CBFA2T3-GLIS2 fusion transcript is a novel common feature in pediatric, cytogenetically normal AML, not restricted to FAB M7 subtype

Riccardo Masetti, Martina Pigazzi, Marco Togni, Annalisa Astolfi, Valentina Indio, Elena Manara, Rita Casadio, Andrea Pession, Giuseppe Basso and Franco Locatelli
**MYELOID NEOPLASIA**

**CBFA2T3-GLIS2** fusion transcript is a novel common feature in pediatric, cytogenetically normal AML, not restricted to FAB M7 subtype

Riccardo Masetti,1 Martina Pigazzi,2 Marco Togni,1 Annalisa Astolfi,3 Valentina Indio,5 Elena Manara,2 Rita Casadio,4 Andrea Pession,1 Giuseppe Basso,2 and Franco Locatelli5

1Department of Pediatrics, Lalla Seràgni, Hematology-Oncology Unit, University of Bologna, Italy; 2Department of Woman and Child Health, Laboratory of Hematology-Oncology, University of Padova, Padova, Italy; 3Giorgio Prodi Cancer Research Center, University of Bologna, 4Biocomputing Group, Department of Biological, Geological and Environmental Sciences, University of Bologna, Bologna, Italy; and 5Department of Pediatric Hematology-Oncology, Istituto di Ricovero e Cura a Carattere Scientifico Ospedale Bambino Gesù, Rome, University of Pavia, Italy

---

**Key Points**

- The **CBFA2T3-GLIS2** fusion transcript is common in pediatric cytogenetically normal AML and not restricted to FAB M7 subtype.
- The **CBFA2T3-GLIS2** fusion transcript is associated with poor prognosis in pediatric patients with AML.

---

**Introduction**

Pediatric acute myeloid leukemia (AML) is a molecularly heterogeneous disease that arises from genetic alterations of pathways that regulate self-renewal and myeloid differentiation. While the majority of patients carry recurrent chromosomal translocations, almost 20% of childhood AMLs do not show any recognizable cytogenetic alteration and are defined as cytogenetically normal AML (CN-AML). Many genetic abnormalities have been identified in AML with normal karyotype, with the most frequent affecting genes such as NPM1, FLT3, CEBPA, and WT1.

Genome-wide analyses have been used with the aim of determining the full array of genetic lesions of CN-AML. Recent studies have provided new insight into the molecular genetics and biology of AML, confirming both the complexity and the heterogeneity of this disease. Novel lesions such as mutations in IDH1 and DNMT3A have been identified. However, these alterations are rare in pediatric AML, with IDH1/IDH2 accounting for 2% to 4% of cases and DNMT3A not even being found mutated in childhood AML. Recently, 2 studies identified a novel recurrent translocation involving CBFA2T3 and GLIS2 in about 30% of children with non-Down syndrome acute megakaryoblastic leukemia (non-DS AMKL, AML French–American–British [FAB] M7). Nevertheless, there are many children with CN-AML in whom no genetic abnormality has been detected. The identification of the different genetic profiles characterizing this subgroup is a primary objective to be pursued. To this end, we performed whole-transcriptome massively parallel sequencing of 7 cases of pediatric CN-AML with the aim of identifying recurrent somatic mutations or genomic rearrangements. Subsequently, we validated our findings in a larger cohort of 230 pediatric CN-AML patients.

---

**Study design**

**Patient samples**

Patients analyzed either in the parallel sequencing screening or in the validation cohort are children with newly diagnosed de novo AML other than promyelocytic leukemia. The patients are enrolled in the Associazione Italiana Ematologia Oncologia Pediatrica 2002/01 Protocol, which was approved by the institutional review board.

---


R.M. and M.P. contributed equally to this study. G.B. and F.L. are co-senior authors.

