PAX5-ESRRB is a recurrent fusion gene in B-cell precursor pediatric acute lymphoblastic leukemia

Genome-wide screening using genotyping microarrays and next-generation sequencing has discovered genomic rearrangements associated with the clinical outcome of pediatric patients with acute lymphoblastic leukemia (ALL).[1] These findings have increased our understanding of pediatric ALL and have contributed to the identification of new molecular markers for diagnosis and risk stratification.[2,3] Recently, we discovered a novel *PAX5-ESRRB* fusion transcript by RNA-sequencing in a single patient with B-cell precursor acute lymphoblastic leukemia (BCP ALL) (GEOID: ALL_619) cytogenetically karyotyped with non-recurrent chromosomal aberrations, denoted "other".[4]

The 5'-partner in this fusion gene is *PAX5* [chr9p13], which encodes an essential B-cell transcription factor with a conserved DNA-binding motif known as the paired box. Paired box transcription factors are crucial regulators of tissue development and early cellular differentiation, and alteration of their expression is thought to promote tumorigenesis.[5,6] Deletions, amplifications and rearrangements involving *PAX5* occur in ~40% of pediatric BCP ALL cases.[7-9]

The Estrogen-Related Receptor Beta gene (*ESRRB*) [chr14q24] is a previously unknown 3'-fusion partner of *PAX5*, which encodes a protein similar to the estrogen receptor. ESRRB is an orphan nuclear receptor with an unknown endogenous ligand. It mediates self-renewal and pluripotency in embryonic stem cells, and is able to reprogram somatic cells to a pluripotent state by upregulating embryonic stem cell-specific genes.[10] Alterations of *ESRRB* have been implicated in impaired placental development and aberrant metabolic function in mice, however little is known regarding its role in cancer and it has not been previously associated with ALL.[11]

Using 450k DNA methylation array data (GEO:GSE49031)[12] available for ALL_619, copy number analysis (calculation of LogR from the probe signal intensities) revealed a deletion of chromosomal region 9p (chr9p) and a duplication of 14q (chr14q) corresponding to the *PAX5-ESRRB* fusion gene breakpoints. This suggested that the fusion resulted from an unbalanced t(9;14) rearrangement. To ascertain whether the *PAX5-ESRRB* fusion is recurrent in ALL, we screened 664 Nordic pediatric BCP ALL cases from the same dataset for concordant copy number changes on chr9p and chr14q. Three additional cases were identified with concordant deletions of chr9p and duplications of chr14q, within *PAX5* and *ESRRB* respectively, totaling <1% (4/664) (Table 1). No additional cases were identified after screening 78 patients with cytogenetically unclassified BCP ALL by RNA-sequencing (data not shown).

To validate the findings from the methylation array, DNA from the four ALL cases was subjected to copy number analysis on HumanOmni2.5M SNP arrays (Illumina, Inc). The breakpoints on chr9p and chr14q within *PAX5* and *ESRRB* were confirmed in all four cases, with the corresponding LogR and B-allele frequency values providing evidence of their presence in the dominant clone. The deletion of chr9p included the terminal region of *PAX5*, while the duplication on the distal region of chr14q suggested an unbalanced translocation on derivative chr9 (*Online Supplementary Figures S1 and S2*). Although there were indications of an abnormal chr9p in ALL_757, the exact alteration could not be detected by G-banding. Using fluorescent *in situ* hybridization analysis, a der(9)t(9;14) rearrangement was verified in ALL_619 and ALL_757 (Figure 1A). Besides our earlier study in which we reported *PAX5-ESRRB* for the first time,[4] a single additional case with der(9)t(9;14)(p13;q24) was found in the literature, but with no information other than the karyotype available.[13] We identified two breakpoints in the *PAX5* gene by copy number analysis. One breakpoint was found in intron 8 of *PAX5* in ALL_619 (breakpoint 1) and a second in intron 5 of *PAX5* in cases ALL_51, ALL_401 and ALL_757 (breakpoint 2). The breakpoint in *ESRRB* was located in intron 4 in all four cases (Figure 1B). Additionally, both ALL_51 and ALL_757 had a duplication of chromosome 5 and ALL_401 and ALL_757 shared a subclonal duplication of chromosome 13 (*Online Supplementary Figures S1* and *S2*). All four cases harbored homozygous deletions of CDKN2A, including loss of a small region on the non-rearranged chromosome 9 homolog. No other alterations known to be associated with ALL including *CRLF2, ETV6, ERG* and *BCR* were identified (*Online Supplementary Table S1*).

