information (SI) Table S1]. We found a total of 9 cases with similar start sites (9p13.2), mapping to the *PAX5* gene (Fig. 2*A* and data not shown). In 2 of these cases, simple abnormalities were detected by SNP-chip: case 514 had only del9p13.2-pter and del7q11.2-pter; case 458 had only del9p13.2-pter and dup3p13-pter (Table S1 and Fig. 2*A*). Three cases (536, 543, 572) had complex abnormalities including del9p13.2-pter and del20q11.21-qter, all with the *C20ORF112* gene within the start site of del20q (Table S1 and Fig. 2*A*). The other 2 cases (659, 767) had complex abnormalities that included *ETV6* on 12p13 (Table S1 and Fig. 2*A*).

Thus, we found four candidate partner genes fused to *PAX5* in seven cases by SNP-chip analysis; *ETV6* on 12p13 (two cases) (12), *C20orf112* on 20q11.1 (three cases), *AUTS2* on 7q11.1 (one case) and *FOXP1* on 3p13 (one case) (Fig. 2*A*). Because these translocations could lead to fusion transcripts between *PAX5* and different partner genes, the presence of the predicted fusion transcript was examined by RT-PCR using the mapping information from the SNP-chip data. RT-PCR and nucleotide sequencing data of the PCR products confirmed that the *PAX5* gene was fused to either the *ETV6* (two cases), *C20orf112* (three cases), *AUTS2* (one case), or *FOXP1* (one case) gene and transcribed into aberrant fusion messages (Fig. 2 *B* and *C*). Each fusion gene was mutually and exclusively detected in the samples studied. In one case with dic(9;20), exon 5 of *PAX5* was fused to exon 8 of *C20orf112*, and in two cases with dic(9;20), exon 8 of *PAX5* was fused to exon 3 of *C20orf112*. *PAX5/ETV6* involved exon 4 of *PAX5* and exon 3 of *ETV6*.

Cellular Localization and DNA Binding Affinity of PAX5 Fusion Products. In the *PAX5/FOXP1* fusion transcript, the amino acid coding frame of the *FOXP1* gene was not identical to that of *PAX5*, leading to a frame-shift and an early termination codon after the fusion point of these two genes (Fig. 2*D*). However, all other fusion genes were in frame and were predicted to encode chimeric proteins. Two proteins (a short and long form) with different breakpoints were predicted from the *PAX5/C20orf112* fusion genes (Fig. 2*D*).

To confirm cellular localization of PAX5-fusion proteins, we transfected vectors encoding wild-type *PAX5* and *PAX5* fusion genes (*PAX5-ETV6*, *PAX5-FOXP1*, *PAX5-C20ORF112S*, and *PAX5-C20ORF112L*) into 293T cells, fractionated the cytoplasmic and nuclear proteins, and examined the wild-type PAX5 and PAX5-fusion proteins by Western blot analysis (Fig. 2*E*). PAX5-ETV6 protein was detected in both the cytoplasm and nucleus; PAX5-FOXP1 and PAX5-C20ORF112L proteins were predom-

