Site- and allele-specific polycomb dysregulation in T-cell leukaemia

Jean-Marc Navarro1,2,3,*, Aurore Touzart4,*, Lydie C. Pradel1,2,3,5,*, Myriam Koubi6, Romain Fenouil1,2,3, Sandrine Le Noir4, Muhammad Ahmad Maqbool1,2,3,†, Ester Morgado1,2,3, Claude Gregoire1,2,3, Sebastien Jaeger1,2,3, Emilie Mamessier1,2,3, Charles Pignon1,2,3, Salima Hacein-Bey-Abina7, Bernard Malissen1,2,3, Marta Gut8, Ivo G. Gut8, Hervé Dombret9, Elizabeth A. Macintyre4, Steven J. Howe10, H. Bobby Gaspar10, Adrian J. Thrasher10, Norbert Ifrah11, Dominique Payet-Bornet1,2,3, Estelle Duprez5,8, Jean-Christophe Andrau1,2,3,†, Vahid Asnafi4,8 & Bertrand Nadel1,2,3,§

T-cell acute lymphoblastic leukaemias (T-ALL) are aggressive malignant proliferations characterized by high relapse rates and great genetic heterogeneity. TAL1 is amongst the most frequently deregulated oncogenes. Yet, over half of the TAL1+ cases lack TAL1 lesions, suggesting unrecognized (epi)genetic deregulation mechanisms. Here we show that TAL1 is normally silenced in the T-cell lineage, and that the polycomb H3K27me3-repressive mark is focally diminished in TAL1+ T-ALLs. Sequencing reveals that >20% of monoallelic TAL1+ patients without previously known alterations display microinsertions or RAG1/2-mediated episomal reintegration in a single site 5′ to TAL1. Using ‘allelic-ChIP’ and CrispR assays, we demonstrate that such insertions induce a selective switch from H3K27me3 to H3K27ac at the inserted but not the germline allele. We also show that, despite a considerable mechanistic diversity, the mode of oncogenic TAL1 activation, rather than expression levels, impact on clinical outcome. Altogether, these studies establish site-specific epigenetic desilencing as a mechanism of oncogenic activation.