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

© 2013 by The American Society of Hematology
board of the Sant’ Orsola-Malpighi Hospital, Bologna, Italy. Patients gave informed consent in accordance with the Declaration of Helsinki. FAB morphological classification and immunophenotypic analysis were centrally established at the laboratory of Pediatric Hematology of the University Hospital in Padova, Italy. Chromosome analysis was performed on bone marrow (BM) aspirates using standard laboratory procedures. Karyotypes were reported according to the International System for Human Cytogenetic Nomenclature. For fluorescence in situ hybridization, an MLL locus-specific dual-color probe for 11q23 (Abbott-Vysis, Downers Grove, IL) was used. All CN patients were also negative for known recurrent genetic abnormalities involving MLL, CBFB, NPM1, and FLT3 genes.

Whole-transcriptome sequencing and bioinformatics analyses

Total RNA was extracted from BM leukemia cells of CN-AML patients by TRIzol, following the manufacturer’s protocol (Invitrogen, Karlsruhe, Germany); 250 to 1000 ng of total RNA was used for the synthesis of cDNA libraries with TruSeq RNA Sample Prep Kit v2 (Illumina, San Diego, CA) according to the manufacturer’s recommendations. Sequencing by synthesis was performed on HiScanSQ sequencer (Illumina) at 75bp in paired-end mode. Reads were aligned with TopHat2/Bowtie2 to the reference human genome hg19/GRCh37. deFuse17 and ChimeraScan18 packages were used to detect sequencing by synthesis was performed on HiScanSQ (Illumina, San Diego, CA) according to the manufacturer (Karlsruhe, Germany); 250 to 1000 ng of total RNA was used for the

Whole-transcriptome sequencing and bioinformatics analyses

Whole-transcriptome massively parallel sequencing in the 7 children with CN-AML yielded an average of 88.3 million mapped reads per patient, thus reaching an average coverage of 36×. Single nucleotide variant calling confirmed the absence of mutations in genes such as NPM1, CEBPA, FLT3, MLL, IDH1, IDH2, and DNMT3A.

Table 1. Clinical features of the CN-AML patients harboring the CBFA2T3-GLIS2 fusion gene