Expression of the *PAX5-ESRRB* fusion was verified by reverse-transcription polymerase chain reaction (RT-PCR) followed by Sanger sequencing in all cases (Figure 1B; *Online Supplementary Figure S3; Online Supplementary Table S2*). Fusion of *PAX5* exon 8 with *ESRRB* exon 5 (breakpoint 1) was confirmed in ALL_619, while a fusion between *PAX5* exon 5 and *ESRRB* exon 5 (breakpoint 2) was confirmed in ALL_51, ALL_401 and ALL_757 (Figure 1B; Table 1). Both gene fusion variants were in-frame and predicted to produce a functional PAX5-ESRRB protein. Regardless of the breakpoint, the fusion protein

Table 1. Clinical and cytogenetic information for four BCP ALL patients with 9p13 and 14q24 breakpoints.

Patient ID	Karyotype[a]	WBC[b] count	Age (years)	NOPHO protocol[c]	Event (time in months)	*PAX5-ESRRB* breakpoint[d]
ALL_619	45,XY,dic(7;12)(p11;p11)	1.4	16	NOPHO-2000, II	CR1 (69.2)	Breakpoint 1
ALL_51	46,XY	3.0	11	NOPHO-92, IR	CR1 (173.9)	Breakpoint 2
ALL_401	46,XX	16.0	7	NOPHO-2000, II	CR1 (97.4)	Breakpoint 2
ALL_757	47,XY,+5,add(9)(p2?1),der(20)t(2;20)(p2?1;q1?2)	1.1	16	NOPHO-2008, SR	CR1 (30.9)	Breakpoint 2

*CR1: continuous first remission. [a]Chromosomal aberrations determined at diagnosis by karyotyping. [b]White blood cell (WBC) count at diagnosis (x10[9]/L). [c]Treatment group according to the Nordic Society for Pediatric Hematology and Oncology (NOPHO) 1992-2008 protocols. II: intermediate intensive; IR: intermediate risk; SR: standard risk. [d]*PAX5-ESRRB *fusion variant. Breakpoint 1; fusion of* PAX5 *exon 8 (chr9:36,882,001) with* ESRRB *exon 5 (chr14:76,928,888), Breakpoint 2; fusion of* PAX5 *exon 5 (chr9:37,002,645) with* ESRRB *exon 5 (chr14:76,928,888). Genomic coordinates are based on GRCh37/hg19.*

