PHF6 mutations in T-cell acute lymphoblastic leukemia

Pieter Van Vlierberghe1–3,21, Teresa Palomero1,4,21, Hossein Khiabanian5, Joni Van der Meulen2, Mireia Castillo4, Nadine Van Roy2, Barbara De Moerloose6, Jan Philippé7, Sara González-García8, María L Toribio8, Tom Taghon7, Linda Zuurbier3, Barbara Cauwelier9, Christine J Harrison10, Claire Schwab10, Markus Pisecker11, Sabine Strehl11, Anton W Langerak12, Jozef Gecz13,14, Edwin Sonneveld15, Rob Pieters3,15, Elisabeth Paietta16, Jacob M Rowe17, Peter H Wiernik16, Yves Benoit6, Jean Soulier18, Bruce Poppe2, Xiaopan Yao19, Carlos Cordon-Cardo4, Jules Meijerink3, Raul Rabadan5, Frank Speleman2,22 & Adolfo Ferrando1,4,20,22

1Institute for Cancer Genetics, Columbia University Medical Center, New York, New York, USA. 2Center for Medical Genetics, Ghent University Hospital, Ghent, Belgium. 3Department of Pediatric Oncology/Hematology, Erasmus MC, Rotterdam, The Netherlands. 4Department of Pathology, Columbia University Medical Center, New York, New York, USA. 5Center for Computational Biology and Bioinformatics, Columbia University, New York, New York, USA. 6Department of Pediatric Hemato-Oncology, Ghent University Hospital, Ghent, Belgium. 7Department of Clinical Chemistry, Immunology and Microbiology, Ghent University Hospital, Ghent, Belgium. 8Centro de Biología Molecular “Severo Ochoa”, Consejo Superior de Investigaciones Científicas (CSIC), Universidad Autónoma de Madrid (UAM), Madrid, Spain. 9Department of Hematology, Hospital St-Jan, Bruges, Belgium. 10Leukaemia Research Cytogenetics Group, Northern Institute for Cancer Research, Newcastle University, Newcastle, UK. 11Children’s Cancer Research Institute, St. Anna Kinderkrebsforschung, Vienna, Austria. 12Department of Immunology, Erasmus MC, Rotterdam, The Netherlands. 13Department of Genetics and Molecular Pathology, University of Adelaide, Adelaide, Australia. 14Department of Pediatrics, University of Adelaide, Adelaide, Australia. 15On behalf of the Dutch Childhood Oncology Group (DCOG), The Hague, The Netherlands. 16Montefiore Medical Center North, Bronx, New York, USA. 17Rambam Medical Center and Technion, Israel Institute of Technology, Haifa, Israel. 18Hematology Laboratory APHP, INSERM U944, Hôpital Saint Louis, Paris, France. 19Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute, Boston, Massachusetts, USA. 20Department of Pediatrics, Columbia University Medical Center, New York, New York, USA. 21These authors contributed equally to this work. 22These authors jointly directed this work. Correspondence should be addressed to A.F. (af2196@columbia.edu).

Abstract

Tumor suppressor genes on the X chromosome may skew the gender distribution of specific types of cancer1,2. T-cell acute lymphoblastic leukemia (T-ALL) is an aggressive hematological malignancy with an increased incidence in males3. In this study, we report the identification of inactivating mutations and deletions in the X-linked plant homeodomain finger 6 (PHF6) gene in 16% of pediatric and 38% of adult primary T-ALL samples. Notably, PHF6 mutations are almost exclusively found in T-ALL samples from male subjects. Mutational loss of PHF6 is importantly associated with leukemias driven by aberrant expression of the homeobox transcription factor oncogenes TLX1 and TLX3. Overall, these results identify PHF6 as a new X-linked tumor suppressor in T-ALL and point to a strong genetic interaction between PHF6 loss and aberrant expression of TLX transcription factors in the pathogenesis of this disease.

T-ALL is an aggressive malignancy in which multiple genetic defects collaborate in the transformation of T-cell progenitors4,5. Notably, T-ALL has a threefold higher incidence in males3, whereas other immature hematological tumors such as precursor B-lineage ALL are equally frequent in males and females3.

To identify a possible X-linked tumor suppressor in T-ALL, we performed an X-chromosome-targeted mutational analysis in tumor DNA samples from 12 males with T-ALL. For each sample, we performed in-solution DNA capture of 7,674 regions encompassing 3,045,708 nucleotides corresponding to 5,215 X-chromosome exons using the Agilent Sure Select oligonucleotide capture system6. DNA samples enriched for X-chromosome exons were then analyzed by next-generation sequencing using the SOLiD 3 platform from Applied Biosystems. This analysis identified 66 candidate previously uncharacterized nonsynonymous single-nucleotide variants and 7 positions with high confidence calls for containing complex variants such as insertions or deletions (Fig. 1a). Dideoxynucleotide DNA sequencing of PCR products encompassing affected exons confirmed the presence of 92% (61/66) of these single-nucleotide variants and 57% (4/7) of the more complex variants, including 2 insertions and 2 deletions (Supplementary Tables 1 and 2). Sequence analysis of paired DNA samples obtained at the time of clinical remission showed that most of these variants corresponded to previously unreported germline polymorphisms. However, and most notably, we also identified three somatically acquired changes corresponding to two nonsynonymous single-nucleotide substitutions (A902G T300A and A990G H330R) and a frameshift-creating insertion of five nucleotides (124_125insAGGCA, H43fs) in the PHF6 gene (Fig. 1a).

