Selective regulation of lymphopoiesis and leukemogenesis by individual zinc fingers of Ikaros

Hilde Schjerven1,2,8, Jami McLaughlin1, Teresita L Arenzana1,2, Seth Frietze3,8, Donghui Cheng1,4, Sarah E Wadsworth1,2, Gregory W Lawson5,6, Steven J Bensinger2,8, Peggy J Farnham3, Owen N Witte1,2,4,7 & Stephen T Smale1,2,7

C2H2 zinc fingers are found in several key transcriptional regulators in the immune system. However, these proteins usually contain more fingers than are needed for sequence-specific DNA binding, which suggests that different fingers regulate different genes and functions. Here we found that mice lacking finger 1 or finger 4 of Ikaros exhibited distinct subsets of the hematological defects of Ikaros-null mice. Most notably, the two fingers controlled different stages of lymphopoiesis, and finger 4 was selectively required for tumor suppression. The distinct defects support the hypothesis that only a small number of genes that are targets of Ikaros are critical for each of its biological functions. The subcategorization of functions and target genes by mutagenesis of individual zinc fingers will facilitate efforts to understand how zinc-finger transcription factors regulate development, immunity and disease.

In the post-genomics era, a critical goal is to define the target genes and mechanisms of action of transcription factors that contribute to the development of multicellular organisms, disease and the response to environmental cues. Although progress in this area has been rapid, a major hurdle is that many factors are involved in several biological pathways and can contribute to multiple steps in a single pathway, which makes it difficult to study key regulatory events in isolation. Ikaros is one transcription factor that regulates numerous biological events. Ikaros-null (Ikzf1null) mice lack B lineage cells, natural killer cells, peripheral lymph nodes and fetal T cells, and they exhibit many other abnormalities1–3. B cells are absent from adult Ikzf1null mice because of an inability of lymphoid-primed multipotent progenitor cells to develop into common lymphoid progenitor cells4, but Ikaros also regulates later stages of B cell development5–9. Moreover, mice with mutations in the gene encoding Ikaros (Ikzf1) develop thymic lymphoma with high penetrance10,11.

In humans, mutations at the IKZF1 locus have been observed in a high percentage of samples of BCR-ABL7. B cell acute lymphoblastic leukemia (B-ALL) and high-risk progenitor B-ALL12,13. IKZF1 was the only gene for which mutations were found to be useful in predicting a poor response to therapy13, yet the mechanisms responsible for tumor suppression by Ikaros remain poorly understood. Ikaros is thought to contribute to both the activation and repression of transcription, with the Mi-2–NuRD complex as a predominant interacting partner14–16. Notably, the target genes responsible for most of the key phenotypes of Ikaros-mutant mice remain unknown7,17–19. The widespread deregulation of gene expression in Ikaros-mutant cells and the finding that Ikaros binds several thousand genomic sites have increased the challenge of identifying the mechanisms by which Ikaros regulates lymphocyte development and leukemogenesis16,20.

Members of the Ikaros family have a conserved DNA-binding domain near the amino (N) terminus, usually with four C2H2 zinc fingers21. Ikaros proteins also have two carboxy (C)-terminal C2H2 zinc fingers dedicated to dimerization and multimerization22–24. C2H2 zinc fingers are found in many ubiquitous transcription factors and in numerous proteins involved in tissue-specific and developmental stage–specific transcription25–27, including PLZF, Bcl-6, Gfi1, Blimp-1 and ThPOK of the immune system.

Stable binding of DNA in vitro by a C2H2 zinc-finger protein usually requires two or three tandem fingers27,28. In Ikaros, fingers 2 and 3 are sufficient for stable binding and recognize the core consensus sequence GGGAA. In contrast, fingers 1 and 4 seem to modulate binding to specific sites29–31. In addition to expressing full-length Ikaros (Ik-1), most or all hematopoietic cells have abundant expression of a smaller isoform that lacks finger 1 (Ik-2). Ikaros isoforms that lack other fingers because of alternative splicing of pre-mRNA have been observed, but these isoforms are generally much less abundant except in transformed cell lines29,32,33.

1Department of Microbiology, Immunology, and Molecular Genetics, University of California, Los Angeles, California, USA.
2Molecular Biology Institute, University of California, Los Angeles, California, USA.
3Department of Biochemistry and Molecular Biology, Norris Comprehensive Cancer Center, University of Southern California, Los Angeles, California, USA.
4Howard Hughes Medical Institute, University of California, Los Angeles, California, USA.
5Division of Laboratory Animal Medicine, University of California, Los Angeles, California, USA.
6Department of Pathology and Laboratory Medicine, University of California, Los Angeles, California, USA.
7Eli and Edythe Broad Center of Regenerative Medicine and Stem Cell Research, University of California, Los Angeles, California, USA.
8Present addresses: Department of Laboratory Medicine, University of California, San Francisco, California, USA (H.S.), and School of Biological Sciences, University of Northern Colorado, Greeley, Colorado, USA (S.F.). Correspondence should be addressed to S.T.S. (smale@mednet.ucla.edu).

Received 9 July; accepted 14 August; published online 8 September 2013; doi:10.1038/ni.2707

© 2013 Nature America, Inc. All rights reserved.
Given that two or three fingers are generally sufficient for stable binding of DNA\(^{28}\), it is striking that most zinc-finger DNA-binding proteins contain more than three tandem fingers. The reasons for this phenomenon have been of interest since the early studies of the nine-finger protein TFIIB\(^{27,34,35}\). All of the tandem fingers may be required for stable binding to key targets in vitro, with the multiple fingers recognizing extended DNA regions with high specificity. Alternatively, defined subsets of fingers may recognize distinct target genes. Initial support for the latter hypothesis emerged from in vitro protein-DNA-interaction studies showing that different combinations of fingers from TFIIB, Ikaros, CTCF and other proteins can bind different DNA sequences\(^{29,30,35–39}\). In addition, studies of ectopically expressed mutant CTCF by chromatin immunoprecipitation followed by deep sequencing (ChIP-seq) have revealed that CTCF-binding sites can be separated into distinct classes on the basis of binding by different finger subsets\(^{40}\).

To determine in a physiological setting whether multifinger DNA-binding proteins use different fingers to regulate different genes and biological functions, or whether all functions require the full complement of fingers, we generated two new Ikaros-mutant mouse strains with germline deletion of \(\text{Ikzf1}\) exons encoding fingers 1 and 4. Phenotypic analysis suggested that the two fingers regulate distinct biological events. Transcriptome profiling confirmed that numbers above bracketed lines indicate percent of live cells (right). Each symbol (right) represents an individual mouse; small horizontal lines indicate the mean (± s.e.m.). Data are representative of three experiments (a), four experiments, (b) fourteen experiments (c) or five to fifteen experiments (d; average and s.e.m., left and middle).

The two fingers regulate distinct sets of genes, and DNA-binding domains exist at least in part to allow the proper regulation of only a small number of genes that are targets of Ikaros.

**RESULTS**

**Generation of mutant mouse strains**

We generated mice lacking Ikaros zinc-finger 1 (\(\text{Ikzf1}^{\Delta F1/\Delta F1}\) or zinc-finger 4 (\(\text{Ikzf1}^{\Delta F4/\Delta F4}\) by germline deletion of exon 4 or exon 6, respectively (Supplementary Fig. 1a). Each deletion generated a mutant protein lacking a zinc finger and a small number of adjacent residues. Exon 4 encodes 87 residues, including the 23 residues of finger 1, 63 residues preceding the finger, and one residue following the finger. Exon 6 encodes only 41 residues. The 24-residue finger 4 is preceded by four residues that constitute the linker between fingers 3 and 4; these residues closely resembles the linker consensus sequence found in most zinc-finger proteins\(^{28}\). After the finger-encoding sequence, exon 6 encodes 13 additional residues; two prolines near the end of this region may separate the DNA-binding domain from an exon 7–encoded domain.

We confirmed correct targeting of the \(\text{Ikzf1}\) exons by Southern blot analysis, PCR (data not shown) and immunoblot analysis (Fig. 1a). \(\text{Ikzf1}^{\Delta F1/\Delta F1}\) thymocytes lacked the full-length Ikaros isoform (Ik-1) but expressed Ik-2, which in wild-type mice is generated by alternative splicing of pre-mRNA\(^2\). In the \(\text{Ikzf1}^{\Delta F4/\Delta F4}\) strain, deletion of exon 6 reduced the size of both Ik-1 and Ik-2 (Fig. 1a). Analysis of thymocyte mRNA by high-throughput sequencing technology (RNA-Seq) further confirmed correct targeting, as ‘reads’ from exons 4 and 6 were absent from the \(\text{Ikzf1}^{\Delta F1/\Delta F1}\) and \(\text{Ikzf1}^{\Delta F4/\Delta F4}\) samples, respectively (Fig. 1b).

