MYELOID NEOPLASIA

**CUX1 is a haploinsufficient tumor suppressor gene on chromosome 7 frequently inactivated in acute myeloid leukemia**

Megan E. McNerney,1,2 Christopher D. Brown,1,3 Xiaoyue Wang,1,3 Elizabeth T. Bartom,4 Subhradip Karmakar,1 Chaitanya Bandlamudi,1,5 Shan Yu,1,5 Jinkyung Ko,6 Barry P. Sandall,6 Thomas Stricker,7 John Anastasi,2 Robert L. Grossman,1,4 John M. Cunningham,6 Michelle M. Le Beau,1,8 and Kevin P. White1,3,9

1Institute for Genomics and Systems Biology, 2Department of Pathology, 3Department of Human Genetics, 4Center for Research Informatics, 5Committee on Genetics, Genomics, and Systems Biology, and 6Department of Pediatrics, Section of Hematology/Oncology and Stem Cell Transplantation, University of Chicago, Chicago, IL; 7Department of Pathology, Microbiology, and Immunology, Vanderbilt University, Nashville, TN; and Sections of 8Hematology/Oncology and 9Genetic Medicine, University of Chicago, Chicago, IL

© 2013 by The American Society of Hematology

**Key Points**

- CUX1 is a transcription factor encoded on a region of chromosome 7 that is frequently deleted in high-risk acute myeloid leukemia.
- Haploinsufficiency of CUX1/cut promotes hematopoietic overgrowth in both Drosophila melanogaster and human xenograft mouse models in vivo.

**Loss of chromosome 7 and del(7q) [−7/del(7q)] are recurring cytogenetic abnormalities in hematologic malignancies, including acute myeloid leukemia and therapy-related myeloid neoplasms, and associated with an adverse prognosis.** Despite intensive effort by many laboratories, the putative myeloid tumor suppressor(s) on chromosome 7 has not yet been identified. We performed transcriptome sequencing and SNP array analysis on de novo and therapy-related myeloid neoplasms, half with −7/del(7q). We identified a 2.17-Mb commonly deleted segment on chromosome band 7q22.1 containing CUX1, a gene encoding a homeodomain-containing transcription factor. In 1 case, CUX1 was disrupted by a translocation, resulting in a loss-of-function RNA fusion transcript. CUX1 was the most significantly differentially expressed gene within the commonly deleted segment and was expressed at haploinsufficient levels in −7/del(7q) leukemias. Haploinsufficiency of the highly conserved ortholog, cut, led to hemocyte overgrowth and tumor formation in Drosophila melanogaster. Similarly, haploinsufficiency of CUX1 gave human hematopoietic cells a significant engraftment advantage on transplantation into immunodeficient mice. Within the RNA-sequencing data, we identified a CUX1-associated cell cycle transcriptional gene signature, suggesting that CUX1 exerts tumor suppressor activity by regulating proliferative genes. These data identify CUX1 as a conserved, haploinsufficient tumor suppressor frequently deleted in myeloid neoplasms. (*Blood. 2013;121(6):975-983*)

**Introduction**

Loss of chromosome 7 and del(7q) [−7/del(7q)] was first recognized as a frequent event in acute myeloid leukemia (AML) nearly 40 years ago.1 −7/del(7q) occurs in 8% of de novo AML2 and 50% of therapy-related myeloid neoplasms (t-MNs).3 −7/del(7q) is also found in myelodysplastic syndromes, AMLs arising from myeloproliferative neoplasms, the blast phase of chronic myelogenous leukemia, Ph+ acute lymphoblastic leukemia, and AMLs associated with inherited syndromes.4,10 −7/del(7q) is an adverse-risk prognostic indicator in myeloid disorders, and the long-term outcome for patients is typically poor. The median overall survival for patients with de novo AML or t-MNs with −7/del(7q) is ~6 months.2,3

Loss of 1 or more tumor suppressor gene(s) (TSGs) is thought to contribute to leukemic growth in myeloid malignancies with −7/del(7q). Several groups have mapped a commonly deleted segment (CDS) of chromosome band 7q22 using polymorphic markers, conventional cytogenetic analysis, and FISH analysis.11-13 In one study of 81 patients with malignant myeloid disorders characterized by chromosome 7 abnormalities, the CDS was mapped to a 2.52-Mb region of 7q22 by FISH using YAC clones.11 However, deletion of a 2-Mb syntenic region in mice did not result in overt myeloid disease.14 Other studies have mapped rearrangements involving 7q22 and identified similar15 or slightly more centromeric intervals (Figure 1).12,13,16