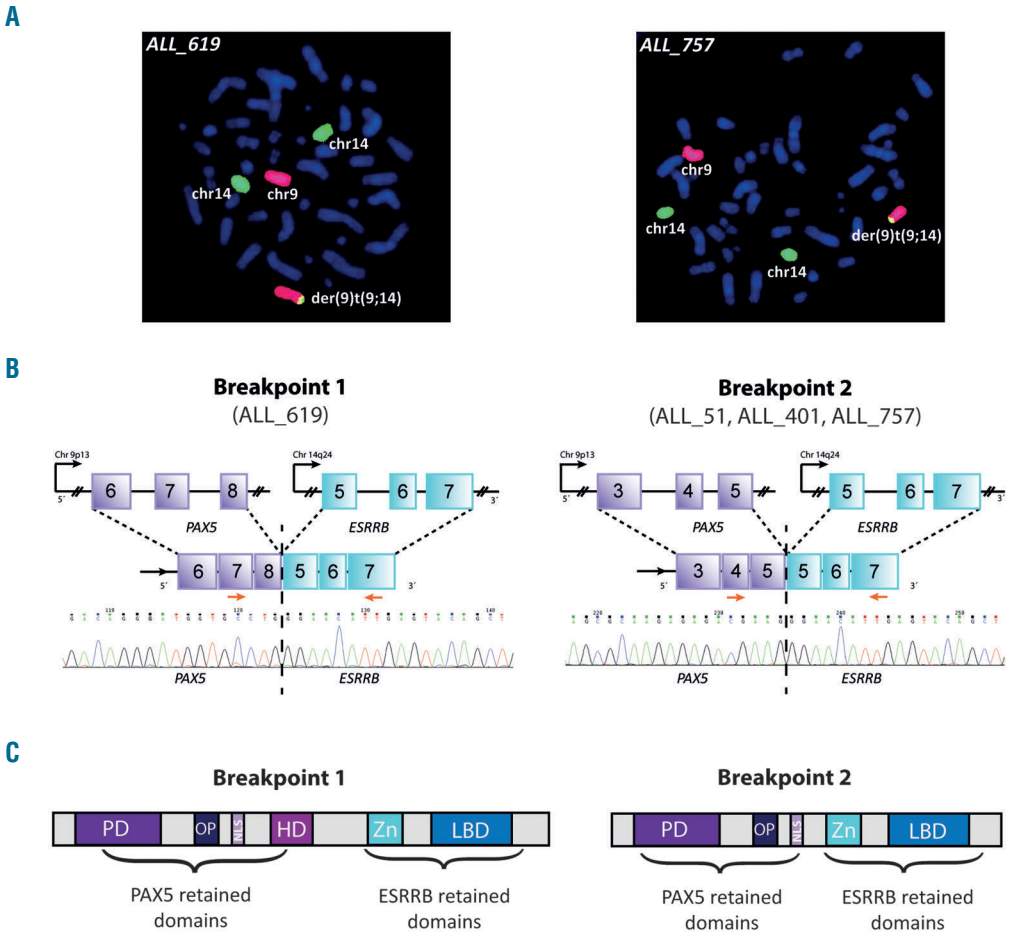

Figure 1. (A) Fluorescent *in situ* hybridization (FISH) analysis using whole chromosome paint for chr9 (red) and chr14 (green) in ALL_619 and ALL_757. The structural rearrangement between chr9 and chr14 giving rise to the der(9)t(9;14)*PAX5-ESRRB* was verified in both cases. (B) Schematic structure of the two breakpoint variants and representative Sanger sequencing traces. Breakpoint 1 occurs as a result of the fusion of exon 8 of *PAX5* (chr9:36882001, hg19) with exon 5 of *ESRRB* (chr14:76928888, hg19) in patient ALL_619. Breakpoint 2 is a fusion between exon 5 of *PAX5* (chr9:37002645, hg19) and exon 5 of *ESRRB* (chr14:76928888, hg19) in patients ALL_51, ALL_401 and ALL_757. PCR primers were designed against *PAX5* exon 7 and *ESRRB* exon 7 for breakpoint 1 and *PAX5* exon 4 and *ESRRB* exon 7 for breakpoint 2, as indicated by red arrowheads. (C) Schematic representation of both PAX5-ESRRB fusion protein variants. PD: paired domain; OP: octapeptide domain; NLS: nuclear localization signal; HD: partial homeodomain; Zn: zinc finger in nuclear hormone receptors, type c4; LBD: ligand binding domain of hormone receptors.

retains functionally important domains, such as the paired box domain, the highly conserved octapeptide motif for protein-protein interactions, and the nuclear localization signal of PAX5 (Figure 1C). Retention of the paired box domain is a key feature of all PAX5 fusion proteins, required for binding PAX5 to its target regions. The fused ESRRB protein retains the ligand binding domain of hormone receptors and a zinc finger motif in the DNA binding domain (Figure 1C). Consistent with the fluorescent *in situ* hybridization analysis demonstrating an unbalanced translocation der(9)t(9;14)*PAX5-ESRRB*, no evidence for expression of the reciprocal *ESRRB-PAX5* fusion gene was found by RNA-sequencing in ALL_619.