**Initial analysis of mutant strains**

We initiated examination of the \(\text{Ikzf1}^{\Delta F1/\Delta F1}\) and \(\text{Ikzf1}^{\Delta F4/\Delta F4}\) phenotypes soon after the mice were generated but confirmed all phenotypes after backcrossing those mice onto the C57BL/6 background through nine or more generations. Efforts to backcross \(\text{Ikzf1}^{\Delta F1/\Delta F1}\) mice have been unsuccessful (K. Georgopoulos and S. Winandy, personal communication), presumably because the null mutation results in embryonic death on the C57BL/6 background. The successful backcrossing of the \(\text{Ikzf1}^{\Delta F1/\Delta F1}\) and \(\text{Ikzf1}^{\Delta F4/\Delta F4}\) strains provided initial evidence that neither finger 1 nor finger 4 was required for all functions of Ikaros.

Initial studies revealed that the development of CD11b\(^+\) myeloid cells was unperturbed in each strain, similar to the phenotype of \(\text{Ikzf1}^{\Delta F1/\Delta F1}\) mice (Fig. 1c and Supplementary Fig. 1b). Conventional B cells were also present in the bone marrow and spleen of each mutant strain, albeit in smaller amounts than those in wild-type mice (Fig. 1d and Supplementary Fig. 1c,d). The presence of B cells was in contrast to the absence of all cells of the B lineage in \(\text{Ikzf1}^{\Delta F1/\Delta F1}\) mice. These results suggested that DNA-binding fingers 2 and 3 were sufficient for the successful progression of a fraction of hematopoietic progenitors through B cell development.
Additional finger-specific functions in hematopoiesis

The developmental abnormalities of T cells were far more severe in Ikzf1^{ΔF4/ΔF4} mice than in Ikzf1^{ΔF1/ΔF1} mice. Both strains exhibited slightly reduced thymic cellularity, with a greater reduction in Ikzf1^{ΔF4/ΔF4} mice, relative to that of their wild-type littermates (Fig. 3a). During the CD4−CD8− double-negative stages 1–4 (DN1–DN4) of thymopoiesis, reduced cell numbers were readily apparent in Ikzf1^{ΔF4/ΔF4} mice but not in Ikzf1^{ΔF1/ΔF1} mice, relative to their abundance in wild-type mice (Fig. 3b and Supplementary Fig. 3a–d).
The DN1 population (CD44(+)CD25(−)) seemed to be smaller in Ikkzf1(−/−) mice than in wild-type mice (Fig. 3b and Supplementary Fig. 3a). However, that reduction was due to greatly reduced numbers of all other cell types that contributed to the DN1 population, including thymic B cells (discussed below). Consistent with that interpretation, the percentage of early thymic progenitor cells was normal in Ikkzf1(−/−) mice but reduced in Ikkzf1(−/−) mice (Fig. 3c). The frequency of DN2 (CD44(DB25(−))) and DN3 (CD44(DB25(−))) cells was also greatly reduced in Ikkzf1(−/−) mice but normal in Ikkzf1(−/−) mice, relative to their abundance in wild-type mice (Fig. 3b and Supplementary Fig. 3b,c). DN4 cells (CD44(DB25(−))) recovered in the Ikkzf1(−/−) strain to percentages and numbers similar to those observed in wild-type mice, presumably by homeostatic proliferation (Fig. 3b and Supplementary Fig. 3d). The ratio of CD4(+)CD8(−) double-positive (DP) thymocytes to CD4(+) or CD8(−) single-positive (SP) thymocytes in Ikkzf1(−/−) mice was similar to that of wild-type mice but was significantly reduced in Ikkzf1(−/−) mice (Fig. 3d and Supplementary Fig. 3e). The reduced number of DP thymocytes in Ikkzf1(−/−) mice correlated with a reduction in the size of the thymic cortex (Fig. 3e). Neither Ikkzf1(−/−) nor Ikkzf1(−/−) mice exhibited the extreme skewing toward the CD4(+) lineage that characterizes Ikkzf1(−/−) mice (Fig. 3d and Supplementary Fig. 3f).

Ikkzf1(−/−) mice have reduced numbers of thymic γδ T cells. These cells were largely unperturbed in number in Ikkzf1(−/−) mice and increased in number in Ikkzf1(−/−) mice, relative to their abundance in wild-type mice (Fig. 3f). Thymocyte analysis also revealed a severe reduction in the number of thymic B cells (CD19(+)CD4(−)) in Ikkzf1(−/−) mice relative to that of wild-type mice (Fig. 3g). This selective and consistent phenotype is difficult to explain, given the similar number of conventional B-2 cells in the spleen of the two mutant strains. Furthermore, examination of the peritoneal cavity revealed reduced numbers of the B-1a population in each mutant strain and a reduced number of B-1b cells only in the Ikkzf1(−/−) strain, relative to that of wild-type mice (Supplementary Fig. 3g,h). Further analysis of these phenotypes may provide insight into the developmental origin, homing mechanism and/or function of thymic B cells and peritoneal B-1 cells.
Although adult mutant mice had substantial numbers of mature B cells and T cells, the *Ikzf1ΔF4/ΔF4* strain lacked cells of both the B lineage and T lineage in the fetus (Fig. 3h–j and Supplementary Fig. 4a–c). In contrast, fetal T cell numbers were normal and fetal B cells were only moderately reduced in the *Ikzf1ΔF1/ΔF1* strain (Fig. 3h–j and Supplementary Fig. 4a–c). Lymphoid tissue–inducer cells (CD45+CD3ε+CD4+IL-7R+) were also selectively absent from the fetal intestinal mesentery of *Ikzf1ΔF4/ΔF4* mice (Fig. 3k). This phenotype correlated with the selective absence of lymph nodes (inguinal and lumbar) and Peyer’s patches in adult *Ikzf1ΔF4/ΔF4* mice (Fig. 3l and Supplementary Fig. 4d), consistent with the important role of fetal lymphoid tissue–inducer cells in the development of these lymphoid structures. Nasal–associated lymphoid tissue, which develops postnatally, was intact (Supplementary Fig. 4e). These developmental defects were reminiscent of those described for *Ikfz1*null mice, which suggested that zinc-finger 4 was required for the proper regulation of genes encoding molecules involved in early stages of fetal lymphopoiesis. Notably, zinc-finger 1 and all other residues encoded by exon 4 seemed to be dispensable for these developmental events.

**Finger-specific DNA binding in vivo**

The generation of the *Ikzf1ΔF1/ΔF1* and *Ikzf1ΔF4/ΔF4* strains was inspired by evidence that fingers 1 and 4 support the binding to different DNA sequences in *vitro*. To determine whether the two fingers also participated in binding to different DNA sequences in *vivo*, we used ChIP-seq analysis of total thymocytes from 4-week-old wild-type, *Ikzf1ΔF1/ΔF1* and *Ikzf1ΔF4/ΔF4* mice. The most prominent ChIP-seq peaks were present at simple repetitive elements that contained multiple copies of the core Ikaros consensus sequence of GGGAA (Fig. 4a–d). Some of these simple repeats were located in close proximity to protein-encoding genes, but these repeats have not been conserved through evolution and therefore may not have important functions. Nevertheless, the ChIP-seq profiles confirmed that fingers 1 and 4 contributed to differences in DNA binding in *vivo*. For example, a repeat region in *Snx25* (which encodes a sorting nexin) exhibited a prominent ChIP-seq peak in wild-type thymocytes and in thymocytes from both mutant strains (Fig. 4a), which suggested that fingers 2 and 3 were sufficient for binding to this region. In contrast, a prominent ChIP-seq peak in *Nr3c2* (which encodes the mineralocorticoid receptor Mr) was present in wild-type and *Ikzf1ΔF4/ΔF4* thymocytes but not in *Ikzf1ΔF1/ΔF1* thymocytes (Fig. 4b). At a repeat region adjacent to *Snx25* (which encodes a sorting nexin) (Fig. 4d), and at a repeat in chromosome 19 that is not linked to an annotated gene (Fig. 4c), ChIP-seq peaks were present in wild-type and *Ikzf1ΔF4/ΔF4* thymocytes but not in *Ikzf1ΔF1/ΔF1* thymocytes.

