Cancer Cell
Article

Integrated Transcript and Genome Analyses Reveal *NKX2-1* and *MEF2C* as Potential Oncogenes in T Cell Acute Lymphoblastic Leukemia

Irene Homminga,[1] Rob Pieters,[1] Anton W. Langerak,[2] Johan J. de Rooi,[3,4] Andrew Stubbs,[3] Monique Verstegen,[1] Maartje Vuerhard,[1] Jessica Buijs-Gladdines,[1] Clarissa Kooi,[1] Petra Klous,[5,6] Pieter van Vlierberghe,[7] Adolfo A. Ferrando,[8] Jean Michel Cayuela,[9] Brenda Verhaaf,[2] H. Berna Beverloo,[10] Martin Horstmann,[11,12,13] Valerie de Haas,[14] Anna-Sophia Wiekmeijer,[15] Karin Pike-Overzet,[15] Frank J.T. Staal,[15] Wouter de Laat,[5,6] Jean Soulier,[9] Francois Sigaux,[9,16] and Jules P.P. Meijerink[1,16,*]

[1]Department of Pediatric Oncology/Hematology, Erasmus MC/Sophia Children's Hospital, Rotterdam 3015GJ, The Netherlands
[2]Department of Immunology
[3]Department of Bioinformatics
[4]Department of Biostatistics
[5]Department of Cell Biology
Erasmus MC, Rotterdam 3015GE, The Netherlands
[6]Hubrecht Institute-KNAW & University Medical Center Utrecht, Utrecht 3584CT, The Netherlands
[7]Center for Medical Genetics, Ghent University Hospital, 9000 Ghent, Belgium
[8]Institute for Cancer Genetics, Columbia University, New York, NY 10032, USA
[9]INSERM U944, Institut Universitaire d'Hématologie Université Paris 7, Hôpital Saint Louis, 75010 Paris, France
[10]Department of Clinical Genetics, Erasmus MC, Rotterdam 3015GE, The Netherlands
[11]Research Institute Children's Cancer Center, 20251 Hamburg, Germany
[12]Clinic of Pediatric Hematology and Oncology, University Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany
[13]Co-operative Study Group for Childhood Acute Lymphoblastic Leukemia (COALL), 20246 Hamburg, Germany
[14]Dutch Childhood Oncology Group (DCOG), The Hague 2545CJ, The Netherlands
[15]Department of Immunohematology and Blood Transfusion, Leiden University Medical Center, Leiden 2333ZA, The Netherlands
[16]These authors contributed equally to this work
*Correspondence: j.meijerink@erasmusmc.nl
DOI 10.1016/j.ccr.2011.02.008

SUMMARY

To identify oncogenic pathways in T cell acute lymphoblastic leukemia (T-ALL), we combined expression profiling of 117 pediatric patient samples and detailed molecular-cytogenetic analyses including the Chromosome Conformation Capture on Chip (4C) method. Two T-ALL subtypes were identified that lacked rearrangements of known oncogenes. One subtype associated with cortical arrest, expression of cell cycle genes, and ectopic *NKX2-1* or *NKX2-2* expression for which rearrangements were identified. The second subtype associated with immature T cell development and high expression of the *MEF2C* transcription factor as consequence of rearrangements of *MEF2C*, transcription factors that target *MEF2C*, or MEF2C-associated cofactors. We propose *NKX2-1*, *NKX2-2*, and *MEF2C* as T-ALL oncogenes that are activated by various rearrangements.

INTRODUCTION

T-lineage acute lymphoblastic leukemia (T-ALL) is a malignancy of thymocytes. T-ALL represents about 15% of pediatric ALL cases but has an inferior outcome compared to B-ALL because approximately 30% of T-ALL cases relapse during therapy or within the first 2 years following treatment and eventually die (Pieters and Carroll, 2008; Pui and Evans, 2006). T-ALL is mostly

Significance

For 40% of pediatric T-ALL cases, underlying oncogenic rearrangements remain unresolved. By combined expression profiling and molecular-cytogenetic techniques, we revealed two T-ALL entities lacking known oncogenic rearrangements and representing ~20% of pediatric T-ALL cases. One subtype associated with cortical thymocytic arrest, and ten out of 12 cases ectopically expressed *NKX2-1/NKX2-2* for which five rearrangement variants were identified in seven cases. The second subtype was associated with high *MEF2C* expression (11 out of 12 cases), and rearrangements involving *MEF2C* or transcription factors and transcription cofactors that directly target *MEF2C* were identified in six cases. Ectopic expression of NKX2-1 or MEF2C was able to transform cells and interfered with T cell differentiation. We propose that *NKX2-1*, *NKX2-2*, and *MEF2C* are oncogenes in leukemia.

characterized by genetic abnormalities that are crucial for T cell pathogenesis (Van Vlierberghe et al., 2008a). Various genetic rearrangements in T-ALL occur in a mutually exclusive pattern (Van Vlierberghe et al., 2008a) in contrast to frequent *CDKN2A/ARF* deletions (Hebert et al., 1994) or *NOTCH1*-activating mutations (Weng et al., 2004). These mutually exclusive rearrangements are considered as driving chromosomal abnormalities that affect the *TAL1*, *LMO2*, *TLX1*, *TLX3*, *MYB*, or *HOXA* oncogenes (Van Vlierberghe et al., 2008a). Based on gene expression data (Ferrando et al., 2002; Soulier et al., 2005; Van Vlierberghe et al., 2008b), these oncogenes have been associated with distinct T-ALL subgroups denoted as the *TAL/LMO*, *TLX1*, *TLX3*, and the *HOXA* subgroups. Initial profiling data also pointed to the existence of an additional immature T-ALL subgroup (Soulier et al., 2005). This entity probably corresponds to the *LYL1* T-ALL subgroup as previously defined (Ferrando et al., 2002) and to the recently described immature T-ALL subset that is characterized by an early T cell precursor (ETP) profile and inferior outcome (Coustan-Smith et al., 2009). For approximately 40% of all patients with T-ALL, including the immature T-ALL entity, the driving chromosomal aberrations have thus far remained elusive.

RESULTS

Cluster Analyses Predict T-ALL Genetic Subgroups

To identify driving oncogenic mechanisms in T-ALL, we performed unsupervised hierarchical cluster analyses based on microarray expression data of 117 diagnostic pediatric T-ALL samples and seven normal bone marrow controls. A total of 77 T-ALL samples was characterized by oncogenic rearrangements, including *TAL1* (n = 24), *TAL2* (n = 1), *LMO1* (n = 1), *LMO1/TAL2* (n = 1), *LMO2* (n = 9), *TLX3* (n = 22), *TLX1* (n = 7), *HOXA*-activating rearrangements (including *CALM-AF10*, Inv(7)(p15q34), *SET-NUP214*; n = 10), or *MYB* translocations (n = 2). No such abnormalities were identified in the remaining 40 T-ALL patient samples. Four robust T-ALL clusters were observed in unsupervised cluster analysis, regardless of the number of genes included or the data normalization methods chosen (Figure 1A; see Figure S1 and Tables S1–S3 available online). The association with clinical and molecular-cytogenetic data, immunophenotypic markers, and expression of *TAL1* and *LYL1* for these four subgroups is given in Table 1 and Figure 1A. Two clusters represented established T-ALL genetic subgroups (Ferrando et al., 2002; Soulier et al., 2005; Van Vlierberghe et al., 2008b), corresponding to abnormalities of *TAL1/LMO2*, and *TLX3/HOXA* transcription factors.

A third cluster included cases that highly expressed *CD1* genes. This corresponded with a CD1a-positive immunophenotype for most cases of this cluster ($p < 0.001$; Table 1), which validated our gene expression data. This cluster also comprised most *TLX1*-translocated cases, a genetic entity that was previously associated with CD1 positivity and cortical developmental arrest (Ferrando et al., 2002), and that may share a similar biology with the other samples present in this cluster. In the unsupervised cluster analysis, this cluster is characterized by expression of genes that are involved in cell cycle regulation (*CDKN3*), G1/S transition (*UHRF1*, *CDC2*), cell cycle progression (*TTK*, *E2F7*, *CDC2*), DNA replication and chromosome condensation (*TOP2A*), the spindle-assembly checkpoint (*NUSAP1*, *MAD2L1*, *KIF15*, *KIF11*), the G2/M checkpoint (*PBK*), and genes whose expression are linked to cell cycle (*RRM2*, *ECT2*). Furthermore, differentially expressed genes for this cluster compared to all other T-ALL cases as identified by t-statistics were enriched for genes that are strongly associated with the cell cycle pathway and spindle assembly (Table S4), and this cluster strongly expressed the proliferation marker *MKI67*. This cluster was accordingly denoted as ''proliferative cluster.'' Most of the cases in this cluster lacked currently known driving mutations, which may point toward involvement of not yet identified T-ALL oncogenes. This was further supported by the fact that most of these unknown samples clustered as a separate entity (12 cases) distinct from established T-ALL genetic subgroups, including the *TLX1*-rearranged cases in a supervised cluster analysis (Figure 1C).

The fourth cluster was enriched for immunophenotypic immature CD4/CD8 double-negative cases ($p = 0.008$; Table 1), and was named the ''immature cluster'' by reference to previous work (Soulier et al., 2005). Samples in this cluster frequently expressed myeloid markers CD13 and/or CD33 ($p = 0.006$), and were characterized by expression of genes associated with protein binding, protein dimerization, and TGFBR1-signal transduction. They expressed low levels of genes associated with cellular proliferation contrary to samples of the proliferative cluster (Figure 1B). This cluster comprised three *HOXA*-activated cases with an immature immunophenotype unlike other *HOXA*-activated cases that usually have a more advanced immunophenotype. Other samples in this immature cluster were devoid of known driving mutations. This cluster may comprise a second molecular-cytogenetic T-ALL entity for which driving oncogenes are unknown, and in support of this notion, most of these samples appeared as a separate subgroup (12 cases) in the supervised principal component analysis (PCA) based on differentially expressed genes among the known four T-ALL genetic subgroups (Figure 1C). This immature cluster largely overlaps with the *LYL1*-positive cluster as described earlier (Ferrando et al., 2002) because it expressed the highest *LYL1* levels (Table 1). Our immature cluster was highly enriched for ETP T-ALL cases, as previously described (Coustan-Smith et al., 2009), because 13 out of 15 immature cases, in contrast to only three out of 102 remaining cases, were predicted as ETPs by prediction analysis for microarrays (PAM) using the 62 probe set profile that defined the ETP group ($p < 0.001$; data not shown). In contrast to that study (Coustan-Smith et al., 2009), the overall survival (OS) for immature cases in our cohort was not extremely poor (5-year OS = 73% ± 11%) but seemed equally low to the outcome of *TAL/LMO* or *TLX* subgroups (5-year OS = 65% ± 6%). The proliferative subgroup seemed to have an improved outcome (5-year OS = 88% ± 8%), albeit not significant ($p = 0.096$; Figure S2).