In a complementary approach we analyzed X-chromosome array comparative genome hybridization (array-CGH) data from 246 primary T-ALL samples (179 male and 67 female) in a multicenter setting. These analyses revealed the presence of recurrent deletions in chromosomal band Xq26 in 8 out of 246 (~3%) T-ALL samples (Table 1). For three del(X)(q26)-positive T-ALL samples, we performed array-CGH analysis against the corresponding remission material, which showed that these Xq26 deletions were somatically acquired leukemia-associated genetic events (Table 1). Reanalysis of all eight del(X)(q26)-positive T-ALL samples on a custom high-resolution X-chromosome oligonucleotide array (Fig. 1b,c) narrowed down the common minimally deleted region to an area of 80 kb containing the PHF6 gene. Consistently, quantitative PCR analysis confirmed loss of the PHF6 locus in the del(X)(q26)-positive cases (Fig. 1d). The convergent findings of our X-chromosome exon mutation analysis and analysis of copy

number alterations by array-CGH thus identified PHF6 as a new tumor suppressor gene mutated and deleted in T-ALL.

PHF6 encodes a plant homeodomain (PHD) factor containing four nuclear localization signals and two imperfect PHD zinc-finger domains7 with a proposed role in controlling gene expression7. Notably, inactivating mutations in PHF6 cause Börjeson-Forssman-Lehmann syndrome (MIM301900), a relatively uncommon type of X-linked familial syndromic mental retardation that has not been associated with increased incidence of T-ALL7–9. Quantitative RT-PCR analysis demonstrated ubiquitous expression of PHF6 transcripts in human tissues, with highest levels of expression in thymus, ovary and thyroid, and moderate levels of expression in spleen, testes and adipose tissue (Supplementary Fig. 1). Consistent with these results, PHF6 was readily detected by immunohistochemistry in mouse thymus (Supplementary Fig. 1). Finally, quantitative RT-PCR analysis of human thymocyte populations at different stages of development showed variable levels of PHF6 expression, with marked upregulation of PHF6 transcripts in CD4/CD8 double-positive cells (Supplementary Fig. 1).

Mutational analysis of PHF6 in an extended panel of pediatric and adult T-ALL primary samples identified truncating or missense mutations in PHF6 in 38% (16/42) of adult and ~16% (14/89) of pediatric T-ALL samples (Fig. 2a and Table 1). In all available cases (7/30), analysis of matched buccal and/or bone marrow remission genomic DNA confirmed the somatic origin of PHF6 mutations (4/21 frameshift mutations and 3/9 missense mutations) (Fig. 2b and Table 1). Finally, no mutations in PHF6 were identified in DNA samples from B-lineage ALL (n = 62), suggesting that mutational loss of PHF6 in lymphoid tumors could be restricted to T-ALL.

Nonsense and frameshift mutations accounted for 70% (21/30) of all PHF6 mutations identified in our series and were evenly distributed throughout the gene. Missense mutations accounted for the remaining 30% (9/30) of PHF6 lesions and recurrently involved codon C215 and the second zinc-finger domain of the protein (Fig. 2a). DNA sequence analysis of PHF6 in a panel of 15 well-characterized T-ALL cell lines (Supplementary Table 3) showed the presence of truncating mutations in PHF6 in the DND41, HPB-ALL and T-ALL1 cell lines. Protein blot analysis and immunohistochemical staining demonstrated robust expression and nuclear localization of PHF6 in PHF6 wild-type tumors and complete loss of PHF6 protein in T-ALL cell lines harboring mutations in PHF6 (Fig. 2c,d).

PHD finger-containing proteins have been implicated in numerous cellular functions, including transcriptional regulation and in some instances as specialized reader modules that recognize the methylation status of histone lysine residues10. In addition, PHF6 has been

reported to be phosphorylated during mitosis11 and by the ATM and ATR kinases upon DNA damage12, which suggests a dynamic regulation of PHF6 during cell cycle and DNA repair. Consistent with this notion, short hairpin RNA (shRNA) knockdown of PHF6 resulted in increased levels of phosphorylated H2AX (γ-H2AX), a post-translational modification associated with the presence of DNA double-strand breaks13 (Fig. 2e).

Sex determination in humans is controlled by differential representation of the X and Y chromosomes, with the presence of an XY pair in males and two copies of the X chromosome in females. The presence of numerous genes in the nonautosomal region of the X chromosome could result in a genetic imbalance between male and female cells, which is compensated for by random chromosomal inactivation of one copy of the X chromosome in female cells14. However, allelic-expression analysis has shown that some genes can escape X-chromosome inactivation in certain tissues1,2,15. To test the possibility that PHF6 could escape X-chromosome inactivation in T-ALL cells, we performed allelic-expression analysis of a silent SNP (rs17317724) located in the 3′ UTR of PHF6 in lymphoblasts from three informative samples from females with T-ALL. In each of these samples, PHF6 was monoallelically expressed, suggesting that biallelic expression of PHF6 is not commonly found in T-ALL (Fig. 3a). Most notably, we found that PHF6 mutations are almost exclusively found in samples from males with T-ALL. PHF6 mutations were present in 32% (29/92) of males and in only ~2.5% (1/39) of females ($P < 0.001$; Fig. 3b and Supplementary Table 4). Moreover, all eight PHF6 deletions identified by array-CGH analysis were found in samples from males with T-ALL, and each of the three cell lines with mutations in PHF6 were derived from males with T-ALL.

Immunohistochemical analysis of PHF6 expression in wild-type primary T-ALL samples showed positive PHF6 immunostaining (n = 5; three males and two females), whereas cases with PHF6-truncating mutations (n = 4) (Fig. 3c) or a point mutation in C215 (C215F) were negative for PHF6 protein expression (Fig. 3c). In contrast, primary T-ALL cells harboring a PHF6 point mutation in the PHD2 domain (T300A) were positive for PHF6 protein expression (Fig. 3c). Overall, these results suggest that truncating mutations and point mutations in C215 impair PHF6 expression, whereas amino acid substitutions in the PHD2 domain of PHF6 may selectively impair the tumor suppressor function of this protein.