In this study, we used single nucleotide polymorphism (SNP) arrays to map the 7q22 CDS with resolution of ~1 kb. Overlaying transcriptome-sequencing with copy number aberrations, we identified the gene encoding the transcription factor, CUX1, to be within the CDS and expressed at haploinsufficient levels. Haploinsufficient levels of the CUX1 ortholog, cut (ct), in Drosophila melanogaster hemocytes led to hemocyte overproliferation and melanotic tumor formation in vivo. Similarly, decreased expression of CUX1 led to an engraftment advantage for human hematopoietic progenitors transplanted into immunodeficient mice. Thus, from invertebrates to humans, CUX1/ct is a conserved, haploinsufficient hematopoietic TSG.
Methods

Patient samples

This research was approved by the University of Chicago Institutional Review Board. Leukemia samples were obtained from the University of Chicago Hematopathology and Cancer Cytogenetics Laboratories (supplemental Table 1, available on the Blood Web site; see the Supplemental Materials link at the top of the online article) with informed consent per the Declaration of Helsinki.

SNP array

DNA was analyzed on Infinium HumanOmni2.5-Quad Version 1.0 DNA Analysis BeadChips (Illumina). Log R ratio and B allele frequencies generated using GenomeStudio (Illumina) were used to identify copy number aberrations by GenoCN.\(^\text{17}\) GEO accession number is GSE42482.

RNA sequencing library preparation

Paired-end libraries were prepared following the standard protocol recommended by Illumina. In brief, purified mRNA (MicroPoly(A) Purist kit, Ambion) was fragmented for 7 minutes at 85°C, and first-strand cDNA generated (Superscript II Reverse Transcriptase, Invitrogen) with random hexamers (Invitrogen), followed by second-strand synthesis with RNaseH and DNA Polymerase I (New England Biolabs). cDNA was repaired and polished with T4 DNA polymerase, Klenow, and T4 PolyNucleotide Kinase (New England Biolabs), followed by adenosine addition with Klenow Fragment 3' → 5' exo^- (New England Biolabs). Paired-end adaptors (Illumina) were ligated with DNA T4 ligase, and libraries were amplified with p5 and p7 primers (Illumina) and Platinum Pfx taq polymerase (Invitrogen) for 18 PCR cycles; 450-bp fragments were gel-extracted using QIAquick Gel Extraction kit (QIAGEN).

Sequence analysis

Paired-end reads of lengths 36-100 bp were generated on the Illumina Genome Analyzer II. Reads were individually trimmed from the 3' end, such that the trimmed 3' bases had an average Phred-scaled quality score < 15. Reads were aligned to the human reference genome hg18 using TopHat (Version 1.3.1)\(^\text{18}\) and output in the BAM format.\(^\text{19}\) Further alignment manipulation was performed with SAMTools (Version 0.1.18),\(^\text{18}\) Picard (Version 1.52, http://picard.sourceforge.net), and custom perl scripts. Read alignments with mapping qualities < 10 were removed. Alignments were refined with the Genome Analysis Toolkit (Version 1.0.5) Realigner-TargetCreator and IndelRealigner tools.\(^\text{20}\) Cufflinks (Version 1.3.0) was used to estimate transcript abundance using RefSeq gene models and output as Fragments Per Kilobase per Million mapped reads (FPKM) with upper-quartile normalization (supplemental Table 2).\(^\text{21}\) Statistical tests were performed on log2-transformed data after addition of 1 pseudocount. RNA-sequencing data are in the Short Read Archive SRA061655.
Möller, and sample T12 defined the distal boundary of the 2.17-Mb CDS on chromosome 7q22. One sample, T12, was found to have an intrachromosomal 2.17-Mb deletion of chromosome 7 and were uninformative for large deletions of chromosome 7 and were uninformative for the study of the borders of the deleted segment.

