Comprehensive Transcriptome Profiling of Cryptic CBFA2T3–GLIS2 Fusion–Positive AML Defines Novel Therapeutic Options: A COG and TARGET Pediatric AML Study

Jenny L. Smith1, Rhonda E. Ries1, Tiffany Hylkema1, Todd A. Alonzo2,3,4, Robert B. Gerbing2, Marianne T. Santaguida5, Lisa Eidsenschink Brodersen6, Laura Pardo1,6, Carrie L. Cummings1, Keith R. Loeb1, Quy Le1, Suzan Imren1, Amanda R. Leonti1, Alan S. Gamis7, Richard Aplenc8, E. Anders Kolb5, Jason E. Farrar10, Timothy J. Triche Jr11, Cu Nguyen12, Daoud Mezrzan12, Michael R. Loken6, Vivian G. Oehler1, Hamid Bolouri13, and Soheil Meshinchi1,2,14

ABSTRACT

Purpose: A cryptic inv(16)(p13.3q24.3) encoding the CBFA2T3–GLIS2 fusion is associated with poor outcome in infants with acute megakaryocytic leukemia. We aimed to broaden our understanding of the pathogenesis of this fusion through transcriptome profiling.

Experimental Design: Available RNA from children and young adults with de novo acute myeloid leukemia (AML; N = 1,049) underwent transcriptome sequencing (mRNA and miRNA). Transcriptome profiles for those with the CBFA2T3–GLIS2 fusion (N = 24) and without (N = 1,025) were contrasted to define fusion-specific miRNAs, genes, and pathways. Clinical annotations defined distinct fusion-associated disease characteristics and outcomes.

Results: The CBFA2T3–GLIS2 fusion was restricted to infants <3 years old (P < 0.001), and the presence of this fusion was highly associated with adverse outcome (P < 0.001) across all morphologic classifications. Further, there was a striking paucity of recurrent cooperating mutations, and transduction of cord blood stem cells with this fusion was sufficient for malignant transformation. CBFA2T3–GLIS2 positive cases displayed marked upregulation of genes with cell membrane/extracellular matrix localization potential, including NCAM1 and GABRE. Additionally, miRNA profiling revealed significant overexpression of mature mir-224 and mir-452, which are intronic miRNAs transcribed from the GABRE locus. Gene-set enrichment identified dysregulated Hippo, TGFβ, and hedgehog signaling, as well as NCAM1 (CD56) interaction pathways. Therapeutic targeting of fusion–positive leukemic cells with CD56-directed antibody–drug conjugate caused significant cytoxicity in leukemic blasts.

Conclusions: The CBFA2T3–GLIS2 fusion defines a highly refractory entity limited to infants that appears to be sufficiently for malignant transformation. Transcriptome profiling elucidated several highly targetable genes and pathways, including the identification of CD56, providing a highly plausible target for therapeutic intervention.

Introduction

Pediatric acute myeloid leukemia (AML) is characterized by significant cytogenetic and molecular heterogeneity. However, the number of patients with AML having clinically actionable variants is limited (1). Although the majority of patients have identifiable chromosomal abnormalities, up to 20% of children have no cytogenetic abnormalities (1, 2). Genomic profiling of megakaryocytic AML (French–American–British classification M7) identified a cryptic CBFA2T3–GLIS2 fusion that was associated with adverse outcome (3, 4). Although CBFA2T3–GLIS2 was subsequently found to be enriched in children with normal cytogenetics (CN-AML), and later, not exclusive to either CN or M7 AML (3, 5–7), the prevalence and prognostic significance of CBFA2T3–GLIS2 in a large heterogeneous cohort of patients have not been fully evaluated.

CBFA2T3 was initially identified as a fusion partner with RUNXI in therapy-related AML and shown to facilitate transcriptional repression, as well as being implicated in hematopoietic stem cell quiescence (3). GLIS2 is closely related to the GLI subfamily and likely functions in modulating the hedgehog signaling pathway (8, 9). The fusion of GLIS2 to CBFA2T3 as a result of inv(16)(p13.3q24.3) leads to increased expression of the GLIS2 DNA-binding domain (3). A number of important features of CBFA2T3–GLIS2 AML have been identified through
Translational Relevance

The CBFA2T3–GLIS2 fusion is a potent fusion oncogene that leads to a highly refractory phenotype in pediatric acute myeloid leukemia (AML). Transcriptome profiling of 1,049 pediatric AML patients defined a unique expression profile for this fusion and identified a number of altered pathways and potential therapeutic targets, including NCAM1 (CD56). We further demonstrated that the CBFA2T3–GLIS2 fusion is sufficient for malignant transformation, causally associated with expression of CD56, and its induction in cord blood stem cells recapitulates human disease. The causal association between the fusion oncogene and surface CD56 expression, as well as the promising results of the CD56–ADC (antibody–drug conjugate) across four independent primary samples, provides a viable target that may be curative for these highly refractory patients. Although at present targeted therapies are not clinically available, the vast body of data generated defines fusion oncoprotein–mediated alterations that can be used for the development of precisely directed targeted therapies.