1 Center of Immunology of Marseille Luminy, Aix-Marseille University, Parc Scientifique de Luminy case 906, 13288 Marseille, France. 2 INSERM U1104, 13288 Marseille, France. 3 CNRS UMR7280, 13288 Marseille, France. 4 Université Paris Descartes Sorbonne Cité, Institut Necker-Enfants Malades (INEM), Institut national de recherche médicale (INSMR) U1115, and Laboratory of Onco-Hematology, Assistance Publique-Hôpitaux de Paris (AP-HP), Hôpital Necker Enfants-Malades, 75015 Paris, France. 5 Laboratoire Hématologie, APHM, 13385 Marseille, France. 6 CRCM Inserm U1068, Institut Paoli Calmettes; Aix-Marseille Université, UM 105; CNRS UMR7258, 13009 Marseille, France. 7 Biotherapy Department, INSERM U429, Hôpital Necker-Enfants-Malades, 149 rue de Sèvres, 75015 Paris, France. 8 Centro Nacional de Análisis Genómico, 08028 Barcelona, Spain. 9 Department of Hematology, AP-HP Hôpital St-Louis, 75010 Paris, France. 10 Centre for Immunodeficiency, Molecular and Cellular Immunology, Institute of Child Health, University College London, London WC1N 1EH, UK. § Joint supervisors. † Present addresses: TAGC/INSERM U1090, 13288 Marseille, France (L.C.P.); Institut de Génétique Moléculaire de Montpellier (IGMM), CNRS UMR5535, 1919 Route de Mende, 34293 Montpellier, France (M.A.M. and J.-C.A.). Correspondence and requests for materials should be addressed to V.A. (email: vahid.asnafi@nck.aphp.fr) or to B.N. (email: nadel@ciml.univ-mrs.fr).
A recurrent epigenetic mechanism of TAL1 deregulation. To determine if similar structural abnormalities occurred recurrently in TAL1 patients, 134 primary T-ALL samples were analysed by high-density Affymetrix SNP array-6 analysis; an ~700-bp region surrounding the Jurkat insertion site was also sequenced in a subset of 93 samples and six cell lines; in parallel, the literature was reviewed for cases with unexplained TAL1 activation. While no macromolecular TAL1 alteration was identified by SNP array (0/134), sequencing revealed seven new cases of similar microinsertions (1–9 bp), all precisely located at the Jurkat insertion breakpoint (Fig. 3a). Such insertions were not present in the germline from 2/134, sequencing revealed seven new cases of similar microinsertions (1–9 bp), all precisely located at the Jurkat insertion breakpoint (Fig. 3a). Such insertions were not present in the germline from 2/134, sequencing revealed seven new cases of similar microinsertions (1–9 bp), all precisely located at the Jurkat insertion breakpoint (Fig. 3a). Such insertions were not present in the germline from 2/134, sequencing revealed seven new cases of similar microinsertions (1–9 bp), all precisely located at the Jurkat insertion breakpoint (Fig. 3a). Such insertions were not present in the germline from 2/134, sequencing revealed seven new cases of similar microinsertions (1–9 bp), all precisely located at the Jurkat insertion breakpoint (Fig. 3a). Such insertions were not present in the germline from 2/134, sequencing revealed seven new cases of similar microinsertions (1–9 bp), all precisely located at the Jurkat insertion breakpoint (Fig. 3a). Such insertions were not present in the germline from 2/134, sequencing revealed seven new cases of similar microinsertions (1–9 bp), all precisely located at the Jurkat insertion breakpoint (Fig. 3a). Such insertions were not present in the germline from 2/134, sequencing revealed seven new cases of similar microinsertions (1–9 bp), all precisely located at the Jurkat insertion breakpoint (Fig. 3a). Such insertions were not present in the germline from 2/134, sequencing revealed seven new cases of similar microinsertions (1–9 bp), all precisely located at the Jurkat insertion breakpoint (Fig. 3a). Such insertions were not present in the germline from 2/134, sequencing revealed seven new cases of similar microinsertions (1–9 bp), all precisely located at the Jurkat insertion breakpoint (Fig. 3a). Such insertions were not present in the germline from 2/134, sequencing revealed seven new cases of similar microinsertions (1–9 bp), all precisely located at the Jurkat insertion breakpoint (Fig. 3a). Such insertions were not present in the germline from 2/134, sequencing revealed seven new cases of similar microinsertions (1–9 bp), all precisely located at the Jurkat insertion breakpoint (Fig. 3a).
surrounding ~700 bp region in tumour samples. Significantly, insertions were exclusively found in TAL1+ patients; moreover, among patients with informative SNPs in the TAL1 3' UTR allowing distinction of mono- from biallelic expression (n = 60), insertions were exclusively found in monoallelic cases (4/19, 20%), in agreement with a cis-mode of TAL1 activation (Fig. 3b).

Oncogenic RAG1/2-mediated episomal reintegration. One additional candidate was recovered by data mining. In this case (patient #OC), TAL1high activation concurred with the insertion of a large piece of chromosome 7 disrupting the TAL1 locus. Strikingly, breakpoint mapping by ligation-mediated PCR revealed that the insertion occurred at the very same insertion site, although this time with few nucleotide deletions on each side of the breakpoint (Fig. 3a). Detailed analysis of the junctions revealed the occurrence of RAG1/2-mediated reinsertion of an ~370-kb TCRβ episomal circle (TRECβ, excised during normal V(D)J recombination, Fig. 4 and Supplementary Fig. 5). This establishes the first example of oncogenic RAG1/2-mediated
reintegration, demonstrating that TREC5s may indeed contribute to oncogenesis17. Owing to the large TREC size, we considered the possibility that, unlike other microinsertions, a promoter located in the episome could have initiated a 47-kb-long fusion transcript encompassing TAL1 (Supplementary Fig. 6). However, reverse transcription-PCR (RT–PCR) exon walking and 5′ RACE assays indicated that transcripts initiated from the TAL1 p4 promoter, excluding this possibility.