Results

CUX1 resides within a commonly deleted segment of chromosome band 7q22

We performed SNP-array analysis of 35 de novo and t-MN leukemia samples and defined a 2.17-Mb CDS on chromosome band 7q22 (Figure 1). Cytogenetic analysis revealed clonal 7q22 in 17 of these. None of the samples showed biallelic deletions of chromosome 7. As seen in previous reports,15–17 most samples had large deletions of chromosome 7 and were uninformative for defining the CDS of 7q22. One sample, T12, was found to have an interstitial deletion encompassing 7q22 not visible by cytogenetic analysis (Figure 1). Samples T12 and T20 (with a del(7)(q22q36)) had similar proximal breakpoints and defined the proximal boundary, and sample T12 defined the distal boundary of the 2.17-Mb CDS (Figure 1). Although this CDS is defined by a small number of patients, this region overlaps with prior studies.

unless otherwise noted, all the experiments were performed in duplicate and the results were expressed as the mean ± SEM. Statistical analysis was performed using GraphPad Prism 5 software (GraphPad, La Jolla, CA).

RT-PCR
cDNA was generated using the High Capacity RNA-to-cDNA Master Mix (Applied Biosystems). CUX1 was amplified with Power SYBR Green PCR Master Mix (Invitrogen) on the ABI Step One Plus (Applied Biosystems). cut was amplified from cDNA from hemocytes pooled from 6 larvae per group, using EXPRESS qPCR Supermix (Invitrogen) and detected with a dual-labeled probe (IDT). Primer sequences are provided (supplemental Table 3). Samples were run in triplicate, and expression levels relative to a housekeeping control gene were determined using the comparative Ct method according to the manufacturer’s software.

Fusion identification
deFuse22 (Version 0.5.0) was used to identify chimeric transcripts requiring > 7 split reads, > 7 spanning reads, split_span P value > 0.2, split position P value > 0.2, breakpoint homology < 11, and breakpoints in coding regions. To identify reads mapping to the fusion breakpoint junction, exon-exon junctions between the 2 genes were constructed and all reads were mapped to this junction library with BWA (Version 0.5.9).23 Reads with < 3 mismatches and > 14 bp anchor region around the junction were identified.

Cell culture
K562, Kasumi-1, KG-1, NB4, Mono7, and U-937 human AML cell lines, and HeLa adenocarcinoma cells were maintained in RPMI 1640-Glutamax (Invitrogen) supplemented with 10% fetal bovine serum (Invitrogen), 1 g/mL streptomycin, and 0.25 g/mL amphotericin B (Invitrogen). Cells were grown at 37°C with 5% CO2.

Western blot
A total of 2 million cells were lysed in RIPA with sodium orthovanadate, protease inhibitor cocktail, and phenylmethylsulfonyl fluoride (Santa Cruz Biotechnology). A total of 40 g of supernatant, quantified by Micro BCA (QIAGEN) was amplified using Taq 5X Master Mix (New England Biolabs) with primers for CUX1 and 0.3 units Shrimp Alkaline Phosphatase (Affymetrix) for 30 minutes at 37°C followed by 15 minutes at 80°C. Sequencing traces were analyzed using MutationSurveyor Version 4.00 software (SoftGenetics).
**CUX1 is expressed at haploinsufficient levels in −7/del(7q) leukemias**

To identify genes that are differentially expressed in samples with −7/del(7q), RNA-seq was performed on 23 of the leukemia samples analyzed by SNP array. Eleven of the sequenced samples have loss of 7q22. Although sample A72 was reported to have −7[15%]/t(7)(p22q22)[75%] by cytogenetic analysis, we categorized it as having diploid 7q22 based on the SNP array (Figure 1). The percentage of blasts was similar in the 2 groups with 41.9% ± 30.1% and 49.7% ± 20.8% blasts in −7/del(7q) and other leukemias, respectively. Alignment statistics are provided in supplemental Table 4.

We performed an unbiased, genome-wide comparison of transcript levels in samples with −7/del(7q) versus other samples to define regions of haploinsufficient gene expression. Across all 20 139 expressed transcripts, 4 of the 5 most significantly differentially expressed transcripts with >1 fold change are located on chromosome band 7q22.1 and include CUX1 (Figure 2A). The RNA-seq dataset thereby provides an independent validation of the CDS defined by SNP-array (Figure 1). The 2.17-Mb CDS contains 49 RefSeq genes, 3 of which exhibited significantly decreased expression after Bonferroni correction (Figure 2B). CUX1 is the most significantly differentially expressed transcript within the CDS.

