Published in final edited form as:

*Nat Immunol*. 2021 July ; 22(7): 893–903. doi:10.1038/s41590-021-00951-z.

# A Variant in Human AIOLOS Impairs Adaptive Immunity by Interfering with IKAROS

**Motoi Yamashita**[1,2], **Hye Sun Kuehn**[3,#], **Kazuki Okuyama**[2,#], **Satoshi Okada**[4], **Yuzaburo Inoue**[5,6], **Noriko Mitsuiki**[1], **Kohsuke Imai**[1,7], **Masatoshi Takagi**[1], **Hirokazu Kanegane**[1,8], **Masahiro Takeuchi**[9,10], **Naoki Shimojo**[5,11], **Miyuki Tsumura**[4], **Aditya K. Padhi**[12], **Kam Y. J. Zhang**[12], **Bertrand Boisson**[13,14,15], **Jean-Laurent Casanova**[13,14,15,16], **Osamu Ohara**[17], **Sergio D. Rosenzweig**[3], **Ichiro Taniuchi**[2,*], **Tomohiro Morio**[1,*]

[1]Department of Pediatrics and Developmental Biology, Graduate School of Medical and Dental Sciences, Tokyo Medical and Dental University, Tokyo, 113-8510, Japan.

[2]Laboratory for Transcriptional Regulation, RIKEN Center for Integrative Medical Sciences, Kanagawa, 230-0045, Japan.

[3]Immunology Service, Department of Laboratory Medicine, National Institute of Health, Bethesda, MD, USA.

[4]Department of Pediatrics, Hiroshima University Graduate School of Biomedical and Health Sciences, Hiroshima, 734-8551, Japan.

[5]Department of Pediatrics, Graduate School of Medicine, Chiba University, Chiba, 260-8677, Japan.

[6]Department of Allergy and Rheumatology, Chiba Children's Hospital, Chiba, 266-0007, Japan.

[7]Department of Community Pediatrics, Perinatal and Maternal Medicine, Tokyo Medical and Dental University, Tokyo, 113-8510, Japan.

[8]Department of Child Health and Development, Tokyo Medical and Dental University, Tokyo, 113-8510, Japan.

[9]Department of Hematology, Chiba University Hospital, Chiba, 260-8677, Japan.

[10]Department of Hematology and Medical Oncology, Chiba Cancer Center, Chiba, 260-8717, Japan.

[11]Center for Preventive Medical Sciences, Chiba University, Chiba, 260-8677, Japan.

[*]Correspondence: Ichiro Taniuchi (ichiro.taniuchi@riken.jp) and Tomohiro Morio (tmorio.ped@tmd.ac.jp).
[#]These authors contributed equally

Author contributions

M.Y. and K.O. analyzed mice and performed experiments using cell lines. H.S.K. performed reporter gene assay, EMSA, and co-localization assay. S.O. performed EMSA. N.M. and O.O. performed genetic analysis. M. Takeuchi, N.S., and Y.I. provided clinical information and samples of the patients. A.K.P and K.Y.J.Z. performed the homology modeling and molecular dynamics simulations. K.I., M. Takagi, H.K., and S.D.R. provided intellectual guidance. B.B. and J.L.C. provided cohort information. T.M. and I.T. supervised project. M.Y., I.T., and T.M. wrote the manuscript.

Competing interests

The authors declare no competing financial interest.

[12]Laboratory for Structural Bioinformatics, RIKEN Center for Biosystems Dynamics Research, Kanagawa, 230-0045, Japan.

[13]St. Giles Laboratory of Human Genetics of Infectious Diseases, The Rockefeller University, New York, NY 10065, USA.3000

[14]Laboratory of Human Genetics of Infectious Diseases, Necker Branch, INSERM U1163, Imagine Institute, Necker Hospital for Sick Children, Paris, France.

[15]University of Paris, Paris, France.

[16]Howard Hughes Medical Institute, New York, NY, USA.

[17]Department of Applied Genomics, Kazusa DNA Research Institute, Chiba 292-0818, Japan.

## Abstract

Here, we report a human-inherited impaired adaptive immunity disorder, which predominantly manifested as a B cell differentiation defect, caused by a heterozygous *IKZF3* missense variant, resulting in a glycine to arginine replacement within the DNA binding domain of the encoded AIOLOS protein. Using mice that bear the corresponding variant and recapitulate the B and T cell phenotypes, we show that the mutant AIOLOS homodimers and AIOLOS–IKAROS heterodimers did not bind the canonical AIOLOS/IKAROS DNA sequence. Additionally, homodimers and heterodimers containing one mutant AIOLOS bound to genomic regions lacking both canonical motifs. However, the removal of the dimerization capacity from mutant AIOLOS restored B cell development. Hence, the adaptive immunity defect is caused by the AIOLOS variant hijacking IKAROS function. Heterodimeric interference is a novel mechanism of autosomal dominance that causes inborn errors of immunity by impairing protein function via the mutation of its heterodimeric partner.