Epigenetic modulation and TAL1 gene expression. Allelic-ChIP was then performed on patients TAMFA and OC. Similar to Jurkat, significant enrichment of the repressive H3K27me3 mark was consistently observed in germline compared with inserted alleles (Fig. 5a,b). The amplitude of the allelic distortion appeared higher in OC than in Jurkat and TAMFA, possibly due to the large difference in the insertion size. Of note, TAL1 transcription levels were also higher in OC (Fig. 3b). Using CrispR DNA editing, we next mimicked site-specific insertion and disruption of the region 7 kb 5′ of TAL1 in the TAL-negative PEER cell line (Fig. 5b). In clone #2.4 recapitulating the 12-bp Jurkat insertion at its 3′ end, an approximately fivefold increase of TAL1 could be observed. While we cannot formally exclude the possibility that the selection cassette contributed to the fivefold

Figure 2 | Site- and allele-specific analysis of histone methylation/acetylation marks at the insertion breakpoint in Jurkat. (a) Allelic-ChIP assay of H3K27me3 marks. Top panel: the assay to discriminate the germline (GL) from the inserted allele (Ins.) by substituting one of the GL primers allowing detection of the GL configuration at the insertion site (Rev. TAL1), with an insertion-specific primer allowing detection of the inserted configuration (Rev. Ins., overlapping the 12-bp insertion). Primer pairs were tested on GL p(GL) or inserted p(Ins) cloned fragments (and on cell lines containing (Jurkat) or not (DND) the insertion) to exclusively amplify each configuration, and do not crossreact. Bottom middle panel: western blot of EZH2 protein content on shMock or EZH2 knockdown conditions. Allelic-ChIP assays were performed in presence of a non-silencing sh-RNA (shControl) or a sh-RNA-targeting EZH2 (shEZH2) (left panel) or after the incubation of Jurkat cells with GSK126 (0.5 μM, 72H) or vehicle (dimethyl sulphoxide, DMSO; right panel). GAPDH and HoxD11 were used as controls for activated/repressed genes, modulated according to the polycomb-dependent H3K27me3 marks. Note that EZH2 knockdown/inhibition triggered only partial decrease of H3K27me3 marks at the PcG-repressed HoxD11 control gene, possibly due to incomplete knockdown/inhibition and/or redundancy of polycomb components in the adult lymphoid lineage42. (b) Enrichment of acetylation marks at the TAL1 locus. H3K27Ac ChIP was performed with Jurkat cells incubated with vehicle (DMSO) or the histone deacetylase inhibitor sodium butyrate (SOB) (5 mM, 4H); DNA was then analysed by ChIP-seq (left panel) or by allelic-ChIP (bottom panel). For the ChIP-Seq, quantification of the number of tag sequences at the insertion point is shown (right panel). ***P<0.001; **P<0.001; *P<0.05, unpaired t-test. Errors bars represent 95% confidence interval.
change in clone #2.4, a 55-fold increase was observed in clone #5.10, in which an ~1.3 kb deletion 5′ of the insertion site mimicked locus disruption in patient #OC; furthermore, this was accompanied by an allelic switch from H3K27 methylation to acetylation. This provides direct evidence for a causal relationship between site mutagenesis, epigenetic modulation and TAL1 gene expression.