Gene expression levels of *PAX5* and *ESRRB* in cases with the der(9)t(9;14)*PAX5-ESRRB* were assessed by RNA-sequencing (only for ALL_619) and by quantitative real-time PCR (qRT-PCR) in all four *PAX5-ESRRB* cases (Figure 2A,B; *Online Supplementary Figure S4; Online Supplementary Table S2*). RNA-sequencing data for *ESRRB* revealed a high expression in ALL_619 compared to that in 41 BCP ALL cases and five normal samples of whole blood sorted CD19$^+$ B cells (Figure 2A; *Online Supplementary Table S3*). The observed *ESRRB* expression in ALL_619 derived from the retained *ESRRB* region in the der(9)t(9;14) translocation as seen in the RNA-sequencing data (*Online Supplementary Figure S5*). Similarly, qRT-PCR analysis of *ESRRB* in the four *PAX5-ESRRB* cases showed higher expression of fused *ESRRB* compared to that in normal CD19$^+$ B cells (Figure 2B). Expression of wild-type *ESRRB* was not detectable by qRT-PCR in any of the *PAX5-ESRRB* cases or in the CD19$^+$ B cells. Collectively, this suggests that expression of fused *ESRRB* is driven by the translocated *PAX5* promoter in cases with the der(9)t(9;14) rearrangement, retaining key functional domains from both proteins.

Next, we used DNA methylation data to establish whether cases with the der(9)t(9;14) acquire specific methylation changes.[12] Methylation data from the four *PAX5-ESRRB* cases was analyzed together with data from 24 additional ALL cases representing six cytogenetic subtypes of BCP ALL; dic(9;20), HeH, t(1;19)*TCF3-PBX1*, t(12;21)*EVT6-RUNX1*, t(9;22)*BCR-ABL1* and 11q23/MLL (*Online Supplementary Table S4*). Six additional BCP ALL cases from the same DNA methylation data set known to