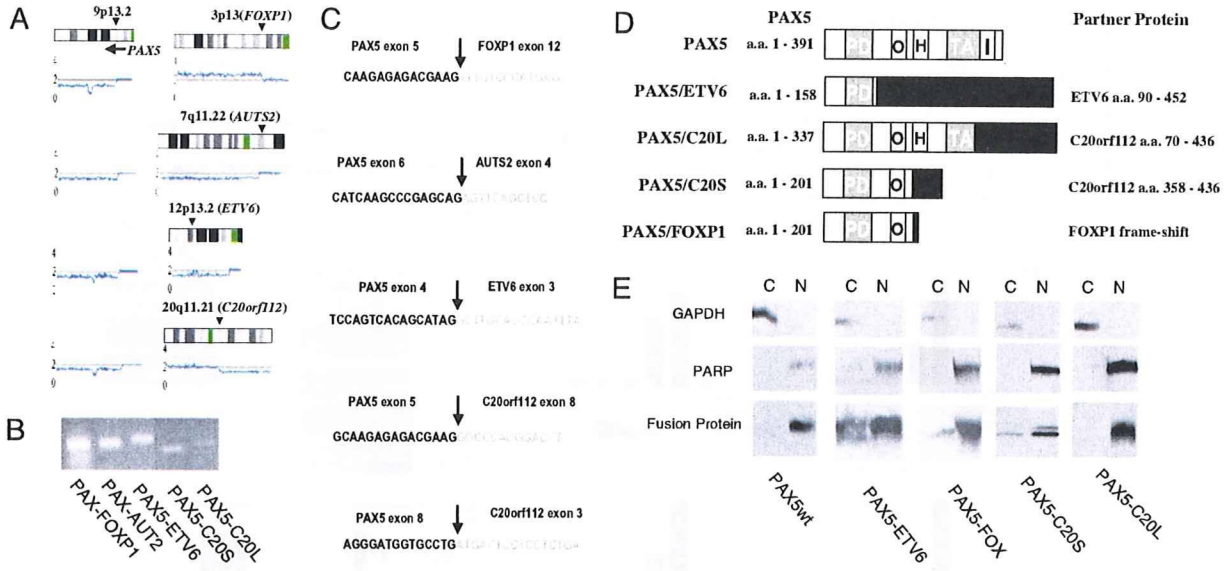

Fig. 2. *PAX5* gene is fused to partner genes. (*A*) Start sites of deletion at 9p13.2 involving the PAX5 gene. (*Left*) SNP-chip data of representative cases with 9p13.2 deletions. A vertical arrow indicates the start sites of 9p deletion that involves the *PAX5* gene. A horizontal arrow shows the direction of transcription of the *PAX5* gene. (*Right*) Chromosomal abnormalities of partner chromosomes. Arrow heads indicate the start sites of duplication or deletions. Genes involved in the start sites are shown. (*B*) Result of RT-PCR. The ALL samples suggesting the presence of PAX5 fusion genes by SNP-chip analysis were examined by RT-PCR using the primers of PAX5 and the respective partner genes. (*C*) Fusion sequences of the *PAX5* and partner genes. Joining sequences of fused transcripts are shown from the indicated exon of the fused gene. (*D*) Schematic structure of wild-type and mutant PAX5. Amino acid positions (aa) of each protein are indicated. PAX5/FOXP1 fusion construct has an early termination codon caused by a frame-shift. PD, paired domain; TA, transcription activation domain; O, octapeptide H, homeodomain-like; I, inhibitory domain. (*E*) Subcellular fractionation of PAX5-fusion proteins. pcDNA vector encoding wild-type PAX5, PAX5-ETV6, PAX5-FOXP1, PAX5-C20ORF112S, or PAX5-C20ORF112L was transfected into 293T cells. Nuclear and cytoplasmic proteins were separated and electrophoresed in the gel. Localization of PAX5-fusion proteins was examined by PAX5 N-terminal specific antibody. Purity of cytoplasmic protein was examined with anti-GAPDH antibody and purity of nuclear proteins with the anti-PARP antibody. C, cytoplasmic fraction; N, nuclear fraction.

inantly localized in the nucleus; and 20% and 80% of PAX5-C20ORF112S proteins were localized in the cytoplasm and the nucleus, respectively (Fig. 2*E*). Localization of the fusion proteins was also confirmed by immunohistochemical staining (data not shown).