molecular experiments and next-generation sequencing analyses. Acute megakaryoblastic leukemia (AMKL) with CBFA2T3–GLIS2 displayed altered expression and activation of hedgehog and bone morphogenic signaling pathways, thereby representing potential therapeutic targets (3, 10, 11). Further genomic interrogation of fusion-positive AMKL cell lines identified overexpression of ERG transcripts and a synergistic relationship between the CBFA2T3–GLIS2 oncoprotein and the transcription factor ERG, which primes the leukemic phenotype and blocks differentiation (12). In addition, CBFA2T3–GLIS2 positive cases were found to overexpress the cell-surface marker CD56 (NCAM1), and chromatin immunoprecipitation demonstrated that the fusion protein binds to the proximal promoter, inducing its overexpression (4, 13).

However, previous genomic and transcriptomic studies of CBFA2T3–GLIS2 have limited the investigation to AMKL and CN-AML patients in their characterization of fusion-positive cohorts, in part due to the low frequency of the fusion in unselected populations (~2%). Our study expands on the known features of fusion-positive AML by including a diverse cohort of patients encompassing all morphologic categories, in addition to normal cytogenetics and megakaryoblastic, to provide increased scope and detail in understanding the morphologic categories, in addition to normal cytogenetics and megakaryoblastic leukemia (AMKL) with CBFA2T3–GLIS2 and CD56, to provide increased scope and detail in understanding the morphologic categories, in addition to normal cytogenetics and megakaryoblastic leukemia (AMKL) with CBFA2T3–GLIS2.

Flow cytometry, cytotoxicity assays, and CBFA2T3–GLIS2 transduction

Myeloid immune marker flow cytometry data were provided by Hematologics for 437 pediatric AML patients enrolled on AAML0531. Methods for flow cytometry were previously reported (18). Cytotoxicity assays were carried out on primary AML samples by Notable Labs translational drug-discovery platform and utilized the CD56 antibody clone m906 as previously described (19). The antibody controls included chKTI-SPP-DM1 (20) and human IgG-SAP and goat IgG-ZAP (Advanced Targeting Systems, #IT-36, #IT-35). Human cord blood samples were obtained with informed consent under Swedish Medical Center Institutional Review Board. CD34+ cord blood cells were transduced with a pRRL lentivirus encoding the CBFA2T3–GLIS2 fusion transcript and GFP (21). Cell-surface antigen expression was assessed by flow cytometry (18).

RNA-sequencing library construction

RNA from primary patient samples was purified using the QIAcube system with AllPrep DNA/RNA/miRNA Universal Kits (QIAGEN, #80224). The mRNA libraries were prepared for 75-bp strand-specific paired-end sequencing using the ribodepletion 2.0 protocol by the British Columbia Genome Sciences Center (BCGSC, Vancouver, BC). Sequenced libraries were aligned to GRCh37, and gene-level coverage was quantified using the BCGSC pipeline v1.1 with Ensembl v69 annotations. The microRNA libraries were produced using the miRNA 3.0 protocol by BCGSC. Sequenced reads (31-bp) were trimmed to remove adapter sequences and aligned to GRCh37. Perfect alignments with no mismatches were retained for mature miRNA quantification using miRBase v20 annotations. Transcriptomic data are available through the dbGap TARGET: AML study (accession: phs000465.v19.p6).

Screening of CBFA2T3–GLIS2 fusion

The CBFA2T3–GLIS2 fusion transcript was detected by RNA-sequencing for AAML1031 using STAR-fusion v1.1.0 and TransAbys v1.4.10 fusion detection software (22, 23) and experimentally verified using RT-PCR. Patients enrolled on prior studies were screened by fragment length analysis (FLA) using the Applied Biosystems 3730xl DNA Analyzer. Primers for RT-PCR verification and FLA are listed in Supplementary Methods.