The mode of TAL1 activation impacts on clinical outcome. Patients with identified insertions were globally of adverse prognosis. We sought to determine if clinical outcome correlated with quantitative or qualitative aspects of TAL1 deregulation. A cohort of 165 adult T-ALL treated prospectively in the GRAALL (Group for Research in Adult Acute Lymphoblastic Leukaemia) trial was split into TAL1 expression quartiles, and compared for disease-free (DFS) and overall survival (OS). The seven patients with identified insertions (three of whom were GRAALL treated), all belonged to the high-expression quartiles (Q3–4). However, no significant difference in survival was observed between the quartiles (Fig. 6a,b), suggesting that quantitative TAL1 expression does not correlate with clinical outcome. We next tested whether cis-mediated TAL1 alterations leading to monoallelic expression (including or not the SIL-TAL1 + cases) affected the clinical outcome compared with trans-mediated events, associated with biallelic TAL1 expression. Clinical outcome was indeed found to be significantly improved in the biallelic group (DFS, \( P = 0.04 \); OS, \( P = 0.03 \); Fig. 6c–f). Although numbers are low, monoallelic cases retained an inferior OS trend in multivariate analysis (including age, leukocytosis; \( P = 0.07 \), Cox analysis). Despite genetic heterogeneity, monoallelic cases also displayed higher blasts counts at diagnosis than biallelic cases and a significantly lower frequency of deregulation of recurrent oncogenes such as TLX1, CALM-AF10 and TLX3 (Supplementary Table 1).

Discussion

Establishing the detailed maps of the complex oncogenic networks involved in T-ALL has contributed to major genetic discoveries, and has been of prime importance for further
therapeutic improvement. Over three decades of intense efforts in genomic research have allowed unravelling the extraordinary diversity of the mechanisms by which oncogenes are deregulated in this disease. Yet, a large number of major oncogene deregulations still remain unexplained to date. Among the diversity of mechanisms involved, V(D)J recombination-mediated alterations (translocations, microdeletions) constitute the hallmarks of T-ALLs. Interestingly, despite arrays of biochemical and functional evidence that the reintegration of excised episomal circles (TRECs) by the V(D)J recombinase (RAG1/2) might constitute a potent source of genomic instability, such events remained so far unreported in human cancer patients. Here we report the first case of such an oncogenic RAG1/2-mediated episomal reintegration, demonstrating that TRECs can indeed contribute to human oncogenesis. This and the other insertional mutagenesis T-ALL cases described here also revealed a novel oncogenic activation.

Figure 4 | Schematic representation of the episomal reintegration in Patient OC. The TCRβ locus is displayed (top lane, not to scale). A functional Vβ7.4-to-Dβ1 rearrangement generating an excised TRECβ, and containing a (Vβ7.4/Dβ1) signal joint (SJ) is represented. The episome might have been open at the SJ by a nick-nick process generating 3’ hydroxyl ends before integration in chromosome 1. The episome is integrated in reverse orientation 10 kb downstream of the STIL gene, and 7 kb upstream of the TAL1 gene (middle lane). A 10-bp deletion (Δ, underlined) occurred at the insertion site. Localization of cryptic RSSs used by illegitimate V(D)J-mediated SIL-TAL deletion, and by t(1;14) TCRδ/TAL1 translocations are indicated by black arrow heads. TAL1 promoters (P1a, P1b, P4) are indicated. The breakpoints sequences (Bkp1/2) are shown (bottom lane).
Figure 5 | Insertional mutagenesis is associated with epigenetic modulation and TAL1 gene expression. (a) Allelic-ChIP analysis of H3K27me3 marks at the insertion breakpoint in primary patients. See legend to Fig. 2a. Marks at the GL alleles in one TAL1^+ biallelic (RENE) and in one TAL1^- (DAV) patients were performed as controls. ***P<0.001, unpaired t-test; Relative-fold plots: Inserted and/or GL allele ChIP values were calculated as fold increases relative to GAPDH (numbers in blue) or HoxD11 (numbers in red), and folds plotted as relative percentages (GAPDH relative folds: blue histograms; HOXD11 relative folds: red histograms). Blue histograms over 50% indicate higher differences with the expressed than the repressed control genes and correspond to (partially) repressed TAL1 expression; conversely, red histograms over 50% indicate higher differences with the repressed than the expressed control genes and correspond to (partially) derepressed TAL1 expression; histograms are ordered according to decreasing TAL1 repression.