Leukemic transformation of immature thymocytes is the result of a multistep process involving numerous genetic abnormalities, which can be associated with different clinical features, including age and prognosis. Notably, PHF6 mutations were significantly more prevalent in adult T-ALL patients (16/42; 38%) than in pediatric patients (14/89; 16%) ($P = 0.005$;

Fig. 4a). Detailed genetic information was available for T-ALL cases treated in Dutch Childhood Oncology Group (DCOG) clinical trials (n = 65) (Supplementary Table 5). In this cohort, PHF6 mutations were significantly associated with the aberrant expression of TLX1 and TLX3 ($P < 0.005$; Fig. 4b and Supplementary Table 5), two related oncogenes activated by chromosomal translocations in T-ALL16–18. No significant associations were observed between PHF6 mutations and NOTCH1, FBXW7 or PTEN mutations in either pediatric (n = 65) or adult (n = 34) T-ALL cohorts (Supplementary Tables 5 and 6). Overall survival in PHF6 wild-type children with T-ALL treated on DCOG protocols19 was 65% (33/51) vs. 71% (10/14) for PHF6-mutated cases (log-rank $P = 0.71$) (Fig. 4c). Overall survival in PHF6 wild-type adults with T-ALL treated in the Eastern Cooperative Oncology Group ECOG-2993 clinical trial was 36% (7/12) vs. 58% (8/22) for PHF6-mutated samples (log-rank $P = 0.24$) (Fig. 4d).

Overall, these results identify PHF6 as a new X-linked tumor suppressor gene and imply a specific interaction between the oncogenic programs activated by aberrant expression of TLX transcription factors and mutational loss of PHF6 in the pathogenesis of T-ALL.

Methods

Clinical samples and cell lines.

Leukemic DNA and cryopreserved lymphoblast samples were provided by collaborating institutions in the United States (ECOG), The Netherlands (DCOG), France (Hôpital Saint-Louis, Paris), Austria (Children's Cancer Research Institute, St. Anna Kinderkrebsforschung, Vienna) and Belgium (Department of Pediatric Hemato-Oncology, Ghent University Hospital, Ghent; Department of Hematology, Hospital St.-Jan, Bruges). All samples were collected in clinical trials with informed consent and under the supervision of local institutional review board committees. The subjects' parents or their legal guardians provided informed consent to use leftover material for research purposes according to the Declaration of Helsinki. T-cell phenotype was confirmed by flow cytometry. Survival analysis was performed in pediatric T-ALL samples from DCOG trials ALL7, ALL8 and ALL9 (ref. 19) and from adults with T-ALL treated in the ECOG2993 clinical trial20.

Jurkat and PF382 cells were obtained from the American Type Culture Collection. The ALL-SIL, HPB-ALL and T-ALL-1 cell lines were from the DSMZ repository (the German national resource center for biological material). The DND41 cell line was a gift from A.T. Look (Dana-Farber Cancer Institute). T-ALL cell lines were cultured in RPMI 1640 medium supplemented with 10% fetal

bovine serum, 100 U/ml penicillin G and 100 µg/ml streptomycin at 37 °C in a humidified atmosphere under 5% CO2.

X chromosome exome capture and next-generation sequencing.

Libraries of synthetic biotinylated RNA oligonucleotides (baits) targeting the X-chromosome exons were obtained from Agilent Technologies. The targeted region includes 5,217 exons for a total of 3 megabases and is designed to capture 85% of the exons on the human X chromosome. Fragment libraries using 2–4 µg of genomic DNA as starting material were prepared following the SOLiD standard library preparation protocol with some modifications, including the use of shortened adaptors and a precapture amplification for six cycles. A total of 500 ng of captured library were hybridized with the baits for 42 h, washed and eluted using the protocol provided by Agilent. The resulting captured DNA was amplified using Herculase II Fusion DNA polymerase (Agilent) for 18 cycles. Enrichment in the targeted regions was calculated by real-time PCR quantification of single exons within four X-chromosome loci (ARSF, OTC, NAP1L3 and SOX3), which showed an average enrichment of 400- to 1,200-fold for the different loci across the 12 different samples. After library quantification by real-time PCR, the amplified captured libraries were subjected to emulsion PCR and sequenced following standard SOLiD 3 protocols by depositing 10–15 million beads per sample using an eight-region mask.

Data analysis SOLiD3 ultradeep sequencing.

A reference genome of the captured regions was created based on the March 2006 human reference sequence (hg18). To map the sequence data into this reference genome, we used the SHRiMP algorithm with its default parameters21. SOLiD platform employs a two-base encoding system, where a single variation in the color space solely indicates a sequencing error and two consecutive variations in the color space point to a base change in the nucleotide-space. In our analysis we included only the reads with a maximum number of two color-space mismatches that are also uniquely mapped to the reference genome. An average 90.1% of the reference genome is covered in the 12 samples, where the mean depth is 42 per base. Less restrictive filtering increases the false-positive rate of candidate genomic variants without improving the coverage to any great extent. We found 66 candidates of exonic nonsynonymous single-nucleotide variation by requiring each variation to be reported in a minimum 75% of at least three reads mapping to its position. These candidates exclude the previously reported SNPs in the human genome. Using ParMap, an algorithm specifically developed to identify small deletions and insertions (along with their nucleotide sequence) through statistical analysis of partially mapped reads22, we identified seven candidates of such complex variations. Validation

of the next-generation sequencing results was done by Sanger sequencing of PCR amplified exons. Overall, 89% of all previously uncharacterized candidate variants were confirmed.

Microarray-based comparative genomic hybridization.

Analysis of X-chromosome array-CGH data was performed in a multicenter setting. Depending on the institution of origin, array-CGH analysis was performed using an oligonucleotide array-CGH platform18,23 (Agilent) or a tiling path BAC array-CGH platform24. To determine the exact size of the recurrent Xq26 deletions, we reanalyzed all eight del(X)(q26)-positive T-ALL cases using a custom high-resolution X-chromosome oligonucleotide array with an average resolution of 3 kb according to the manufacturer's instructions (Agilent). Slides were scanned in a 2565AA DNA microarray scanner (Agilent). Microarray images were analyzed using Feature Extraction software (Agilent), and the data were subsequently imported into array-CGH Analytics software (Agilent).