Materials and Methods

Biological specimens and clinical annotations

Pediatric patients with de novo AML enrolled on Children’s Oncology Group (COG) trials CCG-2961, AAML03P1, AAML0531, and AAML1031, with written informed consent and available clinical and/or molecular annotations, were included for biological study. Trials were conducted in accordance with the Declaration of Helsinki. The Fred Hutchinson Cancer Research Center Institutional Review Board and the COG Myeloid Biology Committee approved the study. Details of treatment protocols have been previously published (14–17). Analyses of outcome and clinical characteristics were limited to patients from AAML0531 and AAML1031. Data were current as of March 31, 2018, for outcome analyses, and cohort characteristics are summarized in Supplementary Tables S1 and S2. Patients with anFLT3-ITD high allelic ratio from AAML1031 were excluded from survival and clinical associations due to enrollment on the phase I sorafenib treatment arm, because a proportion of patients were still on therapy with data under DSMC review. Risk groups were defined as follows: patients with t(8;21), inv(16), NPM1 or CEBPA mutations were considered low risk, whereas patients with monosomy 7, monosomy 5(del5q), or high allelic ratio FLT3-ITD (ITD-AR > 0.4) were considered high risk (16). Remaining patients were considered standard risk. As a comparison, we screened 299 patients enrolled in the SWOG adult AML trial S1006.
Differential expression, hierarchical clustering, gene set enrichment analysis, and miRNA–mRNA interactions

All analyses were completed in the R statistical environment. Differentially expressed genes and miRNAs were identified using Limma v.3.36.1, and those with absolute log2 fold change > 1 and false discovery rate (FDR) < 0.05 were retained. Gene counts were TMM normalized and converted to log2 scale for unsupervised hierarchical clustering. Gene set enrichment analysis was completed using GAGE v2.30.0 (24). Interactions between miRNA–mRNA were investigated by selecting pairs of differentially expressed miRNAs and genes with significant anti-correlation using Spearman’s rho (FDR < 0.05). The miRNA–mRNA pairs were investigated for both predicted and validated interactions using anamiR v1.10.0 and multiMiR v1.4.0 packages. LNA inhibition of CBFA2T3–GLIS2 fusion was also detected in 50.0% morphologic CR rate at the end of induction 1 (EOI1) compared with 76.1% for fusion negative (P < 0.001). Fusion-positive AML had a higher median percentage of residual disease (MRD) at the end of induction 1 and 2 compared with the fusion-negative population; 80% were MRD positive at the end of induction 1 (vs. 26.7%, P < 0.001) and 35.3% at the end of induction 2 (vs. 13.7%, P = 0.005). In fact, 94.3% of CBFA2T3–GLIS2 patients had detectable residual disease after induction 1 (median: 0.7%, range: 0.0%–41.0%). Overall survival (OS) at 5 years after study entry for CBFA2T3–GLIS2 AML was 22.0% versus that of 63.0% in fusion-negative cases (HR = 2.8, P < 0.001, Fig. 1E). CBFA2T3–GLIS2 fusion was also associated with a higher probability of an event (death, relapse, or induction failure) with an event-free survival (EFS) at 5 years from study entry 18.9% versus 46.9% for fusion-negative patients (HR = 2.0, P < 0.001). Associated adverse prognostic impact was observed regardless of whether or not patients had megakaryocytic phenotype. A multivariable analysis including CBFA2T3–GLIS2 status, cytogenetic/molecular risk classification, and M7 morphology demonstrated that CBFA2T3–GLIS2 was an independent prognostic factor for both OS (HR = 2.3, P < 0.001) and EFS (HR = 2.0, P = 0.005; Supplementary Table S3). In addition, the CBFA2T3–GLIS2 fusion has prognostic value independent of age (Supplementary Figs. S2 and S3). Infant AML patients < 1-year-old with the CBFA2T3–GLIS2 fusion (N = 30) had adverse OS compared with the fusion-negative (N = 308) population (27.6% vs. 61.3% at 5 years after enrollment, P = 0.001). Infant CBFA2T3–GLIS2 patients with FAB M7 morphology (N = 17) remained at greater risk of death (OS 25%) compared with other FAB M7 infant AML cases (N = 44, OS 58.9%, P = 0.025) 5 years after study entry.

CBFA2T3–GLIS2 immunophenotype and targeted therapy

Evaluation of leukemic blast by multidimensional flow cytometry (MDF) demonstrated that leukemias with CBFA2T3–GLIS2 had a highly elevated CD36 expression, that in combination with dim or absent expression of HLA-DR, CD38, and CD45 can be characterized as previously described RAM phenotype (18). All patients with CBFA2T3–GLIS2 fusions were evaluated by MDF and were found to have RAM phenotype (100%, P < 0.001). Expression values for 13 cell-surface antigens (myeloid marker panel) by MDF were available for a subset of patients from AAML0531 (N = 437). Unsupervised hierarchical clustering demonstrated that fusion-positive tumors (7/7)
clustered together by mean fluorescence intensity (MFI) of the 13 antigens, indicating that CBFA2T3–GLIS2 tumors have highly distinctive immunophenotypes, with extremely elevated surface CD56 expression ($P < 0.001$, Fig. 2A; Supplementary Fig. S4). We also evaluated the available diagnostic immunophenotype data to determine whether fusion-positive cases could be identified using the characteristic RAM phenotype antigens: CD56, CD45, CD38, and HLA-DR. Unconstrained classic metric multidimensional scaling (MDS) was used to cluster the 437 patients using MFI values of the four RAM antigens, and observed 100% (7/7) of fusion-positive AML patients clustered away from the fusion-negative cohort, providing compelling evidence that CBFA2T3–GLIS2 is associated with a unique cell-surface expression profile that can be characterized by the RAM antigens alone (Fig. 2B).