**CUX1 transcript levels in samples with a single CUX1 allele are present at half the level (46.3% ± 20.2%) of samples with 2 copies of the gene (Figure 3A)** consistent with a recent report. Transcript expression was confirmed by quantitative RT-PCR. Leukemia samples with loss of 7q22 express 45.8% ± 4.7% of the level of CUX1 expressed by other samples (Figure 3B). Loss of 1 copy of CUX1 is also associated with haploinsufficient protein levels in human AML cell lines (Figure 3C).

The **CUX1 gene is disrupted by a translocation resulting in a chimeric transcript**

As an orthogonal dataset to identify the putative TSG on chromosome 7, RNA-seq data were mined for chimeric fusion reads involving genes within the CDS. Nine high-confidence fusion events were identified: 4 validated by Sanger sequencing, 2 were false positives, and there was insufficient patient material to confirm the remaining (Table 1; supplemental Table 5). MLL-MLLT1 and USP42-RUNX1 fusions have been previously described in AML. Three expressed fusion events were identified in sample T20, 1 of which contains the first exon of CUX1 (7q22) upstream of exons 2–4 of CLND7 (chromosome band 17p13.1), in the same orientation (Figure 4). The resulting transcript is predicted to contain the codons for the first 21 amino acids of CUX1, followed by 8 amino acids in the wrong reading frame from CLND7, followed by a premature stop codon, thus truncating 98.6% of the CUX1 transcript. The second allele of CUX1 in T20 probably remains intact based on RNA-seq (data not shown) and RT-PCR.
The full-length p180 protein isoform of CUX1 is shown. The same blot was probed for and have AML cell lines were probed for CUX1 protein by Western blot. UoCM1, Mono7, and KG-1 mean within 1 experiment, A19 and A35 vs A56 and T52). (C) Protein lysates from human experiment of 3 independent experiments is shown. Samples with in 144 AML/t-MNs samples, 39 of which have other RNA-sequenced samples. Sanger sequencing of T55 PTMA CXCR4 2q37 2q21 Insufficient patient material No T55 USP42 RUNX1 7p22 21q22 Insufficient patient material Yes T20 ATP7B TRPC4 13q14 13q13 Yes Yes T20 GNAS PCDH9 20q13 13q21 Insufficient patient material No T19 KEAP1 ACER1 19p13 19p13 Yes Yes A44 MLL MLLT1 11q23 19p13 Yes Yes Sample Gene 1 Gene 2 Gene 1 chromosome Gene 2 chromosome PCR validated Open reading frame Table 1. Fusions identified in RNA-sequenced samples (Figure 4B). The level of CUX1 in T20 is 47.6% of the level in samples without loss of 7q22 as estimated by FPKM. deFuse did not identify CUX1 in fusion events in any of the other RNA-sequenced samples. Sanger sequencing of CUX1 exons in 144 AML/t-MNs samples, 39 of which have −7/del(7q), did not reveal any somatic mutations (data not shown). Similarly, there are no mutations in CUX1 reported in 200 AML samples by the Cancer Genome Atlas (tcga-data.nci.nih.gov, Level 2 DNA sequencing somatic mutations, Version 2.2.0). There is 1 reported case of CUX1 mutation acquired during leukemic transformation of a myeloproliferative neoplasm.69 Thus, CUX1 function is disrupted primarily by chromosome loss or deletion (Figure 1) and infrequently by a translocation event (Figure 4) or mutation.30