<table>
<thead>
<tr>
<th>ID</th>
<th>Age, years</th>
<th>Gender</th>
<th>WBC, × 10⁹/L</th>
<th>FAB</th>
<th>BM blast, percentage at diagnosis</th>
<th>Extramacular involvement</th>
<th>HSCT (type)</th>
<th>CR after induction therapy</th>
<th>Relapse (site)</th>
<th>Disease-free duration (months)</th>
<th>Survival duration (months)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1*</td>
<td>0.8</td>
<td>M</td>
<td>13.63</td>
<td>M2</td>
<td>52</td>
<td>No</td>
<td>Yes (MFD)</td>
<td>Yes</td>
<td>Yes (BM)</td>
<td>46.7</td>
<td>+48.0</td>
</tr>
<tr>
<td>2*</td>
<td>0.3</td>
<td>F</td>
<td>7.52</td>
<td>M5A</td>
<td>88</td>
<td>No</td>
<td>Yes (AUTO)</td>
<td>Yes</td>
<td>Yes (BM)</td>
<td>90.6</td>
<td>+91.4</td>
</tr>
<tr>
<td>3</td>
<td>1.4</td>
<td>F</td>
<td>65.08</td>
<td>M0</td>
<td>98</td>
<td>No</td>
<td>Yes (AUTO)</td>
<td>Yes</td>
<td>Yes (BM)</td>
<td>9.3</td>
<td>13.0</td>
</tr>
<tr>
<td>4</td>
<td>12.7</td>
<td>F</td>
<td>22.8</td>
<td>M1</td>
<td>61</td>
<td>No</td>
<td>Yes (AUTO)</td>
<td>Yes</td>
<td>Yes (BM)</td>
<td>70.1</td>
<td>+74.5</td>
</tr>
<tr>
<td>5</td>
<td>17.0</td>
<td>F</td>
<td>0.91</td>
<td>M1</td>
<td>98</td>
<td>No</td>
<td>Yes (AUTO)</td>
<td>Yes</td>
<td>Yes (BM + skin)</td>
<td>38.7</td>
<td>50.4</td>
</tr>
<tr>
<td>6</td>
<td>13.3</td>
<td>M</td>
<td>7.52</td>
<td>M5</td>
<td>90</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>2.6</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.9</td>
<td>M</td>
<td>7.53</td>
<td>M5</td>
<td>60</td>
<td>No</td>
<td>Yes (MUD)</td>
<td>Yes</td>
<td>Yes (BM)</td>
<td>15.1</td>
<td>19.3</td>
</tr>
<tr>
<td>8</td>
<td>12.1</td>
<td>M</td>
<td>20.82</td>
<td>M0</td>
<td>70</td>
<td>No</td>
<td>Yes (AUTO)</td>
<td>Yes</td>
<td>Yes (BM)</td>
<td>22.7</td>
<td>24</td>
</tr>
<tr>
<td>9</td>
<td>17.2</td>
<td>M</td>
<td>35.7</td>
<td>M0</td>
<td>85</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>3.2</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>16.4</td>
<td>F</td>
<td>26.4</td>
<td>M4</td>
<td>70</td>
<td>No</td>
<td>Yes (AUTO)</td>
<td>Yes</td>
<td>Yes (BM)</td>
<td>8.2</td>
<td>12.1</td>
</tr>
<tr>
<td>11</td>
<td>3.2</td>
<td>F</td>
<td>24.2</td>
<td>M7</td>
<td>70</td>
<td>No</td>
<td>Yes (MUD)</td>
<td>Yes</td>
<td>Yes (BM)</td>
<td>12.2</td>
<td>24.7</td>
</tr>
<tr>
<td>12</td>
<td>4.0</td>
<td>M</td>
<td>12.79</td>
<td>M7</td>
<td>74</td>
<td>No</td>
<td>Yes (MUD)</td>
<td>Yes</td>
<td>Yes (BM)</td>
<td>8.6</td>
<td>10.7</td>
</tr>
<tr>
<td>13</td>
<td>0.7</td>
<td>M</td>
<td>13.6</td>
<td>M7</td>
<td>95</td>
<td>Yes (CNS)</td>
<td>Yes (MUD)</td>
<td>Yes</td>
<td>No</td>
<td>9.4</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>0.7</td>
<td>F</td>
<td>9.6</td>
<td>M7</td>
<td>90</td>
<td>No</td>
<td>Yes (MUD)</td>
<td>No</td>
<td>Yes (BM)</td>
<td>109.2</td>
<td>+111.8</td>
</tr>
<tr>
<td>15</td>
<td>2.0</td>
<td>F</td>
<td>7.4</td>
<td>M7</td>
<td>95</td>
<td>No</td>
<td>Yes (MDF)</td>
<td>Yes</td>
<td>Yes (BM)</td>
<td>2.1</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>1.4</td>
<td>F</td>
<td>128.6</td>
<td>M7</td>
<td>50</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>3.2</td>
<td>5.7</td>
</tr>
<tr>
<td>17*</td>
<td>0.5</td>
<td>F</td>
<td>13.25</td>
<td>M7</td>
<td>90</td>
<td>No</td>
<td>Yes (MUD)</td>
<td>Yes</td>
<td>Yes (BM + CNS)</td>
<td>3.2</td>
<td>5.7</td>
</tr>
<tr>
<td>18</td>
<td>3.0</td>
<td>F</td>
<td>13.63</td>
<td>M7</td>
<td>40</td>
<td>No</td>
<td>Yes (MUD)</td>
<td>Yes</td>
<td>Yes (BM)</td>
<td>37.2</td>
<td>+45.2</td>
</tr>
<tr>
<td>19</td>
<td>1.6</td>
<td>F</td>
<td>115</td>
<td>M7</td>
<td>90</td>
<td>No</td>
<td>Yes (MDF)</td>
<td>Yes</td>
<td>Yes (BM)</td>
<td>17.7</td>
<td>24.7</td>
</tr>
<tr>
<td>20</td>
<td>1.9</td>
<td>F</td>
<td>56.5</td>
<td>M7</td>
<td>30</td>
<td>No</td>
<td>Yes (MDF)</td>
<td>Yes</td>
<td>Yes (BM)</td>
<td>13.5</td>
<td>+15.6</td>
</tr>
</tbody>
</table>

+= patients alive and in CR; AUTO, autologous; CNS, central nervous system; CR, complete remission; HSCT, hematopoietic stem cell transplantation; MFD, matched family donor; MUD, matched unrelated donor; WBC, white blood cells.