We then searched for candidate genes that participate in oncogenic chromosomal abnormalities using several methods, including COPA (Tomlins et al., 2005), SAM (Tusher et al., 2001), and PAM statistics (Tibshirani et al., 2002). Both COPA and PAM analyses identified *NKX2-1* and *MEF2C* as characteristic genes for the proliferative and immature clusters, respectively (Table S5). The *NKX2-1* homologous *NKX2-2* gene was also identified by COPA as outlier gene for the proliferative

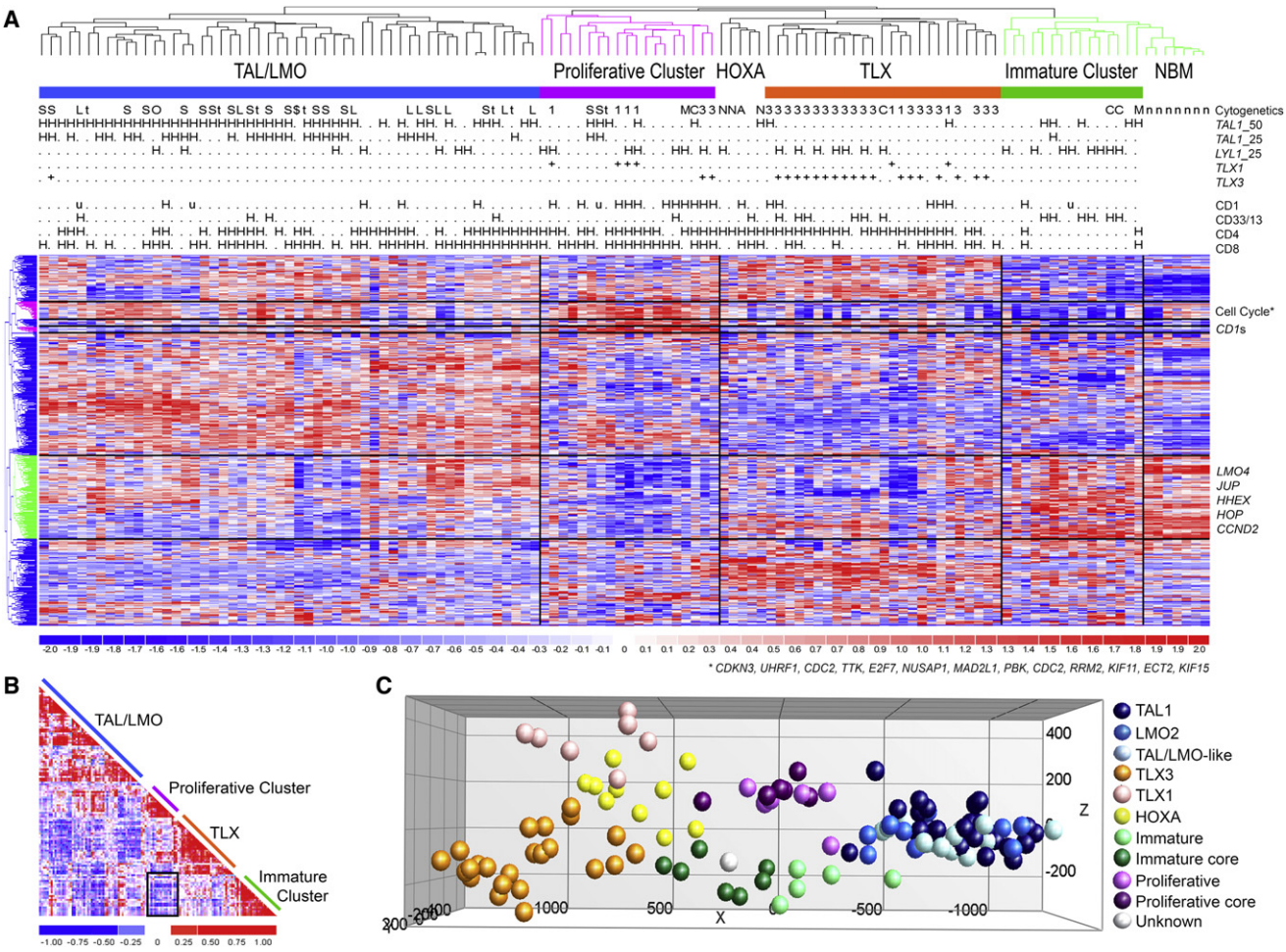


Figure 1. Identification of Two Entities in Pediatric T-ALL that Lack Known Driving Oncogenic Hits

(A) Unsupervised hierarchical cluster analysis by the average linkage method in dChip based on 435 probe sets (Table S3) for RMA-solo (Soulier et al., 2005) normalized U133 plus two Affymetrix data from 117 pediatric T-ALL samples and seven normal bone marrow controls. Cytogenetic rearrangements indicated are: S, *SIL-TAL1*; T, *TAL1*; t, *TAL2*; O, *LMO1*; L, *LMO2* (includes del(11)(p12p13)); $, *TAL2/LMO1*; N, *SET-NUP214*; C, *CALM-AF10*; M, *MYB*; A, Inv(7)(p15q34); 1, *TLX1*; 3, *TLX3*; and n, normal bone marrow controls. The 50th and/or the 25th percentiles of samples with the highest *TAL1* or *LYL1* expression, positivity for *TLX1* and *TLX3* expression as measured by RQ-PCR, and expression of the immunophenotypic markers CD13 and/or CD33, CD4 or CD8 are indicated; u, no data available.

(B) Pearson correlation plot for the patient samples belonging to the four unsupervised *TAL/LMO*, *TLX*, proliferative, and immature clusters.

(C) PCA of patients with pediatric T-ALL based upon the top 100 most significant differentially expressed probe sets among major T-ALL subgroups (i.e., *TAL1/LMO2*, *HOXA*, *TLX1*, and *TLX3* [Table S3]). The immature cluster (12 cases) and the proliferative cluster (12 cases) are indicated by green and purple dots, respectively. Samples repeatedly assigned to the proliferative or immature clusters (i.e., the core samples) in multiple unsupervised analyses on RMA-solo (A), RMA, or VSN normalized data sets (not shown) or the supervised cluster analysis (C) are visualized by dark-green or purple dots. See also Figure S1 and Tables S1–S4.

cluster. High microarray expression levels of *NKX2-1* and *MEF2C* were validated by RQ-PCR (Figure 2) for the proliferative and immature cluster cases, respectively, that lack known oncogenic rearrangements. These cases form separate clusters in the supervised analysis (Figure 1C). *NKX2-1* or *MEF2C* was either absent or expressed at relative low levels in most cases belonging to other supervised clusters. However, some *TLX1*-positive patient samples that are part of the proliferative cluster in the unsupervised analyses express *NKX2-1*. Also, the *CALM-AF10*-positive *HOXA*-activated patient sample No. 1509 that highly expresses *MEF2C* has an immature phenotype and co-clusters in the immature cluster in unsupervised analyses.

Molecular-Cytogenetic Identification of *NKX2-1* Rearrangements

These data formed the start of detailed molecular-cytogenetic analyses on the 12 immature cluster and the 12 proliferative cluster samples that seemed to form two genetic T-ALL entities (Figure 1C), and for which driving oncogenic hits were unknown. We used a variety of molecular-cytogenetic techniques including FISH, array-comparative genomic hybridization (array-CGH), and Chromosome Conformation Capture on Chip (4C) (Simonis et al., 2009) to identify potential deletions, amplifications, and T cell receptor- or *BCL11B*-driven oncogenic events (Table 2; Table S6). The 4C method was originally developed to study

Table 1. Clinical and Biological Characteristics of Unsupervised T-ALL Clusters

	Cohort		TAL/LMO		TLX		Proliferative		Immature		p Value
Total (n)	117		53		30		19		15		
Clinical											
Gender (n)											
Male	83	71%	40	75%	20	67%	15	79%	8	53%	
Female	34	29%	13	25%	10	33%	4	21%	7	47%	0.306[a]
Age at diagnosis (years)											
Median	7.8		9.3		7.7		5.5		10.1		
Range	1.5–17.8		1.6–16.7		3.2–17.8		1.5–16.7		3.1–16.4		0.404[b]
WBC (10×10^{E9}/l)											
Median	115.1		156.9		121.9		64.3		87.6		
Range	1.8–900		16.1–900		1.8–417		27.2–192		2.3–435		0.001[b]
Immunophenotype											
CD34 (n)	111										
Negative	77	69%	37	71%	19	66%	14	93%	7	47%	
Positive	34	31%	15	29%	10	34%	1	7%	8	53%	0.046[a]
CD13/33 (n)	110										
Negative	92	84%	47	92%	19	73%	17	94%	9	60%	
Positive	18	16%	4	8%	7	27%	1	6%	6	40%	0.006[a]
CD1 (n)	113										
Negative	62	55%	33	65%	14	47%	3	17%	12	86%	
Positive	51	45%	18	35%	16	53%	15	83%	2	14%	<0.001[a]
CD4 (n)	115										
Negative	42	37%	22	42%	4	13%	3	17%	13	87%	
Positive	73	63%	30	58%	26	87%	15	83%	2	13%	<0.001[a]
CD8 (n)	115										
Negative	45	39%	14	27%	16	53%	2	11%	13	87%	
Positive	70	61%	38	73%	14	47%	16	89%	2	13%	<0.001[a]
CD4/8 (n)	115										
Negative	61	53%	26	50%	17	57%	5	28%	13	87%	
Positive	54	47%	26	50%	13	43%	13	72%	2	13%	0.008[a]
CD3 (n)	114										
Negative	59	52%	21	40%	19	63%	10	59%	9	60%	
Positive	55	48%	31	60%	11	37%	7	41%	6	40%	0.169[a]
Oncogenes											
TAL1 (% expression of GAPDH $\times 10^{E-2}$)											
Median	3.1		13		0.73		1.4		1.14		
Range	0.09–1820		0.75–1820		0.088–11		0.17–22		0.10–14		<0.001[b]
LYL1 (% expression of GAPDH $\times 10^{E-4}$)											
Median	1.7		1.3		3.1		3.5		8.5		
Range	0–126		0–32		0–16.6		0.28–15.3		0.96–126		0.001[b]

[a] The p values are calculated according to the chi-square test. See also Figure S2.
[b] The p values are calculated according to the Mann-Whitney U test. See also Figure S2.