Because PAX5-fusion proteins were localized in the nucleus, we analyzed DNA binding affinity of these PAX5-fusion proteins *in vitro*. DNA binding affinity of the PAX5 wild-type and fusion proteins expressed in 293T cells was analyzed by electrophoretic mobility shift assay (EMSA), and signals of probes bound to the proteins were plotted graphically (Fig. 3*A*). Binding activity of each protein in the absence of cold competitor oligonucleotide probe was regarded as 1.0, and the binding activity in the presence of cold competitor oligonucleotide probes was measured. All PAX5-fusion proteins showed similar binding activity to the PAX5 recognition sequences as the wild-type PAX5 (Fig. 3*A*).

***PAX5* Fusion Products Suppressed Transcriptional Activity of Wild-Type PAX5 in a Dominant Negative Fashion, Leading to Inhibition of B-Cell Development.** To examine the effect of PAX5-fusion proteins on transcriptional activity of wild-type PAX5, we performed a reporter gene assay using 293T cells. Cotransfection reporter gene assays using wild-type and fusion *PAX5* expression vectors along with a reporter gene driven by the murine CD19 promoter (which contains three repeats of PAX5 binding sequences) showed that the PAX5 fusion products suppressed transcriptional activity of PAX5 in a dominant-negative fashion (Fig. 3*B*). Expression of wild-type PAX5 proteins was minimally affected by coexpression of PAX5-fusion proteins (Fig. 3*C*), suggesting that PAX5-fusion proteins competed with wild-type PAX5 for the PAX5 binding sequences on the reporter gene.

Further, we transfected vectors encoding either PAX5-C20orf112S or PAX5-C20orf112L, each coexpressing the GFP marker, into Nalm 6 cells (a human B cell ALL cell line, which expresses endogenous PAX5) (data not shown). After transfection, GFP-positive cells were sorted by FACS and expression of PAX5-downstream genes was examined by semiquantitative RT-PCR (Fig. 3*D* and data not shown). We examined 10 downstream target genes (seven positively regulated direct target genes and three negatively regulated genes) of PAX5 (10–12) and found that four, including *ATP1B1*, *BLK*, *NEDD5* and *TCF7L2*, were down-regulated by induction of either PAX5-C20orf112S or PAX5-C20orf112L protein. However, expression of other reported PAX5 downstream target genes, including three positively regulated direct target genes (*IRF8*, *BST1*, *CD19*) and three negatively regulated genes (*CCR2*, *CCR5*, *NOTCH1*) were not affected by the induction of expression of the fusion proteins in these cells.

To examine the effect of PAX5 fusion protein on binding of wild-type PAX5 to the direct target gene *BLK* in the leukemic cells, we performed chromatin-immunoprecipitation (ChIP) assay using Nalm 6 cells transfected with either an empty vector or a construct encoding PAX5-C20orf112S. We used a PAX5 antibody detecting the C-terminal region of the protein, which could detect wild-type PAX5, but not PAX5-C20orf112S, as the C-terminal end of PAX5 was replaced by C20orf112S in this fusion protein. Although wild-type specific PAX5 antibody precipitated the promoter region of *BLK* after transfection of the empty vector, the amount of DNA of the *BLK* promoter region bound to wild-type PAX5 was reduced after transfection of the PAX5-C20orf112S gene (Fig. 3 *E* and *F*).

To examine the effect of PAX5-fusion proteins on B cell development in murine hematopoietic cells, we infected murine

