Integrated genome-wide genotyping and gene expression profiling reveals *BCL11B* as a putative oncogene in acute myeloid leukemia with 14q32 aberrations

Saman Abbas,[1] Mathijs A. Sanders,[1] Annelieke Zeilemaker,[1] Wendy M.C. Geertsma-Kleinekoort,[1] Jasper E. Koenders,[1] Francois G. Kavelaars,[1] Zabiollah G. Abbas,[1] Souad Mahamoud,[1] Isabel W.T. Chu,[1] Remco Hoogenboezem,[1] Justine K. Peeters,[1] Ellen van Drunen,[2] Janneke van Galen,[2] H. Berna Beverloo,[2] Bob Löwenberg,[1] and Peter J.M. Valk[1]

[1]Department of Hematology, Erasmus University Medical Center, Rotterdam; and [2]Department of Clinical Genetics, Erasmus University Medical Center, Rotterdam, the Netherlands

 doi:10.3324/haematol.2013.095604
Manuscript received on August 18, 2013. Manuscript accepted on January 16, 2014.
Correspondence: p.valk@erasmusmc.nl

SUPPLEMENTARY

Integrated genome-wide genotyping and gene expression profiling reveals *BCL11B* as a putative oncogene in acute myeloid leukemia with 14q32 aberrations

Saman Abbas[1], Mathijs Sanders[1], Annelieke Zeilemaker[1], Wendy M.C. Geertsma-Kleinekoort[1], Jasper E. Koenders[1], Francois G. Kavelaars[1], Zabiollah G.Abbas[1], Souad Mahamoud[1], Isabel W.T. Chu[1], Remco Hoogenboezem[1], Justine K. Peeters[1], Ellen van Drunen[2], Janneke van Galen[2], H. Berna Beverloo[2], Bob Löwenberg[1] and Peter J.M. Valk[1]

METHODS

Patients samples

After informed consent, bone marrow aspirates or peripheral blood samples of a representative cohort of AML patients were collected (MEC-2004-030). Eligible patients had a diagnosis of primary AML, confirmed by cytological examination of blood and bone marrow. All patients were treated according to the HOVON (Dutch-Belgian Hematology-Oncology Co-operative group) protocols (http://www.hovon.nl). Blasts and mononuclear cells were purified by Ficoll-Hypaque (Nygaard, Oslo, Norway) centrifugation and cryopreserved. The AML samples contained 80-100 percent blast cells after thawing, regardless of the blast count at diagnosis.

Genome-wide genotyping and gene expression profiling

Genome-wide genotyping data sets of 48 patients with various subtypes of AML were generated using Affymetrix 500K *Nsp*I/*Sty*I DNA mapping arrays and 89 patients with cytogenetically normal AML using Affymetrix 250K *Nsp*I or *Sty*I DNA Mapping arrays. High-molecular-weight DNA was isolated using the standard high salt procedures and the Affymetrix mapping arrays were used according to the protocol of the

manufacturer (Affymetrix, Santa Clara, CA). In brief, 250 ng of genomic DNA was digested with *Nsp*I or *Sty*I and ligated to an *Nsp*I or *Sty*I adapter using T4 DNA ligase (New England Biolabs, Ipswich, MA, USA). Samples were then amplified by PCR using TITANIUM Taq polymerase (Clontech, Mountain View, CA, USA). PCR products were pooled and purified using the Clontech purification kit and subjected to fragmentation using DNaseI. The DNA fragments were subsequently biotin-labeled with terminal deoxynucleotidyl transferase, hybridized on the array in a GeneChip® Hybridization Oven 640, and washed and stained in a GeneChip® Fluidics Station 450. Data was obtained using the GeneChip Scanner 3000 7G. Genotypes were calculated using BRLMM and copy numbers were assessed using dChip SNP [1]. The copy numbers of all AML samples were calculated using diploid references, i.e., 15 normal karyotype AML samples that were included in the series of 137 AML samples analyzed.

Gene expression profiles of the same AMLs were generated using Affymetrix HG-U133 plus 2.0, as described elsewhere [2]. These data sets are accessible at the NCBI Gene Expression Omnibus through GEO Series accession number GSE6891. Pearson correlation analyses was performed as described previously[3].