(b) Epigenetic modulation and TAL1 gene expression by DNA editing mimicking insertional mutagenesis. Left panel: schematic representation of the CRISPR design for homologous recombination at the TAL1 locus, and configuration of two edited clones in the PEER cell line. The locations of PCR primers (plain arrows) for detecting successful targeted events and for genome walking are indicated. Bottom left panel: successful homologous recombination was confirmed by PCR of the expected genome-donor and donor-insert boundaries. Top right panel: RQ-PCR analysis of TAL1 expression after editing. Transcripts were normalized to ABL and reported as relative values to non-edited PEER cells. Four PCR replicates were performed on 1 (clone 2.4, due to impaired growth) or 2 (clone 5.10) independent RNA extractions. Bottom right panel: allelic-ChIP assays of H3K27me3 and H3K27ac marks in edited clone 5.10. See legend to Fig. 2a. ***P<0.001, unpaired t-test.
pathway, whereby a genetic alteration drives a site-specific and monoallelic epigenetic deregulation. We demonstrate that such insertions drove a switch from H3K27me3 to H3K27ac deposition, leading to the maintenance and/or re-expression of TAL1 expression through T-cell differentiation. Interestingly, the difference in TAL1 expression levels observed in mutants from the gene editing assay (Fig. 5) suggest that while small insertions might be sufficient to prevent the deposition of PcG repressive marks during T-cell lineage specification (thus permitting H3K27ac switch and maintenance of TAL1 expression), further disruption of the region 5' of the insertion site (by deletion or insertional uncoupling) might be necessary to impose desilencing once TAL1 extinction is established in the T-cell lineage (Supplementary Fig. 1). These kinetics are coherent with thymocyte ontogeny in patient #OC, in which TREC rearrangement/reintegration (DN2-3) likely occurred after TAL1 silencing (DN1-2). Altogether, our data are in line with current models of permanent gene extinction of transcription factors during T-lineage commitment (Supplementary Fig. 1) and further identify locus control regions involved in deposition and/or maintenance of TAL1 silencing. Their genetic disruption constitute a recurrent epigenetic mechanism of TAL1 deregulation in T-ALLs, contributing to a substantial fraction (>20%) of the TAL1+ monoallelic 'unresolved cases', and associated with adverse prognostic. That a cis-deregulation regrouping as diverse mechanisms as SIL-TAL deletions, translocations or insertional desilencing impact more on prognosis than TAL1 expression levels underlines the fundamental oncogenic difference between a deregulation targeting a single locus, and the wider effect of trans-acting factors. Transcription factors indeed often bind to a large number of target genes (hundreds to thousands) and their deregulation (whether gain or loss) will likely affect a complex set of cellular functions, some of which might antagonize tumour progression, or resistance to treatment. Recently, reports identifying loss-of-function mutations in polycomb-related components19–21,33 have provided the framework by which global epigenetic modification might trigger the indirect (and biallelic) activation of numerous target genes, likely including a complex and conflicting set of oncogenes and tumour suppressors. In humans, PcG are recruited

Figure 6 | Survival analysis. Kaplan–Meier analysis showing DFS and OS of 165 protocolar patients treated in the GRAALL trial according to: (a,b) TAL1 expression quartiles; (c–f) the mode of TAL1 expression. *P* values are indicated, log-rank (Mantle–Cox) test.
to and repress specific regions in the genome through as yet undefined set(s) of DNA-binding transcription factors and long non-coding RNAs. The insertional mutagenesis described here identifies a site- and allele-specific switch from H3K27me3 recruitment/maintenance to H3K27ac, providing new avenues to recruit/maintenance to H3K27ac, providing new avenues to identify a site- and allele-specific switch from H3K27me3 enrichment in ChIP-seq between alleles with and without insertion.