Next, considering that a majority of CBFA2T3–GLIS2 patients express the RAM immunophenotype, we investigated whether transcript-level expression would support the observations of low-to-dim expression of CD45, CD38, and HLA-DR. Using rRNA-depleted RNA-sequencing data from 24 fusion-positive cases, it was found both CD38 and CD45 transcripts were significantly down-regulated >2-fold compared with fusion-negative AML ($N = 1,025$), as well as the HLA-DR genes: HLA-DRA, HLA-DRB1, HLA-DRB5, and HLA-DRB6 (FDR < 0.001). Investigation of miRNA–mRNA interactions using paired miRNA-sequencing data revealed CD38 mRNA had significant anti-correlation (Spearman’s rho) with its predicted target binding sites for the overexpressed miR-203a-3p (rs $\leq 0.62$, FDR < 0.02). For CD45, there were significant associations with highly expressed mature miR-452-3p, miR-425-5p, and both miR-224 species ($r_{\leq 0.74}$, FDR $< 0.003$). HLA-DR genes also had relationships with overexpressed miRNAs: HLA-DRA had predicted target sites for miR-135a-3p ($r_{\leq 0.62}$, FDR = 0.02) and HLA-DRB5 with miR-5683 ($r_{\leq 0.53}$, FDR < 0.05), thus implicating miRNAs in negative regulation of the cell-surface markers that define the RAM phenotype.
We next evaluated CD56 as a potential therapeutic target using an anti-CD56 antibody–drug conjugate (m906-PBD-ADC) developed by the NCI (19, 20). Leukemic blasts and control lymphocytes from a patient with relapsed fusion-positive AML were incubated with varying doses of m906, and cytotoxicity was assessed after 72 hours. The CD56-ADC exhibited a dose-dependent and CD56-specific toxicity on leukemic blasts. There was 87% cytotoxicity of CD56+ blasts at 3 nmol/L concentration, suggesting that CD56 might prove to be a therapeutic target in this high-risk cohort of patients (P < 0.001, Fig. 2C). The CD56-ADC was further evaluated in an additional 3 independent CBFA2T3–GLIS2 patient samples; the dose-dependent toxicity was observed in all samples. In total, 3 of 4 CBFA2T3–GLIS2 primary samples showed >75% cytotoxicity of leukemic blasts at 3 nmol/L concentration of the CD56-ADC (Fig. 2D; Supplementary Fig. S5).

We next evaluated CD56 as a potential therapeutic target using an anti-CD56 antibody–drug conjugate (m906-PBD-ADC) developed by the NCI (19, 20). Leukemic blasts and control lymphocytes from a patient with relapsed fusion-positive AML were incubated with varying doses of m906, and cytotoxicity was assessed after 72 hours. The CD56-ADC exhibited a dose-dependent and CD56-specific toxicity on leukemic blasts. There was 87% cytotoxicity of CD56+ blasts at 3 nmol/L concentration, suggesting that CD56 might prove to be a therapeutic target in this high-risk cohort of patients (P < 0.001, Fig. 2C). The CD56-ADC was further evaluated in an additional 3 independent CBFA2T3–GLIS2 patient samples; the dose-dependent toxicity was observed in all samples. In total, 3 of 4 CBFA2T3–GLIS2 primary samples showed >75% cytotoxicity of leukemic blasts at 3 nmol/L concentration of the CD56-ADC (Fig. 2D; Supplementary Fig. S5).