Real-time quantification of DNA copy number.

Chromosome Xq26.3 deletions were confirmed with real-time quantitative DNA PCR using the Fast SYBR Green Master Mix (Applied Biosystems) and the LightCycler 480 Real-Time PCR System (Roche Diagnostics) as described25 using TIE2 as control gene. Data were analyzed using the comparative ΔΔCT method (Applied Biosystems). The primers used for the quantitative PCR analysis of the PHF6 locus are shown in Supplementary Table 7.

PHF6 mutation analysis.

PHF6 mutations were analyzed by PCR amplification of PHF6 exons 2–10 followed by direct bidirectional DNA sequencing. The primers used for PHF6 mutation detection are summarized in Supplementary Table 7.

Protein blot and immunohistochemistry.

Protein blot analysis was performed using a rabbit polyclonal antibody specific to PHF6 (1:10,000; Novus Biologicals) recognizing a C-terminal (amino acids 315–365) epitope in PHF6; a mouse monoclonal antibody (1:1,000) recognizing an N-terminal (amino acids 120–140) PHF6 epitope26; an antibody specific to γ-H2AX (1:1,000; Cell Signaling Technologies) and a mouse monoclonal antibody specific to GAPD (1:1000; Santa Cruz Biotechnology) using standard procedures.

Immunohistochemistry analysis was performed as described following the standard avidin-biotin immunoperoxidase staining procedure using an N-terminal antibody to PHF6. Briefly,

PHF6 immunostaining of formalin-fixed paraffin-embedded tissue sections was performed after heat-induced epitope retrieval in a microwave in citrate buffer, pH 6.0. Subsequently, sections were incubated in 10% normal goat serum for 30 min, followed by primary antibody incubation (amino acids 1–94 of rabbit polyclonal antibody to PHF6; Sigma Prestige Antibodies, dilution 1:100) overnight at 4 °C. Next, slides were incubated with biotinylated immunoglobulins specific to rabbit at a 1:1000 dilution (Vector Laboratories) for 30 min, followed by avidin-biotin peroxidase complexes at a 1:25 dilution (Vector Laboratories) for 30 min. Diaminobenzidine was used as the chromogen and hematoxylin as a nuclear counterstain.

Quantitative real-time PCR.

Thymocyte populations were isolated from human thymi as described before27. Total RNA was extracted using the Trizol method (Invitrogen) following the manufacturer's instructions. Total RNA from 20 different normal human tissues was obtained from the FirstChoice Human Total RNA Survey Panel (Applied Biosystems). Complementary DNA (cDNA) was generated with the ThermoScript RT-PCR system (Invitrogen) and analyzed by quantitative real-time PCR using the SYBR Green RT-PCR Core Reagents kit (Applied Biosystems) and the 7300 Real-Time PCR System (Applied Biosystems). PHF6 expression levels were calculated using GAPDH as a reference gene. Primers used for PHF6 expression analysis are shown in Supplementary Table 7.

SNP genotyping.

Genotyping of SNP rs17317724 located in the 3’ UTR region of PHF6 was performed using a TaqMan SNP Genotyping Assay (Assay ID C__34812972_10, Applied Biosystems) according to manufacturer's instructions in DNA samples from females with wild-type PHF6 T-ALL. Genotyping was confirmed by direct DNA sequencing of PCR products encompassing the 3’ UTR of PHF6. Allelic expression analysis was performed in cDNA samples from heterozygous females with T-ALL so as to evaluate monoallelic and biallelic PHF6 expression. Before cDNA synthesis, RNA samples were treated with DNase I using the DNA-free DNase Treatment kit (Applied Biosystems) to remove any traces of genomic DNA. The PHF6 3’ UTR-specific primers used for amplification of SNP rs17317724 are also summarized in Supplementary Table 7.

PHF6 shRNA knockdown.

We produced lentiviral particles driving the expression of a shRNA directed against PHF6 (target sequence CAGAATTTGGAGACTTTGA) using the pGIPZ Lentiviral shRNAmir vector system (V2LHS_138602, Open Biosystems) as described28. We infected HEK293T cells with viral supernatants generated with pGIPZ PHF6 or an empty pGIPZ control using spinoculation in the

presence of polybrene. PHF6 knockdown was evaluated by protein blot analysis at 72 h as described above.

Statistical analysis.

Fisher's exact test was used to compare the frequency of PHF6 mutations between clinical and genetic groups of T-ALL. Bar graphs represent mean values ± s.e.m. Therapeutic outcome was analyzed in leukemia patients treated in DCOG trials ALL7, ALL8 and ALL9 and in ECOG trial 2993 according to overall survival. Kaplan-Meier curves were used to assess survival, and differences between groups were compared by the log-rank test.

Acknowledgments

This study was supported by the Fund for Scientific Research (FWO) Flanders (postdoctoral grants to P.V.V. and T.T., PhD grant to J.V.d.M., senior clinical investigator award to B.P. and project grants G.0198.08 and G.0869.10N to F.S.); the GOA-UGent (grant no. 12051203); the IWT-Vlaanderen (SBO grant no. 060848); the Children Cancer Fund Ghent (F.S.); Leukemia Research UK (C.J.H.); the Stichting Kinderen Kankervrij (KiKa; grant no. KiKa 2007-012 to L.Z.); the Belgian Program of Interuniversity Poles of Attraction; the Belgian Foundation against Cancer; the Austrian Ministry of Science and Research (GEN-AU Child, GZ 200.136/1-VI/1/2005 to S.S.), the US National Library of Medicine (1R01LM010140-01 to R.R. and H.K.); the ECOG and DCOG tumor banks; grants from Spain's Plan Nacional (BFU 2007-60990 and PlanE2009-0110 to M.L.T.), Comunidad de Madrid (S-SAL0304-2006 to M.L.T.), Fundación MM (M.L.T.), Instituto de Salud Carlos III (RECAVA RD06/0014/1012 to M.L.T.), an Institutional Grant from the Fundación Ramón Areces (M.L.T.), the Alex's Lemonade Stand Foundation Young Investigator Award (T.P.); a US Northeast Biodefense Center ARRA award (U54-AI057158 to R.R.); the US National Institutes of Health (R01CA120196 and R01CA129382 to A.F.); the Rally across America Foundation (A.F); the Swim across America Foundation (A.F.); the Golfers against Cancer Foundation (A.F.); and a Leukemia and Lymphoma Society Scholar Award (A.F.). We thank the Pediatric Cardiosurgery Units from Centro Especial Ramón y Cajal and Ciudad Sanitaria La Paz (Madrid, Spain) for thymus samples.