Inborn errors of immunity (IEIs), also known as primary immunodeficiency, are characterized by immune system defects resulting in increased susceptibility to infections or dysregulated immune responses. Most IEIs are monogenic, and more than 430 causative genes have been reported[1]. Although classical IEIs are recessive and caused by homozygous, compound heterozygous, or hemizygous loss-of-function or hypomorphic mutations, recent studies have identified several heterozygous variants as causal for dominant IEIs through mechanisms of gain-of-function[2–5], negative dominance[6–8], or haploinsufficiency[9].

The IKAROS zinc finger (IKZF) protein family is a group of transcription factors that play essential roles in hematopoiesis. Domains that define this family are the zinc fingers (ZFs) in the middle of the molecule that bind to DNA and two ZFs located at the C-terminus for homo- and hetero-dimerization with other IKZF members. Early mouse studies demonstrated that the deletion of Ikaros, a pioneer member of the IKZF family encoded by *Ikzf1*, and the dominant-negative Ikaros mutation result in defective development of T cells, B cells, and natural killer (NK) cells[10,11]. B cell deficiency in these mice indicates that the function of Ikaros is crucial for B cell development. Conditional deletion of Ikaros in the early developmental stage further blocked development at the pro-B cell stage[12]. Recently, germline heterozygous *IKZF1* variants were identified in patients with B cell deficiency that could manifest as common variable immunodeficiency (CVID) or combined

immunodeficiency, among other phenotypes, through haploinsufficiency, dominant-negative, or dimerization defect mechanisms[6,9,13–18]. Variable autoimmune disorders, such as systemic lupus erythematosus and immune thrombocytopenic purpura, were also reported in these patients[6,9,13–18]. Besides the regulation of lymphocyte development and immune tolerance, IKAROS is essential to the suppression of hematologic malignancies. Somatic mutations and deletions to *IKZF1* have been reported in pediatric B cell precursor acute lymphoblastic leukemia (B-ALL)[19,20]. Some *IKZF1* germline variants also predispose B cell precursors to malignant transformation (B-ALL)[21]. However, the molecular pathogenesis of developmental and autoimmune disorders and tumorigenesis caused by heterozygous *IKZF1* mutations is not well understood.

AIOLOS, another member of the IKZF family, is encoded by *IKZF3*. Although deletions to *IKZF3* in BCP-ALL have been reported[19], the relevance of AIOLOS in the early development of B lymphocytes is underrated since primary B cell development in bone marrow (BM) is not disrupted by Aiolos deficiency in mice[22]. In this study, we describe a new hereditary error of adaptive immunity with a predominant B cell defect caused by a heterozygous missense variant in *IKZF3*, which generates the AIOLOS$^{G159R}$ variant. The use of a mouse model harboring a human variant showed that AIOLOS$^{G159R}$ variant causes B cell developmental defects and T cell abnormalities by interfering with IKAROS function in heterodimers as well as the dysfunction of AIOLOS homodimers.

## Results

### A family of patients with B-cell deficiency

The study subjects were a family of patients with B cell deficiency and susceptibility to Epstein–Barr virus (EBV) infection. P1 (II.3) and P2 (III.2) experienced recurrent sinopulmonary infections (Fig. 1a), and both developed B cell lymphoma in adulthood. Another family member (P3, III.1) also exhibited susceptibility to EBV infection. (Fig. 1a, Supplementary Table 1, and Supplementary Clinical Information). In all affected patients, B cells were undetectable in the peripheral blood (Fig. 1b and Supplementary Table 1). Thus, P1 underwent a BM examination, which revealed low proportions of B cell lineage-committed progenitors ($CD10^+CD19^+$) and immature B cells ($CD34^-CD10^-CD19^+IgM^+$) fractions (Fig. 1b).