The genome-wide genotyping and gene expression profiling data sets were examined using SNPExpress [4]. SNPExpress is an easily accessible software tool to accurately analyze Affymetrix and Illumina SNP genotype calls, copy numbers, polymorphic copy number variations (CNVs) and Affymetrix gene expression in a combinatorial and efficient way (freely available at http://www.erasmusmc.nl/hematologie/SNPExpress).

Fluorescence In Situ Hybridization (FISH)

Freshly prepared slides from stored fixed cytogenetic suspensions with methanol/ acetic acid (3:1) (at -20 C) were used to carry out dual color fluorescence *in situ* hybridization (FISH) with BAC clones RP11-431B1, RP11-876E22, RP11-830F3, RP11-782I5, RP11-450C22, RP11-57E12, RP11-1069L3 and RP11-242A7 covering the BCL11B region and regions up- and downstream (BACPAC resources, Oakland, USA). Clone isolation and labeling were performed using biotin-16-dUTP and digoxigenin-11-dUTP (Roche Diagnostics Belgium, Vilvoorde, Belgium) according to

the manufacturer's protocol. The FISH analysis was performed as previously described [5]. Slides were embedded in DABCO/Vectashield containing 4', 6-diamidino-2-phenylindole (DAPI) as counterstain. In addition, whole chromosome paints for chromosome 6 and 8 (Cambio, Sanbio, Uden, The Netherlands) were used to verify the chromosomal identification as was first determined through the banding pattern provided by the DAPI counterstain. Slides were embedded in DABCO/Vectashield containing 40, 6-diamidino-2-phenylindole (DAPI) as counterstain. Fluorescent signals were visualized with an epi-fluorescence microscope (Zeiss, Axio-Imager Z1, Zeiss, Sliedrecht, the Netherlands) using Metasystems Ikaros software (Metasystems, Altlussheim, Germany). Both interphase nuclei and metaphase analysis was carried out to determine the presence or absence of the fluorescent signal and to determine its localization.

Targeted sequencing of the 14q32 genomic region

Library preparation and targeted resequencing was performed following the protocols as described previously [6]. In brief, high molecular weight DNA of AML #2301 and #7073 were sheared using a Covaris E210 waterbath sonificator. The *BCL11B* 14q32 - tel genomic region (chr14:93930247-105928955 (hg19)) was captured with a Roche/Nimblegen SeqCap EZ Choice XL Library. The captured region were subsequently paired-end sequenced using the Illumina HiSeq2000. The data has been analyzed using an in-house pipeline, including visualization with IGV, which identifies single nucleotide variants, small and large indels and copy number variations. The chromosomal breakpoints were identified using Breakdancer [7] in the 14q32 region and the partner chromosome. The genomic fusions were subsequently confirmed by Sanger sequencing.

Western blot analyses

For Western blot analyses, 100 micrograms of whole cell extracts (lysis buffer: 20 mM Tris pH=8, 137 mM NaCl, 50 mM NaF, 10 mM EDTA, 1% NP40, 10% Glycerol) were heated prior to electrophoresis for 5 minutes at 95^{o}C. Samples were separated on by 10% SDS-PAGE and transferred to Protran BA83 Nitrocellulose membranes (Whatman, Dessel, Germany). Membranes were then blocked in block-buffer (10% BSA (0.6%), 0.5M EDTA (1mM), TBS/0.05% Tween) and immunoblotted with affinity-purified rabbit polyclonal anti-BCL11b antibody (Novus Biologicals, Littleton, USA).

Immune complexes were detected by binding anti-mouse IgG conjugated to horseradish peroxidase (DAKO, Heverlee, Belgium) followed by the enhanced chemiluminescence assay (Amersham Bioscience, Piscataway, NJ) according to the manufacturer's recommendations. GAPDH was stained with primary affinity-purified rabbit polyclonal antibody (α-GAPDH FL-335) (Santa Cruz Biotechnology, California, USA).