the three-dimensional structure of DNA (Simonis et al., 2006), but it was recently shown that it robustly identifies chromosomal rearrangements, in particular inversions and translocations, even when they are balanced (Simonis et al., 2009). In the proliferative cluster, two out of 12 samples were characterized by *MYB* translocations, a rearrangement considered as a driving oncogenic hit (Clappier et al., 2007). No further *MYB* translocations were identified in the remaining ten cases by FISH (Table S6). We identified five rearrangements of *NKX2-1* or *NKX2-2* genes in seven out of 12 patient samples that, to our knowledge, were not observed before in human cancer (Table 2, and Figures 3A–3E; Figure S3). The *NKX2-1* gene was inverted to the T cell receptor gene *TRA@* in two cases (Nos. 1446 and 9247), inverted to the immunoglobulin heavy-chain gene *IGH@* in one case (No. 9919), and translocated to the *TRB@* locus (t(7;14)(q34;q13)) in one other case (No. 9989), as identified by 4C analyses (Figure 3A). *NKX2-1*

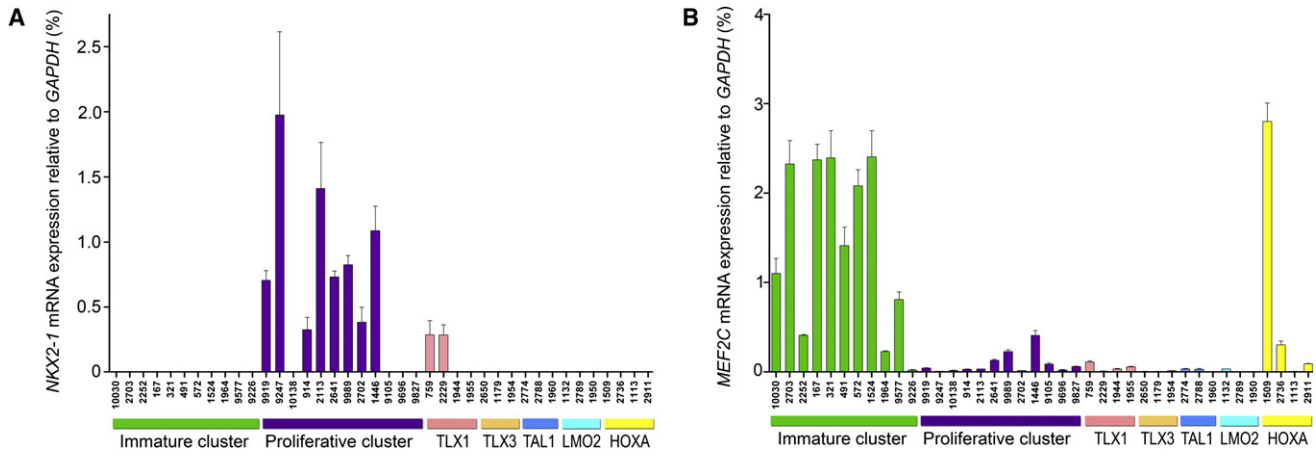


Figure 2. Validation of Elevated *NKX2-1* and *MEF2C* Levels in Patients from Proliferative and Immature Supervised Clusters

Relative expression levels of (A) *NKX2-1* or (B) *MEF2C* are determined by RQ-PCR. *NKX2-1* and *MEF2C* expression levels are indicated for 11 out of 12 immature cluster patient samples (green) and 12 proliferative cluster samples (purple) according to the supervised analysis (Figure 1C) compared to cases of other T-ALL molecular-cytogenetic subgroups. The SEMs are shown. See also Table S5.

rearrangements in these patients were validated by FISH (Figure 3B). The der(7) chromosomal breakpoint for this t(7;14)(q34;q13) in patient No. 9919 was cloned (Figure 3C). A fifth patient (No. 2641) had a *NKX2-1* rearrangement based on FISH results (Table 2; data not shown), whereas a sixth patient (No. 2702) had an amplification at 14q13 based on array-CGH (Figure 3D), presumable due to a *NKX2-1* duplication or an insertion into another chromosome (data not shown). These patients highly expressed NKX2-1 protein levels (Figure 3F, representative cases are shown). A seventh case (No. 10138) had a translocation between the homologous *NKX2-2* gene and the *TRD@* locus, for which both reciprocal breakpoint regions were cloned (Figure 3E). This patient highly expressed NKX2-2 protein levels (data not shown). For the *TLX1*-rearranged cases that co-cluster with these *NKX2-1/NKX2-2* rearranged cases in unsupervised cluster analysis that also expressed *NKX2-1* (Figure 2A), we did not find evidence for *NKX2-1* rearrangements by FISH (data not shown). This indicates that *TLX1* and *NKX2-1/NKX2-2* oncogenes may exert identical or closely related pathogenic mechanisms.

Molecular-Cytogenetic Identification of *MEF2C* and *MEF2C*-Activating Rearrangements

We subsequently investigated the 12 immature cluster cases lacking known driving oncogenic hits, and identified chromosomal abnormalities that converge on the activation of the *MEF2C* gene in at least five cases. Two cases had chromosomal copy number loss of the 5q14-qter chromosomal arm with breakpoint in a 0.5–2 Mb proximity telomeric of *MEF2C*. A similar deletion was also identified in T-ALL cell line LOUCY (Figure 4A). These 5q14-qter deletions were not identified in 90 other T-ALL cases as included in our profiling study for which array-CGH data were available (Figure S4A). For patient No. 1964, this 5q14-qter deletion was part of an unbalanced chromosomal translocation between chromosomal bands 5q14 and 4q27 fusing the telomeric *MEF2C* region to the telomeric region ~0.6 Mb distal of the *PITX2* gene on chromosome 4 (Figure 4B). In contrast to other genes in the 5q14 region, *MEF2C* is highly upregulated in both patients indicating that *MEF2C* represents the target of these 5q rearrangements (Figures S4B ad S4C).

A *NKX2-5/BCL11B* translocation was identified by FISH in a third case (Figure 4C), and this case highly expressed *NKX2-5*. This rare translocation has been reported in T-ALL before (Nagel et al., 2003). Knockdown of NKX2-5 levels by siRNA molecules in the *NKX2-5* translocation-positive cell line PEER lowered *MEF2C* levels (Figures 5A–5C), indicating that NKX2-5 controls *MEF2C*. Chromatin immunoprecipitation (ChIP) experiments confirmed that NKX2-5 directly binds in the promoter region of *MEF2C* (Figure 5D).

A fourth case harbored a *BCL11B* translocation to *SPI1*, which encodes for PU.1 (Figure 4D). This patient uniquely expressed *SPI1* compared to the other T-ALL cases in this study (Figure 4E). PU.1 was recently identified as important regulator for *MEF2C* expression in normal lymphoid development (Stehling-Sun et al., 2009), and this patient with T-ALL highly expressed *MEF2C* (Figure 2B and Table 2).

A fifth case (No. 1524) harbored a t(8;12)(q13;p13) as identified by FISH (Figures S4D and S4E), resulting in reciprocal *ETV6-NCOA2* fusion products, and both reciprocal breakpoints were cloned for this patient (Figure 4F). Similar fusions were recently identified in biphenotypic T-ALL (Strehl et al., 2008). NCOA2 is a known coregulator of MEF2C (Chen et al., 2000), and *MEF2C* was found consistently upregulated in selected *ETV6–NCOA2* rearranged cases (Figure 4G).

A sixth immature case with high *MEF2C* levels had a karyotypic t(2;21) that involved the *RUNX1/AML1* gene (Figure S4F). For this patient we cloned reciprocal in-frame *RUNX1-AFF3* and *AFF3-RUNX1* fusion products as consequence of this translocation (Figure 4H). How RUNX1 fusion products could upregulate *MEF2C* expression remains to be determined.

To investigate whether MEF2C could indeed regulate the expression of various genes from the immature signature, *MEF2C* stable-transfected clones and mock-transfected controls were generated for the cell line Jurkat (Figure 5E) that does not have an immature signature (data not shown). As shown in Figure 5F, the *MEF2C*-transfected Jurkat clone 2B3,

Table 2. Identified Rearrangements in Patient Samples of the Proliferative and Immature Clusters