Lymphocyte subset analysis of P1 revealed abnormalities in T cell subsets as well as a marked reduction in the number of B cells, whereas $CD4^+$ T cells were skewed to the memory phenotype (Extended Data Fig. 1a and Supplementary Table 2). $CD4^+$ and $CD8^+$ T cells exhibited the activated phenotype, characterized by the expression of CD38 and HLA-DR. Additionally, the proportion of $T_H17$ cells was decreased, whereas that of $T_H1$* cells ($CD3^+CD4^+CD45RO^+CD161^+CCR6^+CXCR3^+$) was increased. Meanwhile, the T cells had decreased expression levels of CD3 and T cell receptor α/β (TCRα/β) and slightly decreased CD8 expression (Extended Data Fig. 1b). Similarities in the phenotypes suggested that the patients developed IEIs with autosomal dominant inheritance.

### Heterozygous missense variant of *IKZF3*

Whole-exome sequencing of P1, P2, and healthy family members was performed to identify causative germline genetic variants (Extended Data Fig. 2a). The pedigree suggested autosomal dominant inheritance; thus, heterozygous non-synonymous or splice mutations shared by the patients were selected (Extended Data Fig. 2b). Among the 13 candidates, we focused on a heterozygous missense variant in *IKZF3* (p.G159R, c.475G>C in NM_012481.4), given that *IKZF1* variants are known to be associated with B cell deficiency. Subsequent Sanger sequencing identified the same variant in a deceased family member (P3) who had the same phenotype as P1 and P2 (Fig. 1a). The *IKZF3* c.475G>C variant is not listed in the Single Nucleotide Polymorphism Database, Human Genetic Variation Database, or Bravo database and was not detected in an in-house B cell deficiency cohort of more than 100 patients or a cohort of 5,000 patients with severe infections. Nevertheless, G159 of *IKZF3* is highly conserved across species (Extended Data Fig. 2c) and G159R variant was deemed as damaging by combined annotation-dependent depletion analysis (Phred score = 31). These data suggested that heterozygous G159R variant in the *IKZF3* gene is causal for the impaired immunity in the patients.

### Altered DNA binding specificity of AIOLOS$^{G159R}$

*IKZF3* encodes AIOLOS, one of the five members of the IKZF family, which are expressed at various levels during T cell development. However, *IKZF1* and *IKZF3* are predominantly expressed in progenitor and mature B cells[23]. Analysis of the expression levels of the IKZF family of transcription factors during different developmental stages of human B cells confirmed the induction of *IKZF3* during the transition of pro-B cells to the large pre-B stage with constitutive expression of *IKZF1* (Extended Data Fig. 2d). The c.475G>C variant generates the AIOLOS$^{G159R}$ variant harboring a glycine to arginine replacement in the second ZF domain, which is involved in DNA binding (Fig. 1c). Homology modeling analysis with the use of a template of a six-finger fragment of human PR/SET domain 9 allele C (PDB ID: 5V3G)[24] predicted that G159 of AIOLOS directly interacts with DNA and the G159R substitution results in steric clashes with DNA (Fig. 2a–c). Computer simulations of molecular dynamics indicated that the G159R substitution reduces the binding capacity of the AIOLOS$^{G159R}$ mutant by pushing DNA away from the second ZF (Fig. 2b, c). The results of an electrophoretic mobility shift assay (EMSA) with the use of a DNA probe containing the canonical AIOLOS sequence (IK-BS4)[25] and nuclear extracts of human embryonic kidney (HEK) 293T cells transfected with wild-type (WT) AIOLOS or the AIOLOS$^{G159R}$ construct revealed that WT AIOLOS formed a complex with the DNA probe, whereas the AIOLOS$^{G159R}$ mutant failed to bind to the probe (Fig. 2d). Importantly, when AIOLOS$^{G159R}$ was co-expressed with WT AIOLOS at a 1:1 ratio, the binding of WT AIOLOS to the probe was significantly reduced (Fig. 2e). These observations indicate not only that AIOLOS$^{G159R}$ failed to recognize the canonical AIOLOS target DNA sequence but also that the mutant functions as a dominant-negative form to alter the DNA binding capacity of WT AIOLOS (Fig. 2e).

Next, the genome-wide binding ability of the AIOLOS$^{G159R}$ variant was assessed using the human pre-B cell line NALM-6 (Fig. 3a). After the isolation of a NALM-6 clone lacking AIOLOS expression by CRISPR-Cas9 genome editing, FLAG-tagged WT AIOLOS