Malignant transformation of cord blood stem cells by CBFA2T3–GLIS2 fusion

Given the paucity of cooperating events in a highly aggressive disease in early life, we questioned whether this fusion may be sufficient for malignant transformation. To this end, we transduced human CD34+ cord blood stem cells (CBSC) with a lentivirus encoding the CBFA2T3–GLIS2 fusion transcript and GFP or mock control; expression of the fusion was confirmed by RT-PCR. Fusion-transduced CBSCs were more proliferative compared with mock-transduced CBSCs (Supplementary Fig. S6, P = 0.21). After 12 weeks in culture, we evaluated immunophenotypic and morphologic features of CBFA2T3–GLIS2–transduced cells. CBFA2T3–GLIS2+ cells had immature morphology with high CD117 expression and absence of CD11b, CD36, and CD64 (Fig. 3A and C). Of significance, fusion-transduced cells had remarkably high CD56 expression (3,556 MFI vs. 43 MFI for the control), similar to that observed in patients with fusion-positive RAM phenotype (2,040 MFI), suggesting that CD56 expression is causally linked to the expression of the fusion transcript. In addition, morphologic evaluation of CBFA2T3–GLIS2+ cells revealed multinucleated cells, prominent nucleoli, and abundant focally basophilic and vacuolated cytoplasm with cytoplasmic blebs, which are morphologic features suggestive of megakaryocytic differentiation. Megakaryocytic lineage was confirmed by demonstration of high expression of CD41 and CD61 on the surface of transduced cells (Fig. 3D and E).
Gene-expression profiling of CBFA2T3–GLIS2 patients

Of 39 pediatric CBFA2T3–GLIS2 AML patients identified, 24 of 39 fusion-positive and 1,025 fusion-negative cases had material available for high depth ribodepleted mRNA-sequencing and paired miRNA-sequencing. Transcriptome data from CBFA2T3–GLIS2-positive and -negative tumors were contrasted to identify fusion-specific pathways, genes, and miRNAs (Fig. 4A). Gene set enrichment analysis (24) identified a substantial activation of Hippo (FDR < 0.001) and TNF signaling pathways (FDR < 0.001), which have been implicated in various cancers by promoting cellular functions that enhance tumor growth, migration, and proliferation (26, 27). In addition, TGFB/BMP and Hedgehog signaling pathways showed significant activation, as previously identified (refs. 3, 6; FDR < 0.001, Fig. 4B and C). Furthermore, the cell-surface NCAM1 (CD56) interaction pathway was highly positively enriched (FDR < 0.001), whereas gene-ontology (GO) term analysis revealed significant upregulation of numerous cell adhesion and cell-surface marker genes, including extracellular matrix binding, cell-adhesion molecule binding, and integrin binding genes (FDR < 0.001).

Considering that a characteristic feature of CBFA2T3–GLIS2 AML was positive regulation of cell adhesion and extracellular matrix binding (ECM), overexpressed genes whose protein products have demonstrated localization to the plasma membrane and ECM were examined (28). Of 3,711 differentially expressed genes (DEG), 789 genes had cell adhesion, plasma membrane, and ECM associations (FDR < 0.001, Fig. 5A; Supplementary Fig. S7). Plasma membrane and ECM genes included 10 NCAM1 interaction pathway genes, overexpressed hematopoietic lineage adhesion molecules CD44, ITGA2, and ITGA2B, as well as numerous receptor tyrosine kinases. CD44, expressed at a level more than 2-fold greater than fusion-negative patients, is aberrantly expressed on leukemic stem cells and a marker of HSCs (29, 30). Similarly, ITGA2 is a cell adhesion molecule on hematopoietic cells and used to classify subsets of HSCs (29). Upregulated receptor tyrosine kinases (RTK) that localize to the cell periphery and cell membrane included ROR1, MET, and two of the TAM family kinases, TYRO3 and AXL, were noted [log2 fold change (LFC) ≥ 2.0, FDR < 0.001]. The TAM family of RTKs has been shown to increase tumor cell survival, migration, and chemoresistance (31), whereas others have known roles in cancer progression (32, 33). In addition, genes affected by BMP and hedgehog signaling cross-talk (10) with demonstrated ECM localization were upregulated, including BMP2, WNT9A, and WNT11. A concomitant overexpression of WNT11 and ERG transcripts (LFC ≥ 1.5, FDR < 0.001) was observed, and it has been shown that the ERG transcription factor directly binds the WNT11 locus to promote cancerous transformation in AML (34).