Inspection of the repeats revealed regions containing the Ikaros consensus sequence GGGAA but with different flanking nucleotides. Motif analysis of these regions in *Snx29* and *Nr3c2* revealed two repetitive sequences that may support the binding of Ikaros, while one repetitive sequence that may support the binding of Ikaros was present in the regions in *Snx25* and chromosome 19 (Fig. 4). In *vivo* binding-site-selection experiments have suggested that finger 1 recognizes a pyrimidine-rich sequence downstream of the Ikaros core (S.E.W. and S.T.S., data not shown). On the basis of those studies, finger 1 would be predicted to bind the TCC sequence in one of the *Nr3c2* repeats (Fig. 4b, bottom right). The other *Nr3c2* consensus (Fig. 4b, bottom left) and the two *Snx29* consensus sequences (Fig. 4a) may also contact finger 1, but perhaps more weakly because the nucleotides that would be contacted by finger 1 include a purine. In contrast, we predict that the repeats in *Snx25* and chromosome 19 (Fig. 4c,d) have little ability to interact with finger 1 because the GGGAA core is followed by purine residues.

---

**Figure 4** Differences in the binding of Ikaros to DNA in *Ikzf1ΔF1/ΔF1* and *Ikzf1ΔF4/ΔF4* thymocytes. (a–d) ChIP-seq analysis of Ikaros binding to thymocytes from 4-week-old wild-type, *Ikzf1ΔF1/ΔF1* and *Ikzf1ΔF4/ΔF4* mice, presented as UCSC Genome Browser tracks for four genomic regions with repetitive Ikaros-recognition motifs (which yields unusually strong ChIP-seq peaks), showing interactions that do not depend on finger 1 or finger 4 (*Snx29*) or that depend on finger 1 (*Nr3c2*) or finger 4 (chromosome (Chr) 19 and *Snx25*). Input, sequencing of input genomic DNA (negative control). Below, repetitive genomic sequences, with putative Ikaros-recognition sites (yellow highlight, GGGAA or GGA; green highlight, TTCC); bottom, position weight matrices of the putative Ikaros-recognition sequences that were repeated, with the zinc fingers (F1–F4) predicted to contact or lie in close proximity to the nucleotides flanking the GGGAA core recognition sequence. (e) EMSA with oligonucleotide probes for three repetitive sequences and recombinant proteins containing Ikaros fingers 1–4, 1–3 or 2–4 (above). Below, zinc fingers (F1 or F4) predicted to contact or lie in close proximity to the nucleotides flanking the GGGAA core recognition sequence. (f) Bioinformatics analysis of a wild-type Ikaros ChIP-seq data set, showing the localization of the core GGGAA sequence relative to the center of each ChIP-seq peak. Data are representative of three experiments (a–d), three or more experiments (e) or two experiments (f).
Figure 5  Deregulation of distinct sets of genes in *Ikzf1*ΔF1/ΔF1 and *Ikzf1*ΔF4/ΔF4 DP thymocytes. (a) RNA-Seq analysis of mRNA from DP thymocytes sorted by flow cytometry from 4-week-old mice, presented as a Venn diagram of genes upregulated (Up) or downregulated (Down) at least threefold (*P* ≤ 0.001) in either *Ikzf1*ΔF1/ΔF1 or *Ikzf1*ΔF4/ΔF4 mice relative to their expression in wild-type mice (for genes with an RPKM value of 2× in at least one of the six samples). (b) Genes in a with an increased (Up) or decreased (Down) mRNA abundance of at least tenfold (*P* ≤ 0.001) in one of the mutant strains. (c) Gene-ontology analysis of genes upregulated more than threefold (*P* ≤ 0.001) in *Ikzf1*ΔF4/ΔF4 DP thymocytes. (d) Distribution of RPKM values (mRNA abundance) for all annotated genes (left) and all genes upregulated by at least threefold (*P* ≤ 0.001) in *Ikzf1*ΔF4/ΔF4 mice (middle) or *Ikzf1*ΔF4/ΔF4 mice (right). Data are representative of two experiments (duplicates).

The binding site–selection data and additional findings (S.E.W. and S.T.S., data not shown) failed to reveal a strong DNA sequence ‘preference’ for finger 4. Instead, finger 4 seemed to stabilize binding in a DNA sequence–independent manner, with its DNA contacts particularly important at sites not stably bound by fingers 2 and 3 alone (S.E.W. and S.T.S., data not shown). We hypothesize that finger 4 may be important for the repeats in Snx25 and chromosome 19 because the DNA-binding energy of fingers 2 and 3 is weakened by the presence of a purine just downstream of the GGGAA core; we have found that a pyrimidine at this position is important for stable protein-DNA interactions by fingers 2 and 3 (S.E.W. and S.T.S., data not shown).

To determine whether the difference in binding observed by ChIP-seq was due to intrinsic DNA-binding ‘preferences’, we did electrophoretic mobility-shift assays (EMSA) with recombinant proteins containing different combinations of zinc fingers. Consistent with the ChIP-seq results, we observed substantial binding to one of the Snx29 sequences with a protein containing all four fingers, as well as with proteins containing only fingers 1–3 or 2–4 (Fig. 4e). The second Snx29 repeat (with a GaGAA core instead of the GGGAA core, where the lowercase ‘a’ indicates the difference from the consensus core sequence) was unable to bind any of the proteins (Supplementary Fig. 5a).

We observed substantial binding to one of the Nr3c2 sequences with proteins containing fingers 1–4 or 1–3 but not with a protein containing fingers 2–4 (Fig. 4e), which demonstrated finger 1–dependent binding. Finally, we observed substantial binding to the Snx25 repetitive sequence with proteins containing fingers 1–4 or fingers 2–4 but not with a protein containing fingers 1–3 (Fig. 4e). Thus, the intrinsic in vitro binding capacity mirrored the in vivo ChIP-seq results.

In addition to identifying the prominent interactions noted above (Fig. 4), the ChIP-seq results revealed thousands of other Ikaros-binding sites throughout the genome. More than 60% of the ChIP-seq peaks coincided with the Ikaros core sequence of GGGAA (Fig. 4f and Supplementary Fig. 5b). In experiments with wild-type or mutant thymocytes, the number of ‘called peaks’ varied from about 1,000 to 17,000. Those numbers are in the same range as the 7,000 Ikaros peaks in thymocytes reported before16, with extensive overlap observed between the peaks in our experiments and those reported before. Zfp64, Zfp260, Cd4, Notch1, Hdac7, and Bcl11b are examples of genes exhibiting the same Ikaros peaks in our experiments and in published reports16,18,19 (Supplementary Fig. 5c). Motif-enrichment analysis revealed that the core Ikaros recognition sequence of GGGAA was highly prevalent at peaks that overlapped in the wild-type data set and in the data sets from both mutant strains (Supplementary Fig. 5d). Furthermore, the sequence GGGAAAGGGAA was prevalent at peaks selectively absent from the *Ikzf1*ΔF4/ΔF4 sample (Supplementary Fig. 5d). That sequence is very similar to the sequences noted above (Fig. 4c,d) that exhibited finger 4 dependence, which suggested that finger 4 would be broadly important for binding to such sequences throughout the genome. Thus, the ChIP-seq data supported a model in which the different functions of fingers 1 and 4 are due at least in part to their ability to facilitate binding to distinct genomic sites (discussed below).