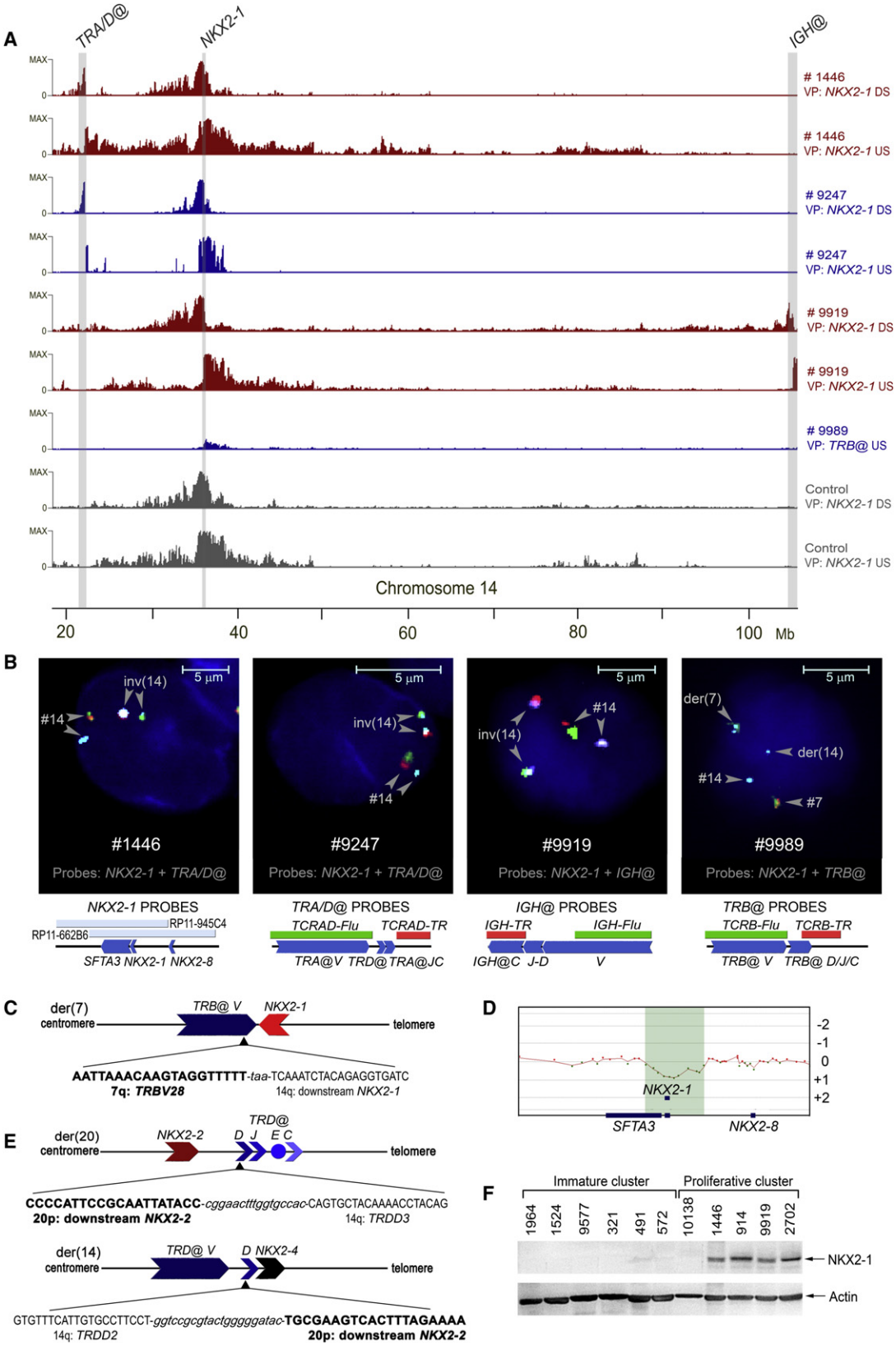
Proliferative Cluster					
Patient Number	*NKX2-1* Expression[a]	Aberration	*Partner Gene 1*	*Partner Gene 2*	Methods
9919[b]	+	inv(14)(q13q32.33)	*IGH@*	*NKX2-1*	FISH, 4C
9247[b]	+	inv(14)(q11.2q13)	*TRA@*	*NKX2-1*	FISH, 4C
10138[b]	+[c]	t(14;20)(q11;p11)	*TRD@*	*NKX2-2*	FISH, LM-PCR
914	+	t(6;7)(q22-23;q34)	*TRB@*	*MYB*	FISH
2113	+	–	–	–	–
2641	+	Rearrangement	*?*	*NKX2-1*	FISH
9989[b]	+	t(7;14)(q34;q13)	*TRB@*	*NKX2-1*	FISH, 4C
2702[b]	+	dup(14)(q13.3q13.3) or ins(?)(?q13.3)	*?*	*NKX2-1*	Array-CGH, FISH
1446[b]	+	inv(14)(q11.2q13)	*TRA@*	*NKX2-1*	FISH, 4C
9105	+	t(6;7)(q22-23;q34)	*TRB@*	*MYB*	FISH
9696		–	–	–	–
9827		–	–	–	–
Immature Cluster					
Patient Number	*MEF2C* Expression[a]	Aberration	*Partner Gene 1*	*Partner Gene 2*	Methods
10030[d]	+	–	–	–	–
2703	+	–	–	–	–
2130		–	–	–	–
2252	+	t(11;14)(p11.2;q32.2)	*BCL11B*	*SPI.1*	FISH, 4C
167[d]	+	–	–	–	–
321[d]	+	–	–	–	–
491[d]	+	del(5)(q14)	–	*MEF2C*	FISH, array-CGH
572[d]	+	t(2;21)(q11.2-12;q22.3)	*RUNX1*	*AFF3*	Karyotype, 3′-RACE
1524[d]	+	t(8;12)(q13;p13)	*ETV6*	*NCOA2*	RT-PCR, FISH
1964[d]	+	der(5)t(4;5)(q26;q14)	4q26	*MEF2C*	4C, array-CGH
9577	+	t(5;14)(q34;q32.2)	*BCL11B*	*NKX2-5*	FISH
9226	±	–	–	–	–
Cell Lines					
LOUCY	+	t(5;14)(q34;q32.2)	*BCL11B*	*NKX2-5*	(Przybylski et al., 2006)
PEER	+	del(5)(q14)	–	*MEF2C*	(Nagel et al., 2008)

[a] *NKX2-1* or *MEF2C* expression based on expression array and/or RQ-PCR results.
[b] Core proliferative cases repeatedly assigned in unsupervised and supervised analyses to the immature or proliferative clusters, respectively. See also Table S6 and Figure S6.
[c] Sample No. 10138 expresses the *NKX2-1* homologous *NKX2-2* gene.
[d] Core immature cases repeatedly assigned in unsupervised and supervised analyses to the immature or proliferative clusters, respectively. See also Table S6 and Figure S6.

but not the mock-transfected control 3G4, highly activates five out of six selected immature signature genes (*PSCD4*, *HHEX*, *FAM46A*, *LMO2*, and *LYL1*), indicating that MEF2C may function as a transcriptional regulator for many genes that are highly expressed in immature T-ALL cases. For the reciprocal setting in cell line PEER, knockdown of NKX2-5 using siRNA molecules that reduced *MEF2C* expression (Figures 5A–5C) also led to reduced levels of *LMO2*, *LYL1*, and *HHEX* (Figure 5G). Oncogenic rearrangements of *LMO2*, *LYL1*, and the *LYL1* homologous *TAL1* gene are exclusively found in the *TAL/LMO* subgroup but have never been observed in immature T-ALL cases (this work; Ferrando et al., 2002). Activation of *LMO2* and *LYL1* through MEF2C may be crucial to prime early-committed T cells for leukemogenesis. By using ChIP we demonstrated that MEF2C directly binds to the promoter of *HHEX* as well as to the distal and proximal promoters of *LMO2* in the immature cell line LOUCY. This could also be demonstrated for diagnostic leukemic cells of three patients that belong to the immature cluster (Nos. 491, 321, and 167; data not shown), but not in the control cell line Jurkat (Figure 5H). The *MN1* gene, which is targeted by chromosomal alterations in inv(16) M4EO AML subtype (Buijs et al., 2000; Grosveld, 2007), was also identified as a highly activated gene for the immature cluster (Table S5). As for *HHEX*, we did not find evidence for chromosomal rearrangements of *MN1* by FISH in immature T-ALL cases (Table S6).

Oncogenic Activity of *NKX2-1* and *MEF2C*

To substantiate potential oncogenic activity for *NKX2-1* and *MEF2C*, we tested whether both genes had transforming


A
TRA/D@
NKX2-1
IGH@
1446 VP: NKX2-1 DS
1446 VP: NKX2-1 US
9247 VP: NKX2-1 DS
9247 VP: NKX2-1 US
9919 VP: NKX2-1 DS
9919 VP: NKX2-1 US
9989 VP: TRB@ US
Control VP: NKX2-1 DS
Control VP: NKX2-1 US
Chromosome 14
20
40
60
80
100 Mb
B
inv(14)
#14
#1446
Probes: NKX2-1 + TRA/D@
#9247
#9919
Probes: NKX2-1 + IGH@
der(7)
der(14)
#7
#9989
Probes: NKX2-1 + TRB@
NKX2-1 PROBES
RP11-662B6
RP11-945C4
SFTA3 NKX2-1 NKX2-8
TRA/D@ PROBES
TCRAD-Flu
TCRAD-TR
TRA@V TRD@ TRA@JC
IGH@ PROBES
IGH-TR
IGH-Flu
IGH@C J-D V
TRB@ PROBES
TCRB-Flu
TCRB-TR
TRB@ V TRB@ D/J/C
C
der(7)
centromere
TRB@ V
NKX2-1
telomere
AATTAAACAAGTAGGTTTTT-taa-TCAAATCTACAGAGGTGATC
7q: TRBV28
14q: downstream NKX2-1
D
-2
-1
0
+1
+2
NKX2-1
SFTA3
NKX2-8
E
der(20)
centromere
NKX2-2
TRD@
D J E C
telomere
CCCCATTCCGCAATTATACC-cggaactttggtgccac-CAGTGCTACAAAACCTACAG
20p: downstream NKX2-2
14q: TRDD3
der(14)
centromere
TRD@ V
D NKX2-4
telomere
GTGTTTCATTGTGCCTTCCT-ggtccgcgtactgggggatac-TGCGAAGTCACTTTAGAAAA
14q: TRDD2
20p: downstream NKX2-2
F
Immature cluster
Proliferative cluster
1964
1524
9577
321
491
572
10138
1446
914
9919
2702
NKX2-1
Actin

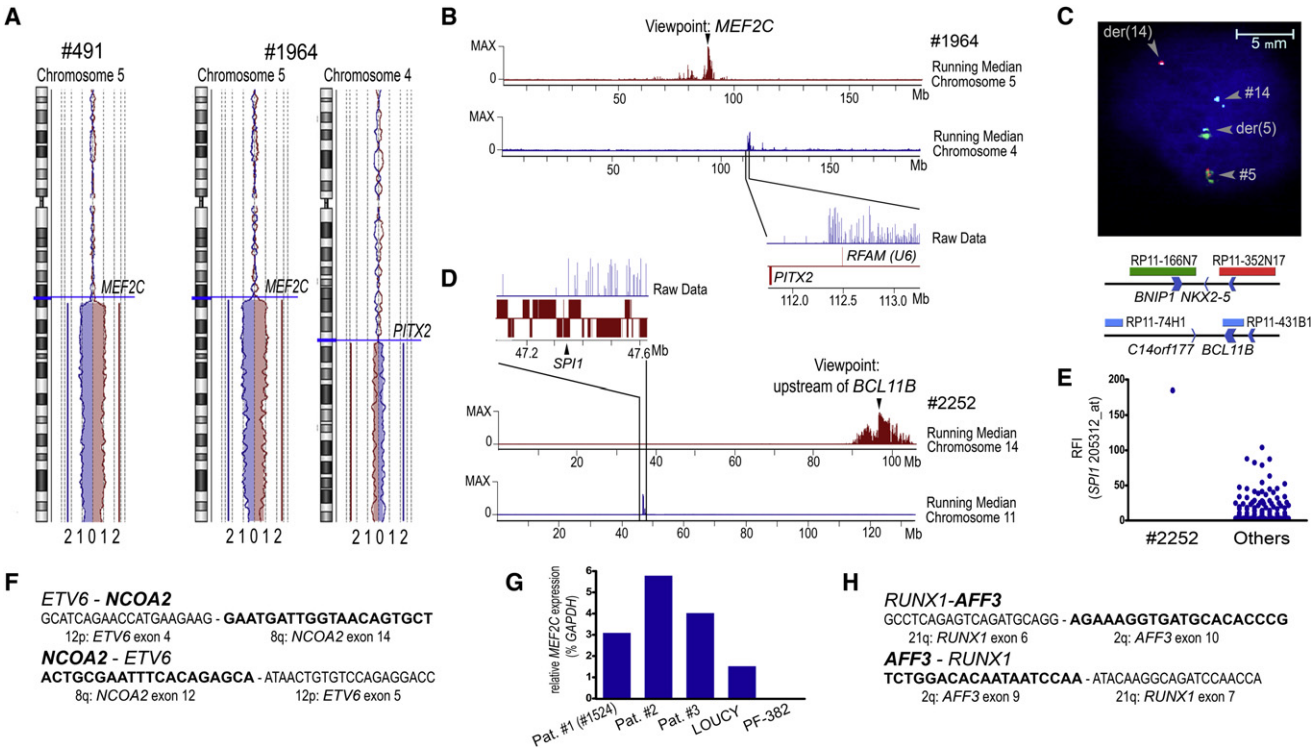


Figure 4. *MEF2C*-Activating Rearrangements for Immature Cluster Samples

(A) Array-CGH results for chromosomes 4 and/or 5 for patient Nos. 491 and 1964. Blue and red tracings represent two independent, dye-swopped experiments. Positions of *MEF2C* and *PITX2* have been indicated.