Figure 3. CBFA2T3–GLIS2 induces overexpression of CD56 (NCAM1) in healthy normal cord blood cells. A, Surface immunophenotype of mock-transduced cord blood cells cultured for 6 weeks (green); B, CBFA2T3–GLIS2 transduced cord blood cells cultured for 10 weeks (purple), and (C) patient with CBFA2T3–GLIS2 and RAM phenotype (red). Patient data for CD64 and CD36 are not available. Morphology of CBFA2T3–GLIS2-transduced cells includes (D) megakaryocytic features, such as cytoplasmic blebs, noted with arrow, and (E) cells express megakaryocytic surface markers CD41 and CD61.
Unsupervised hierarchical clustering based on the most highly expressed (90th percentile LFC, $N = 79$ genes) cell surface-associated genes found CBFA2T3–GLIS2 AML exhibits a unique expression profile with a distinct set of dysregulated genes: NCAM1 (CD56), the NCAM1-interacting partner CACNB2, and GABRE receptor gene (Fig. 5B). Hierarchical clustering on the cell adhesion/ECM-associated markers also revealed a number of fusion-negative AML patients ($N = 66$) that clustered closely with the fusion-positive cohort. This cohort shared similar characteristics with fusion-positive patients: they tend to be very young, with a median age of 1.8 years (range, 0.09–17.6 years) and 85% of cases are < 3 years old. They were enriched in M7 AML ($N = 37/54$, 68%), including the majority of NUP98–KDM5A ($N = 12/17$, 70.5%) and RBM15–MKL1 ($N = 10/10$, 100%) cases. Patients with similar cell-surface associated gene expression also had poor outcomes compared with the fusion-negative population for OS (47.5% vs. 64.8%, $P < 0.001$), though not EFS ($P = 0.17$). Identification of similar patients allows the inclusion of many AML patients that otherwise would be missed in a targeted investigation of therapeutic biomarkers. However, these patients have some distinct differences, sharing 57 of 79 (72%) of the most upregulated cell

---

**Figure 4.**

A cohort of 24 CBFA2T3–GLIS2 AML patient samples was used in transcriptional profiling for mRNA and miRNA differential expression analysis and gene set enrichment analysis. A, Schematic representation of the data types and transcriptomic analyses in this study. B, Heat map of positively and negatively (blue) enriched pathways from gene set enrichment analysis of the CBFA2T3–GLIS2 cohort. Pathways (rows) have Benjamini–Hochberg–adjusted $P$ values ≤ 0.001. The enrichment statistic for the gene set enrichment analysis is denoted in each tile and was derived from a two-sample $t$ test for differential expression of gene sets. C, Differentially expressed genes involved in Hedgehog signaling or its cross-talk found to be upregulated in CBFA2T3–GLIS2 tumors.
surface–associated genes in common with CBFA2T3–GLIS2 (Supplementary Table S4).

**MicroRNA profiling of CBFA2T3–GLIS2 patients**

The 1,049 patients with whole-transcriptome data had paired miRNA-sequencing completed. Contrasting CBFA2T3–GLIS2 AML with the fusion-negative cohort (N = 1,025) revealed 134 differentially expressed mature miRNAs, with the magnitude of differential expression being much greater for the upregulated miRNAs (Fig. 6A). The most highly expressed miRNAs were miR-224-5p, miR-452-3p, miR-452-5p, and miR-6507-5p (Fig. 6B). Human miR-224 and miR-452 have both been implicated as oncomirs in malignant melanoma, as well as colorectal cancers, and are intronic miRNAs transcribed from the GLIS2 locus, one of the most highly aberrantly expressed cell membrane–associated genes in CBFA2T3–GLIS2 (35, 36). Conversely, the most downregulated miRNAs were miR-6503-3p, miR-196b-3p, miR-196b-5p, miR-133a-3p, and miR-5584-5p. The mature miRNA miR-196b functions as a tumor suppressor during lung carcinogenesis (37), and well-characterized miR-133a has been found to function in a tumor-suppressive manner in as many as five cancers, including breast and colorectal cancer (38).

To examine miRNA regulatory networks in CBFA2T3–GLIS2, miRNA–mRNA interactions were investigated by selecting significantly correlated differentially expressed miRNA–gene pairs (Spearman’s rho, FDR < 0.05). The miRNA–gene pairs were queried across independent target-prediction and experimental evidence databases, identifying 1,727 miRNA–mRNA interacting partners. Predicted interactions were only considered if two or more algorithms identified the miRNA binding-site homology or the interaction was supported by experimental evidence. The most highly overexpressed miRNAs in CBFA2T3–GLIS2 tumors, miR-224 and miR-452, were associated with numerous mRNA targets. GO enrichment analysis of miR-224 target genes (N = 107 genes) revealed that these were involved in immune response (FDR = 0.002), as well as myeloid leukocyte differentiation (FDR = 0.001), and immune system processes (FDR = 0.005), indicating these pathways are inhibited in fusion-positive AML and high expression of miR-224 potentially contributes to the pathways’ modulation. The overexpressed mature miR-452 associated with 127 target genes, which are involved in leukocyte differentiation (FDR = 0.001), and immune system processes (FDR = 0.001), as well as myeloid leukocyte differentiation (FDR = 0.02). A perturbation experiment using LNA miRNA inhibitors for (i) hsa-miR-224-5p, (ii) hsa-miR-224-5p, (iii) hsa-miR-452-5p, or (iv) hsa-miR-452-3p was performed. The CBFA2T3–GLIS2 cell line M07E was exposed to inhibitor for 72 hours, resulting in a knockdown efficiency > 90% for each miRNA (Supplementary Fig. S8). Proliferation rates were assessed for the following 4 days (7 days of continuous LNA inhibitor exposure). The inhibition of each miRNA alone produced minimal to no morphologic or immunophenotypic alterations and led to subtle increase in proliferation compared with negative control inhibitors (Supplementary Fig. S9; P ≤ 0.01).