**Selective misregulation of gene expression**

Although thousands of Ikaros-binding sites were identified by ChIP-seq, the highly selective phenotypes of the *Ikzf1*ΔF1/ΔF1 and *Ikzf1*ΔF4/ΔF4 strains suggested that DNA-binding fingers 1 and 4 may contribute to the regulation of distinct sets of genes. To examine the roles of fingers 1 and 4 in transcriptional control, we obtained mRNA from CD4+CD8+ thymocytes sorted from 4-week-old wild-type, *Ikzf1*ΔF1/ΔF1, and *Ikzf1*ΔF4/ΔF4 mice and analyzed the mRNA by RNA-Seq. The results revealed misregulation of small number of genes in each of the mutant strains, relative to their expression in wild-type mice. In *Ikzf1*ΔF4/ΔF4 thymocytes, only 110 genes and 12 genes (with RPKM (reads per kilobase per million mapped reads) values of 2×) were upregulated and downregulated, respectively, by more than threefold, with only 24 and 21 genes upregulated and downregulated, respectively, by this magnitude in *Ikzf1*ΔF1/ΔF1 thymocytes (Fig. 5a and Supplementary Fig. 6). Notably, only 15 genes were upregulated in both mutant strains and 1 gene was downregulated in both strains. Three genes upregulated in *Ikzf1*ΔF4/ΔF4 thymocytes were downregulated in *Ikzf1*ΔF1/ΔF1 thymocytes. Notably, much smaller numbers of genes were upregulated or downregulated by tenfold or more (Fig. 5b and Supplementary Fig. 6). Although it is difficult to accurately compare RNA-Seq data sets to microarray data sets, many of the genes found by RNA-Seq to be misregulated in *Ikzf1*ΔF1/ΔF1 or *Ikzf1*ΔF4/ΔF4...
thymocytes have been found by microarray to be misregulated in $\text{Ikzf1}^{\text{null}}$ thymocytes. Thus, misregulated genes in $\text{Ikzf1}^{\text{null}}$ thymocytes seem to represent a composite of the genes misregulated in $\text{Ikzf1}^{\text{F1/F1}}$ and $\text{Ikzf1}^{\text{F4/F4}}$ thymocytes.

Gene-ontology analysis of genes upregulated threefold or more in $\text{Ikzf1}^{\text{F4/F4}}$ thymocytes revealed enrichment for genes encoding molecules involved in cell adhesion, cell communication and signal transduction (Fig. 5c). Consistent with the invasive properties of the thymic lymphomas that arose in older $\text{Ikzf1}^{\text{F4/F4}}$ mice (reported below), several genes encoding molecules involved in tumor invasion and metastasis were among the genes upregulated the most, including $\text{Mmp14}$ (which encodes a matrix metalloprotease), $\text{Cmmnd1}$ (which encodes $\beta$-catenin) and $\text{Dock1}$ (which encodes a guanine nucleotide–exchange factor) (Supplementary Fig. 6e,g).

Notably, genes that were upregulated in either $\text{Ikzf1}^{\text{F1/F1}}$ or $\text{Ikzf1}^{\text{F4/F4}}$ thymocytes generally had very low expression in wild-type thymocytes (Fig. 5d). We assessed the distribution by the expression (RPKM) in thymocytes of all annotated genes and the RPKM distribution in wild-type cells of genes upregulated by at least threefold in $\text{Ikzf1}^{\text{F1/F1}}$ or $\text{Ikzf1}^{\text{F4/F4}}$ thymocytes (Fig. 5d). As the vast majority (>96%) of upregulated genes were upregulated from very low expression (RPKM < 5), these results were consistent with evidence that Ikaros often functions as a transcriptional repressor.

**Zinc-finger requirements for tumor suppression**

Published Ikaros-mutant strains develop thymic lymphoma. We never observed lymphoma in $\text{Ikzf1}^{\text{F4/F4}}$ mice, but $\text{Ikzf1}^{\text{F4/F4}}$ mice developed thymic lymphoma with a penetrance similar to that reported for other Ikaros-mutant strains (Fig. 6a). Also similar to other Ikaros-mutant strains, the lymphomas in $\text{Ikzf1}^{\text{F4/F4}}$ mice were aggressive (highly invasive and metastatic) (Fig. 6b,c) and they displayed variable expression of the coreceptors CD4 and CD8 (Fig. 6d), clonal rearrangement of the gene encoding the T cell antigen receptor $\beta$-chain, and aberrant expression of the intracellular domain of the signaling receptor Notch (data not shown).

**Ikzf1 mutations in a mouse model of BCR-ABL+ B-ALL**

Although $\text{Ikzf1}$ mutations invariably give rise to T cell malignancies in mice, human malignancies of T cell origin rarely have $\text{IKZF1}$ mutations. Instead, human $\text{IKZF1}$ mutations are frequently associated with BCR-ABL+ B-ALL and other progenitor–B cell malignancies. To investigate this species difference, and with the additional goal of developing models with which to study progenitor–B cell malignancies associated with $\text{IKZF1}$ mutations, we first made use of a well-established in vitro culture assay for BCR-ABL+ B-ALL. We obtained bone marrow cells from wild-type, $\text{Ikzf1}^{\text{F1/F1}}$ and $\text{Ikzf1}^{\text{F4/F4}}$ mice and transduced those cells with a retrovirus expressing the oncoprotein tyrosine kinase BCR-ABL, then monitored colony numbers in the progenitor B cell cultures over the course of approximately 3 weeks. Transduced $\text{Ikzf1}^{\text{F4/F4}}$ cells proliferated much more rapidly than did their wild-type counterparts (Fig. 7a). Notably, transduced $\text{Ikzf1}^{\text{F4/F4}}$ cells proliferated more slowly than did their wild-type counterparts and eventually stopped growing (Fig. 7a and data not shown).

We next used an in vivo model of BCR-ABL+ B-ALL in which we transplanted bone marrow cells transduced with the retrovirus expressing BCR-ABL into irradiated wild-type C57BL/6 recipient mice. In this assay, transduced $\text{Ikzf1}^{\text{F4/F4}}$ cells yielded much more aggressive malignancies than did their wild-type or $\text{Ikzf1}^{\text{F4/F4}}$ counterparts (Fig. 7b). These results demonstrated that an $\text{Ikzf1}$ mutation was able to contribute to B cell malignancy in mice when combined with BCR-ABL expression. Furthermore, these results established an important parallel between the tumor-suppressor function of Ikaros in thymocytes and progenitor cells–B cells; in both cell types, finger 4 was essential for tumor suppression, whereas finger 1 was dispensable.

To approach an understanding of the tumor-suppressor function of Ikaros, we transduced cells with retrovirus expressing BCR-ABL, obtained mRNA from the cells after 21 or 28 d of culture and analyzed the mRNA by RNA-Seq. We also analyzed sorted pro-B cells established an important parallel between the tumor-suppressor function of Ikaros in thymocytes and progenitor cells–B cells; in both cell types, finger 4 was essential for tumor suppression, whereas finger 1 was dispensable.
**Figure 7** Selective synergy between BCR-ABL and the *Ikzf1*ΔF4/ΔF4 mutation in vitro and in vivo. (a) In vitro growth of bone marrow cells from wild-type, *Ikzf1*ΔF1/ΔF1, and *Ikzf1*ΔF4/ΔF4 mice transduced with a BCR-ABL-expressing retrovirus and grown under B-ALL culture conditions. (b) Survival of irradiated recipient mice given transplantation of 1 × 10^6 BCR-ABL-transduced bone marrow cells from wild-type mice (n = 15), *Ikzf1*ΔF1/ΔF1 mice (n = 8) or *Ikzf1*ΔF4/ΔF4 mice (n = 14), presented as Kaplan-Meier curves. (c) RNA-Seq analysis of mRNA from sorted pro-B cells (pro-B) and pre-BI plus large pre-BII cells (pre-BI-II), as well as from B-ABL-transformed cells at day 21 or 28 of culture, for genes whose mRNA abundance differed by threefold or more between any two samples among the 12 samples analyzed (presented as k-means clustering). (d) Genes of the two clusters with genes selectively upregulated or downregulated in BCR-ABL-transformed *Ikzf1*ΔF4/ΔF4 cells, identified by k-means clustering of the data sets from B-ABL-transformed cultures; expression data from pro-B cells and pre-BI cells plus large pre-BII cells were aligned after the cluster analysis was completed. (e) Flow cytometry of the in vitro cultures in a, gated on cells expressing yellow fluorescent protein as a marker of BCR-ABL transduction. (f) Expression of *Il2ra* mRNA (encoding CD25) and *Kit* mRNA (encoding c-Kit), from the RNA-Seq data in c (n = 2 samples per genotype). Data are from one experiment representative of three experiments (a) or three or more experiments (e) or are from three (b) or two (d–f) combined experiments (error bars, s.e.m.).