Allelic-ChIP. Input and IP genomic DNAs were analysed by RT–PCR using power SybrGreen on a 7,500 Fast Real-time PCR system (Applied Biosystems). IgG control cycle over the threshold (Ct) was determined using the following formula: 

\[ D = 2^{\text{Ct}_{\text{IP}} - \text{Ct}_{\text{Input}}} \]

A non-silencing sh-RNA (pTRIPZ-NS) was used as control. Jurkat cells were electroporated and cells containing the pTRIPZ were selected on puromycine. Knockdown of EZH2 was obtained by the addition of doxycycline (2 μg·mL⁻¹) to the cells during 10 days. Western blot was performed using the anti-EZH2 BD Biosciences # 612666.

Sequencing and SNP array. For Jurkat mapping, a region of 10 kb 5' of TAL1 exon 1 was mapped on both alleles by LRPCR/cloning and standard Sanger sequencing (see Supplemental Table 2 for details). Identification of allelic variants (SNPs versus somatic indels/mutations) was performed with vector NTI using alignment against reference TAL1 sequence and variants (http://www.ncbi.nlm.nih.gov/SNP; http://projects.tcag.ca/variation, Supplementary Fig. 8). For the sequencing screen on T-ALL patients and cell lines, a region of ~700 bp surrounding the Jurkat insert site was directly PCR/sequenced on both strands in a subset of 93 samples and six cell lines. Heterogeneous sequences (ambiguous reading due to allelic differences) were systematically cloned, sequenced and analysed as above. For SNP array, hybridization on Affymetrix GenomeWide SNP Array-6 was performed according to the manufacturer’s recommendations. Data analysis was performed with Chromosome Analysis Suite software using the following settings: the CGH log2 copy number ratio for heterozygous deletion was defined as 0.5 to 1.5, whereas log2 copy number ratios <1.5 were defined as homozygous deletions. Gene copy number (GCN) aberrations were compared with the Database of Genomic variants (http://projects.tcag.ca/variation) to study only non-variant GCN aberrations.

Patients. Diagnostic samples from a consecutive series T-ALLs from 165 adults (older than 16 years) included in GRAALL-03/05 trial (registration #NCT00327678 and #NCT00222027) were analysed for TAL1 expression. Sample collection and analyses were approved by the local ethical committee. Informed consent was obtained from the patients or relatives in accordance with the Declaration of Helsinki, with the institutional review board approval of all involved hospitals. Diagnosis of T-ALL was based on the World Health Organization 2008 criteria, defined by expression of cytoplasmic and/or surface CD3, and negativity of CD19 and MPO, as reported. To be included, samples had to contain at least 80% of lymphoblasts. Immuno-genotypying and oncogenic quantification were performed as previously described.

Cell lines. Cell lines used in this study were purchased from the ATCC collection and were mycoplasma free.

RQ-PCR. RNA was reverse transcribed using MMLV (Invitrogen). We used a TaqMan assay to quantify TAL1 transcript with the following primers: TAL1 F: 5′-ACA-ATC-GAG-TGA-AGA-GGA-GAC-CTT-C-3′, TAL1 Probe: fam-5′-ACA-ATC-GAG-TGA-AGA-GGA-GAC-CTT-C-3′, TAL1 R: 5′-ACC-GCG-CAC-AAT-TGT-GGT-G-3′, 40 cycles were run on ABI 7500HT (Applied Biosystem) as described. TAL1 transcript quantification was performed after normalization with the housekeeping gene ARL using the ΔCt method and results calculated according to the following formula: 

\[ \text{FC} = 2^{-\Delta\Delta\text{Ct}} \]

TAL1 allelic expression analysis. Allelic expression was performed as previously described. In brief, polymorphic markers in the 3′ UTR of the TAL1 gene were identified by PCR amplification and direct sequencing of 100 ng of genomic DNA. Allelic expression analysis was performed by PCR amplification and by direct sequencing of RT–PCR products from heterozygous patient samples. Three different RT–PCR products were made to cover nine most frequent SNP among the 11 SNP previously described.