CUX1 is associated with a promitotic transcriptional gene signature in leukemia

CUX1 is a compelling candidate TSG: it is a homeodomain-containing transcription factor that regulates cell cycle progression and apoptosis and has been implicated in tumorigenesis of solid tumors.31 Consistent with a TSG, we uncovered a CUX1-associated cell cycle program within the RNA-sequencing data. Of the 662 genes significantly associated with CUX1 expression (FDR < 0.20, P < .0083), there is a substantial enrichment for the mitotic cell cycle pathway (DAVID, GO:0000278, Benjamini P = 2.57 × 10⁻³).32 CUX1-associated genes replicate in an independent microarray dataset of 38 subtypes of normal human hematopoietic cells.33 There is a significant overlap (P = .0063, permutation test) between the top 1000 CUX1-associated genes (FDR < 1.58 × 10⁻¹⁰) in normal hematopoietic cells and CUX1-associated genes in leukemia cells (2300 genes at P < .05). 73.8% of the 126 overlapping genes have a concordant direction (Pearson χ² test P = 6.24 × 10⁻¹²). CUX1-associated genes in normal cells are also enriched for cell cycle (GO:0007049, Benjamini P = .017). Thus, the CUX1-associated cell growth transcriptional profile is present in both leukemia and normal cells.

This transcriptional program is probably a result of direct regulation by CUX1, as many of these genes are bound by CUX1. We overlaid the CUX1-correlated leukemia gene set with previously reported CUX1 genome-wide promoter occupancy data from 8 human cancer cell lines, 5 of which are hematopoietic.34 There is a significant enrichment for direct targets of CUX1 (defined as binding with a P < .0005 in any cell type) within the leukemia CUX1-associated genes (P = .0016, permutation test). The 125 overlapping genes are also enriched for cell cycle genes (GO:0022403, Benjamini P = .05, supplemental Table 6). Nine of the 10 cell cycle genes in the overlapping gene set are inversely correlated with CUX1 expression, further implicating CUX1 as a negative regulator of hematopoietic proliferation.

The CUX1 ortholog, cut, is a conserved hematopoietic TSG in D melanogaster

The hematopoietic system in Drosophila is composed of hemocytes with functions similar to human myeloid cells.35 The development of hemocytes involves signaling molecules and transcription factors that are highly conserved in Drosophila and humans, such as GATA2, RUNX1, ZFPM1 (FOG-1), and NOTCH-WNT-family orthologs.35 Several genetic lesions in human leukemia result in hemocyte hyperproliferation in Drosophila, including gain-of-function mutations of RAS- and JAK-family genes, and expression of the human RUNX1-RUNX1T1 fusion protein.35 Most human TSGs have Drosophila orthologs with
evolutionarily conserved networks that regulate proliferation and differentiation; indeed, much of our understanding of TSGs comes from genetic screens in Drosophila. Ct is highly conserved between humans and Drosophila, and ectopic expression of human CUX1 can rescue ct-deficient phenotypes in Drosophila.

To test the hypothesis that ct is a TSG in vivo, ct, which is expressed in hemocyte progenitors, was targeted by RNAi specifically in developing Drosophila hemocytes. Expression of ctRNAi A in developing hemocytes led to the development of melanotic pseudotumors in both third instar larvae (Figure 5A) and pupae (data not shown). In a small number of cases, tumors were seen as early as the second instar (data not shown). A total of 23.8% (10 of 42) of larvae developed tumors, whereas none (0 of 30) of the control larvae did. To confirm that tumor formation was not the result of an off-target effect of the ct shRNA, we tested a second RNAi construct. Expression of ctRNAi B in hemocytes also led to tumor formation in 16.1% (10 of 62) of larvae.

Melanotic tumors are a phenotype of hemocyte overgrowth; however, we confirmed that there are increased numbers of circulating hemocytes. ctRNAi A and ctRNAi B knockdown larvae have significantly increased numbers of hemocytes compared with nonspecific shRNA-expressing larvae (Figure 5B). In addition, we tested a second hemocyte-specific GAL4-driver. Nine of 16 ct-GAL4/UAS-ctRNAi A larvae developed tumors and had significantly increased numbers of circulating hemocytes (33 737 ± 5719 SE in ct-GAL4/UAS-ctRNAi A, n = 7, vs 18 826 ± 1984 ct-GAL4 wild-type, n = 6, P = .042, data not shown). Residual expression of ct transcripts in hemocytes was 38.3% ± 3.8% (ctRNAi A, Figure 5C) and 28.9% ± 6.8% (ctRNAi B) of control, indicating that hemocyte overgrowth was a result of haploinsufficient levels of ct.