or AIOLOS$^{G159R}$ was expressed using a retroviral vector (Fig. 3b and Extended Data Fig. 3a, b). AIOLOS$^{G159R}$ is apparently not simply a loss-of-function mutant as a substantial number of AIOLOS$^{G159R}$ peaks were detected via FLAG-chromatin immunoprecipitation followed by next-generation sequencing (ChIP-seq) analysis. 14715 peaks were detected in 3 out of 4 replicates of WT AIOLOS (4465, 25214, 34706, and 18066 peaks for each replicates); and 52738 peaks were detected in 3 out of 4 replicates of AIOLOS$^{G159R}$ (85290, 67057, 69456, and 81714 peaks for each replicates). Physiological binding of AIOLOS (represented by WT AIOLOS peaks) was lost or often undetectable at sites with the AIOLOS consensus sequence (Fig. 3c and Extended Data Fig. 3c, *EIPR1* and *MIR4303* loci). Conversely, novel genome regions bound by AIOLOS$^{G159R}$ were detected (Fig. 3c and Extended Data Fig. 3c, *LOC283683* and *RNF212* loci). The summary heatmap analysis of quadruplicate ChIP-seq data clearly differentiated genome-wide binding of AIOLOS$^{G159R}$ from that of WT AIOLOS (Fig. 3d, e). Quantitatively, 2146 and 45709 peaks were shown to have significantly strong binding by WT AIOLOS and AIOLOS$^{G159R}$, respectively, with a false discovery rate (FDR) cut-off of <0.05 (Fig. 3f). Concordant with the EMSA data, AIOLOS$^{G159R}$ lost the ability to bind to the AIOLOS canonical binding motif (GGGAA and TGGAA) (Fig. 3g). Surprisingly, motif analysis revealed that AIOLOS$^{G159R}$ gained the ability to bind to previously unreported binding sequence for AIOLOS (GGAGC, GGAGG, and CCCAGA). These aberrant motifs were enriched in the highest AIOLOS$^{G159R}$ peaks and differentially bound peaks between WT AIOLOS and AIOLOS$^{G159R}$ (Fig. 3h and Extended Data Fig. 3d). In fact, some of the highest AIOLOS$^{G159R}$ peaks were often enriched with these aberrant sequences, whereas the WT AIOLOS peaks were enriched with the AIOLOS consensus sequence (Fig. 3c). Direct binding of AIOLOS$^{G159R}$ to the aberrant sequence was confirmed via EMSA with a probe sequence taken from GGAGC repeat containing genome region with high AIOLOS$^{G159R}$ ChIP-seq peak (Extended Data Fig. 3e). Collectively, these data suggest that AIOLOS$^{G159R}$ is not simply a loss-of-function mutant but also gained the ability to specifically bind to a novel DNA sequence.

## B cell developmental defect in Aiolos$^{G158R}$ mutant mice

Considering that animal models are powerful tools to elucidate the pathogenesis of human diseases[26], an *Ikzf3* knock-in mouse model that carries a glycine to arginine substitution (G158R) corresponding to that of the patient was generated via CRISPR-Cas9 genome editing (Fig. 4a). During this process, an *Ikzf3* knock-out mouse line (*Ikzf3*$^{-/-}$) was also established by introducing a frameshift mutation in exon 5 of *Ikzf3* (Fig. 4a, Extended Data Fig. 4a). Although a slight reduction in the *Ikzf3* mRNA level was observed in *Ikzf3*$^{G158R/G158R}$ pre-B cells, *Ikzf3* expression in *Ikzf3*$^{+/G158R}$ pre-B cells was comparable with that of the WT control (Extended Data Fig. 4b). Aiolos and Ikaros expression levels in the thymus of *Ikzf3*$^{+/G158R}$ and *Ikzf3*$^{G158R/G158R}$ mice, as assessed via immunoblot analysis, were comparable with those of the WT control (Extended Data Fig. 4c, d). Meanwhile, *Ikzf3* expression of the *Ikzf3*$^{G158R}$ mutant allele in pre-B cells was comparable with that of the WT control mice (Extended Data Fig. 4e), indicating that protein expression is not altered in the *Ikzf3*$^{G158R}$ mutant. The expression of other *Ikzf* family genes, and most importantly that of Ikaros was unaffected by the *Ikzf3*$^{G158R}$ mutant (Extended Data Fig. 4b).