In contrast to gaining proliferative ability, tumors’ modulation of apoptotic genes is often involved in cancer progression (39). Two such
genes, CASP1 and TRAIL-R2, were significantly downregulated, as well as the tumor suppressors GLIPR1 and PER2 (39–41). CASP1 was significantly anti-correlated with, and a putative target of the overexpressed miR-181b-5p ($r_s = -0.76$, FDR $= 0.002$, Fig. 6C). TRAIL-R2 associated with upregulated mature miR-425-5p, and the regulatory interaction had experimentally validated evidence ($r_s = -0.80$, FDR $< 0.001$). GLIPR1 mRNA decreased with increasing miR-130a-3p transcript levels ($r_s = -0.72$, FDR $= 0.003$) and PER2 displayed significantly anti-correlated expression levels with miR-181b-5p ($r_s = 0.82$, FDR $< 0.001$) and possessed experimental evidence of the interaction.

Finally, we investigated regulatory interactions among the most significantly downregulated miRNAs in CBFA2T3–GLIS2 AML. Mature miR-196b-5p had an identified regulatory interaction with the cell adhesion molecule ITGA2, whose expression was 22-fold greater in fusion-positive AML ($r_s = -0.75$, FDR $= 0.03$). ITGA2 is expressed on HSCs (29), and silencing of miR-196b appears to induce abnormal regulation of this marker. In addition, miR-196-5p had target
binding sites identified within DNA methyltransferase DNMT3B transcripts (rs = −0.73, FDR = 0.03). DNMT3B was upregulated in CBFA2T3–GLIS2 tumors and has been associated with adverse outcome in pediatric AML (42). The mir-199a/b family members, characterized as tumor suppressors in multiple cancers including AML (43, 44), had a predicted regulatory relationship with the massively overexpressed plasma-membrane SLITRK5 gene (over 200-fold more highly expressed in fusion-positive cases, FDR < 0.001). SLITRK5, expressed on both HSCs and aberrantly on CD34+ leukemic cells (45), had a predicted binding site for mir-199b-3p and mir-199a-3p (rs ≤ −0.73, FDR ≤ 0.03), thus providing one putative mechanism contributing to the overexpression of this AML-associated gene transcript (Supplementary Fig. S10).

**Discussion**

This comprehensive study represents a large uniformly treated cohort of patients to be evaluated by transcriptome sequencing, in order to define the biological implications of the CBFA2T3–GLIS2 fusion in children with de novo AML. We established that although the prevalence of CBFA2T3–GLIS2 is ~2% across an unselected pediatric cohort, it is highly enriched in younger AML populations (5, 46, 47), with nearly 12% of 1 year-old infants harboring this fusion. In fact, the CBFA2T3–GLIS2 fusion was not detected in those older than 3 years of age. We also confirmed the CBFA2T3–GLIS2 fusion is not limited to those with AMKL or normal karyotypes (3–7), with predilection to patients with African ancestry. Further, a majority of CBFA2T3–GLIS2 positive patients have nonrecurrent karyotypic alterations with unexplained enriched truncation of Trisony 3, and the overwhelming majority express the unique RAM immunophenotype, which provides a seamless mechanism for rapid identification of this high-risk population (18).

Furthermore, the lack of recurrent fusions or somatic mutations in a very young population suggests that this cryptic fusion may be sufficient for leukemic transformation without requirement for additional cooperating events. We confirmed this hypothesis by demonstrating that transduction of CBSC1 leads to malignant transformation that recapitulates infant AML with enhanced proliferation, maturation arrest, as well as morphologic and immunophenotypic alterations consistent with megakaryocytic AML. This is in contrast to the disease in older patients where cooperation between class I and II variants is required for malignant transformation (48). The fact that transduction of CBSC led to induction of NCAM1 (CD56) expression, in combination with both HSCs and aberrantly on CD34+ leukemic cells (45), had a predicted binding site for mir-199b-3p and mir-199a-3p (rs ≤ −0.73, FDR ≤ 0.03), thus providing one putative mechanism contributing to the overexpression of this AML-associated gene transcript (Supplementary Fig. S10).