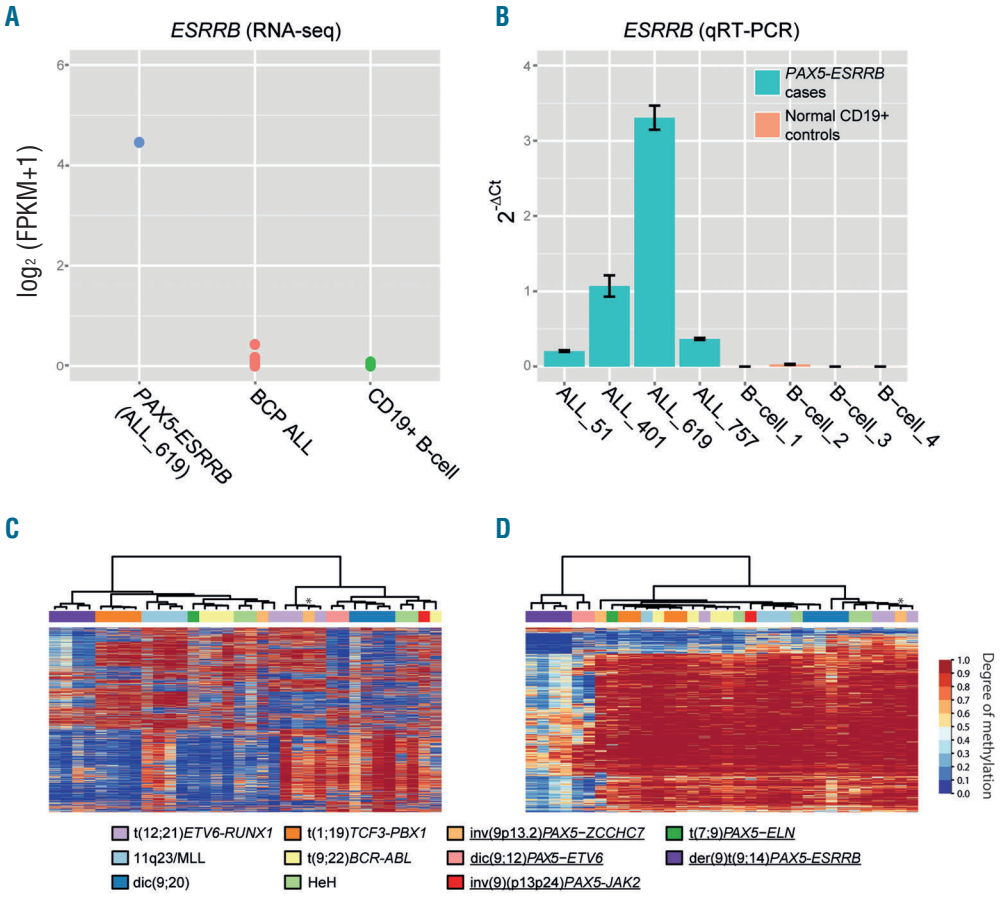

Figure 2. (A) RNA-sequencing expression levels of *ESRRB* in ALL_619 compared to 41 additional BCP ALL cases and five normal CD19⁺ B cells. All samples were sequenced on the HiSeq2000/2500 (Illumina, Inc) to a minimum of 20x10⁶ read pairs. Alignment of sequence reads and quantification of expression levels were performed with TopHat and Cufflinks. Expression levels were measured as "fragments per kilobase of exon model per million mapped reads" (FPKM) using a pseudocount of +1 added to the data set, and plotted on the y-axis as $\log_2$(FPKM+1). The observed *ESRRB* expression in ALL_619 represents the fused *ESRRB* region, as a result of the der(9)t(9;14)*PAX5-ESRRB*. Very low or no expression of *ESRRB* was detected in the BCP ALL group and in the normal controls. (B) qRT-PCR analysis of fused *ESRRB* expression in the four *PAX5-ESRRB* cases compared to normal CD19⁺ B cells. Samples were run in replicates and the average expression level ($2^{-\Delta Ct}$) of *ESRRB* in each sample was calculated after normalization to *GAPDH* expression (ΔCt = Ct(*ESRRB*)-Ct(*GAPDH*)). (C) Unsupervised hierarchical clustering based on DNA methylation levels of the four cases with the der(9)t(9;14)*PAX5-ESRRB* translocation, 24 individuals representing six additional BCP ALL subtypes and six cases with different *PAX5* fusions (underlined in the color key below panels C and D). The *PAX5-ESRRB* cases form a distinct group separated from the other BCP ALL subtypes and the *PAX5* chimeras. The 2000 most variable CpG sites are shown in the heatmap in which each row represents one CpG site and each column represents one ALL sample. The color key for the subtype of individual ALL cases is shown below the heatmap and the color key for the DNA methylation levels is shown to the right of panel (D). (D) Heatmap of 823 differentially methylated CpG sites identified between the *PAX5-ESRRB* group (n=4) and patients belonging to six common BCP ALL subtypes (n=24). Clustering of *PAX5* fusion positive cases (underlined in the color key) and the six BCP ALL subtypes is based on the identified DMC. *PAX5-ESRRB* reveal a hypomethylated pattern compared to the six BCP ALL subtypes and *PAX5-ELN* (n=1), *PAX5-JAK2* (n=1), *PAX5-ETV6* (n=2) and *PAX5-ZCCHC7* (n=2). The asterisk in the hierarchical clustering dendrogram highlights one *PAX5-ZCCHC7* case (ALL_299) that also expresses t(12;21)*ETV6-RUNX1*. This case subsequently clustered together with the other four t(12;21)*ETV6-RUNX1* cases. Each row represents a CpG site and each column represents a single ALL case.

harbor *PAX5-ELN* (n=1), *PAX5-JAK2* (n=1), *PAX5-ETV6* (n=2) or *PAX5-ZCCHC7* (n=2), were also included. Using unsupervised hierarchical clustering across the 2000 most variable CpG sites, the der(9)t(9;14)*PAX5-ESRRB* cases formed a distinct cluster separated from the other subtypes including the other *PAX5* chimeras (Figure 2C). We then interrogated differentially methylated CpG sites (DMC) genome-wide between the *PAX5-ESRRB* group and the six BCP ALL subtypes. Based on a maximum false discovery rate of 10% and an absolute mean methylation difference >20%, 825 CpG sites annotated to 648 unique genes were differentially methylated (Figure 2D; *Online Supplementary Table S5*). The majority of the detected DMC (96.8%) were hypomethylated in the der(9)t(9;14)*PAX5-ESRRB* group compared to the six BCP ALL subtypes (Figure 2D). Similarly, the identified DMC were more hypomethylated in the *PAX5-ESRRB* group than in *PAX5-ELN*, *PAX5-JAK2* and *PAX5-ZCCHC7*, which clustered with the other BCP ALL subtypes. The *PAX5-ETV6* cases were more hypomethylated and clustered close to the *PAX5-ESRRB* cases. Genes containing DMC were enriched to the Wnt/Beta catenin pathway (P value=2.25x10^{-5}), affecting genes associated with leukemia such as *CREBBP, LEF1* and *TCF7L2* (*Online Supplementary Figure S6*). Functional annotation of DMC to nearby genes revealed enrichment towards genes involved in hematopoietic functions; furthermore, binding motif analysis revealed an enrichment of the DMC in relation to the leukemia-associated transcription factors EGR1 (P value=3.68x10^{-5}) and EBF1 (P value=6.05x10^{-5})