Next, B cell development was evaluated in *Ikzf3*$^{G158R}$ knock-in mutant mice. Profound B cell developmental defects were observed in *Ikzf3*$^{G158R/G158R}$ mice, as manifested by a significant block at the pre-pro-B cell (B220$^{+}$AA4.1$^{+}$IgM$^{-}$CD43$^{+}$CD24$^{-}$) to pro-B cell (B220$^{+}$AA4.1$^{+}$IgM$^{-}$CD43$^{int}$) and pre-B cell (B220$^{+}$AA4.1$^{+}$IgM$^{-}$CD43$^{-}$CD24$^{+}$) stages (Fig. 4b). In *Ikzf3*$^{G158R/G158R}$ mice, immature and recirculating IgM$^{+}$ cells were nearly undetectable in the BM and there was a marked reduction of all B220$^{+}$CD19$^{+}$ B cell subsets in the spleen (Fig. 4b–d, Extended Data Fig. 4f, g). Heterozygous *Ikzf3*$^{+/G158R}$ mice displayed a milder but similar phenotype (i.e., developmental blockade at the pre-pro-B cell to pro-B cell and pre-B cell stages) accompanied by a significant decrease in recirculating B cells in the BM and notable decreases in all B cell subsets in the spleen. The proportion of follicular B cells, defined as B220$^{+}$AA4.1$^{-}$CD23$^{+}$CD21$^{-}$, was particularly decreased by the *Ikzf3*$^{G158R}$ mutant in a gene dosage-dependent manner (Fig. 4b, d), while IgM/IgD expression pattern was not altered (Extended Data Fig. 4g). Notably, the defect in early B cell development in *Ikzf3*$^{+/G158R}$ mice became more severe with age (Fig. 4e). However, there was no clear impairment in early B cell development or a reduction in the number of splenic B cells in *Ikzf3*$^{-/-}$ mice, as previously reported[22], indicating that the Aiolos$^{G158R}$ mutant is likely to inhibit B cell development through mechanisms other than the loss of Aiolos function.

To determine whether the B cell developmental defect caused by the *Ikzf3*$^{+/G158R}$ mutation is intrinsic, competitive BM transplantation was performed. While the generation of T, NK, and myeloid cells was comparable between *Ikzf3*$^{+/G158R}$ mice and the WT competitor, the *Ikzf3*$^{+/G158R}$ mice had significantly fewer BM B cell progenitors and splenic B cells (Fig. 4f). Within the population of BM B cell progenitors, the proportion of pre-pro-B cells was similar between the *Ikzf3*$^{+/G158R}$ mice and WT controls. The ratio of *Ikzf3*$^{+/G158R}$ cells started to decrease in comparison with that of the WT competitor at the pro-B cell stage, which was accompanied by significant decreases at the pre-B cell and later B cell stages (Fig. 4f). These observations demonstrate cell-intrinsic impairment of B cell differentiation at the pre-pro-B cell to pro-B and pre-B cell stages in *Ikzf3*$^{G158R/G158R}$ and *Ikzf3*$^{+/G158R}$ mice.

In T cells, similar to those of patient P1, TCRβ expression of post-selection thymocytes and CD3ε expression of peripheral T cells of *Ikzf3*$^{+/G158R}$ and *Ikzf3*$^{G158R/G158R}$ mice were lower than those of the thymocytes and T cells of *Ikzf3*$^{+/+}$ mice (Extended Data Fig. 5a, b). The *Ikzf3*$^{G158R/G158R}$ mice were also characterized by the reduction of CD4SP thymocytes and the emergence of CD4$^{lo}$CD8$^{+}$ thymocytes (Extended Data Fig. 5a, c). CD4$^{+}$ T cells of *Ikzf3*$^{G158R/G158R}$ mice were skewed to the memory phenotype, which also resembled the immunological phenotype of patient P1 (Extended Data Fig. 5a). Notably, the proportion of CD8$^{+}$ T cells was decreased, whereas that of CD4$^{-}$CD8$^{-}$ T cells was increased in the lymph nodes of *Ikzf3*$^{G158R/G158R}$ and *Ikzf3*$^{+/G158R}$ mice (Extended Data Fig. 5a).

These results indicate that the G158R substitution in the mouse recapitulates the B and T cell phenotypes of patients carrying the AIOLOS$^{G159R}$ mutation, strongly suggesting that the human G159R variant causes disease. These findings also prompted us to delve further into the molecular pathogenesis of how the Aiolos$^{G158R}$ mutant impairs the differentiation of cells in adaptive immunity.

## Aiolos$^{G158R}$ interferes with Ikaros function

To identify genes dysregulated by the presence of Aiolos$^{G158R}$, the gene expression profiles of pre-B cells in *Ikzf3*$^{+/+}$ and *Ikzf3*$^{+/G158R}$ mice were subjected to RNA-seq analyses. The transcriptome of *Ikzf3*$^{+/G158R}$ pre-B cells was clearly differentiated from that of WT pre-B cells (Fig. 5a). Of a total of 25237 genes, 1842 and 1559 were down-regulated and up-regulated, respectively, in the pre-B cells of *Ikzf3*$^{+/G158R}$ mice as compared with those of *Ikzf3*$^{+/+}$ mice, with a FDR cut-off of <0.05 (Fig. 5b, c). Gene ontology analysis was subsequently performed using genes with a fold change in the expression of >$\log_2$0.5 and an average log count per million reads of >3. For gene ontology analysis, the thresholds were set to include 600–800 genes with modest differences in expression patterns between genotypes. The analysis indicated that genes related to B cell differentiation were down-regulated in *Ikzf3*$^{+/G158R}$ pre-B cells (Fig. 5d).