Statistical analysis. ChIP. The power of t-test was estimated a priori using pwR-package, and the expected variations between conditions evaluated. Differences in ChIP data between inserted and GL TAL1 alleles were analysed by unpaired t-test. Samples collections constituted of five or six technical replicates were first checked for normal distribution using Kolmogorov–Smirnov test and the equality of variances was tested using F test. Results of t-test are shown as two-tailed P values. The statistical power of executed t-tests was at least 80%. Errors bars on histograms represent s.e.m.

NATURE COMMUNICATIONS | DOI: 10.1038/ncomms7094 | www.nature.com/naturecommunications

© 2015 Macmillan Publishers Limited. All rights reserved.
RT-PCR exon walking and 5’ RACE. Total RNA was extracted using a column-based system RNEasy mini kit (Qiagen) according to the manufacturer’s instructions. Reverse transcription was performed with SuperScript III Reverse transcriptase (Invitrogen) and random primers (Applied Biosystem). cDNAs were analysed by real-time quantitative PCR (RT–PCR) using power SybrGreen on an ABI-Prism 7500 (Applied Biosystems). All PCR were performed in duplicate. 5’-RACE was performed using 2 μg of total RNA and the 5’/3’ RACE kit, 2nd generation (Roche). Modifications from the instruction manufacturer were the generation a poly(G) tailing of first strand cDNA and the use of an oligo (dC) anchor primer. PCR was performed using the Pfu Ultra II fusion HS DNA polymerase (Agilent technologies).

Genome editing in T-ALL cell line by type II CRISPR system. PEER T-ALL cells line were cultured in RPMI medium (Life Technologies) containing 20% fetal calf serum, 1% L-glutamine, 1% sodium pyruvate and 100 U/ml - penicillin/streptomycin (Life Technologies) at 37°C in the presence of 5% CO2. The day of transfection, 1 million cells were nucleofected according to the manufacturer’s instruction (Lonza), with 500 ng DNA donor sequence containing Neomycin-resistance and two of the Cas9RNA expression vector (Addgene #42230). The chimeric guide RNA targeted TAL1 insertion site, and was cloned according to Cong et al.41. One day after nucleofection, cells were plated in 96-wells plate at 104 cells per well and incubated in presence of 1,200 μg/ml genetin G418 (Life Technologies) for 2 weeks. After selection and growing, a PCR was conducted to amplify the targeted region with genomic DNA derived from the surviving clones, and amplicons were separated on a 1% agarose gel then extracted with GEL/PCR clean up wizard (Promega) and sequenced (MWG-Biotech). CRISPR guide RNA: 5’-GAAAGGCTAAGCCCTACCTCC-3’. Primers list is available on request.

References
1. Tettell, M. A. & Pandolfi, P. P. Molecular genetics of acute lymphoblastic leukemia. Annu. Rev. Pathol. 4, 175–198 (2009).
2. Van Vlierberghe, P., Pieters, R., Beverloo, H. B. & Meijerink, J. P. Biallelic transcriptional activation of oncogenic specific enhancers and promoters. Nature 449, 840–845 (2007).
3. Van der Meulen, J., van Roy, N., Van Vlierbergh, P. & Speelman, F. The epigenetic landscape of T-cell acute lymphoblastic leukemia. Int. J. Biochem. Cell Biol. (2014).
4. Fenouil, R. et al. CpG islands and GC content dictate nucleosome depletion in a transcriptional-independent manner at mammalian promoters. Genome Res. 22, 2379–2388 (2012).
5. Roulland, S. et al. Follicular lymphoma-like B cells in healthy individuals: a novel intermediate step in early lymphomagenesis. J. Exp. Med. 203, 2425–2431 (2006).
6. Asnafi, V. et al. Analysis of TCR, pT alpha, and RAG-1 in T-acute lymphoblastic leukemias improves understanding of early human T-lymphoid lineage commitment. Blood 107, 2693–2703 (2006).
7. Asnafi, V. et al. Age-related phenotypic and oncogenic differences in T-acute lymphoblastic leukemias may reflect thymic atrophy. Blood (2004).
8. Bergeron, J. et al. Diagnostic and oncogenic relevance of TLX1/HOX11 expression level in T-ALLs. Blood 110, 2324–2330 (2007).
9. Cong, L. et al. Multiplex genome engineering using CRISPR/Cas systems. Science 339, 819–823 (2013).
10. Mochizuki-Kashio, M. et al. Dependency on the polycomb gene Ezh2 distinguishes fetal from adult hematopoietic stem cells. Blood 118, 6553–6561 (2011).
43. Delabesse, E. et al. TAL1 expression does not occur in the majority of T-ALL blasts. Br. J. Haematol. 102, 449–457 (1998).
44. Neiditch, M. B., Lee, G. S., Huye, L. E., Brandt, V. L. & Roth, D. B. The V(D)J recombinase efficiently cleaves and transposes signal joints. Mol. Cell 9, 871–878 (2002).