(B) Visualization of an unbalanced chromosomal translocation t(4;5)(q26;q14) for patient No. 1964 by 4C analysis. The *MEF2C* VP is indicated by an arrow. Running median of probe set intensities for chromosomes 5 and 4 are indicated in red and blue, respectively.

(C) Validation of a chromosomal translocation between *NKX2-5* and *BCL11B* in patient No. 9577 by FISH. Schematic positions of FISH probes are shown.

(D) Identification of the t(11;14)(p11.2;q32.2) chromosomal translocation between *SPI1* and *BCL11B* in patient No. 2252 by 4C. The VP is positioned ~0.6 Mb upstream of *BCL11*B, as indicated by an arrow.

(E) Ectopic *SPI1* expression in patient No. 2252 compared to 116 additional T-ALL patient samples. Raw fluorescent intensities of probe set 205312_at are shown.

(F) Cloned fusion areas for reciprocal *ETV6-NCOA2* and *NCOA2-ETV6* fusion transcripts in patient No. 1524.

(G) Relative *MEF2C* expression by RQ-PCR in three selected *ETV6-NCOA2* rearranged T-ALL patients (Nos. 1–3). Cell lines LOUCY and PF382 are positive and negative controls for *MEF2C* expression, respectively.

(H) Cloned fusion areas for reciprocal *RUNX1-AFF3* and *AFF3-RUNX1* fusion transcripts for patient No. 572. See also Figure S4.

capacity by using cellular transformation assays in NIH 3T3 (Figure 6A) or BJ-EHT cells (Figure 6B). Transfecting *NKX2-1* or *MEF2C* expression constructs into the cells was insufficient to drive cellular transformation. We then tested cellular transformation of *MEF2C* and *NKX2-1* when combined with *RAS* or *MYC*, two oncogenes that are frequently activated in T-ALL through RAS or NOTCH1-activating mutations (Kawamura et al., 1999; Palomero et al., 2006; Weng et al., 2006). *NKX2-1* and *MEF2C* were both able to synergize with *RAS* or *MYC* genes in driving cellular transformation (Figures 6A and 6B).

We then further tested the importance of *MEF2C* for T cell pathogenesis for which we had a cell line model available. In normal human T cell development subsets, *MEF2C* is exclusively expressed at the pre-DN1 and DN1 stages, after which it is downregulated (Figure S5). We knocked down MEF2C expression in T-ALL cell line LOUCY using siRNA molecules. MEF2C knockdown induced cellular differentiation as LOUCY cells became positive for membrane CD3 and TCR$\gamma\delta$ expression (Figures 6C–6E). This indicates that MEF2C can block T cell differentiation at a very immature stage.

Figure 3. *NKX2-1* and *NKX2-2* Rearrangements in Proliferative Cluster Patient Samples

(A) 4C results obtained from *NKX2-1* or *TRB@* VPs. Positions of *TRA@*, *NKX2-1*, and *IGH@* loci are shown by gray vertical bars. 4C results for a normal control are shown in gray. Higher magnifications of the reciprocal breakpoint regions are given in Figure S3.

(B) Validation of *NKX2-1* rearrangements by FISH. Schematic positions of FISH probes are shown.

(C) Schematic representation of the der(7) breakpoint region and breakpoint sequence of the unbalanced t(7;14)(q34;q13) for patient No. 9989.

(D) Visualization of a single-copy *NKX2-1* amplification (green box) in patient No. 2702 as identified by array-CGH.

(E) Schematic representation of t(14;20)(q11;p11) breakpoint regions and cloned breakpoint sequences for patient No. 10138 with the *NKX2-2* rearrangement.

(F) NKX2-1 protein expression in representative proliferative cluster and immature cluster patient samples as shown by western blot. Actin was used as loading control.

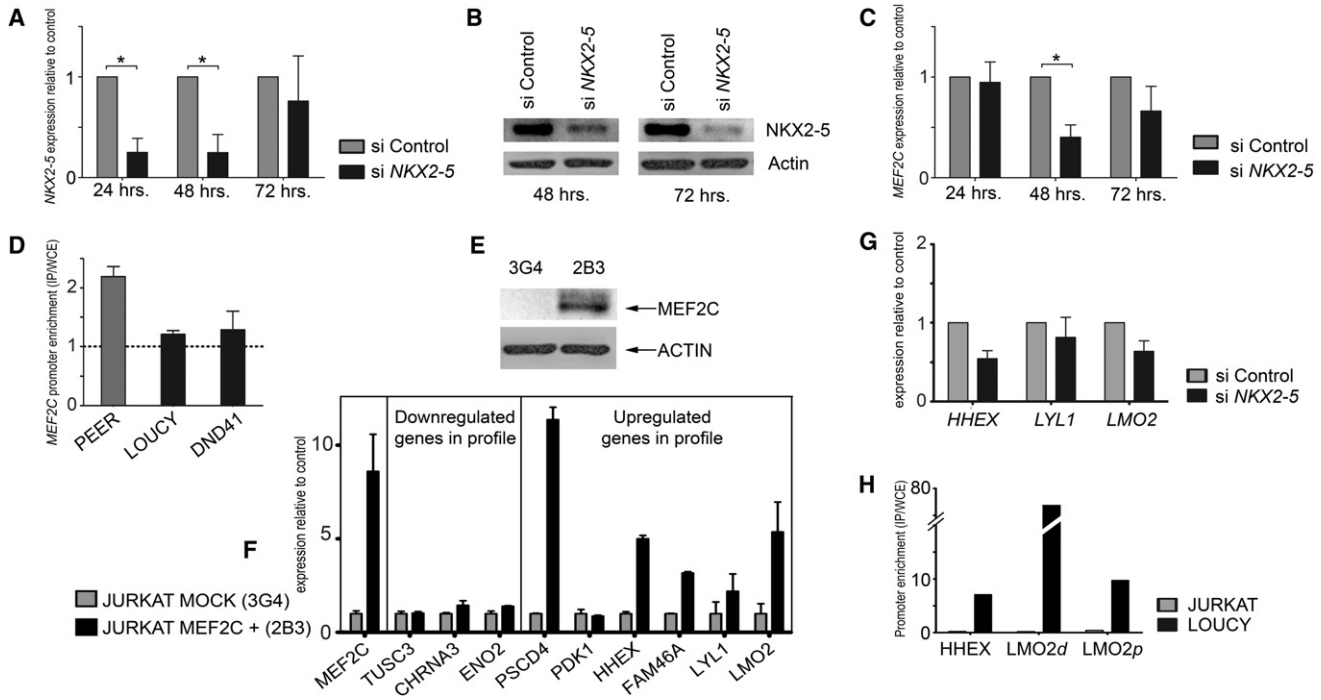


Figure 5. NKX2-5 Controls *MEF2C* Expression

(A) RQ-PCR results of *NKX2-5* mRNA expression levels or (B) NKX2-5 protein levels in cell line PEER at indicated time points following electroporation with siRNAs directed against *NKX2-5* (black bars) relative to control siRNA-treated cells (gray bars). For western blot analysis, actin was used as a loading control.

(C) RQ-PCR results of *MEF2C* mRNA expression levels at indicated time points following electroporation with anti-*NKX2-5* siRNA molecules (black bars) relative to controls (gray bars).

(D) Enrichment of MEF2C promoter sequences in NKX2-5 ChIP analysis in the *NKX2-5* translocated cell line PEER, but not in negative control lines LOUCY or DND41.

(E) Ectopic MEF2C expression in the MEF2C stably transfected Jurkat clone 2B3, as shown by western blot analysis. The mock-transfected Jurkat clone 3G4 served as negative control.

(F) RQ-PCR results for MEF2C-positive Jurkat clone 2B3 clone or the mock-transfected control (3G4) for *MEF2C* and random selected immature signature genes that are relatively downregulated (*TUSC3*, *CHRNA3*, *ENO2*) or upregulated (*PSCD4 (CYTH4)*, *PDK1*, *HHEX*, *FAM46A*, *LYL1*, *LMO2*) in immature T-ALL cases compared to other cluster samples.

(G) Relative expression results for *HHEX*, *LYL1*, and *LMO2* in the cell line PEER 72 hr after electroporation with siRNAs directed against *NKX2-5* (black bars) relative to control siRNA-treated PEER cells (gray bars).

(H) Enrichment of *HHEX* promoter and the distal and proximal *LMO2* promoters upon MEF2C ChIP analysis in the immature cell line LOUCY, but not in the negative control line Jurkat. For all panels the SDs are shown. Significant differences ($p < 0.05$) in relative expression levels are indicated by an asterisk (*).

Validation of the Immature and Proliferative Clusters in Independent T-ALL Cohorts

We then confirmed our T-ALL clustering (Figures 1A and 1C) and molecular-cytogenetic findings (Table 2) in two independent validation cohorts, i.e., a French data set comprising 107 pediatric and adult T-ALL cases (Clappier et al., 2007; Soulier et al., 2005), and a second Rotterdam cohort comprising 108 pediatric and adult T-ALL cases. Upon testing the comparability of the initial Rotterdam cohort and the French data set (Figure S6A), the proliferative and immature clusters could be reproduced in a combined unsupervised cluster analysis (Figure S6B). Based on the unsupervised clustering of our initial Rotterdam cohort, PAM statistics predicted various proliferative cluster cases as well as immature cluster cases in the Rotterdam validation cohort (data not shown). A total of 26 proliferative cluster cases were identified, of which various samples highly expressed *NKX2-1* (Figure S6C). *NKX2-1* translocations/inversions could be demonstrated using FISH in three cases (Figures S6G–S6I). Eight out of ten *TLX1*-rearranged cases were part of the proliferative cluster as well (Figure S6D; data not shown), further supporting the notion that *NKX2-1* and *TLX1* oncogenic rearrangements may share common pathogenic mechanisms. Again, some of these *TLX1*-rearranged cases also expressed *NKX2-1* at low levels (Figure S6C), whereas none of these samples had *NKX2-1* rearrangements. We also validated high *MEF2C* expression for the 24 cases that were assigned to the immature cluster by PAM analysis (Figure S6E), and these samples expressed the highest levels of its downstream target *LYL1* (Figure S6F).