Characterization via EMSA and ChIP-seq indicated that the AIOLOS$^{G159R}$ mutant had the minimal binding capacity to canonical AIOLOS sequences but rather acquired the ability to bind to aberrant sequences, suggesting that other molecules regulated by or cooperating with Aiolos are involved in the pathogenesis of B cell defects caused by the Aiolos$^{G158R}$ mutant. Among the IKZF family genes, *Ikzf1* was stably expressed during human B cell development from the B cell progenitor stage, whereas *Ikzf3* expression was initiated at the pre-B cell stage. Thus, published Ikaros ChIP-seq data of murine pre-B cells[27] were integrated with our RNA-seq data to determine whether the changes to the transcriptome of *Ikzf3*$^{G158R}$ mutant pre-B cells were related to transcriptional regulation by Ikaros. Among 25,237 Ikaros peaks identified in pre-B cells, 7,504 overlapped the gene promoter regions. Transcriptome analysis of pre-B cells revealed 3,401 differentially expressed genes between *Ikzf3*$^{+/+}$ and *Ikzf3*$^{+/G158R}$ pre-B cells. Ikaros bound to the promoter regions of more than half (1801/3401) of the genes, which is substantially higher than the background ratio (7504/25237) (Fig. 5e). These cumulative data suggest that Ikaros function was compromised by the Aiolos$^{G158R}$ mutation via dysfunctional heterodimers in an Aiolos$^{G158R}$-dependent manner.

Co-immunoprecipitation (Co-IP) assay confirmed that the human AIOLOS$^{G159R}$ mutant retained the ability to form dimers with IKAROS to an extent similar to that of WT AIOLOS (Fig. 6a). WT AIOLOS exhibited a speckled staining pattern in the nucleus, indicating the formation of pericentromeric heterochromatin (PC-HC) foci, as previously reported[28], whereas AIOLOS$^{G159R}$ exhibited a diffuse nuclear pattern and defective PC-HC formation (Fig. 6b). Loss of PC-HC targeting by AIOLOS$^{G159R}$ was also supported by the AIOLOS ChIP-seq data. AIOLOS$^{G159R}$ failed to bind to the pericentromeres, where repetitive TGGAA sequences are comprised[28], and were preferentially bound by WT AIOLOS (Extended Data Fig. 6). Next, the effect of AIOLOS$^{G159R}$ on IKAROS localization was examined. The results showed that both WT AIOLOS and AIOLOS$^{G159R}$ colocalized with IKAROS. However, the speckled staining pattern of IKAROS was retained with WT AIOLOS but changed to a diffuse pattern with AIOLOS$^{G159R}$, suggesting defective PC-HC formation (Fig. 6b). Similar results were observed with fluorescent protein-fused AIOLOS and IKAROS expression in COS-7 cells (Fig. 6b). The mean proportions of cells with speckled patterns of WT AIOLOS and AIOLOS$^{G159R}$ obtained from three independent

experiments were 70% and 46% ($p < 0.05$, determined by *t*-test), respectively. These results suggest that the localization of IKAROS to heterochromatic regions was altered in the presence of AIOLOS$^{G159R}$.

