A novel gene, NSD1, is fused to NUP98 in the t(5;11)(q35;p15.5) in de novo childhood acute myeloid leukemia

Rina J. Jaju, Carrie Fidler, Oskar A. Haas, Amanda J. Strickson, Fiona Watkins, Kevin Clark, Nicholas C. P. Cross, Jan-Fang Cheng, Peter D. Aplan, Lyndal Kearney, Jacqueline Boultonwood, and James S. Wainscoat

The recurrent translocation t(5;11)(q35;p15.5) associated with a 5q deletion, del(5q), has been reported in childhood acute myeloid leukemia (AML). We report the cloning of the translocation breakpoints in de novo childhood AML harboring a cryptic t(5;11)(q35;p15.5). Fluorescence in situ hybridization (FISH) analysis demonstrated that the nucleoporin gene (NUP98) at 11p15.5 was disrupted by this translocation. By using 3′-rapid amplification of complementary DNA ends (3′-RACE) polymerase chain reaction, we identified a chimeric messenger RNA that results in the in-frame fusion of NUP98 to a novel gene, NSD1. The NSD1 gene has 2596 amino acid residues and a 85% homology to the murine Nsd1 with the domain structure being conserved. The NSD1 gene was localized to 5q35 by FISH and is widely expressed. The reciprocal transcript, NSD1-NUP98, was also detected by reverse transcriptase–polymerase chain reaction. This is the first report in which the novel gene NSD1 has been implicated in human malignancy.

© 2001 by The American Society of Hematology

Introduction

Recurring chromosome translocations are common in a wide spectrum of hematologic malignancies.1,3 We have reported the identification of a recurrent cryptic translocation t(5;11)(q35;p15.5) associated with a deletion of the long arm of chromosome 5, del(5q), in de novo childhood acute myeloid leukemia (AML).4 Using fluorescence in situ hybridization (FISH), we localized the breakpoint on 11p15.5 between Harvey ras-1 related gene complex (HRC) and radixin pseudogene (RDPX1). The nucleoporin 98 gene (NUP98) lies inside this region, and its rearrangement has been documented in both de novo and therapy-related myelodysplastic syndrome, AML, and acute lymphoblastic leukemia in children and adults.5–13 We speculated that NUP98 was a good candidate gene for the involvement in the t(5;11) AML.

We now confirm that the chromosome 11 breakpoint gene is NUP98, and we report the cloning of its novel fusion partner, nuclear receptor-binding Su(var), Enhancer of zeste [E(z)], and Trithorax (Trx) (SET) domain protein 1 (NSD1), as the chromosome 5 breakpoint gene in the recurrent t(5;11)(q35;p15.5) in childhood AML.

Study design

Donor samples

We analyzed samples from 2 patients with childhood AML-M2 with t(5;11)(q35;p15.5) and a del(5q). The clinical and cytogenetic data of these 2 cases have been previously reported as patients nos. 1 and 3.4 Detailed molecular analysis was performed on the leukemic sample of patient no. 3 because RNA was available only from this patient.

FISH analysis

For FISH studies PAC 1173K1 (containing exons 10-20 of the NUP98 gene), plasmids p9R1 (containing exons 10-12 of the NUP98 gene),14 and p6G2 (containing exons 13-14 of the NUP98 gene)14 from 11p15.5 were used. A 5p probe, cos113-1,15 was used to identify both chromosomes 5. A 1.2-kilobase (kb) complementary DNA (cDNA) clone of the human NSD1 (IMAGE clone 2621998) was also used as a probe for FISH studies. FISH analysis was performed on bone marrow metaphases as previously described.5

Nucleic acid isolation

Total RNA was extracted from cryopreserved leukemic cell suspensions using the Totally RNA kit (Ambion, Austin, TX). Plasmid DNA was extracted using QIagen reagents and protocols (Qiagen miniprep kit, QIagen, Crawley, United Kingdom).