motifs (*Online Supplementary Figure S7*). Taken together, these analyses demonstrate that the der(9)t(9;14) *PAX5-ESRRB*, like other recurrent groups of BCP ALL, is associated with distinct molecular features.

This is the first study to describe the recurrence of a *PAX5* fusion involving the estrogen related receptor *ESRRB* and its molecular characteristics in ALL. Although the cause of frequent *PAX5* associated rearrangements is still unclear, it has become evident that *PAX5* chimeras harbor diverse molecular features driven by the 3´partner gene.[14] BCP ALL with *PAX5-JAK2* fusions are known to display a different expression signature compared to other *PAX5* fusions, characterized by upregulation of JAK-STAT target genes, suggesting that *PAX5* fusions interact not only with *PAX5* target genes but also various downstream targets depending on the 3´partner.[2]

Previously, we have demonstrated the utility of RNA-sequencing to detect novel fusion genes such as *PAX5-ESRRB* in childhood ALL.[4] Recently, the recurrence of a novel fusion gene, *EP300-ZNF384*, was first detected using RNA-sequencing and subsequently screened for by RT-PCR.[15] By combining RNA-sequencing with copy number analysis data from high-density single nucleotide polymorphism arrays and Human Methylation 450 Bead Chips, we identified a group of rare cases harboring the same gene fusion within a large ALL cohort, in which screening by RT-PCR would not be feasible. While genotyping arrays can be useful in screening for unbalanced rearrangements, such as the der(9)t(9;14), they are not suitable for the detection of balanced rearrangements.

In conclusion, we describe the recurrent chromosomal rearrangement der(9)t(9;14)(p13.2;q24.3), an unbalanced translocation resulting in an in-frame fusion transcript involving *PAX5* and *ESRRB* in four BCP ALL cases. We show that der(9)t(9;14)*PAX5-ESRRB* is associated with a distinct hypomethylated pattern compared to other BCP ALL subtypes including other *PAX5* fusions analyzed herein. Whether patients with der(9)t(9;14)*PAX5-ESRRB* share additional biological or clinical features important for diagnosis and therapy warrants further investigation.

Yanara Marincevic-Zuniga,[1] Vasilios Zachariadis,[2] Lucia Cavelier,[3] Anders Castor,[4] Gisela Barbany,[2] Erik Forestier,[5] Linda Fogelstrand,[6,7] Mats Heyman,[8] Jonas Abrahamsson,[9] Gudmar Lönnerholm,[10] Ann Nordgren,[2] Ann-Christine Syvänen,[1] and Jessica Nordlund[1]

YM-Z and VZ contributed equally to this work.

[1]Department of Medical Sciences, Molecular Medicine and Science for Life Laboratory, Uppsala University; [2]Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm; [3]Department of Immunology, Genetics and Pathology, Uppsala University; [4]Pediatric Oncology/Hematology, Skåne University Hospital, Lund; [5]Department of Medical Biosciences, University of Umeå; [6]Department of Clinical Chemistry and Transfusion Medicine, Institute of Biomedicine, Sahlgrenska Academy at University of Gothenburg; [7]Laboratory of Clinical Chemistry, Sahlgrenska University Hospital, Gothenburg; [8]Childhood Cancer Research Unit, Karolinska Institutet, Astrid Lindgren Children's Hospital, Karolinska University Hospital, Stockholm; [9]Department of Pediatrics, Queen Silvia Children's Hospital, Gothenburg; and [10]Department of Women's and Children's Health, Pediatric Oncology, Uppsala University, Sweden

Funding: this work was supported by grants from the Swedish Foundation for Strategic Research (RBc08-008; ACS, GL), the Eric, Karin, and Gösta Selanders Foundation (JN), the Swedish Cancer Society (CAN2010/592; ACS), the Swedish Childhood Cancer Foundation (11098; ACS), the Swedish Research Council for Science and Technology (90559401; ACS), and joint funding from the Swedish Research Councils Forte, Formas, Vinnova and VR (259-2012-23; ACS) for epigenetics.