*Patients identified in the RNA-seq screening.

Screening of CBFA2T3-GLIS2 fusion transcript in the validation patient cohort

Table 1) we identified new putative fusion transcripts. In 3 of 7 patients (1, 2, 17; Table 1) we identified a chimeric transcript involving CBFA2T3 and GLIS2, resulting from a cryptic inversion of the telomeric region of chromosome 16 that fuses the 5′ region of CBFA2T3 with 3′ region of GLIS2. CBFA2T3, also known as MTG16 or ETO2, is a member of the myeloid translocation gene family that is fused with GLIS2, while the third showed a chimeric transcript involving CBFA2T3 and MTG16 in the t(16;21) (q24;q22) translocation that is identified in therapy-related AML. CBFA2T3, GLIS2, and IDH2, and DNMT3A.

Two algorithms, ChimeraScan and deFuse, were used to identify new putative fusion transcripts. In 3 of 7 patients (1, 2, 17; Table 1) we identified a chimeric transcript involving CBFA2T3 and GLIS2, resulting from a cryptic inversion of the telomeric region of chromosome 16 that fuses the 5′ region of CBFA2T3 in frame with the 3′ region of GLIS2. CBFA2T3, also known as MTG16 or ETO2, is a member of the myeloid translocation gene family that is fused to AML1 in the t(16;21) (q24;q22) translocation that is identified in therapy-related AML. CBFA2T3, GLIS2, and IDH2, and DNMT3A.

The following primers were used: forward 5′-CGAAGGGCCTCAGC TACAGT-3′, reverse 5′-AGCCACTGCCTATTTGGAT-3′.

Results and discussion

Identification of CBFA2T3-GLIS2 fusion transcript in children with CN-AML by whole-transcriptome sequencing

Whole-transcriptome massively parallel sequencing in the 7 children with CN-AML yielded an average of 88.3 million mapped reads per patient, thus reaching an average coverage of 36×. Single nucleotide variant calling confirmed the absence of mutations in genes such as NPM1, CEBPA, FLT3, MLL, IDH1, IDH2, and DNMT3A.

Two algorithms, ChimeraScan and deFuse, were used to identify new putative fusion transcripts. In 3 of 7 patients (1, 2, 17; Table 1) we identified a chimeric transcript involving CBFA2T3 and GLIS2, resulting from a cryptic inversion of the telomeric region of chromosome 16 that fuses the 5′ region of CBFA2T3, while the third showed CBFA2T3, exon 10 fused to exon 2 of GLIS2 (Figure 1A). These data established a chimeric transcript involving CBFA2T3 and GLIS2, resulting from a cryptic inversion of the telomeric region of chromosome 16 that fuses the 5′ region of CBFA2T3.

Screening of CBFA2T3-GLIS2 fusion transcript in the validation patient cohort

Total RNA was extracted from BM leukemia cells of all samples using TRIzol. CBFA2T3-GLIS2 fusion transcript was detected with reverse-transcription polymerase chain reaction (RT-PCR) and sequenced with the BigDye terminator v3.1 Cycle Sequencing kit (PE Applied Biosystems, Foster City, CA) on an Applied Biosystems 310 analyzer. The following primers were used: forward 5′-CGAAGGGCCTCAGC TACAGT-3′, reverse 5′-AGCCACTGCCTATTTGGAT-3′.
Probability of 5-year EFS in pediatric CN-AML with or without CBFA2T3-GLIS2 fusion, stratified according to FAB subgroups (M7 vs non-M7): EFS of non-M7 CN-AML with CBFA2T3-GLIS2 fusion transcript is a novel common feature of pediatric CN-AML, predicting poorer outcome.