The remaining three clusters (4–6) included genes with higher expression in BCR-ABL-transformed *Ikzf1*ΔF1/ΔF1 cells than in their wild-type or *Ikzf1*ΔF4/ΔF4 counterparts (Fig. 7c). Those same genes also had higher expression in untransformed *Ikzf1*ΔF4/ΔF1 pre-BI cells plus large pre-BII cells (Fig. 7c), which suggested that the *Ikzf1*ΔF1/ΔF1 cells that proliferated in culture following BCR-ABL transformation were maintained at the pre-BI cell–to–large pre-BII cell developmental stage, a stage at which many genes were misregulated in *Ikzf1*ΔF1/ΔF1 mice. Notably, most genes in all six clusters had similar expression in wild-type and *Ikzf1*ΔF4/ΔF4 cells (Fig. 7c), despite the greatly enhanced proliferation of BCR-ABL-transformed *Ikzf1*ΔF4/ΔF4 cells.

To identify genes selectively misregulated in BCR-ABL-transformed *Ikzf1*ΔF4/ΔF4 cells, we did an additional cluster analysis with the six data sets obtained with transformed cells; we then aligned the expression of the respective genes in untransformed cells after we defined the clusters. This analysis revealed 155 genes that were selectively upregulated and 133 genes that were selectively downregulated in *Ikzf1*ΔF4/ΔF4 cultures relative to their expression in wild-type and *Ikzf1*ΔF1/ΔF1 cultures, at both day 21 and day 28 (Fig. 7d). Among the genes with higher expression in *Ikzf1*ΔF4/ΔF4 cells was *Kit* (which encodes the stem cell factor receptor c-Kit) (Fig. 7d–f), which is known to be silenced during B cell maturation. Among the genes with lower expression were *Il2ra* (which encodes the IL-2 receptor α-chain CD25) and *Enpep* (which encodes the cell surface marker BP-1), which are activated during B cell maturation (Fig. 7d–f). Flow cytometry confirmed misregulation of the proteins encoded by *Kit* (c-Kit) and *Il2ra* (CD25) (Fig. 7e).

These findings suggested that the *Ikzf1*ΔF4/ΔF4 mutation may influence the developmental stage of cells transformed by BCR-ABL. However, most of the other genes that were selectively misregulated in BCR-ABL-transformed *Ikzf1*ΔF4/ΔF4 cells are not developmentally regulated, and several, such as Dock1 and Cimnd1, correspond to genes that were also misregulated in *Ikzf1*ΔF4/ΔF4 thymocytes. This identification of a limited set of genes selectively misregulated in transformed cells in the context of the *Ikzf1*ΔF4/ΔF4 mutation provides an important step toward delineating the elusive mechanisms responsible for the tumor-suppressor function of Ikaros.

**DISCUSSION**

We have created two new mouse strains in which exons encoding individual zinc fingers of Ikaros were deleted. Our results have demonstrated that different zinc fingers in a single DNA-binding domain can participate in the regulation of distinct sets of genes and contribute to distinct biological functions. The disruption of individual fingers may be a generally valuable strategy for delineating the complex biological functions and mechanisms of action of zinc-finger transcription factors.

When designing the mutant strains, we considered three different strategies: deletion of the exons encoding fingers 1 and 4;
deletion of DNA sequences encoding fingers 1 and 4, with retention of the remaining sequences of exons 4 and 6; or mutagenesis of specific nucleotides encoding key residues of fingers 1 and 4 involved in DNA binding. Each strategy had notable advantages and limitations. An advantage of the second and third strategies is that it would be possible to exclude the possibility of involvement of non-finger residues in the resulting phenotypes. However, a major limitation of those strategies is that the mutant proteins do not normally exist in mouse cells, which would raise the possibility of dominant-negative or gain-of-function activities. Consistent with that, a point mutation in the Iksf1 region encoding finger 3 has been found to have much stronger dominant-negative properties in mice than a deletion mutant that results in complete removal of exons encoding fingers 1, 2 and 3 (ref. 51). Furthermore, preliminary data have suggested that exon 4-encoded residues immediately upstream of finger 1 autoregulate the DNA-binding activity of finger 1 (S.E.W. and S.T.S., data not shown); retention of those residues with deletion of finger 1 would have an uncertain outcome.

Because of the limitations noted above, we deleted exons 4 and 6, which encode fingers 1 and 4, respectively, as well as the small number of additional residues encoded by these exons. With this strategy, it remains possible that some of the hematological defects are due to the loss of activities other than the DNA-binding activity of the deleted finger. However, we were able to conclude with confidence that finger 1 and finger 4 were fully and differently dispensable for several biological functions of Ikaros, and our data suggested that at least a subset of the defects of each strain was due to loss of finger-dependent DNA interactions. Furthermore, the Iksf1Δf1/Δf1 strain revealed the specific functions of full-length Ikaros (Ik-1), as this strain lacked the Ik-1 isoform while retaining the naturally occurring Ik-2 isoform. Moreover, since the long-term goal is to delineate the highly selective phenotypes of the mutant strains for the purpose of understanding how Ikaros regulates lymphopoiesis and leukemogenesis, we can proceed with relatively little concern about the possibility of aberrant gain-of-function activities.

We unexpectedly found that the two fingers regulated different biological functions and even different steps in the same developmental pathway. We originally anticipated that Ikaros would regulate numerous genes involved in each developmental step in which it participates, such that each mutant strain would exhibit phenotypes similar to that of Iksf1null mice. That expectation was based on evidence that gene expression is substantially altered in Iksf1null cells and that Ikaros binds to several thousand genomic sites, as shown by ChIP-seq experiments. Instead, we found that finger 1 was largely or fully dispensable for many biological functions of Ikaros, with finger 4 being dispensable for several other functions.

That selectivity provides support for a hypothesis in which only one target gene or a small number of target genes is (are) essential for each biological event in which Ikaros participates. It is difficult to envision that Ikaros would be a critical regulator of hundreds of genes required for each biological event, with finger 1 being dispensable for the regulation of all target genes involved in some events and finger 4 being dispensable for the regulation of all genes involved in other events. Instead, it seems much more likely that each event requires the proper regulation of only one or a few genes that are targets of Ikaros, with either finger 1 or finger 4 being critical for the regulation of those genes. Although we favor a model in which Ikaros is an essential regulator of only a few key genes involved in each biological event, an alternative possibility is that fingers 2 and 3 are sufficient for the regulation of almost all critical target genes involved in each of its biological functions, with fingers 1 or 4 essential for the regulation of only a few key genes.

A longstanding challenge in the transcription field has been the identification of direct targets of transcription factors that can explain their biological functions. Efforts to identify such targets have benefited from the development of gene-expression profiling methods and methods for examining DNA binding in vivo on a genome-wide scale (such as ChIP-seq). By merging ChIP-seq data with the gene-expression profiles of wild-type and mutant cells, insights into the direct targets of a transcription factor can be obtained. However, a major limitation of this approach is that many transcription factors may bind genomic sites at which they do not function. Moreover, because some transcription factors may ‘preferentially’ bind sequences associated with the relatively open chromatin found at active promoters and enhancers, it can be difficult to conclude with confidence that the merging of gene-expression and ChIP data sets has successfully identified the direct functional targets.

In our RNA-Seq studies, we found that most misregulated genes in Iksf1Δf1/Δf1 and Iksf1Δf4/Δf4 thymocytes had low expression in wild-type cells, whereas Ikaros-binding sites defined by ChIP-seq were distributed among genes at all levels of expression (data not shown). Furthermore, the number of ChIP-seq peaks greatly exceeded the number of misregulated genes. One possible and perhaps likely interpretation of these results is that most Ikaros-binding sites identified by ChIP-seq are not functionally relevant. However, many other possibilities must be considered. For example, Ikaros may contribute to the proper regulation of a much larger number of genes, but its absence may alter the level of expression of most target genes by a magnitude that fails to reach the threefold cutoff used for our analyses. Another possibility is that Ikaros acts redundantly with other members of the Ikaros family at many of its target genes.

We originally were hopeful that ChIP-seq analysis of our mutant strains would reveal a distinct loss of Ikaros binding to a well-defined subset of sites, which would lead to improved correlation between binding and transcriptional misregulation of nearby genes. However, although we observed reduced numbers of ChIP-seq peaks for Iksf1Δf1/Δf1 and Iksf1Δf4/Δf4 thymocytes (data not shown), we have been unable to do a meaningful analysis of the binding events that require or are enhanced by these fingers. A chief reason for this difficulty is that the ChIP-seq peaks were not clearly polarized, in that we did not observe a distinct subset of peaks that were entirely dependent on finger 4 and another subset of peaks that were entirely dependent on finger 1. Instead, we observed a continuum of effects, with different yet highly variable degrees of dependence on finger 1 or finger 4.