DISCUSSION

In this study we have identified *NKX2-1*, its related family member *NKX2-2*, and *MEF2C* as potential oncogenes for T-ALL. Supervised cluster analyses based on genes uniquely associated with the known genetic *TAL/LMO*, *TLX3*, *TLX1*, and *HOXA* subgroups revealed that samples with high expression

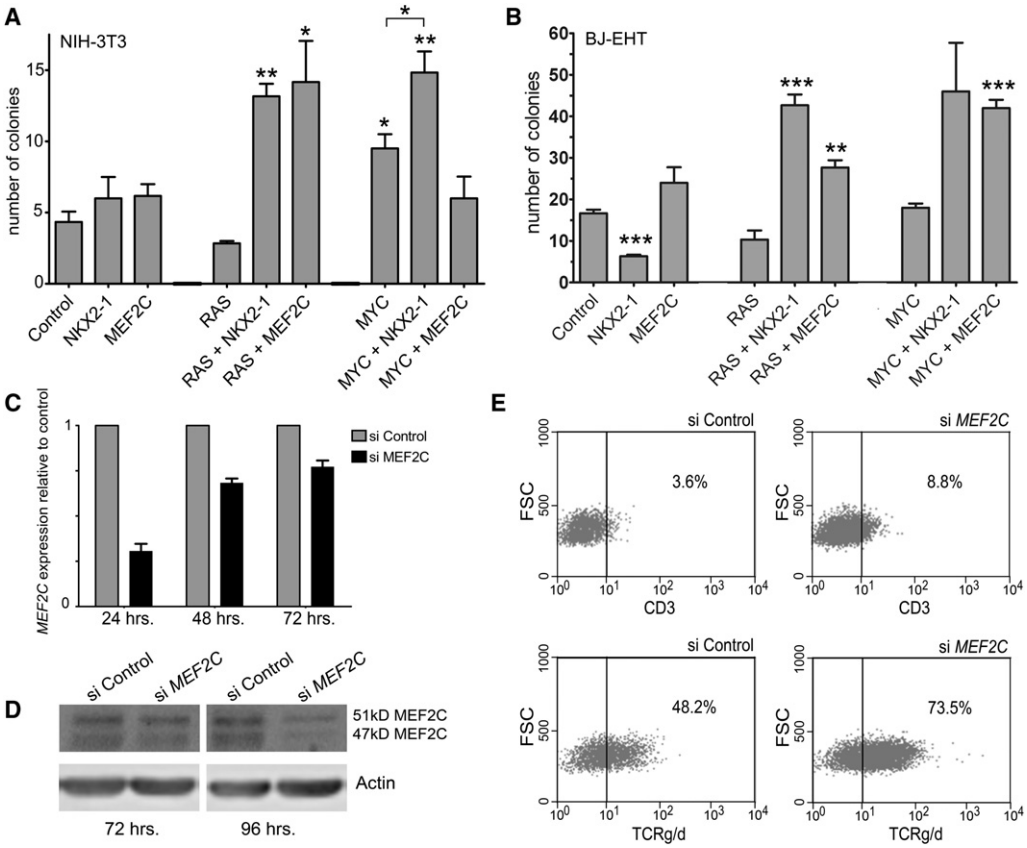


Figure 6. Cellular Transformation by *MEF2C* and *NKX2-1*

Cellular transformation of (A) NIH 3T3 or (B) BJ-EHT cells upon transfection of *MEF2C*, *NKX2-1*, *MYC*, and/or *RAS* expression vectors as indicated. Significance levels for colony number differences between indicated expression construct combinations relative to the empty vector control are indicated (*$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$).

(C) *MEF2C* expression knockdown as measured by RQ-PCR for the *MEF2C*-positive cell line LOUCY at indicated time points following electroporation with *MEF2C*-specific siRNA molecules (black bars) relative to control siRNA-treated cells (gray bars).

(D) Downregulation of MEF2C protein following treatment with *MEF2C*-specific siRNA molecules as validated by western blot. Based on the protein size, the predominant $\alpha 1\beta$ (47 kDa) and the $\alpha 1\beta\gamma$ (51.2 kDa) MEF2C isoform (Zhu and Gulick, 2004) are indicated.

(E) Increase of mCD3 ($p = 0.0032$) and TCR$\gamma\delta$ ($p = 0.023$) expression as demonstrated by FACS analysis in LOUCY cells, 96 hr following treatment with *MEF2C*-specific siRNA molecules. A representative example from three independent experiments is shown. For all panels the SEMs are shown. See also Figure S5.

of *NKX2-1/NKX2-2* or *MEF2C* characterize two T-ALL clusters for which no driving oncogenic hits have been identified so far. Both clusters represent about 20% of all T-ALL cases.

Variant rearrangements for *NKX2-1* and *NKX2-2* to T cell receptor genes (*TRAD@*, *TRB@*) were identified, and one case had an inversion to the *IGH@* locus. The *IgH* enhancer seems functional in this patient with T-ALL, and *IgH* enhancer (Eμ) driven oncogene expression in a T cell context has been described before, both for human T-ALL (Nguyen-Khac et al., 2010) as well as in transgenic mouse models (Katsumata et al., 1992; Strasser et al., 1991). This patient did not express B cell markers, therefore excluding it as a biphenotypic leukemia. *NKX2-1* was able to transform NIH 3T3 and BJ-EHT cells in synergism with *RAS* or *MYC*, two genes that become activated through RAS or NOTCH1-activating mutations in approximately 15% and 60% of T-ALL cases, respectively (Kawamura et al., 1999; Palomero et al., 2006; Weng et al., 2006). Therefore, our data strongly support that *NKX2-1/NKX2-2* may represent oncogenes in T-ALL. *NKX2-1* is not expressed during normal T cell development based on expression data by microarray for flow-sorted thymic subsets (Dik et al., 2005; Soulier et al., 2005).

NKX2-1 and *NKX2-2* have been associated with other types of cancer before: *NKX2-1* is amplified in human lung cancer (Weir et al., 2007); and *NKX2-2* is a target of the EWS/FLI fusion product in Ewing's sarcoma (Smith et al., 2006). NKX2-1 and NKX2-2 are 59% identical for the homeodomain region, indicating that both proteins may exert identical oncogenic roles in T-ALL. This is further supported by the fact that rearrangements for both genes were identified in samples that tightly cluster together in unsupervised and supervised analyses. NK-like homeobox transcription factors play important roles in T-ALL because *NKX2-5* was previously identified as part of an oncogenic rearrangement in T-ALL (Nagel et al., 2003). The NK-like homeobox transcription factor *NKX3-1* has been found to be highly activated in *TAL1*-rearranged cases (Soulier et al., 2005), as a direct TAL1 target gene (Kusy et al., 2010). The homeodomains of NKX2-5 and NKX3-1 are only distantly related (37% identity) and only 48% and 47% identical to the homeodomain of

NKX2-1, respectively. This may explain why *NKX2-5*, *NKX3-1*, and *NKX2-1/NKX2-2* are associated with different T-ALL subgroups: ectopic *NKX3-1* expression in the *TAL/LMO* subgroup (Soulier et al., 2005); *NKX2-1/NKX2-2* rearrangements with the proliferative T-ALL cluster (this study); and *NKX2-5* translocations with immature T cell development (this study) that activates *MEF2C* (this study; Nagel et al., 2008).

In unsupervised analyses, *NKX2-1/NKX2-2* rearranged cases cluster together with *TLX1*-rearranged cases to form the proliferative cluster. This indicates that *NKX2-1/NKX2-2* and *TLX1*-rearranged T-ALLs are biologically related. This is further supported by the fact that *NKX2-1* and *TLX1*-rearranged cases share a similar immunophenotypic makeup consistent with cortical arrest as well by the fact that various *TLX1*-rearranged cases express *NKX2-1* in the absence of *NKX2-1* rearrangements, albeit at low levels. One of the explanations may be that TLX1 controls *NKX2-1* expression. In addition several other cases that are part of the proliferative cluster lack *TLX1*, *NKX2-1*, or *NKX2-2* rearrangements, indicating that an additional oncogenic rearrangement awaits identification for this cluster.

The second cluster had a very immature immunophenotype, with most cases expressing CD34 and frequently coexpressing the CD13 and/or CD33 myeloid markers. We identified various rearrangements that directly or indirectly activate *MEF2C*. *MEF2C* is a member of the MADS-box transcription factor family that includes the four *MEF2A-D* genes that are important regulators of skeletal muscle development (Grounds, 1991). Immature T-ALL subgroups have been identified before (Coustan-Smith et al., 2009; Ferrando et al., 2002; Soulier et al., 2005), and our immature cluster cases could also be predicted based on an ETP expression signature (Coustan-Smith et al., 2009). We now conclude that *MEF2C* is the driving oncogene for immature (ETP) T-ALL cases. Our immature cases also have the highest *LYL1* expression and highly express *LMO2* (Ferrando et al., 2002; this study). *LYL1* and *LMO2* are members of the basic-helix-loop-helix (bHLH) family and the LIM-domain only family, respectively. Apart from *LYL1* and *LMO2*, the immature cases also highly express the homeobox gene *HHEX*. We have now shown that *HHEX*, *LYL1*, and *LMO2* are being regulated by MEF2C, and it was proven that MEF2C directly binds in the promoter regions of at least *HHEX* and *LMO2*. This may support a pathogenic role for established oncogenes such as *LYL1* and *LMO2* in MEF2C-deregulated early-committed T cells. To what extend *LMO2* and/or *LYL1* as MEF2C targets will be sufficient to drive a leukemogenic program in these early-committed T cells is presently unclear. To our knowledge, oncogenic rearrangements of *LMO2* and *LYL1* have not been observed in immature T-ALL (Ferrando et al., 2002; this work) but are exclusive for the *TAL/LMO* subgroup that also includes rearrangements of the *LYL1*-homolog *TAL1*. Therefore, MEF2C may elicit a more comprehensive transcriptional program characteristic for ETP T-ALLs than aberrant expression of LMO2 or LYL1 alone.

MEF2C is a key regulator for lymphoid development that is activated by PU.1 (Stehling-Sun et al., 2009). In B cell development, MEF2C is activated by calcineurin following BCR triggering and warrants for cell viability and proliferation (Wilker et al., 2008). MEF2C has been implicated in human oncogenesis: in myeloid leukemias of *MLL-AF9* transgenic mice, *Mef2c* has been identified as a HoxA9 target gene that regulates self-renewal of leukemic stem cells (Krivtsov et al., 2006). *MEF2C* is also highly expressed in human *MLL*-rearranged AML that is characterized by upregulation of *HOXA* genes, including *HOXA9* (Schwieger et al., 2009). *Mef2c* is further identified as potential oncogene in insertional mutagenesis studies (Du et al., 2005; Schwieger et al., 2009) and can provoke myeloid leukemias (Schwieger et al., 2009). Also, the related family member *MEF2D* is involved in the *MEF2D-DAZAP1* fusion that has been identified in ALL (Prima and Hunger, 2007).