3′-RACE

The 3′-RACE (3′-rapid amplification of cDNA ends) was performed using the SMART-RACE cDNA amplification kit and protocol (Clontech, Palo Alto, CA). Briefly, first-strand cDNA was reverse transcribed from 1 μg total RNA using Superscript II and the 3′-RACE cDNA synthesis primer (3′-CDS) from the kit. An aliquot of the first-strand cDNA was then amplified using a NUP98 gene–specific forward primer (NUP98-1, Table 1) and a universal primer mix (SMART-RACE kit). Polymerase chain reaction (PCR) conditions were as described by the manufacturer. A nested PCR reaction using the nested universal primer (SMART-RACE kit) as the reverse primer and NUP98-2 (internal to NUP98-1, Table 1) as the forward primer was used.
NUP98 gene in-frame to the 3\(^{-}\) of the partner gene. Previous studies have shown the principal NUP98, (p6G2 and p9R1) confirmed that the breakpoint was within \(t(5;11)(q35;p15.5)\) AML cases (patients nos. 1 and 3) (Figure 1A).

### Results and discussion

In the current study, we identified NUP98 as the chromosome 11 breakpoint gene by FISH using a PAC clone 1173K1 in 2 \(t(5;11)(q35;p15.5)\) AML cases (patients nos. 1 and 3) (Figure 1A). Further, FISH analyses using plasmids specific for exons of NUP98 (p6G2 and p9R1) confirmed that the breakpoint was within NUP98, in the intron between exons 12 and 13 of this gene in patients nos. 1 and 3. Previous studies have shown the principal transcript in other NUP98 translocations fuses the 5\(^{\prime}\) end of the NUP98 gene in-frame to the 3\(^{-}\) end of the partner gene. The 3\(^{-}\)-RACE was performed on RNA isolated from patient no. 3 to identify the fusion partner of NUP98.

RACE-PCR products were purified and subcloned into plasmid vectors. A total of 88 recombinant clones were screened by hybridization using standard protocols. An oligonucleotide probe of 116 base pairs (bp) (NUP-ex12) generated using primers NUP98-3 and NUP98-4 was used for screening (Table 1). All positive plasmid clones were selected for sequencing.

### Reverse transcriptase–PCR

Reverse transcriptase (RT)-PCR was performed using the RT one-step RT-PCR kit (ABgene, Surrey, United Kingdom) with sense NUP98-5 and antisense NSD1-1 primers (Table 1). The PCR thermal cycling protocol was performed according to the manufacturer’s instructions with an annealing temperature of 58°C. Similarly, RT-PCR was performed using sense NSD1-2 and antisense NUP98-6 primers and the same PCR conditions as above (Table 1). The RT-PCR products of both reactions were subcloned and sequenced.

### Sequence analysis

Plasmid clones were sequenced using the Autoread Cy-5 sequencing kit and ALF express automated sequencer (Amer sham-Pharmacia Biotech, Amersham, United Kingdom). Sequence analysis was performed using the Wisconsin Package Version 10.1 (Genetics Computer Group, Madison, WI).

### Northern blot analysis

A cDNA clone (Image clone 2621298) of the human NSD1 gene was used to probe human multiple tissue Northern blots according to the manufacturer’s instructions (Clontech).