Acknowledgements
This work was supported by grants to BN laboratory from la Fondation pour la Recherche Médicale (#INE2003114116), Le Conseil Général des Bouches du Rhône, le Cancéropôle PACA, CALYM consortium, l’Institut National du Cancer (INCa PLBIO09), la Fondation de France (#2008001490) and institutional grants from INSERM, CNRS and AMU. M.L. is a recipient of a fellowship from INCa (#ASC12035ASA). B.N. is the recipient of a CHRT INSERM/Assistance Publique-Hôpitaux de Marseille (HP-HM). Work in the Necker laboratory was supported by grants from the ‘Association Laurette Fugain’, the Institut National du Cancer and the ‘Comité départemental de la Ligue Contre le Cancer’. S.L.N. was supported by the Institut National du Cancer (INCa-DHOS) and a grant from the ‘Association pour la Recherche sur le Cancer (ARC)’. R.F. is supported by a Centre National de la Recherche Scientifique grant. M.A.M. was supported by an ANR chromaTin and by the Fondation pour la Recherche Médicale (PRM AJE20130728183). The sequencing costs for ChIP-seq experiments were partially granted by an ESGI consortium grant from the Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 262055. E.D. gratefully acknowledges support from the Association pour la Recherche sur le Cancer and the Fondation pour la Recherche Médicale. We thank participants in the GRAALL-2003 and -2005 study groups for collecting and providing biological samples and data sets. The GRAALL was supported by grants P0200701 and P030425/ AOM03081 from the Programme Hospitalier de Recherche Clinique, Ministère de l’Emploi et de la Solidarité, France and the Swiss Federal Government in Switzerland. Samples were collected and processed by the AP-HP ‘Direction de Recherche Clinique’ Tumour Bank at Necker-Enfants Malades. We are grateful to V. Libritter for her help in collecting clinical samples, I. Spinelli for statistical advice, T. Guignard for Macs2 analyses and C. Pignons for technical help. We are grateful to Dr T. Cathomen for expert advice in DNA editing.

Author contributions
ChIP-seq experiments were performed and analysed by L.C.P., R.F., M.A.M., M.G., I.G.G. and J.-C.A.; allelic-ChIP experiments were performed and analysed by J.-M.N., L.C.P., M.L., M.K., C.P., D.P.-B. and E.D.; T-ALL sample characterizations were performed and analysed by A.T., J.-M.N., L.C.P., S.L.N., S.J., I.S., E.A.M. and V.A.; TALEN/CRISPR assays were performed and analysed by J.-M.N., M.L. and D.P.-B. with the advice of C.G. and B.M.; episomal reintegration characterization was performed by J.-M.N., E.A.M., S.H.-B.A., S.J.H., H.B.G. and A.J.T.; survival analysis was carried out by A.T., H.D., E.A.M., N.I. and V.A.; V.A. and B.N. conceived and directed the project, and B.N. wrote the paper.

Additional information
Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Navarro, J.-M. et al. Site- and allele-specific polycomb dysregulation in T-cell leukaemia. Nat. Commun. 6:6094 doi: 10.1038/ncomms7094 (2015).

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/