The potential value of targeting individual zinc fingers can be summarized as follows. First, the Iksf1Δf1/Δf1 and Iksf1Δf4/Δf4 mice exhibited select subsets of the hematopoietic defects of Iksf1null mice. Second, when we observed a phenotype, it was as robust or nearly as robust as that of Iksf1null mice, which makes the phenotype amenable to further analysis. Third, despite the robust phenotypes, gene-expression changes were quite limited, which narrowed the list of potential direct and indirect target genes responsible for the phenotype. These findings pave the way for detailed analyses of each phenotype of the mutant strains.

Our results raise the question of why Ikaros and perhaps other zinc-finger transcription factors acquired the ability to bind DNA through different subsets of their fingers. One possibility is that this strategy allows Ikaros to recognize a larger number of DNA sequences, so that each target gene does not need to have a sequence that matches a single, rigid consensus. Although that simple scenario is possible, we favor a model in which recognition of DNA through different combinations of fingers supports different functions. One possibility
is that binding to DNA through different combinations of zinc fingers leads to conformational differences in the bound protein that may influence coregulatory interactions. The zinc-finger DNA-binding protein CTCF has similarly been suggested to carry out different functions when bound to DNA through different subsets of its fingers, yet this hypothesis has not yet been tested through the disruption of individual CTCF fingers.

Finally, it can be argued that the most important function of Ikaros to understand is its tumor-suppressor function. The correlation between IKZF1 mutations and the therapeutic response of high-risk progenitor B-ALL suggests that understanding of the tumor-suppressor function of Ikaros may be critical for the development of new therapies. The establishment of a mouse model for BCR-ABL+ background in which many target genes of Ikaros involved in its other functions are disrupted will make this strain particularly valuable for future studies, as it exhibits only a subset of the hematopoietic defects of Ikaros−/− mice. This feature will allow studies of the tumor-suppressor function on a background in which many target genes of Ikaros involved in its other functions are relatively unperturbed.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. GEO: RNA-Seq and ChIP-seq data, GSE33693.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank P. Allahamad, D. Bhatt, K. Dorshkind, C. Li and E. Montecino-Rodriguez for advice and/or critical reading of the manuscript; the Division of Laboratory Animal Medicine of the University of California, Los Angeles, for ongoing care of mice; and H. Mak, T. Jacob, C. Garcia, J. Flores and J. Lorenzano for assistance with the mouse colony. ChIP-seq and RNA-seq libraries were sequenced at the Epigenome Data Production Facility of the University of Southern California, and the Broad Stem Cell Research Center High Throughput Sequencing Core of the University of California, Los Angeles. Supported by the US National Institutes of Health (ROI1DK043726 to S.T.S. and U54HG004558 to P.J.F.). O.N.W. is an Investigator of the Howard Hughes Medical Institute.

AUTHOR CONTRIBUTIONS

H.S., J.M., T.L.A., S.E. D.C., S.E.W. and G.W.L. designed and did experiments and analyzed data; S.J.B. provided intellectual input and experimental advice; P.J.F., O.N.W. and S.T.S. supervised research and analyzed data; and H.S. and S.T.S. wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Reprints and permissions information is available online at http://www.nature.com/ reprints/index.html.

1. Wang, J.H. et al. Selective defects in the development of the fetal and adult lymphoid system in mice with an Ikaros null mutation. Immunity 5, 537–549 (1996).

2. Georgopoulos, K. Haematopoietic cell-fate decisions, chromatin regulation and Ikaros. Nat. Rev. Immunol. 2, 162–174 (2002).

3. Yoshida, T., Ng, S.Y. & Georgopoulos, K. Awakening lineage potential by Ikaro-mediated transcriptional priming. Curr. Opin. Immunol. 22, 154–160 (2010).

4. Yoshida, T., Ng, S.Y., Zuniga-Pflucker, J.C. & Georgopoulos, K. Early hematopoietic lineage restrictions directed by Ikaros. Nat. Immunol. 7, 382–391 (2006).

5. Kirstetter, P., Thomas, M., Dierich, A., Kastner, P. & Chan, S. Ikaros is critical for the regulation of cell differentiation and formation of the thymus. Eur. J. Immunol. 32, 7932–7940 (2002).

6. Thompson, E.C. et al. Ikaros DNA-binding proteins as integral components of B cell developmental-stage-specific regulatory circuits. Immunity 26, 335–344 (2007).

7. Reynaud, D. et al. Regulation of B cell fate commitment and immunoglobulin heavy-chain gene rearrangements by Ikaros. Nat. Immunol. 9, 927–936 (2008).

8. Trager, D. et al. Pre-B cell receptor-mediated cell cycle arrest in Philadelphia chromosome-positive acute lymphoblastic leukemia requires IKAROS function. J. Exp. Med. 206, 1739–1753 (2009).

9. McNeil, A. et al. Ikaros and Aiolos inhibit B-cell proliferation by directly suppressing c-Myc expression. Mol. Cell Biol. 30, 4149–4158 (2010).

10. Winandy, S., Wu, P. & Georgopoulos, K. A dominant mutation in the Ikaros gene leads to rapid development of leukemia and lymphoma. Cell 83, 289–299 (1995).

11. Dumortier, A. et al. Notch activation is an early and critical event during T-cell leukaemogenesis in Ikaros-deficient mice. Mol. Cell Biol. 26, 209–220 (2006).

12. Mullighan, C. et al. BCR-ABL1 lymphoblastic leukaemia is characterized by the deletion of Ikaros. Nature 453, 110–114 (2008).

13. Mullighan, C. et al. Deletion of IKZF1 and prognosis in acute lymphoblastic leukemia. N. Engl. J. Med. 360, 470–480 (2009).

14. Kim, J. et al. Ikaros DNA-binding proteins direct formation of chromatin remodeling complexes in lymphocytes. Immunity 10, 345–355 (1999).

15. Sridharan, R. & Smale, S.T. Predominant interaction of both Ikaros and Helios with the NuRD complex in immature thymocytes. J. Biol. Chem. 282, 30227–30238 (2007).

16. Zhang, J. et al. Harnessing of the nucleosome-remodeling-deacetylase complex controls lymphocyte development and prevents leukemogenesis. Nat. Immunol. 13, 86–94 (2012).

17. Harker, N. et al. The CD8a gene locus is regulated by the Ikaros family of proteins. Mol. Cell 10, 1403–1415 (2002).

18. Naito, T., Gomez-del Arco, P., Williams, C.J. & Georgopoulos, K. Antagonistic interactions between Ikaros and the chromatin remodeler Mediator determine silencer activity and Cd4 gene expression. Immunity 27, 723–734 (2007).

19. Gomez-del Arco, P. et al. Alternative promoter usage at the Notch1 locus supports Igand-independent signaling in T cell development and leukemogenesis. Immunity 33, 685–698 (2010).

20. Ng, S.Y., Yoshida, T., Zhang, J. & Georgopoulos, K. Genome-wide lineage-specific transcriptional networks underscores Ikaros-dependent lymphoid priming in hematopoietic stem cells. Immunity 30, 493–507 (2009).

21. John, L.B. & Ward, A.C. The Ikaros gene family: transcriptional regulators of hematopoiesis and immunity. Mol. Immunol. 48, 1272–1278 (2011).

22. Sun, L., Liu, A. & Georgopoulos, K. Zinc finger-mediated protein interactions modulate Ikaros activity, a molecular control of lymphocyte development. EMBO J. 15, 5358–5369 (1996).

23. Trinh, L.A. et al. Down-regulation of TDT transcription in CD4+CD8+ thymocytes by Ikaros proteins in direct competition with an Ets activator. Genes Dev. 15, 1817–1832 (2001).

24. McCarthy, A.S., Kleiger, G., Eisenberg, D. & Smale, S.T. Selective dimerization of a C2H2 zinc finger subfamily. Mol. Cell 11, 459–470 (2003).

25. Tupper, R., Perini, G. & Green, M.R. Expressing the human genome. Nature 409, 832–833 (2001).

26. Ravai, T. et al. Systematic characterization of the zinc-finger-containing proteins in the mouse transcriptome. Genome Res. 13, 1430–1442 (2003).

27. Klug, A. The discovery of zinc fingers and their applications in gene regulation and genome manipulation. Annu. Rev. Biochem. 79, 213–231 (2010).