Many oncogenic hits as identified in this study involve early hematopoietic transcription factors, including NKX2-5, PU.1, and presumably RUNX1. These factors are important for normal T cell development (Rothenberg, 2007). All these factors converge on MEF2C in immature T-ALL, and it is tempting to speculate that MEF2C is a central regulator for normal early T cell development. MEF2C may need to become downregulated to facilitate maturation beyond this immature stage, and we indeed demonstrated that knockdown of MEF2C expression in T-ALL cell line LOUCY provoked differentiation. In support of these notions, MEF2C is expressed in normal human thymocyte pre-DN1 and DN1 subsets, but expression is dramatically decreased beyond the DN2 stage (Figure S5). A similar downregulation of *MEF2C* expression could be validated from gene expression data for equivalent flow-sorted thymic subsets, as published (Dik et al., 2005; data not shown). *MEF2C* may represent the central oncogene for immature T-ALL cases that seems to provide a T cell differentiation block at the immature stage, as demonstrated in this article. This was further supported by our transformation assay results in which *MEF2C* transformed NIH 3T3 and BJ-EHT cells in combination with *RAS* or *MYC*. We also observed that several genes from the TGFBR1 pathway were upregulated, including *TGFBR1*, *ZEB2*, *SMAD7*, *SMURF2*, and *RUNX3*, or downregulated (*SMAD1*). Because both activators (*TGFBR1*) and inhibitors (like *SMURF2*, *SMAD7*) are overexpressed while the activator *SMAD1* is underexpressed, it is difficult to anticipate the functional consequences of this pathway for the immature T-ALL cases.

In conclusion we used a strategy integrating molecular genetics with large-scale expression profiling and identified two oncogenic subgroups and eight genomic rearrangements that, to our knowledge, have not been identified before in human T-ALL or other cancer types. We have shown that these proliferative and immature subtypes reflect different biological entities: the proliferative cluster strongly expresses proliferation genes and is associated with aberrations and ectopic expression of *NKX2-1* or *NKX2-2*, and expression of CD1. In contrast the immature cluster was characterized by immature T cell development, activation of genes involved in protein binding and dimerization, expression of components of the TGFBR1 pathway, and high expression of the MADS transcription factor *MEF2C* due to abnormalities of *MEF2C*, transcription factors that regulate *MEF2C*, or MEF2C-associating cofactors. We conclude that *NKX2-1*, *NKX2-2*, and *MEF2C* define oncogenic pathways in T-ALL.

EXPERIMENTAL PROCEDURES

Patient Samples

Viably frozen diagnostic bone marrow or peripheral blood samples from 117 patients with pediatric T-ALL and corresponding clinical and

immunophenotypic data were provided by the German Co-operative study group for childhood Acute Lymphoblastic Leukemia (COALL) and the Dutch Childhood Oncology Group (DCOG). The patients' parents or their legal guardians provided informed consent to use leftover material for research purposes according to the Declaration of Helsinki. This study was approved by the Institutional Review Board of the Erasmus MC, Rotterdam. Leukemic cells were isolated and enriched from these samples as previously described (Van Vlierberghe et al., 2006). All resulting samples contained ≥90% leukemic cells, as determined morphologically by May-Grünwald-Giemsa-stained cytospins (Merck, Darmstadt, Germany). Patients were assigned to specific molecular-cytogenetic T-ALL subgroups based on FISH results for *TAL1*, *TAL2*, *LMO1*, *LMO2*, *TLX1*, *TLX3*, *CALM-AF10*, *SET-NUP214*, *MLL*, *MYB*, or Inv(7)(p15;q34)), and positivity by RT-PCR for *SIL-TAL1*, *TLX1*, *TLX3*, *CALM-AF10*, or *SET-NUP214* as described before (van Grotel et al., 2006; Van Vlierberghe et al., 2006, 2008b).

4C

4C was performed as described before (Simonis et al., 2006). Briefly, DNA and protein in approximately 10 million viable cells were crosslinked in a 2% formaldehyde solution to conserve the physical proximity of DNA regions. Cells were lysed, and DNA was digested with HindIII. After dilution of DNA, restriction fragments were ligated. This way, DNA fragments that are physically near each other in the viable cell can be ligated. The sample was subsequently de-crosslinked by an overnight incubation at 65°C. DNA was purified and digested with the frequent cutter DpnII. Samples were diluted and ligated to allow circularization of individual restriction fragments. Following linearization with ScaI (located between both inverse PCR primers), DNA sequences ligated to the fragment of interest were amplified by inverse PCR, labeled, and hybridized on a microarray (Nimblegen, Madison, WI, USA) containing probes that roughly represent individual HindIII fragments in the genome. Raw fluorescence intensities are visualized as the running median per 30 neighboring probes, each representing a HindIII restriction fragment. The viewpoint (VP) is the HindIII restriction fragment where 4C PCR primers are located. Data are visualized with SignalMap software (Nimblegen) (NCBI, Build 36). Inverse PCR primer sets developed for *NKX2-1*, *BCL11B*, and *MEF2C* are listed in the Supplemental Experimental Procedures.

Gene Expression Microarray, Data Extraction, and Normalization

Integrity of patient samples total RNA was checked using the Agilent 2100 Bioanalyzer (Santa Clara, CA, USA). Copy DNA and ccRNA syntheses from total RNA, hybridization to Humane Genome U133 plus2.0 oligonucleotide microarrays (Affymetrix, Santa Clara, CA, USA), and washing steps were performed according to the manufacturers' protocol. Probe set intensities were extracted from CEL files in the statistical data analysis environment *R*, version 2.8.0 (Bioconductor Affy package). All arrays had a 3′–5′ GAPDH ratio lower than 3-fold. Probe intensities were normalized in *R* using RMA-solo, RMA (Irizarry et al., 2003), or VSN (Huber et al., 2002) methods.

Biostatistical Analyses

Biostatistical analyses have been described in detail in the Supplemental Experimental Procedures. Briefly, unsupervised cluster analyses were performed in dChip (Li and Wong, 2001). Identification of differentially expressed genes with FDR control was done by various methods including Wilcoxon statistics ("Multtest" in R), SAM statistics (Tusher et al., 2001) (BRB tools, version 3.7, R. Simon & A.P. Lam), and COPA statistics (Tomlins et al., 2005) for outlier analysis using a R routine. Prediction of identified subtypes was done using various algorithms embedded in BRB tools including Diagonal Linear Discriminant Analysis, 1-nearest neighbor, 3-nearest neighbor, and nearest centroid, as well as tested by PAM (Tibshirani et al., 2002). PCA based on the top100 most significant differentially expressed genes for the major T-ALL subgroups (i.e., the supervised analysis) was performed using GeneMath XT 1.6.1. software (Applied Maths, Inc., Austin, TX, USA). To validate findings from the Rotterdam data set, this data set was combined with the French (Paris) Affymetrix U133A data set (Soulier et al., 2005). Data for overlapping probe sets were extracted from both data sets, RMA-solo normalized, and corrected for batch effects using the Combat Method (Johnson et al., 2007). Profiles for similar T-ALL subgroups in both data sets were tested for comparability by using various methods, including the OrderedList method using the Bioconductor package "OrderedList" in R (Lottaz et al., 2006) as well as the subclass method (Hoshida et al., 2007). Additional methods and materials are described in the Supplemental Experimental Procedures.

ACCESSION NUMBERS

Rotterdam and French microarray data sets are available at http://www.ncbi.nlm.nih.gov/geo/ and the EBI database at http://www.ebi.ac.uk/arrayexpress under accession numbers GSE26713 and E-MEXP-313, respectively.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, and six tables and can be found with this article online at doi:10.1016/j.ccr.2011.02.008.

ACKNOWLEDGMENTS

I.H. is financially supported by the Dutch Cancer Society (KWF-EMCR 2006-3500). C.K. and M.Ve. are financially supported by the Children Cancer Free Foundation (Stichting Kinderen Kankervrij, KiKa; Grant No. KiKa 2008-029). We also would like to thank the German Jose Carreras Leukemia Foundation (Grant No. SP 04/03), Children's Cancer Center Support Community Hamburg, and the French program Carte d'Identité des Tumeurs (CIT) from the Ligue Contre le Cancer for financial support.

Received: March 25, 2010
Revised: November 15, 2010
Accepted: February 4, 2011
Published: April 11, 2011

REFERENCES

Buijs, A., van Rompaey, L., Molijn, A.C., Davis, J.N., Vertegaal, A.C., Potter, M.D., Adams, C., van Baal, S., Zwarthoff, E.C., Roussel, M.F., and Grosveld, G.C. (2000). The MN1-TEL fusion protein, encoded by the translocation (12;22)(p13;q11) in myeloid leukemia, is a transcription factor with transforming activity. Mol. Cell. Biol. *20*, 9281–9293.

Chen, S.L., Dowhan, D.H., Hosking, B.M., and Muscat, G.E. (2000). The steroid receptor coactivator, GRIP-1, is necessary for MEF-2C-dependent gene expression and skeletal muscle differentiation. Genes Dev. *14*, 1209–1228.

Clappier, E., Cuccuini, W., Kalota, A., Crinquette, A., Cayuela, J.M., Dik, W.A., Langerak, A.W., Montpellier, B., Nadel, B., Walrafen, P., et al. (2007). The C-MYB locus is involved in chromosomal translocation and genomic duplications in human T-cell acute leukemia (T-ALL), the translocation defining a new T-ALL subtype in very young children. Blood *110*, 1251–1261.

Coustan-Smith, E., Mullighan, C.G., Onciu, M., Behm, F.G., Raimondi, S.C., Pei, D., Cheng, C., Su, X., Rubnitz, J.E., Basso, G., et al. (2009). Early T-cell precursor leukaemia: a subtype of very high-risk acute lymphoblastic leukaemia. Lancet Oncol. *10*, 147–156.

Dik, W.A., Pike-Overzet, K., Weerkamp, F., de Ridder, D., de Haas, E.F., Baert, M.R., van der Spek, P., Koster, E.E., Reinders, M.J., van Dongen, J.J., et al. (2005). New insights on human T cell development by quantitative T cell receptor gene rearrangement studies and gene expression profiling. J. Exp. Med. *201*, 1715–1723.

Du, Y., Spence, S.E., Jenkins, N.A., and Copeland, N.G. (2005). Cooperating cancer-gene identification through oncogenic-retrovirus-induced insertional mutagenesis. Blood *106*, 2498–2505.

Ferrando, A.A., Neuberg, D.S., Staunton, J., Loh, M.L., Huard, C., Raimondi, S.C., Behm, F.G., Pui, C.H., Downing, J.R., Gilliland, D.G., et al. (2002). Gene expression signatures define novel oncogenic pathways in T cell acute lymphoblastic leukemia. Cancer Cell *1*, 75–87.

Grosveld, G.C. (2007). MN1, a novel player in human AML. Blood Cells Mol. Dis. *39*, 336–339.

Grounds, M.D. (1991). Towards understanding skeletal muscle regeneration. Pathol. Res. Pract. *187*, 1–22.

Hebert, J., Cayuela, J.M., Berkeley, J., and Sigaux, F. (1994). Candidate tumor-suppressor genes MTS1 (p16INK4A) and MTS2 (p15INK4B) display frequent homozygous deletions in primary cells from T- but not from B-cell lineage acute lymphoblastic leukemias. Blood *84*, 4038–4044.