**CUX1** is a haploinsufficient TSG in human hematopoietic stem cells

To determine whether CUX1 is a conserved TSG in humans, we targeted CUX1 for knockdown in human hematopoietic progenitors. CUX1 is normally expressed highly in multipotent hematopoietic stem cells (Figure 6A). CUX1 expression decreases in common myeloid progenitors and varies significantly in terminally differentiated myeloid cells (Figure 6A).

We used 2 different CUX1-targeting shRNAs, shRNA A and shRNA B. K562 cells expressing shRNA A or shRNA B retain 24.1% ± 3.7% or 55.3% ± 10.9% residual CUX1 protein, respectively (Figure 6B). These shRNAs were transduced in a GFP-expressing lentiviral vector into lineage-depleted human cord blood progenitors and transplanted into NOD scid γ (NSG) mice. CUX1 shRNA-transduced progenitors retained 47.0%–66.9% residual CUX1 transcripts at the time of transplantation (Figure 6C). Transduction efficiency was similar in all groups (Figure 6D). Total engraftment of all human cells was not significantly different in CUX1-targeted recipient mice (Figure 6E). However, there was a significant expansion of the GFP+ population in the CUX1-knockdown xenografts (Figure 6E). Compared with nonspecific shRNA-transduced cells, there is a 40% increase in the percentage of CUX1-shRNA–transduced cells. All lineages appeared to be equally expanded in the peripheral blood (Figure 6F), suggesting that CUX1 TSG activity is not restricted to myeloid cells.

**Discussion**

Loss of part or all of chromosome 7 is a frequent event in myeloid malignancies and is associated with a poor prognosis, yet it is
unknown how these deletions contribute to leukemogenesis. We used high-density SNP arrays and transcriptome sequencing to identify a transcription factor gene, CUX1, as a haploinsufficient TSG within a CDS of chromosome 7. We also characterized a t-MNs sample with an inactivated CUX1 gene resulting from a translocation (Figure 4). CUX1 is normally highly expressed in multipotent hematopoietic progenitors and exhibits dynamic dosage changes during the course of differentiation (Figure 6A). Deletion of a single allele of CUX1 is associated with haploinsufficient transcript and protein levels (Figure 3). To test the TSG activity of CUX1, we first used *Drosophila* to demonstrate that haploinsufficient levels of the highly conserved *Drosophila* homolog, *ct*, leads to hemocyte overproliferation and melanotic tumor formation in developing larvae (Figure 5). In a second in vivo model, partial knockdown of *CUX1* in human hematopoietic progenitors leads to increased hematopoietic engraftment on transplantation into NSG mice (Figure 6). These findings indicate that CUX1 is a conserved, haploinsufficient TSG that is essential for the regulation of normal hematopoietic cell growth. A proliferative gene signature was identified within CUX1-associated genes, a finding that was replicated in an independent dataset of normal hematopoietic cells (data not shown). A significant number of these genes are direct targets of CUX1 binding. This suggests that CUX1 may act as a TSG in myeloid cells by regulating cell cycle components at the transcriptional level.

Deletions of chromosome 7 are typically large with heterogeneity in the breakpoints in myeloid diseases, making it difficult to map the CDS. It is probable that multiple TSGs on chromosome 7 cooperate in leukemogenesis. There may be at least 3 different CDS on chromosome 7, including an interval at 7q36 containing the *EZH2* gene. A precedent for this is seen in 5q deletions, which are also heterogeneous.

We have not excluded the possibility that other differentially expressed genes within the CDS of 7q22.1 may also play a role in disease pathogenesis. However, CUX1, which was implicated as early as 1996, was identified in a narrower CDS in 2 recent studies of AML arising from myeloproliferative neoplasms. Four cryptic deletions were identified: 3 samples had deletions that spanned just the CUX1 and *SH2B2* genes, and 1 sample had a deletion containing only CUX1. The CDS defined in the current study is larger, most probably because of smaller sample size.