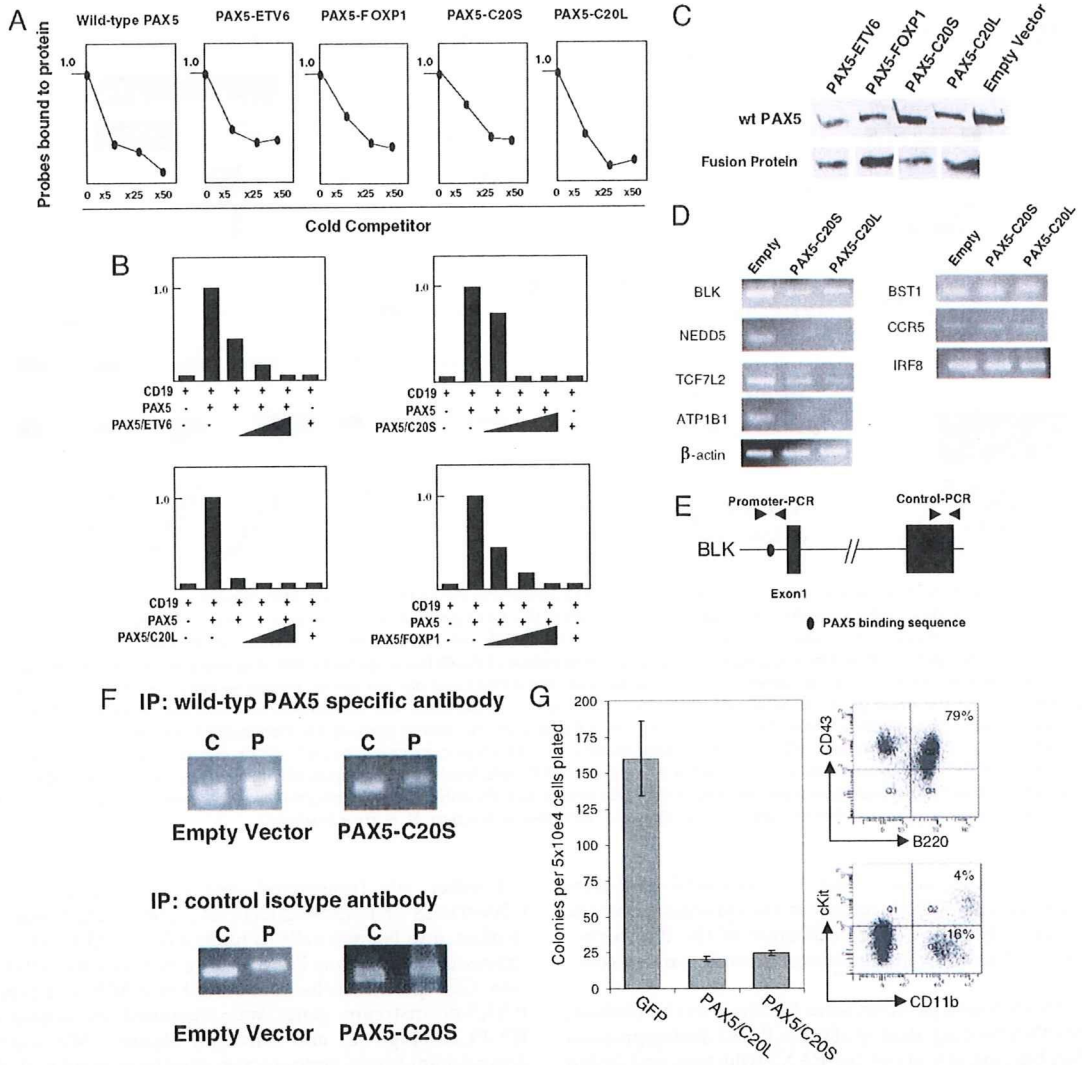

Fig. 3. *PAX5*-fusion proteins suppress transcriptional activity of PAX5 in a dominant-negative fashion and block the growth of B cells. (*A*) Result of EMSA: Wild-type PAX5 and PAX5 fusion were expressed in 293T cells, and nuclear proteins were purified. The purified nuclear proteins were mixed with radioisotope labeled double-strand oligonucleotide DNA, in either the presence or absence of cold competitor oligonucleotides (5-, 25-, and 50-fold cold competitor probes). Intensity of each shifted band indicating DNA–protein complex was measured and plotted graphically. Intensity of shifted bands in the absence of cold competitor probes was regarded as 1.0. (*B*) Reporter gene assay. Wild-type and mutant PAX5 were mixed at a various ratios (1:0, 1:0.3, 1:1, 1:3, respectively, 1 = 500 ng of construct) and transfected. Forty-eight hours later, relative activity of firefly luciferase was measured and plotted. Results represent the mean values of the three experiments. CD19, PAX5 luciferase reporter construct having PAX5 binding region of CD19 promoter; PAX5, wild-type PAX5; PAX5/ETV6, PAX5/ETV6 fusion; PAX5/C20L, long form of PAX5/C20orf112 fusion in which PAX5 exon 8 is fused to C20orf112 exon 3; PAX5/C20S, short form of PAX5/C20orf112 fusion in which PAX5 exon 5 is fused to C20orf112 exon 8; PAX5/FOXP1, PAX5/FOXP1 fusion with an early termination codon caused by a frame-shift after the site of fusion. (*C*) Results of expression of wild-type PAX5 and PAX5-fusion proteins. After cotransfection of equal amounts of vector encoding either wild-type or fusion PAX5 genes into 293T cells, the expression of respective proteins was examined by Western blot. Levels of expression of wild-type PAX5 protein were minimally affected by coexpression of the PAX5-fusion proteins. (*D*) Semiquantitative RT-PCR of downstream target genes of PAX5. Expression of PAX5 downstream target genes was examined by semiquantitative RT-PCR. Nalm 6, a human B cell ALL cell line expressing endogenous PAX5, was transfected with pMSCV-GFP (Empty), pMSCV-GFP-PAX5-C20orf112S (PAX5-C20S), or pMSCV-GFP-PAX5-C20orf112L (PAX5-C20L). GFP-positive cells were sorted and subject to semiquantitative RT-PCR. Optimal cycle numbers to semiquantify the expression of respective genes are as follows; BLK: 25 cycles; Nedd5; 25 cycles; TCF7L2: 25 cycles; ATP1B1: 25 cycles; β-actin: 22 cycles; CCR2: 25 cycles; CCR8: 30 cycles; IRF8: 30 cycles. (*E*) Structure of human *BLK* gene. Structure of *BLK* and primers used for ChIP assay within the 5′ regulatory region (Promoter-PCR) and 3′ end (Control-PCR) of the *BLK* gene is schematically shown. PAX5 binding site in the promoter region is indicated. (*F*) ChIP analysis of the PAX5 binding site in the *BLK* gene promoter. pMSCV-GFP (empty vector) or pMSCV-GFP-PAX5-C20S (PAX5-C20S) was transfected into human Nalm 6 B cell leukemia cells expressing endogenous PAX5. GFP-positive cells were subject to ChIP assay. The cells were fixed in formaldehyde solution and sonicated by ultrasound. DNA-protein complex was incubated with wild-type PAX5 specific antibody, which detected the C-terminal region of PAX5 but not the PAX5-C20orf112S protein (*Upper*). As a control, the DNA-protein complex was reacted with isotype nonspecific antibody (*Lower*). Immnoprecipitated DNA was subjected to PCR to amplify either the *BLK* promoter region containing PAX5 binding sequence (P) or, as an internal control, the 3′ end of the *BLK* gene (C). (*G*) Retrovirus infection experiments. Murine bone marrow cells were collected at 5 days after injection of 5FU. The hematopoietic cells were infected by retrovirus containing pMSCV-GFP empty vector (GFP), pMSCV-GFP-C20orf112L (PAX5/C20L), or pMSCV-GFP-C20orf112S (PAX5/C20S). GFP-positive murine hematopoietic cells were sorted and plated at 5×10^4 cells per plate in methylcellulose containing mSCF, mIL7, and hFL. At 8 days after the plating, the colony numbers were counted (*Left*; results represent means and SD of three experiments). Cell surface antigens on the GFP-positive cells infected with pMSCV-GFP (GFP) at Day 11 were examined by FACS using antibodies against CD43 and B220 (*Upper Right*), c-kit and CD11b (*Lower Right*) antibodies, to confirm the development of B cells.