**CBFA2T3-GLIS2 fusion transcript is recurrent in pediatric CN-AML**

To assess the prevalence of CBFA2T3-GLIS2 fusion, we then examined a validation cohort of 230 children with newly diagnosed de novo CN-AML, also negative for known recurrent genetic abnormalities involving MLL, CBFB, NPM1, and FLT3 genes. Globally, the CBFA2T3-GLIS2 rearrangement was detected in 20 of 237 cases (8.4%) with CN-AML. RT-PCR analysis and Sanger sequencing confirmed that all positive cases in the validation cohort carried the CBFA2T3 exon 11-GLIS2 exon 3 fusion. Fifty percent (N = 10) of the positive patients belonged to the M7 FAB subgroup, while the remaining patients (N = 10) were distributed among the other FAB classes (see Table 1). These results indicate that the CBFA2T3-GLIS2 fusion transcript, recently described as a distinctive feature of pediatric non-DS AMKL,12,13 should be more broadly considered as a genetic abnormality that is shared with other FAB subgroups of pediatric CN-AML.

**CBFA2T3-GLIS2 fusion transcript identifies a subset of childhood CN-AML with poor outcome**

We evaluated whether the presence of CBFA2T3-GLIS2 fusion product influences patients’ outcome. The 5-year event-free survival (EFS) of the 20 patients with the CBFA2T3-GLIS2 fusion gene was significantly worse than that of the 217 pediatric CN-AML patients not harboring the translocation (27.4%, standard error [SE] 10.5 vs 59.6%, SE 3.6; P = .01; Figure 1C). We also stratified the patients with respect to FAB subgroups (M7 vs non-M7; Figure 1D). The 5-year EFS of FAB M7 patients with or without the CBFA2T3-GLIS2 fusion gene was significantly different: 26.6% (SE 15.0) and 60.7% (SE 8.3; P = .04), respectively. Similar results were obtained in non-M7 patients; the 5-year EFS of patients assigned to other FAB categories with or without the CBFA2T3-GLIS2 fusion transcript was 30.0% (SE 14.4) and 59.4% (SE 3.5; P = .04), respectively. No statistically significant difference in EFS of non-M7 and M7 patients harboring the CBFA2T3-GLIS2 fusion transcript (30.0%, SE 14.4 vs 26.6%, SE 15.0; P = .91) was found, suggesting that FAB classification does not interact with the CBFA2T3-GLIS2 fusion product in influencing outcome. Taken together, these data indicate that the CBFA2T3-GLIS2 fusion transcript is a novel common feature in pediatric CN-AML that is not restricted to the FAB M7 subtype, predicting poor outcome.

**Acknowledgments**

This work was supported in part by grants from Fondazione Umberto Veronesi, Milan (R.M.); Fondazione Istituto di Ricerca Pediatrica-Città della Speranza Padova (M.P., E.M., G.B.); and Associazione Italiana Ricerca sul Cancro Special Project 5 × mille (F.L.).

**Authorship**

Contribution: R.M. and M.P. equally coordinated the work, analyzed data, performed statistical analyses, and wrote the paper; M.T. and A.A. performed the whole-transcriptome massively
parallel sequencing; V.I. and R.C. performed bioinformatics analyses; E.M. performed mutation analyses on the validation cohort; and A.P., G.B., and F.L. designed and supervised the research. A.P. and F.L. equally contributed to the critical revision and writing of the manuscript. All authors read and approved the final version of the manuscript.

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

Correspondence: Andrea Pession, Paediatric Oncology and Haematology Unit, Lalla Serágnoli, Sant’Orsola-Malpighi Hospital, Via Massarenti 11, 40137 Bologna, Italy; e-mail: andrea.pession@unibo.it.

References