We note that biallelic inactivation of CUX1 appears to be infrequent (Figure 1). The remaining CUX1 allele is unlikely to be silenced by epigenetic modifications, as samples with −7/del(7q) express CUX1 at 46.3% ± 20.2% of the level of samples with both copies of CUX1 (Figure 3), consistent with a gene dosage effect and expression from the remaining allele. However, rare instances of homozygous deletions of the CUX1 locus and a CUX1 homeodomain mutation in a region of acquired uniparental disomy have been reported in AML arising from myeloproliferative neoplasms. Mutations in CUX1 were not demonstrated by Sanger sequencing of CUX1 exons in 144 AML/t-MNs samples, 39 with −7/del(7q) (data not shown). A previous study also failed to find CUX1 mutations in pediatric myeloid neoplasias with −7. CUX1 is therefore most often inactivated by chromosome loss or deletion but can infrequently be inactivated by mutation or translocation (Figure 4). As biallelic inactivation of CUX1 occurs only rarely in myeloid diseases, CUX1 appears to be a haploinsufficient TSG.

Studies in murine models support the hypothesis that CUX1 is a myeloid TSG in mammals. Cux1−/−/− mice, which have a Cux1 truncation, have increased myeloid cell numbers and myeloid colony-forming units. Transgenic expression of the p75 isoform of CUX1, containing 1 Cut repeat and the homeodomain, leads to a myeloproliferative disease in 33% of mice. It is not clear whether the myeloid expansion is the result of increased Cux1 activity or a dominant negative effect of the transgene. CUX1 is essential for lineage specification in multiple tissues. Consistent with a developmental role in several cell types, altered Cux1 levels led to abnormal growth of multiple cell types in murine models. In humans, CUX1 mutations have been reported in 6.7% of lung squamous cell carcinoma and 2.8% of colorectal carcinoma; however, the functional significance of these mutations remains to be tested. The TSG activity of CUX1 may have several mechanisms, including regulation of differentiation, cell cycle progression, genomic stability, and apoptosis, and the mechanism may be cell type specific. In one report, deletion of a *Drosophila* cancer model promoted invasive tumorigenesis because of both a block in differentiation and increased proliferation.

In conclusion, this is the first biologic confirmation of a haploinsufficient myeloid TSG on chromosome band 7q22. The concentrations of hematopoietic transcription factors play a critical role in hematopoietic cell fate, regulating lineage decisions, proliferative signals, and leading to leukemia when concentrations are above or below crucial thresholds. Examples of the detrimental consequences of abnormal gene product dosage can be found in many master regulators of hematopoiesis, including *SPI1, RUNX1*, and *CBFB*. Interestingly, these genes are typically not completely inactivated when mutated in leukemia but instead retain some level of activity. **CUX1** levels during normal human hematopoiesis
appear to change significantly, depending on the developmental stage and lineage (Figure 6A). A gradient of Ct has been shown to be important in neural development in *Drosophila*: high, medium, and low levels of Ct, measured at the single-cell level, lead to distinct classes of neural cells. Thus, changes in CUX1 levels may have important phenotypes during hematopoiesis and deleterious consequences when altered.

### Acknowledgments

The authors thank Kevin Shannon for critical reading of the manuscript, Ronald Hause for assistance with statistical analysis, and the TReP at Harvard Medical School for transgenic RNAi fly stocks. Next-generation sequencing was performed at the University of Chicago High-throughput Genome Analysis Core. Illumina SNP array services were performed at the Northwestern University Genomics Core. Sanger sequencing was performed at the University of Chicago Comprehensive Cancer Center Genomics Core.

This work was supported by a Leukemia & Lymphoma Society Fellow award (M.E.M), the Cancer Research Foundation, National Institutes of Health (CA40046; M.M.L.), and the Chicago Cancer Genomes Project.

### Authorship

**Contribution:** M.E.M. designed research, performed experiments, analyzed and interpreted data, and wrote the manuscript; C.D.B. designed research, analyzed, and interpreted RNA-sequence data, and edited the manuscript; X.W. performed *Drosophila* crosses and hemocyte counting; E.T.B. assisted in analyzing RNA-sequence data; S.K. assisted in generating sequencing libraries; C.B. identified fusion genes; S.Y. assisted in analyzing SNP data and fusion validation; B.P.S., J.K., and J.M.C. designed and performed xenograft experiments; T.S. assisted in designing research; J.A. performed morphologic analysis and collected biospecimens; M.M.L. performed cytogenetic analysis of leukemia...
samples, designed research, collected biospecimens, interpreted data, and edited the manuscript; R.L.G. provided computational infrastructure and bioinformatics support; and K.P.W. designed research, interpreted data, and edited the manuscript.

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

References