Author contributions

P.V.V. performed array-CGH and mutation analysis of PHF6 and wrote the manuscript. T.P. performed exon capture and next-generation sequencing of T-ALL samples and wrote the manuscript. H.K. analyzed next-generation sequencing data. J.V.d.M. performed additional array-CGH analysis and PHF6 mutation screening in T-ALL and BCP-ALL samples. T.T., N.V.R. and A.W.L. performed experiments. M.C. and C.C.-C. performed and analyzed histological and immunohistochemical staining. J.P. collaborated on PHF6 mutation screening in BCP-ALL samples. C.J.H. and C.S. collaborated on additional screening for genomic PHF6 deletions in T-ALL. Y.B., B.D.M. and B.C. collaborated on the PHF6 mutation screening. R.P., M.P., S.S. and J.S. collaborated on the multicenter array-CGH study. S.G.-G. and M.L.T. performed the isolation of T-cell progenitor cells for expression analysis of PHF6. X.Y. performed survival analysis of ECOG T-ALL patients. J.G. provided critical reagents and discussion. E.S. provided samples and

correlative clinical data from DCOG. E.P., J.M.R. and P.H.W. provided samples and correlative clinical data from ECOG. J.M. and L.Z. collaborated on the multicenter array-CGH study and PHF6 mutation analysis, provided molecular data on the characterization of T-ALL and performed survival analysis of PHF6 mutations in the DCOG series. R.R. designed and directed the analysis of next-generation sequencing results. F.S. and B.P. designed the studies and directed research. A.F. designed the studies, directed research and wrote the manuscript.

References

1. Carrel, L., Cottle, A.A., Goglin, K.C. & Willard, H.F. A first-generation X-inactivation profile of the human X chromosome. Proc. Natl. Acad. Sci. USA 96, 14440–14444 (1999).

2. Carrel, L. & Willard, H.F. X-inactivation profile reveals extensive variability in X-linked gene expression in females. Nature 434, 400–404 (2005).

3. Goldberg, J.M. et al. Childhood T-cell acute lymphoblastic leukemia: the Dana-Farber Cancer Institute acute lymphoblastic leukemia consortium experience.

J. Clin. Oncol. 21, 3616–3622 (2003).

4. Aifantis, I., Raetz, E. & Buonamici, S. Molecular pathogenesis of T-cell leukaemia and lymphoma. Nat. Rev. Immunol. 8, 380–390 (2008).

5. Pui, C.H., Robison, L.L. & Look, A.T. Acute lymphoblastic leukaemia. Lancet 371, 1030–1043 (2008).

6. Gnirke, A. et al. Solution hybrid selection with ultra-long oligonucleotides

for massively parallel targeted sequencing. Nat. Biotechnol. 27, 182–189 (2009).

7. Lower, K.M. et al. Mutations in PHF6 are associated with Borjeson-Forssman-Lehmann syndrome. Nat. Genet. 32, 661–665 (2002).

8. Borjeson, M., Forssman, H. & Lehmann, O. An X-linked, recessively inherited syndrome characterized by grave mental deficiency, epilepsy, and endocrine disorder. Acta Med. Scand. 171, 13–21 (1962).

9. Turner, G. et al. The clinical picture of the Borjeson-Forssman-Lehmann syndrome in males and heterozygous females with PHF6 mutations. Clin. Genet. 65, 226–232 (2004).

10. Baker, L.A., Allis, C.D. & Wang, G.G. PHD fingers in human diseases: disorders arising from misinterpreting epigenetic marks. Mutat. Res. 647, 3–12 (2008).

11. Dephoure, N. et al. A quantitative atlas of mitotic phosphorylation. Proc. Natl. Acad. Sci. USA 105, 10762–10767 (2008).

12. Matsuoka, S. et al. ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. Science 316, 1160–1166 (2007).

13. Lowndes, N.F. & Toh, G.W. DNA repair: the importance of phosphorylating histone H2AX. Curr. Biol. 15, R99–R102 (2005).

14. Payer, B. & Lee, J.T. X chromosome dosage compensation: how mammals keep the balance. Annu. Rev. Genet. 42, 733–772 (2008).

15. Carrel, L. & Willard, H.F. Heterogeneous gene expression from the inactive X chromosome: an X-linked gene that escapes X inactivation in some human cell lines but is inactivated in others. Proc. Natl. Acad. Sci. USA 96, 7364–7369 (1999).

16. Ferrando, A.A. et al. Gene expression signatures define novel oncogenic pathways in T cell acute lymphoblastic leukemia. Cancer Cell 1, 75–87 (2002).

17. Soulier, J. et al. HOXA genes are included in genetic and biologic networks defining human acute T-cell leukemia (T-ALL). Blood 106, 274–286 (2005).

18. Van Vlierberghe, P. et al. The recurrent SET-NUP214 fusion as a new HOXA activation mechanism in pediatric T-cell acute lymphoblastic leukemia. Blood 111, 4668–4680 (2008).

19. van Grotel, M. et al. 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 (2006).