28. Wolfe, S.A., Nekludova, L. & Pabo, C.O. DNA recognition by Cys2His2 zinc finger proteins. Annu. Rev. Biochem. Biomol. Struct. 29, 183–212 (2000).

29. Molnar, A. & Georgopoulos, K. The Ikaros gene encodes a family of functionally diverse zinc finger DNA-binding proteins. Mol. Cell Biol. 14, 8292–8303 (1994).

30. Cobb, B.S. et al. Targeting of Ikaros to pericentromeric heterochromatin by direct DNA binding. Genes Dev. 14, 2162–2170 (2000).

31. Koipally, J., Heller, E.J., Seavitt, J.R. & Georgopoulos, K. Unconventional potentiation of gene expression by Ikaros. J. Biol. Chem. 277, 13007–13015 (2002).

32. Hehl, M. et al. Helios, a T cell-restricted Ikaros family member that quantitatively associates with Ikaros at centromeric heterochromatin. Genes Dev. 12, 782–796 (1998).

33. Payne, K.J. et al. Ikaros isoform x is selectively expressed in myeloid differentiation. J. Immunol. 170, 3091–3098 (2003).

34. Miller, J., McLachlan, A.D. & Klug, A. Repetitive zinc-binding domains in the protein transcription factor IIIA from Xenopus oocytes. EMBO J. 4, 1609–1614 (1985).

35. Shadry, B.S. Transcription factor IIIA (TFIIIA) in the second decade. J. Cell Sci. 109, 535–539 (1996).

36. Filippova, G.N. et al. An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian c-myc oncogenes. Mol. Cell Biol. 16, 2802–2813 (1996).

37. Ohlsson, R., Renkawitz, R. & Lobanenkov, V. CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease. Trends Genet. 17, 520–527 (2001).
38. Renda, M. et al. Critical DNA binding interactions of the insulator protein CTCF: a small number of zinc fingers mediate strong binding, and a single finger-DNA interaction controls binding at imprinted loci. J. Biol. Chem. 282, 33336–33345 (2007).
39. Nurmemmedov, E., Yengo, R.K., Uysal, H., Karlsson, R. & Thunnissen, M.M. New insights into DNA-binding behavior of Wilms tumor protein (WT1)–a dual study. Biophys. Chem. 145, 116–125 (2009).
40. Nakahashi, H. et al. A genome-wide map of CTCF multivalency redefines the CTCF code. Cell Rep. 3, 1678–1689 (2013).
41. Hardy, R.R., Carmack, C.E., Shinton, S.A., Kemp, J.D. & Hayakawa, K. Resolution and characterization of pro-B and pre-pro-B cell stages in normal mouse bone marrow. J. Exp. Med. 173, 1213–1229 (1991).
42. Rolink, A., Grawunder, U., Winkler, T.H., Karasuyama, H. & Melchers, F. IL-2 receptor α chain (CD25, TAC) expression defines a crucial stage in pre-B cell development. Int. Immunol. 6, 1257–1264 (1994).
43. Rothenberg, E.V., Moore, J.E. & Yui, M.A. Launching the T-cell-lineage developmental programme. Nat. Rev. Immunol. 8, 9–21 (2008).
44. Randall, T.D., Carragher, D.M. & Rangel-Moreno, J. Development of secondary lymphoid organs. Annu. Rev. Immunol. 26, 627–650 (2008).
45. Juric, D. et al. Differential gene expression patterns and interaction networks in BCR-ABL-positive and -negative adult acute lymphoblastic leukemias. J. Clin. Oncol. 25, 1341–1349 (2007).
46. Lu, Q. delta-Catenin dysregulation in cancer: interactions with E-cadherin and beyond. J. Pathol. 222, 119–123 (2010).
47. Wang, H. et al. The role of Crk/Dock180/Rac1 pathway in the malignant behavior of human ovarian cancer cell SKOV3. Tumour Biol. 31, 59–67 (2010).
48. Perentes, J.Y. et al. Cancer cell-associated MT1-MMP promotes blood vessel invasion and distant metastasis in triple-negative mammary tumors. Cancer Res. 71, 4527–4538 (2011).
49. Marcias, A. et al. Genetic inactivation of Ikaros is a rare event in human T-ALL. Leuk. Res. 34, 426–429 (2010).
50. Wong, S. et al. Sole BCR-ABL inhibition is insufficient to eliminate all myeloproliferative disorder cell populations. Proc. Natl. Acad. Sci. USA 101, 17456–17461 (2004).
51. Papathanasiou, P. et al. Widespread failure of hematolymphoid differentiation caused by a recessive niche-filling allele of the Ikaros transcription factor. Immunity 19, 131–144 (2003).
52. Lickwar, C.R., Mueller, F., Hanlon, S.W., McNally, J.G. & Lieb, J.D. Genome-wide protein-DNA binding dynamics suggest a molecular clutch for transcription factor function. Nature 484, 251–255 (2012).
53. Virely, C. et al. Haploinsufficiency of the IKZF1 (IKAROS) tumor suppressor gene cooperates with BCR-ABL in a transgenic model of acute lymphoblastic leukemia. Leukemia 24, 1200–1204 (2010).
ONLINE METHODS

Mice. Germine-targeted deletion of exon 4 (encoding zinc finger 1) or exon 6 (encoding zinc finger 4) of *Ikzf1* was achieved by homologous recombination in embryonic stem cells. Embryonic stem cells were transfected by electroporation (Molecular Genetic Technology Center of the University of California, Los Angeles (UCLA)) and were screened for correct targeting by Southern blot analysis. Mutant embryonic stem cell lines were injected into 129 blastocysts by the UCLA Molecular Genetic Technology Center. After germline transmission was achieved, the flox-flanked neomycin-resistance cassette introduced into the genome during homologous recombination was removed by crossing of the mice with *Ella-Cre* mice on a C57BL/6 background. Germline transmission of the deletions was confirmed after backcrossing with wild-type C57BL/6 mice. The mutant strains were subsequently backcrossed back more than ten generations with wild-type C57BL/6. Mice were excluded from analysis of hematopoietic development if they were runted or if they had developed thymic lymphoma (*Ikzf1*<sup>-/-4/4</sup> strain). No randomization or ‘blinding’ was used for the animal studies. Mice of both sexes were used, at E18.5 (fetal hematopoiesis), 4 weeks (thymic development) and 6–8 weeks of age (peripheral lymph nodes, peritoneal B cells, spleen and bone marrow). All phenotypes described were analyzed with a minimum sample size of n = 5, with wild-type litters as controls and are represented by at least three separate experiments with mice from different litters. Animals were housed in the vivaria of the UCLA Division of Laboratory Animal Medicine. All experiments were approved by the UCLA Animal Research Committee and were done according to guidelines of the UCLA Institutional Animal Care and Use Committee. Wild-type C57BL/6 and *Ella-Cre* mice were from The Jackson Laboratory.

Cell preparation and flow cytometry. Cell suspensions were prepared from adult hematopoetic tissues and were filtered through 70-µm nylon cell strainers (BD Biosciences). Fetal hematopoietic cells were prepared as described from mesentery<sup>54</sup> and fetal liver<sup>55</sup>. All antibodies for flow cytometry were from BD Biosciences or eBioscience, except for the antibody to cytoplasmic immunoglobulin-µ, which was from Southern Biotech (*Supplementary Table 1*). All antibodies were initially used at a dilution of 1:200 and were individually ‘titrated’ when necessary. Intracellular staining of cytoplasmic immunoglobulin-µ was done with reagents from the Fop3 Staining Buffer set (eBioscience). A FACSCalibur, BD LSR I, or FACSaria II SOP (BD Biosciences) was used for flow cytometry or sorting, and data were analyzed with Flowjo 7.5 and FACSDiva 6.1.1 software.

Histological analysis and visualization of lymph nodes and Peyer’s patches. Dissected tissues were fixed for >72 h in 10% formalin (4% formaldehyde) (Fisher Scientific), decalcified when needed, embedded in paraffin, sectioned and stained with hematoxylin and eosin by the Translational Pathology Core Laboratory at UCLA. Inguinal and lumbar lymph nodes and Peyer’s patches were visualized as described<sup>54,56</sup>.