Acknowledgments: we would like to thank Christofer Bäcklin and Johan Dahlberg for assistance with data analysis and our colleagues from NOPHO and the patients and families who contributed to this study.

Correspondence: jessica.nordlund@medsci.uu.se
doi:10.3324/haematol.2015.132332

Key words: ESRRB, PAX5, *pediatric acute lymphoblastic leukemia, RNA-sequencing, fusion genes.*

Information on authorship, contributions, and financial & other disclosures was provided by the authors and is available with the online version of this article at www.haematologica.org.

References

1. Pui CH. Genomic and pharmacogenetic studies of childhood acute lymphoblastic leukemia. Front Med. 2015;9(1):1-9.
2. Schinnerl D, Fortschegger K, Kauer M, et al. The role of the Janus-faced transcription factor PAX5-JAK2 in acute lymphoblastic leukemia. Blood. 2015;125(8):1282-1291.
3. Roberts KG, Li Y, Payne-Turner D, et al. Targetable kinase-activating lesions in Ph-like acute lymphoblastic leukemia. N Engl J Med. 2014;371(11):1005-1015.
4. Nordlund J, Backlin CL, Zachariadis V, et al. DNA methylation-based subtype prediction for pediatric acute lymphoblastic leukemia. Clin Epigenetics. 2015;7(1):11.
5. Nutt SL, Heavey B, Rolink AG, Busslinger M. Commitment to the B-lymphoid lineage depends on the transcription factor Pax5. Nature. 1999;401(6753):556-562.
6. Muratovska A, Zhou C, He S, Goodyer P, Eccles MR. Paired-box genes are frequently expressed in cancer and often required for cancer cell survival. Oncogene. 2003;22(39):7989-7997.
7. Mullighan CG, Goorha S, Radtke I, et al. Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. Nature. 2007;446(7137):758-764.
8. Nebral K, Denk D, Attarbaschi A, et al. Incidence and diversity of PAX5 fusion genes in childhood acute lymphoblastic leukemia. Leukemia. 2009;23(1):134-143.
9. Denk D, Bradtke J, Konig M, Strehl S. PAX5 fusion genes in t(7;9)(q11.2;p13) leukemia: a case report and review of the literature. Mol Cytogenet. 2014;7(1):13.
10. Feng B, Jiang J, Kraus P, et al. Reprogramming of fibroblasts into induced pluripotent stem cells with orphan nuclear receptor Esrrb. Nat Cell Biol. 2009;11(2):197-203.
11. Luo J, Sladek R, Bader JA, Matthyssen A, Rossant J, Giguere V. Placental abnormalities in mouse embryos lacking the orphan nuclear receptor ERR-beta. Nature. 1997;388(6644):778-782.
12. Nordlund J, Backlin CL, Wahlberg P, et al. Genome-wide signatures of differential DNA methylation in pediatric acute lymphoblastic leukemia. Genome Biol. 2013;14(9):r105.
13. Heerema NA, Nachman JB, Sather HN, et al. Deletion of 7p or monosomy 7 in pediatric acute lymphoblastic leukemia is an adverse prognostic factor: a report from the Children's Cancer Group. Leukemia. 2004;18(5):939-947.
14. Fortschegger K, Anderl S, Denk D, Strehl S. Functional heterogeneity of PAX5 chimeras reveals insight for leukemia development. Mol Cancer Res. 2014;12(4):595-606.
15. Gocho Y, Kiyokawa N, Ichikawa H, et al. A novel recurrent EP300-ZNF384 gene fusion in B-cell precursor acute lymphoblastic leukemia. Leukemia. 2015 May 6. [Epub ahead of print].