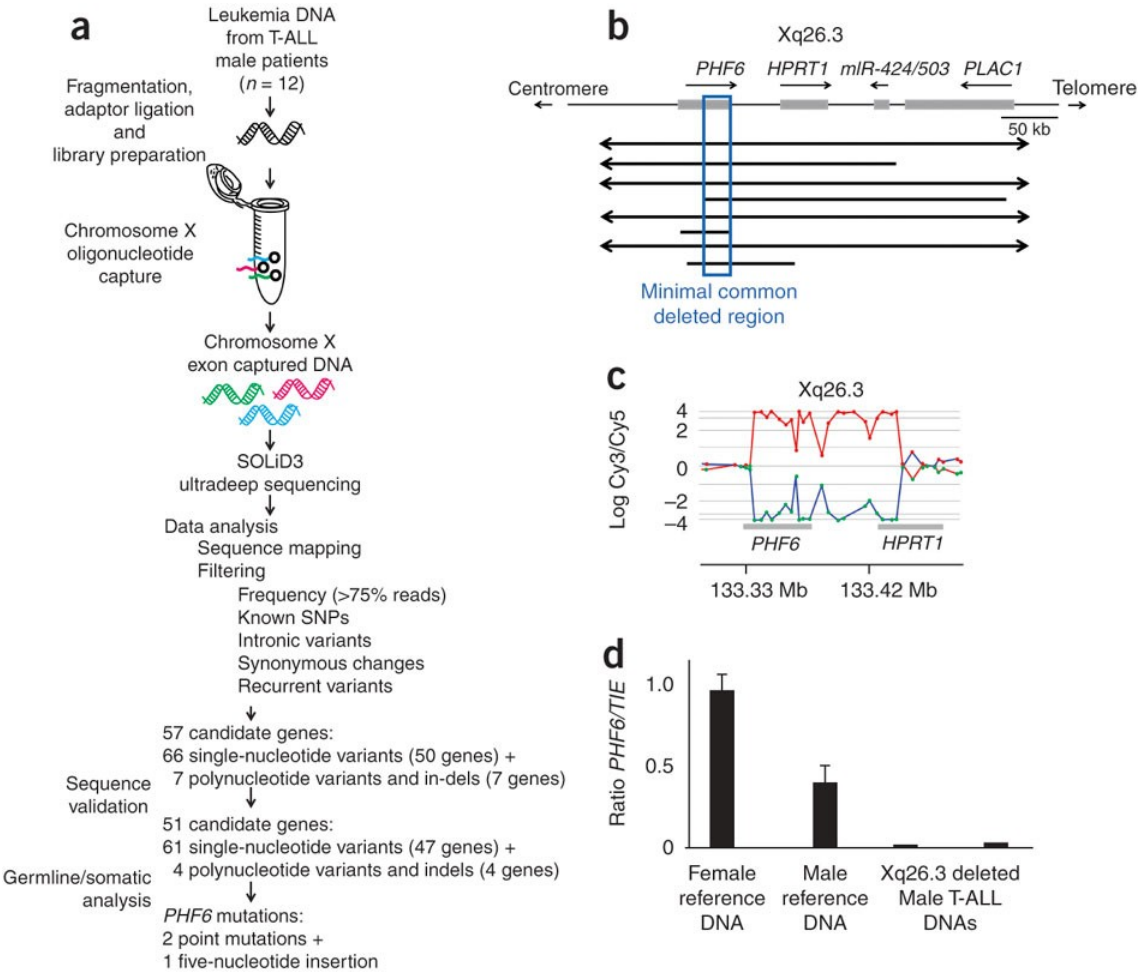

Figure 1.

Next generation sequencing and array CGH analysis of the X chromosome identifies PHF6 mutations in human T-ALL. (a) Overview of mutation screening approach of the human X chromosome exome in a panel of tumor DNA samples from 12 male T-ALL cases using oligonucleotide sequence capture and next generation sequencing with SOLiD3. After filtering and confirmation of high throughput sequencing data, analysis of corresponding remission DNA samples led to the identification of three somatically acquired changes in the PHF6 gene. (b) Schematic overview of the recurrent genomic deletions involving chromosomal band Xq26.3 in 8 human T-ALL samples. Specific genes located in Xq26.3 are shown. (c) Detailed view of a representative oligo array-CGH plot of leukemia DNA/control DNA ratios (blue tracing) versus the dye-swap experiment (red tracing) in a patient harboring an Xq26.3 deletion. (d) DNA quantitative PCR analysis of PHF6 copy number dose in female and male reference genomic DNAs and 2 primary samples from male T-ALL cases harboring Xq26.3 deletions.

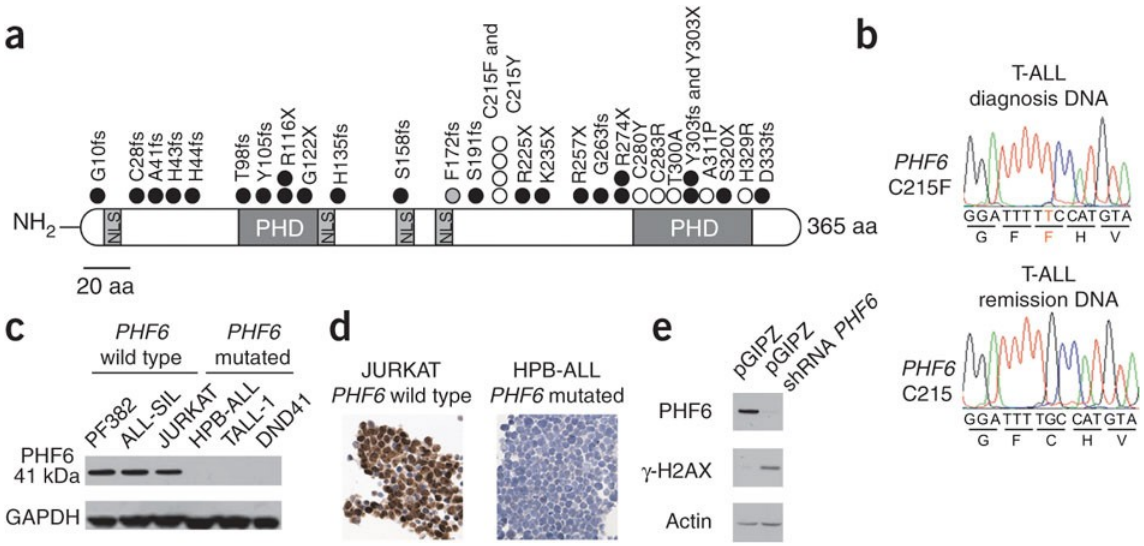

Figure 2.