Immunoblot analysis. Whole-cell extracts were prepared by resuspension of cells in one volume water in the presence of protease inhibitor ‘cocktail’ (Roche) and by immediate direct lysis by the addition of one volume of 2x SDS sample buffer. Samples were separated by SDS-PAGE and transferred to nitrocellulose membrane, and blots were probed with an antibody raised against the N-terminal region of Ikaros (residues 1–80)<sup>52</sup> or antibody to SNAP70 (loading control)<sup>57</sup>.

Retroviral transduction, cell culture and bone marrow transplantation. Retroviral supernatants for expression of BCR-ABL were produced with the plasmid pMSCV-YFP-IRES-p185, and bone marrow cells were transduced as described<sup>50</sup>, except that whole bone marrow from untreated mice was used for transduction. For *in vitro* cell culture, 5 × 10<sup>6</sup> transduced cells were plated on top of pre-established feeder cells from wild-type C57BL/6 bone marrow stroma. Cells were split and counted and then reseeded at a density of 1 × 10<sup>5</sup> to 2 × 10<sup>5</sup> cells per ml in a volume of 5 ml every 2–3 d for growth analysis. For *in vivo* experiments, 1 × 10<sup>6</sup> transduced whole bone marrow cells were injected intravenously into irradiated wild-type C57BL/6 recipient mice. Mice were monitored for development of B-ALL, with the endpoint of complete paralysis of hind legs or a moribund condition.

RNA purification and RNA-Seq. RNA was prepared with TRI Reagent (Molecular Research Center), followed by purification with an RNeasy kit (Qiagen), with on-column treatment with RNase-free DNase I. For RNA-Seq analysis, MicroPoly(A) Purist Kit (Ambion) or the Tru-Seq RNA Sample Prep Kit (Illumina) was used to isolate mRNA according to the manufacturer’s protocol. cDNA was prepared for the Illumina Sequencing platform by the fragmented double-stranded cDNA protocol<sup>58</sup> or an Illumina TruSeq kit. Samples were sequenced on an Illumina HiSeq instrument with 50–base pair single-end reads, at the UCLA Broad Stem Cell Research Center High Throughput Sequencing Core. Raw data were uploaded to the Galaxy website (Pennsylvania State University) and were mapped (with the TopHat fast splice junction mapper for RNA-Seq) to the July 2007 annotation of the mouse (*Mus musculus*) genome (the mma9 assembly of the mouse genome of the National Center for Biotechnology Information), with filtering for uniquely mapped ‘reads’<sup>59–63</sup>. Genome coverage visualization files were created with the BEDTools flexible software suite of utilities for the comparison of genomic features and were uploaded to University of California at Santa Cruz (UCSC) Genome Browser<sup>64</sup>. Relative mRNA expression (RPKM)<sup>65</sup> was calculated on the basis of exonic ‘reads’ with SeqMonk software (Babraham Bioinformatics) and reference genome annotations from the National Center for Biotechnology Information (the mm9 assembly of the mouse genome). For RNA-Seq analysis of DP thymocytes, exonic ‘reads’ were analyzed by the Bioconductor DESeq program<sup>66</sup> for the identification of genes with statistically significant different expression in two populations. Clustering analysis (k-means) was done on log<sub>2</sub>-transformed, mean centered RPKM values, with Cluster 3.0 software<sup>67</sup> and visualized with Java Treeview software<sup>69</sup>. Gene ontology analysis was performed using Panther software<sup>70</sup>.

ChIP and ChiP-seq analysis. Primary thymocytes were fixed for 10 min at 25 °C in 1% formaldehyde, then were washed in PBS and ‘snap frozen’. ChIP of Ikaros with an antibody directed against the N-terminal region of Ikaros (residues 1–80)<sup>52</sup> was followed by PCR analysis of enrichment at published Ikaros-binding sites<sup>18,19</sup> and unpublished (Zfp260) Ikaros-binding sites identified in initial ChiP-seq experiments. Input DNA was prepared from wild-type thymocytes as a negative control. Libraries were prepared for sequencing on the Illumina platform as described<sup>72</sup> and were sequenced at the University of Southern California Epigenome Data Production Facility and the UCLA Broad Stem Cell Research Center High Throughput Sequencing Core. Raw sequences were mapped (with Bowtie software for the alignment of short DNA sequences) to the July 2007 annotation of the mm9 assembly, and significant peaks over background were ‘called’ with Sole-Search software<sup>72,73</sup>. Genome coverage visualization files were created with BEDTools and were uploaded to the UCSC Genome Browser<sup>64</sup>. Homer software<sup>74</sup> was used for de novo motif analysis, and MEME software<sup>75</sup> was used for position weight matrix analysis.

Recombinant protein preparation and EMSA. *Ikzf1* sequences encoding DNA-binding zinc fingers 1–4 were amplified by PCR from cDNA and were subcloned into the bacterial expression plasmid pGEX-4T-1 (GE Healthcare) by standard methods. Recombinant proteins were expressed in Rosetta (DE3) competent cells (Novagen), were induced with isopropyl β-d-thiogalactopyranoside, and purified using B-PER GST Fusion Protein Purification Kit (Thermo Scientific). The glutathione S-transferase tags were removed and proteins were purified by on-column thrombin cleavage. EMSA of Ikaros was done as described<sup>15</sup>, except that bacterially expressed recombinant Ikaros proteins were used instead of nuclear extracts. Sequences of probes are in *Supplementary Table 2*.

Statistical analysis. GraphPad Prism software was used for statistical analyses (unpaired, two-tailed, Student’s t test, 95% confidence intervals).

54. Z. *et al*. Requirement for RORgamma in thymocyte survival and lymphoid organ development. *Science* **288**, 2369–2373 (2000).
55. Montecino-Rodriguez, E., Leathers, H. & Dorshkind, K. Identification of a B-1 B cell-specified progenitor. *Nat. Immunol.* 7, 293–301 (2006).
56. Aliabadi, P., de la Torre, B. & Kaye, J. Shared dependence on the DNA-binding factor TOX for the development of lymphoid tissue-inducer cell and NK cell lineages. *Nat. Immunol.* 11, 945–952 (2010).
57. Pandya-Jones, A. & Black, D.L. Co-transcriptional splicing of constitutive and alternative exons. *RNA* 15, 1896–1908 (2009).
58. Nagalakshmi, U., Waern, K. & Snyder, M. RNA-Seq: a method for comprehensive transcriptome analysis. *Curr. Protoc. Mol. Biol.* 89, 4.11.1–4.11.13 (2010).
59. Giardine, B. et al. Galaxy: a platform for interactive large-scale genome analysis. *Genome Res.* 15, 1451–1455 (2005).
60. Trapnell, C., Pachter, L. & Salzberg, S.L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 25, 1105–1112 (2009).
61. Blankenberg, D. et al. Galaxy: a web-based genome analysis tool for experimentalists. *Curr. Protoc. Mol. Biol.* 89, 19.10.1–19–10–21 (2010).
62. Fujita, P.A. et al. The UCSC Genome Browser database: update 2011. *Nucleic Acids Res.* 39, D876–D882 (2010).
63. Goecks, J., Nekrutenko, A. & Taylor, J. Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biol.* 11, R86 (2010).
64. Kent, W.J. et al. The human genome browser at UCSC. *Genome Res.* 12, 996–1006 (2002).
65. Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* 5, 621–628 (2008).
66. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* 11, R106 (2010).
67. Eisen, M.B., Spellman, P.T., Brown, P.O. & Botstein, D. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA* 95, 14863–14868 (1998).
68. de Hoon, M.J., Imoto, S., Nolan, J. & Miyano, S. Open source clustering software. *Bioinformatics* 20, 1453–1454 (2004).
69. Saldana, A.J. Java Treeview - extensible visualization of microarray data. *Bioinformatics* 20, 3246–3248 (2004).
70. Thomas, P.D. et al. Applications for protein sequence-function evolution data: mRNA/protein expression analysis and coding SNP scoring tools. *Nucleic Acids Res.* 34, W645–W650 (2006).
71. O’Geen, H., Frietze, S. & Farnham, P.J. Using ChIP-seq technology to identify targets of zinc finger transcription factors. *Methods Mol. Biol.* 649, 437–455 (2010).
72. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S.L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10, R25 (2009).
73. Blahnik, K.R. et al. Sole-Search: an integrated analysis program for peak detection and functional annotation using ChIP-seq data. *Nucleic Acids Res.* 38, e13 (2010).
74. Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* 38, 576–589 (2010).
75. Bailey, T.L. et al. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* 37, W202–W208 (2009).