Next, the effect of the Aiolos$^{G158R}$ mutant on the genome-wide binding of Ikaros was investigated using ChIP-seq of *Ikzf3*$^{+/+}$ and *Ikzf3*$^{G158R/G158R}$ thymocytes. At regions co-bound by Ikaros and Aiolos, such as the *Rbms3*, *Dpp6*, and *Epb41l2* loci, which are putative binding sites for the Ikaros–Aiolos heterodimer, Ikaros binding became weaker in *Ikzf3*$^{G158R/G158R}$ thymocytes (Fig. 6c). By contrast, the regions specifically enriched for Aiolos$^{G158R}$ binding in *Ikzf3*$^{G158R/G158R}$ were often co-bound with Ikaros. And Ikaros binding peaks at those Aiolos$^{G158R}$ binding regions were higher than the peaks in *Ikzf3*$^{+/+}$ thymocytes (Fig. 6c). The number of Ikaros-exclusive peaks (presumably bound by the Ikaros–Ikaros homodimer) was lower in *Ikzf3*$^{G158R/G158R}$ thymocytes than in *Ikzf3*$^{+/+}$ thymocytes, and most of the Ikaros peaks were co-occupied by Aiolos$^{G158R}$ (presumably bound via Ikaros–Aiolos$^{G158R}$ heterodimers) (Fig. 6d). Notably, the genome-wide binding sites of Ikaros clustered with Aiolos$^{G158R}$ binding sites in *Ikzf3*$^{G158R/G158R}$ thymocytes rather than Ikaros binding sites in *Ikzf3*$^{+/+}$ thymocytes (Fig. 6e). Motif discovery analysis detected Ikaros and Aiolos consensus binding sequences (GGGAA and AGGAA) in *Ikzf3*$^{+/+}$ thymocytes (Fig. 6f). Highly enriched binding motif of Aiolos$^{G158R}$ in *Ikzf3*$^{G158R/G158R}$ thymocyte was non-Ikaros-consensus sequence (GGAGC and CCCAGA), which was also enriched at the AIOLOS$^{G159R}$-specific peaks in the NALM-6 human pre-B cell line (Fig. 6f). The binding motif of Ikaros was also affected in *Ikzf3*$^{G158R/G158R}$ thymocytes, characterized by high enrichment of the GGAGC motif, which was preferentially bound by the Aiolos$^{G158R}$ mutant (Fig. 6f). Interestingly, motif discovery analysis showed that Ikaros consensus sequences (GGGAA and AGGAA) were associated with Ikaros in *Ikzf3*$^{G158R/G158R}$ thymocytes with higher *p*-values than the aberrant GGAGC motif (Fig. 6f). Ikaros-exclusive peaks (likely bound by Ikaros–Ikaros homodimers), which were decreased, but not absent, in *Ikzf3*$^{G158R/G158R}$ thymocytes, were associated with the Ikaros consensus sequences (Fig. 6d, f). These data suggest that the remaining Ikaros homodimers were likely to bind the Ikaros consensus motif in the *Ikzf3*$^{G158R/G158R}$ mutant.

The transcriptional activity of AIOLOS$^{G159R}$ and its effects on WT AIOLOS and IKAROS were assessed using the luciferase reporter gene assay. The ability of AIOLOS to repress transcriptional activity was abrogated by the co-expression of AIOLOS$^{G159R}$ in a dominant-negative manner (Fig. 6g). Similarly, the ability of IKAROS to repress transcriptional activity was inhibited by the co-expression of AIOLOS$^{G159R}$ but not of WT AIOLOS.

These data indicate that the Aiolos$^{G158R}$ mutant inhibits DNA binding of Ikaros to its consensus motif and instead recruits Ikaros to other genomic regions with non-Ikaros consensus sites, resulting in the hijacking of Ikaros binding from its physiological target in the genome where it exerts transcriptional activities.

## Dimerization defect cancels phenotype of Aiolos$^{G158R}$ mutant

The results suggest that sequestration of Ikaros by the Aiolos$^{G158R}$ mutant via heterodimer formation is the cause of B cell deficiency and T cell abnormalities in *Ikzf3*$^{G158R}$ mutant mice. To prove this concept and to explore potential therapeutic applications for patients

with the *IKZF3*$^{G159R}$ variant, we examined whether the abrogation of the ability of Aiolos$^{G158R}$ to form dimers would restore the immune phenotype in *Ikzf3*$^{G158R}$ mutant mice (Fig. 7a). First, the Co-IP assay confirmed that deletion of the fifth and sixth ZFs from Aiolos abolished heterodimer formation with Ikaros (Fig. 7b). Then, the C-terminal ZF structure in the Aiolos$^{G158R}$ mutant was disrupted by CRISPR-Cas9 technology (Fig. 7a). The founder offspring bearing a ZF5/6-deleting mutation of the *Ikzf3*$^{G158R}$ allele (p.D461fs, designated as *Ikzf3*$^{G158R:\Delta c\text{-}ZF}$ allele) (Extended Data Fig. 7a) were crossed with *Ikzf3*$^{+/+}$ mice to generate *Ikzf3*$^{+/G158R:\Delta c\text{-}ZF}$ mice. The reduction of BM B cell progenitors and the relative increase in the pre-pro-B cell fraction in *Ikzf3*$^{+/G158R}$ mice were normalized in *Ikzf3*$^{+/G158R:\Delta c\text{-}ZF}$ mice (Fig. 7c, d). Additionally, the proportion of splenic B cells was normalized with the recovery of follicular B cells in *Ikzf3*$^{+/G158R:\Delta c\text{-}ZF}$ mice (Fig. 7c, e). T cell phenotypes of *Ikzf3*$^{+/G158R}$ mice were also rescued in *Ikzf3*$^{+/G158R:\Delta c\text{-}ZF}$ mice. The reduced proportion of CD4SP thymocytes, lower expression levels of CD3ε and TCRβ, increased proportion of CD4$^{-}$CD8$^{-}$ αβT cells, and skewing to memory-like CD4$^{+}$ T cells in *Ikzf3*$^{+/G158R}$ mice were rescued in *Ikzf3*$^{+/\,G158R:\Delta c\text{-}ZF}$ mice (Fig. 7c, f and Extended Data Fig. 7b). These findings demonstrate that the Aiolos$^{G158R}$ mutant exerts pathogenicity via the dimerization domain.