Hoshida, Y., Brunet, J.P., Tamayo, P., Golub, T.R., and Mesirov, J.P. (2007). Subclass mapping: identifying common subtypes in independent disease data sets. PLoS ONE *2*, e1195.

Huber, W., von Heydebreck, A., Sultmann, H., Poustka, A., and Vingron, M. (2002). Variance stabilization applied to microarray data calibration and to the quantification of differential expression. Bioinformatics *18* (*Suppl 1*), S96–S104.

Irizarry, R.A., Hobbs, B., Collin, F., Beazer-Barclay, Y.D., Antonellis, K.J., Scherf, U., and Speed, T.P. (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics *4*, 249–264.

Johnson, W.E., Li, C., and Rabinovic, A. (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics *8*, 118–127.

Katsumata, M., Siegel, R.M., Louie, D.C., Miyashita, T., Tsujimoto, Y., Nowell, P.C., Greene, M.I., and Reed, J.C. (1992). Differential effects of Bcl-2 on T and B cells in transgenic mice. Proc. Natl. Acad. Sci. USA *89*, 11376–11380.

Kawamura, M., Ohnishi, H., Guo, S.X., Sheng, X.M., Minegishi, M., Hanada, R., Horibe, K., Hongo, T., Kaneko, Y., Bessho, F., et al. (1999). Alterations of the p53, p21, p16, p15 and RAS genes in childhood T-cell acute lymphoblastic leukemia. Leuk. Res. *23*, 115–126.

Krivtsov, A.V., Twomey, D., Feng, Z., Stubbs, M.C., Wang, Y., Faber, J., Levine, J.E., Wang, J., Hahn, W.C., Gilliland, D.G., et al. (2006). Transformation from committed progenitor to leukaemia stem cell initiated by MLL-AF9. Nature *442*, 818–822.

Kusy, S., Gerby, B., Goardon, N., Gault, N., Ferri, F., Gerard, D., Armstrong, F., Ballerini, P., Cayuela, J.M., Baruchel, A., et al. (2010). NKX3.1 is a direct TAL1 target gene that mediates proliferation of TAL1-expressing human T cell acute lymphoblastic leukemia. J. Exp. Med. *207*, 2141–2156.

Li, C., and Wong, W.H. (2001). Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. Proc. Natl. Acad. Sci. USA *98*, 31–36.

Lottaz, C., Yang, X., Scheid, S., and Spang, R. (2006). OrderedList—a bioconductor package for detecting similarity in ordered gene lists. Bioinformatics *22*, 2315–2136.

Nagel, S., Kaufmann, M., Drexler, H.G., and MacLeod, R.A. (2003). The cardiac homeobox gene NKX2-5 is deregulated by juxtaposition with BCL11B in pediatric T-ALL cell lines via a novel t(5;14)(q35.1;q32.2). Cancer Res. *63*, 5329–5334.

Nagel, S., Meyer, C., Quentmeier, H., Kaufmann, M., Drexler, H.G., and MacLeod, R.A. (2008). MEF2C is activated by multiple mechanisms in a subset of T-acute lymphoblastic leukemia cell lines. Leukemia *22*, 600–607.

Nguyen-Khac, F., Barin, C., Chapiro, E., Macintyre, E.A., Romana, S., and Bernard, O.A. (2010). Cyclin D3 deregulation by juxtaposition with IGH locus in a t(6;14)(p21;q32)-positive T-cell acute lymphoblastic leukemia. Leuk. Res. *34*, e13–e14.

Palomero, T., Lim, W.K., Odom, D.T., Sulis, M.L., Real, P.J., Margolin, A., Barnes, K.C., O'Neil, J., Neuberg, D., Weng, A.P., et al. (2006). NOTCH1 directly regulates c-MYC and activates a feed-forward-loop transcriptional network promoting leukemic cell growth. Proc. Natl. Acad. Sci. USA *103*, 18261–18266.

Pieters, R., and Carroll, W.L. (2008). Biology and treatment of acute lymphoblastic leukemia. Pediatr. Clin. North Am. *55*, 1–20.

Prima, V., and Hunger, S.P. (2007). Cooperative transformation by MEF2D/DAZAP1 and DAZAP1/MEF2D fusion proteins generated by the variant t(1;19) in acute lymphoblastic leukemia. Leukemia *21*, 2470–2475.

Przybylski, G.K., Dik, W.A., Grabarczyk, P., Wanzeck, J., Chudobska, P., Jankowski, K., von Bergh, A., van Dongen, J.J., Schmidt, C.A., and Langerak, A.W. (2006). The effect of a novel recombination between the homeobox gene NKX2-5 and the TRD locus in T-cell acute lymphoblastic leukemia on activation of the NKX2-5 gene. Haematologica *91*, 317–321.

Pui, C.H., and Evans, W.E. (2006). Treatment of acute lymphoblastic leukemia. N. Engl. J. Med. *354*, 166–178.

Rothenberg, E.V. (2007). Regulatory factors for initial T lymphocyte lineage specification. Curr. Opin. Hematol. *14*, 322–329.

Schwieger, M., Schüler, A., Forster, M., Engelmann, A., Arnold, M.A., Delwel, R., Valk, P.J., Löhler, J., Slany, R.K., Olson, E.N., and Stocking, C. (2009). Homing and invasiveness of MLL/ENL leukemic cells is regulated by MEF2C. Blood *114*, 2476–2488.

Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., de Wit, E., van Steensel, B., and de Laat, W. (2006). Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). Nat. Genet. *38*, 1348–1354.

Simonis, M., Klous, P., Homminga, I., Galjaard, R.J., Rijkers, E.J., Grosveld, F., Meijerink, J.P.P., and de Laat, W. (2009). High-resolution identification of balanced and complex chromosomal rearrangements by 4C technology. Nat. Methods *6*, 837–842.

Smith, R., Owen, L.A., Trem, D.J., Wong, J.S., Whangbo, J.S., Golub, T.R., and Lessnick, S.L. (2006). Expression profiling of EWS/FLI identifies NKX2.2 as a critical target gene in Ewing's sarcoma. Cancer Cell *9*, 405–416.

Soulier, J., Clappier, E., Cayuela, J.M., Regnault, A., Garcia-Peydro, M., Dombret, H., Baruchel, A., Toribio, M.L., and Sigaux, F. (2005). HOXA genes are included in genetic and biologic networks defining human acute T-cell leukemia (T-ALL). Blood *106*, 274–286.

Stehling-Sun, S., Dade, J., Nutt, S.L., DeKoter, R.P., and Camargo, F.D. (2009). Regulation of lymphoid versus myeloid fate 'choice' by the transcription factor Mef2c. Nat. Immunol. *10*, 289–296.

Strasser, A., Harris, A.W., and Cory, S. (1991). bcl-2 transgene inhibits T cell death and perturbs thymic self-censorship. Cell *67*, 889–899.

Strehl, S., Nebral, K., Konig, M., Harbott, J., Strobl, H., Ratei, R., Struski, S., Bielorai, B., Lessard, M., Zimmermann, M., et al. (2008). ETV6-NCOA2: a novel fusion gene in acute leukemia associated with coexpression of T-lymphoid and myeloid markers and frequent NOTCH1 mutations. Clin. Cancer Res. *14*, 977–983.

Tibshirani, R., Hastie, T., Narasimhan, B., and Chu, G. (2002). Diagnosis of multiple cancer types by shrunken centroids of gene expression. Proc. Natl. Acad. Sci. USA *99*, 6567–6572.

Tomlins, S.A., Rhodes, D.R., Perner, S., Dhanasekaran, S.M., Mehra, R., Sun, X.W., Varambally, S., Cao, X., Tchinda, J., Kuefer, R., et al. (2005). Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. Science *310*, 644–648.

Tusher, V.G., Tibshirani, R., and Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. Proc. Natl. Acad. Sci. USA *98*, 5116–5121.

van Grotel, M., Meijerink, J.P., Beverloo, H.B., Langerak, A.W., Buys-Gladdines, J.G., Schneider, P., Poulsen, T.S., den Boer, M.L., Horstmann, M., Kamps, W.A., et al. (2006). The outcome of molecular-cytogenetic subgroups in pediatric T-cell acute lymphoblastic leukemia: a retrospective study of patients treated according to DCOG or COALL protocols. Haematologica *91*, 1212–1221.

Van Vlierberghe, P., van Grotel, M., Beverloo, H.B., Lee, C., Helgason, T., Buijs-Gladdines, J., Passier, M., van Wering, E.R., Veerman, A.J., Kamps, W.A., et al. (2006). The cryptic chromosomal deletion, del(11)(p12p13), as a new activation mechanism of LMO2 in pediatric T-cell acute lymphoblastic leukemia. Blood *108*, 3520–3529.

Van Vlierberghe, P., Pieters, R., Beverloo, H.B., and Meijerink, J.P. (2008a). Molecular-genetic insights in paediatric T-cell acute lymphoblastic leukaemia. Br. J. Haematol. *143*, 153–168.

Van Vlierberghe, P., van Grotel, M., Tchinda, J., Lee, C., Beverloo, H.B., van der Spek, P.J., Stubbs, A., Cools, J., Nagata, K., Fornerod, M., et al. (2008b). The recurrent SET-NUP214 fusion as a new HOXA activation mechanism in pediatric T-cell acute lymphoblastic leukemia. Blood *111*, 4668–4680.

Weir, B.A., Woo, M.S., Getz, G., Perner, S., Ding, L., Beroukhim, R., Lin, W.M., Province, M.A., Kraja, A., Johnson, L.A., et al. (2007). Characterizing the cancer genome in lung adenocarcinoma. Nature *450*, 893–898.

Weng, A.P., Ferrando, A.A., Lee, W., Morris, J.P., 4th, Silverman, L.B., Sanchez-Irizarry, C., Blacklow, S.C., Look, A.T., and Aster, J.C. (2004). Activating mutations of NOTCH1 in human T cell acute lymphoblastic leukemia. Science *306*, 269–271.

Weng, A.P., Millholland, J.M., Yashiro-Ohtani, Y., Arcangeli, M.L., Lau, A., Wai, C., Del Bianco, C., Rodriguez, C.G., Sai, H., Tobias, J., et al. (2006). c-Myc is an important direct target of Notch1 in T-cell acute lymphoblastic leukemia/lymphoma. Genes Dev. *20*, 2096–2109.

Wilker, P.R., Kohyama, M., Sandau, M.M., Albring, J.C., Nakagawa, O., Schwarz, J.J., and Murphy, K.M. (2008). Transcription factor Mef2c is required for B cell proliferation and survival after antigen receptor stimulation. Nat. Immunol. *9*, 603–612.

Zhu, B., and Gulick, T. (2004). Phosphorylation and alternative pre-mRNA splicing converge to regulate myocyte enhancer factor 2C activity. Mol. Cell. Biol. *24*, 8264–8275.