DNA constructs and generation of BCL11B expressing 32D/GCSFR cells

A murine Bcl11b cDNA (kindly donated by Dorina Avram, Albany Medical Center, Albany, NY) was subcloned into pLXSN expression vector under control of 5' long terminal repeat (LTR) of the Moloney murine sarcoma virus (MoMSV) (Clontech, Mountainview, USA). Vector constructs were confirmed by nucleotide sequencing and retrovirally transfected into 32D cells that stably express human granulocyte colony-stimulating factor receptor (GCSF-R) [8] using Fugene transfection reagent (Roche, Indianapolis, USA). Cells were cultures in RPMI 1640 medium with 10% fetal calf serum (FCS), penicillin and streptomycin and supplemented with interleukin-3 (IL3, 25ng/ml) and selected with neomycin. To study proliferation and differentiation, 32D cells, transduced with murine *Bcl11b*, were plated in IL3 (25ng/ml) or GCSF (25 ng/ml), counted and assessed for granulocytic differentiation by morphological criteria. Morphological analysis was performed by microscopy on May-Grünwald-Giemsa-stained cytospins (Shandon Holland, Amsterdam, The Netherlands) using a Zeiss Axioskop microscope with 63x plan-apochromat objective. Pictures were taken with a Leica DC 500 camera.

References

1. Lin M, Wei LJ, Sellers WR, Lieberfarb M, Wong WH, Li C. dChipSNP: significance curve and clustering of SNP-array-based loss-of-heterozygosity data. Bioinformatics. 2004;20:1233-1240.
2. Verhaak RG, Wouters BJ, Erpelinck CA, Abbas S, Beverloo HB, Lugthart S, et al. Prediction of molecular subtypes in acute myeloid leukemia based on gene expression profiling. Haematologica. 2009;94:131-134.
3. Valk PJ, Verhaak RG, Beijen MA, Erpelinck CA, Barjesteh van Waalwijk van Doorn-Khosrovani S, Boer JM, et al. Prognostically useful gene-expression profiles in acute myeloid leukemia. N Engl J Med. 2004;350:1617-1628.

4. Sanders MA, Verhaak RG, Geertsma-Kleinekoort WM, Abbas S, Horsman S, van der Spek PJ, et al. SNPExpress: integrated visualization of genome-wide genotypes, copy numbers and gene expression levels. BMC Genomics. 2008;9:41.
5. von Bergh AR, van Drunen E, van Wering ER, van Zutven LJ, Hainmann I, Lonnerholm G, et al. High incidence of t(7;12)(q36;p13) in infant AML but not in infant ALL, with a dismal outcome and ectopic expression of HLXB9. Genes Chromosomes Cancer. 2006;45:731-739.

Supplementary table 1:
Top 50 genes correlating with *BCL11B* expression in AML

#	Probe set	Correlation coefficient	Gene
1	205456_at	0.83	CD3E
2	206804_at	0.82	CD3G
3	210915_x_at	0.81	TRBV19
4	213193_x_at	0.80	TRBV19
5	203828_s_at	0.80	IL32
6	204891_s_at	0.77	LCK
7	211796_s_at	0.77	TRBV21-1
8	204890_s_at	0.75	LCK
9	210031_at	0.71	CD3Z
10	206545_at	0.70	CD28
11	212400_at	0.70	C9orf132
12	205254_x_at	0.69	TCF7
13	205831_at	0.69	CD2
14	210439_at	0.69	ICOS
15	205798_at	0.68	IL7R
16	211893_x_at	0.67	CD6
17	205255_x_at	0.67	TCF7
18	213958_at	0.66	CD6
19	211339_s_at	0.66	ITK
20	208602_x_at	0.66	CD6
21	210479_s_at	0.64	RORA
22	205590_at	0.64	RASGRP1
23	211900_x_at	0.64	CD6
24	204777_s_at	0.64	MAL
25	220418_at	0.64	UBASH3A
26	217838_s_at	0.63	EVL
27	210426_x_at	0.62	RORA
28	219442_at	0.62	MGC3020
29	202761_s_at	0.61	SYNE2
30	210972_x_at	0.61	TRA@
31	203413_at	0.61	NELL2
32	210607_at	0.61	FLT3LG
33	206337_at	0.60	CCR7
34	209671_x_at	0.60	TRA@
35	206485_at	0.60	CD5
36	214470_at	0.58	KLRB1
37	202524_s_at	0.58	SPOCK2
38	213539_at	0.58	CD3D
39	210948_s_at	0.58	LEF1
40	214032_at	0.58	ZAP70
41	221558_s_at	0.58	LEF1
42	211902_x_at	0.58	TRA@
43	220485_s_at	0.57	SIRPB2
44	206966_s_at	0.57	KLF12
45	209670_at	0.56	TRAC
46	209604_s_at	0.56	GATA3
47	207651_at	0.56	GPR171
48	213906_at	0.56	MYBL1
49	205259_at	0.56	NR3C2
50	204638_at	0.55	ACP5