PHF6 mutations and expression in T-ALL lymphoblasts. (a) Structure of the PHF6 protein including four nuclear localization signals and two imperfect PHD zinc finger domains. Overview of all PHF6 mutations identified in primary T-ALL samples and T-ALL cell lines. Filled circles represent nonsense and frameshift mutations, whereas missense mutations are depicted as open circles. Circles filled in gray indicate mutations identified in female T-ALL cases. (b) Representative DNA sequencing chromatograms of paired diagnosis and remission genomic T-ALL DNA samples showing a somatic mutation in exon 7 of PHF6. (c) Western blot analysis of T-ALL cell lines revealed complete loss of PHF6 protein expression in the PHF6 mutated T-ALL cell lines. (d) PHF6 immunostaining in the Jurkat and HPB-ALL, wild-type and mutant T-ALL cell lines, respectively. (e) Western blot analysis of PHF6 and gamma-H2AX expression in HEK293T cells upon PHF6 shRNA knockdown. Actin levels are shown as loading control.

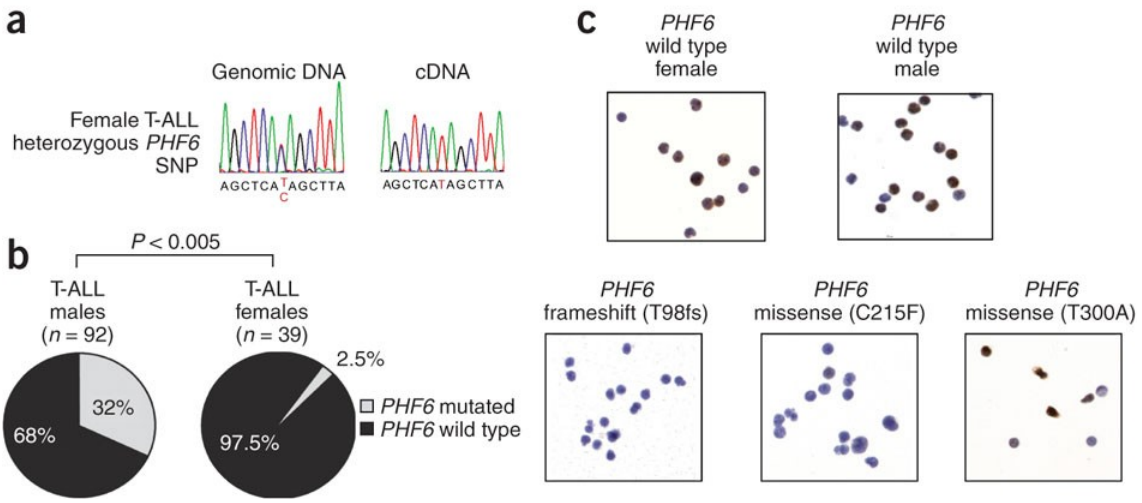

Figure 3.

PHF6 expression in T-ALL lymphoblasts. (a) Sequence analysis of paired genomic DNA and cDNA samples shows monoallelic expression of PHF6 SNP rs17317724 in lymphoblasts from a wild-type PHF6 female T-ALL case. (b) Differential distribution of PHF6 mutations in T-ALL samples from male and female cases. (c) Immunohistochemical analysis of PHF6 expression in wild type and mutant T-ALL lymphoblasts.