In order to unravel the complex regulatory networks that are involved in the transcriptional dysregulation of CBFA2T3–GLIS2, we investigated the expression of mature miRNAs in fusion-positive AML. The most highly upregulated miRNAs were mir-224 and mir-452, which have been shown to be associated with poor outcomes across six independent colorectal cancer cohorts, and both are overexpressed in malignant melanoma (35, 36). Interestingly, target genes of mir-224 and mir-452 were associated with GO enrichment of immune system processes and leukocyte differentiation, suggesting a possible mechanism of an immunosuppressive environment in fusion-positive AML. It was also demonstrated that mir-224/mir-452 genes are located in a miRNA cluster and are induced by transactivation of GABRE, which was one of the most highly upregulated genes in CBFA2T3–GLIS2 (LFC = 10.6, FDR<0.001; ref. 35). However, the evidence suggested overexpression of E2F1 drove transactivation, yet CBFA2T3–GLIS2 AML lacked an overexpression of E2F1, suggesting an alternative mechanism of regulation. Transcription factor binding sites in K562 ChIP-seq from the ENCODE project showed that TAL1 and RFXO2, both upregulated >2-fold in CBFA2T3–GLIS2 patients, bind to the promoter region of GABRE, indicating alternative avenues for the observed overexpression (49). The results of the knockdown experiments are more difficult to interpret, and may indicate that these miRNAs may not be limited to enhancing cell-division/mitotic potential, but offer an alternative clonal advantage, possibly related to cell quiescence considering their apparent negative regulation of proliferation. In addition, it cannot be discounted that inhibition of these closely related miRNAs individually may not account for potential synergistic effects. Combinatorial inhibition experiments may provide improved elucidation of the regulatory mechanisms to which the four miRNAs contribute.

Alternatively, many upregulated miRNAs in fusion-positive AML were found to target apoptotic and tumor suppressor genes; mir-181b-5p was highly elevated in CBFA2T3–GLIS2 tumors and had predicted target sites on the downregulated apoptotic genes CASP1 and TRAIL-R2, and the circadian rhythm gene PER2. Meanwhile, tumor-suppressive miRNAs, including mir-196a/b, mir-135a, and mir-196b/b, had significantly reduced expression in fusion-positive AML, suggesting another mode for tumor progression. Upregulation of ITGA2 and DNMT3B, which have been shown to be independent prognostic indicators associated with poor outcomes in AML when highly expressed, were associated with the loss of mir-196b expression (42, 50). Taken together, this suggests another means for leukemic tumor cells to deregulate normal signals for cell growth arrest.

In summary, CBFA2T3–GLIS2 fusion is a highly potent fusion oncogene that is sufficient for malignant transformation and induction of highly refractory leukemia. Identification of this fusion early at diagnosis would allow more appropriate risk adaptive therapeutic intervention. Although at present targeted therapies are not available for these highly refractory patients, identification of causal association of the fusion oncogene and surface CD56 expression provides a viable target that may be curative. The promising results of the CD56-ADC across 4 independent primary samples demonstrated the potency and specificity of this therapeutic strategy. Targeting the CD56 antigen offers the most immediately actionable target in this cohort of high-risk patients, and an effective CD56-ADC with limited off-target toxicity is available for continued refinement toward this goal. Additionally, this study provides comprehensive transcriptome profiling of CBFA2T3–
GLIS2 patients who defines an expression signature and dysregulated pathways and networks in this distinct and highly lethal AML of infancy. The vast body of data generated defining fusion oncogene–mediated alterations that lead to therapeutic resistance can be used for the development of precisely directed targeted therapies.

Disclosure of Potential Conflicts of Interest

M.T. Santaguida is an employee/paid consultant for Notable Labs. L.E. Brodersen is an employee/paid consultant for Hematologics. L. Pardo is an employee/paid consultant for Hematologics. M.R. Loken is an employee/paid consultant for and holds ownership interest (including patents) in Hematologics. No potential conflicts of interest were disclosed by the other authors.

Authors’ Contributions


References

Transcriptome Profiling of the Cryptic CBFA2T3–GLIS2 Fusion

Positive AML Defines Novel Therapeutic Options: A COG and TARGET Pediatric AML Study


Comprehensive Transcriptome Profiling of Cryptic CBFA2T3–GLIS2 Fusion –Positive AML Defines Novel Therapeutic Options: A COG and TARGET Pediatric AML Study

Updated version
Access the most recent version of this article at:

Supplementary Material
Access the most recent supplemental material at:
http://clincancerres.aacrjournals.org/content/suppl/2019/11/12/1078-0432.CCR-19-1800.DC1

Cited articles
This article cites 49 articles, 12 of which you can access for free at:
http://clincancerres.aacrjournals.org/content/26/3/726.full#ref-list-1

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, use this link
http://clincancerres.aacrjournals.org/content/26/3/726.
Click on "Request Permissions" which will take you to the Copyright Clearance Center’s (CCC) Rightslink site.