<table>
<thead>
<tr>
<th>Primer</th>
<th>Oligonucleotide sequence (5(^{\prime}) → 3(^{\prime}))</th>
<th>Gene</th>
<th>Nucleotides</th>
</tr>
</thead>
<tbody>
<tr>
<td>NUP98-1</td>
<td>GGGACTCTTGGAACTGGGCTTGGTGCA</td>
<td>NUP98</td>
<td>1372-1399</td>
</tr>
<tr>
<td>NUP98-2</td>
<td>GGGAGGCTTGGTGCTAAGAGAGTGG</td>
<td>NUP98</td>
<td>1462-1490</td>
</tr>
<tr>
<td>NUP98-3</td>
<td>GTCTGAGACAGCTTCTTG</td>
<td>NUP98</td>
<td>1418-1437</td>
</tr>
<tr>
<td>NUP98-4</td>
<td>TCCAAAGGGAAATGCTGTTG</td>
<td>NUP98</td>
<td>1533-1514</td>
</tr>
<tr>
<td>NUP98-5</td>
<td>TCTTGTACAGAACCTTTG</td>
<td>NUP98</td>
<td>1470-1489</td>
</tr>
<tr>
<td>NSD1-1</td>
<td>TCCAAAGACCTTCTGGTG</td>
<td>NSD1</td>
<td>3557-3537</td>
</tr>
<tr>
<td>NSD1-2</td>
<td>GGACAGTTAAAATTTG</td>
<td>NSD1</td>
<td>3400-3419</td>
</tr>
<tr>
<td>NUP98-6</td>
<td>GACATCGGATTCCGGAAGAG</td>
<td>NUP98</td>
<td>1664-1644</td>
</tr>
</tbody>
</table>

Nucleotide coordinates refer to GenBank accession numbers U41815 (NUP98) and AF332907 (NSD1).
RT-PCR on material from patient no. 3 was performed using primers flanking the NUP98-NSD1 junction and gave a product of expected size (135 bp) (Figure 2C). Sequence analysis of RT-PCR products confirmed that NUP98 and NSD1 mRNA were fused in-frame at the same NUP98 exon as reported in the literature.5-8,10,11,13 A reciprocal fusion mRNA transcript NSD1-NUP98 was also detected by RT-PCR analysis with an expected product size of 200 bp (Figure 2D). Sequence analysis showed that the mRNA was fused in-frame joining nucleotides 3503 of NSD1 to nucleotide 1553 of NUP98 (Figure 2B). Expression of both NUP98-NSD1 and NSD1-NUP98 transcripts suggests that both may have a biological role.

The NUP98 gene encodes a 98-kd component of the nuclear pore complex and localizes to the nucleoplasmic side of the nuclear pore complex.18 It is thought to function as a docking protein through the N-terminal domain of the protein, which contains the conserved multiple phenylalanine-glycine (FG) repeats.5-13 The FG repeats were also retained in the NUP98-NSD1 fusion transcript in the patient reported here.

The fusion partner of NUP98 in the t(5;11) AML is NSD1, which is reported here for the first time and is the human homolog of murine Nsd1. Murine Nsd1 contains 2 distinct nuclear receptor interaction domains and also contains distinct activation and repression domains, and it has been suggested that it may define a novel class of bifunctional transcriptional intermediary factors.17 By comparison to the murine Nsd1 protein, the human NSD1 protein has the domain structure conserved and these domains are found in proteins involved in the epigenetic control of transcription; it therefore has been suggested that Nsd1 is involved in some aspects of transcriptional control at the chromatin level.17 It is thought that the PHD domain is a protein-protein interaction domain and that SET/SAC domains at least in one protein serve as
a histone methyltransferase. The schematic representations of NUP98-NSD1 and NSD1-NUP98 fusion proteins are shown in Figure 2E.

In summary, we have cloned the recurrent t(5;11) chromosomal translocation, which results in the previously undescribed in-frame fusion between NUP98 and a novel gene NSD1 in de novo childhood AML. Future studies based on the structure-function relationship of the NSD1 gene and functional analysis of the 2 predicted chimeric proteins will provide an insight into the mechanism of leukemogenesis.

Acknowledgment

The authors thank Robin Roberts-Gant for assistance with the figure preparations.

References

A novel gene, *NSD1*, is fused to *NUP98* in the t(5;11)(q35;p15.5) in de novo childhood acute myeloid leukemia

Rina J. Jaju, Carrie Fidler, Oskar A. Haas, Amanda J. Strickson, Fiona Watkins, Kevin Clark, Nicholas C. P. Cross, Jan-Fang Cheng, Peter D. Aplan, Lyndal Kearney, Jacqueline Boulwood and James S. Wainscoat