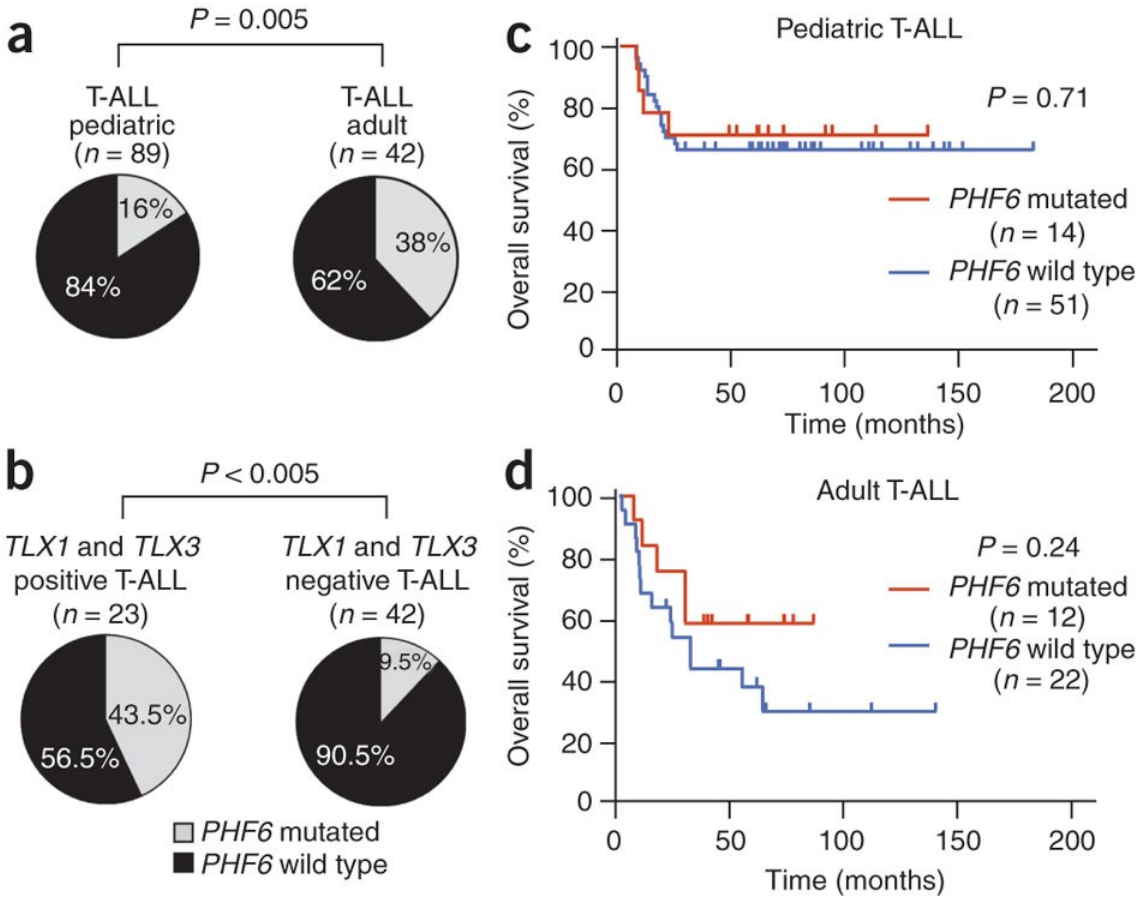

Figure 4.

Clinical and biological characteristics associated with PHF6 mutations in T-ALL. (a) Frequencies of PHF6 mutations in pediatric and adult T-ALL samples. (b) Differential distribution of PHF6 mutations in TLX1/TLX3 positive and negative T-ALL samples. (c) Kaplan-Meier curve of overall survival in pediatric T-ALL patients from DCOG trials ALL7, ALL8 and ALL9 with and without PHF6 mutations. (d) Kaplan-Meier survival curve in adult T-ALL patients with and without mutations in PHF6 treated in ECOG clinical trial ECOG2993.

ID	Sex	Age	WBC (× 10^9 l^{-1})	Immuno-phenotype	Genetic subtype	*NOTCH1*	Type of alteration	Deletion size or predicted *PHF6* protein lesion	Germline or somatic
1	M	Ped	77	Cortical	*TLX3*	Mut	Deletion	0.55 Mb	NA
2	M	Ped	46	Pre-T	*TLX3*	WT	Deletion	0.23 Mb	NA
3	M	Ped	31	Pre-T	*TLX3*	NA	Deletion	1.50 Mb	NA
4	M	Ped	2	Pre-T	Unknown	NA	Deletion	0.27 Mb	NA
5	M	Ped	NA	Cortical	*HOXA*	NA	Deletion	1.90 Mb	Somatic
6	M	Ped	NA	Cortical	Unknown	NA	Deletion	0.20 Mb	Somatic
7	M	Ped	NA	Cortical	*TLX1*	NA	Deletion	0.08 Mb	Somatic
8	M	Adult	NA	Pre-T	Unknown	NA	Deletion	0.11 Mb	NA
9	M	Ped	185	Cortical	*TLX3*	WT	Nonsense	G122X	NA
10	M	Ped	417	Pre-T	*TLX3*	Mut	Nonsense	R116X	NA
11	F	Ped	280	Pre-T	*TLX1*	Mut	Frameshift	F172fs	NA
12	M	Ped	405	Pre-T	*TLX3*	Mut	Frameshift	Y303fs	NA
13	M	Ped	159	Pre-T	*TLX1*	Mut	Nonsense	K235X	NA
14	M	Ped	500	Pre-T	*TLX3*	Mut	Frameshift	A41fs	NA
15	M	Ped	347	Cortical	*HOXA*	Mut	Nonsense	K274X	NA
16	M	Ped	129	Cortical	Unknown	Mut	Frameshift	D333fs	NA
17	M	Ped	174	Cortical	*TLX3*	Mut	Nonsense	R225X	NA
18	M	Ped	27	Cortical	*TLX1*	WT	Nonsense	R116X	NA
19	M	Ped	310	Cortical	*TAL1*	Mut	Missense	C283R	NA
20	M	Ped	189	Cortical	*TAL1*	Mut	Frameshift	C28fs	NA
21	M	Adult	170	Cortical	*TLX3*	WT	Frameshift	H44fs	NA
22	M	Adult	21	Cortical	*TLX1*	Mut	Frameshift	H43fs	Somatic
23	M	Adult	NA	Pre-T	*TLX3*	Mut	Frameshift	T98fs	Somatic
24	M	Adult	14	Pre-T	*TLX3*	WT	Frameshift	Y105fs	NA
25	M	Adult	28	Mature	*TLX3*	WT	Frameshift	S158fs	NA
26	M	Adult	NA	Cortical	*TLX3*	Mut	Missense	C215Y	NA
27	M	Adult	NA	Pre-T	Unknown	Mut	Missense	C215F	NA
28	M	Adult	31	Cortical	Unknown	Mut	Missense	C215Y	NA
29	M	Adult	NA	Mature	*TLX3*	Mut	Missense	T300A	Somatic
30	M	Adult	21	Cortical	*TLX1*	WT	Missense	A311P	NA
31	M	Adult	30	Mature	Unknown	WT	Missense	C280Y	NA
32	M	Adult	23	Cortical	Unknown	WT	Missense	H329R	Somatic
33	M	Ped	NA	NA	*TAL1*	NA	Nonsense	R257X	Somatic
34	M	Ped	NA	NA	*TLX1*	NA	Frameshift	S191fs	Somatic
35	M	Adult	NA	NA	*TLX1*	NA	Missense	C215F	Somatic
36	M	Adult	NA	NA	*TLX1*	NA	Nonsense	Y303X	NA
37	M	Adult	NA	NA	Unknown	NA	Nonsense	R274X	NA
38	M	Adult	NA	NA	Unknown	NA	Frameshift	H135fs	NA

Mb, megabases; Mut, mutated; NA, not available; Ped, pediatric; WT, wild-type; X, stop codon.

Table 1

Characteristics of 38 primary T-ALL samples showing PHF6 inactivation.