## Discussion

We report an IEI caused by a heterozygous missense variant of *IKZF3*, which manifested as B cell deficiency with T cell abnormalities. The AIOLOS$^{G159R}$ variant is characterized as a loss-of-function mutation regarding DNA binding to canonical AIOLOS sequences, although the G159R substitution concurrently confers novel DNA binding specificity. A mouse model harboring the AIOLOS variant of the patient showed that the *Ikzf3*$^{G158R}$ mutation affected the binding of not only Aiolos but also Ikaros. The results of a reporter gene assay showed that the AIOLOS$^{G159R}$ variant altered the transcriptional activity of IKAROS. While the differentiation of the earliest stage of B cell progenitors were not impaired, pre-B cell and later stage of B cells were decreased in the *Ikzf3*$^{G158R}$ mutant mice.

Given that *IKZF3* is expressed from pre-B cell stage, such a stage specific effect of AIOLOS$^{G159R}$ variant is likely to stem from increase of mutant AIOLOS$^{G159R}$ protein during B cell development. Supported by the fact that heterozygous *IKAROS* variants have been found in patients with B cell deficiencies and CVID[6,9,13–18], we concluded that the interfering effect of the AIOLOS mutant against IKAROS is responsible for the B cell-deficient phenotype of the patients and knock-in mice. We termed this novel pathogenesis as heterodimeric interference, in which a mutant protein in a heterodimer hijacks the function of the normal partner protein. To date, several heterozygous missense variants of transcription factors have been identified as the causes of IEIs[1]. The pathogenesis of IEIs has been identified as gain-of-function[2–5], negative dominance[6–8], and haploinsufficiency[9] of the products of genes harboring causative variants. Collectively, our data demonstrate a novel mechanism, heterodimeric interference, underlying the pathogenesis of autosomal dominant diseases.

According to Muller's classification, the AIOLOS$^{G159R}$ variant can be classified as a *neomorph*. It is plausible that functional abrogation of IKAROS by AIOLOS$^{G159R}$ via

heterodimer formation is caused by the impediment of the binding of the IKAROS–AIOLOS$^{G159R}$ heterodimer to its physiological binding site due to defective DNA binding by AIOLOS$^{G159R}$. As an alternative mechanism, IKAROS activity is inhibited by sequestration of IKAROS to aberrant loci by AIOLOS$^{G159R}$ and the acquisition of novel DNA binding specificity as a result of a *neomorphic* mutation. However, these results do not formally exclude the possible involvement of aberrant transcriptional regulation by the novel binding of the IKAROS–AIOLOS$^{G159R}$ heterodimer and/or AIOLOS$^{G159R}$ homodimer in disease development. Hence, further research is required to investigate the contribution of the loss of AIOLOS binding ability and the gain of novel DNA binding motif specificity by the AIOLOS$^{G159R}$ variant on disease pathogenesis.

AIOLOS is considered less important than IKAROS for lymphocyte development. This underrated conclusion is based primarily on the finding that Aiolos-null mice did not exhibit a significant immunodeficiency phenotype[22]. IKZF family proteins are known to be functionally redundant[23,29]; thus, the complete loss of Aiolos might be compensated by Ikaros. The importance of heterodimerization among IKZF family proteins is implicated by different phenotypes among the three *Ikzf1* mutant mouse strains. The deletion of *Ikzf1* exon 8 results in the absence of Ikaros (Ikaros-null), and B and NK cells, but not T cells, are absent in Ikaros-null mice, albeit T cell development is delayed [13]. Conversely, the deletion of the first three ZFs of Ikaros (dominant-negative [DN] mutation) results in the absence of T, B, and NK cells[10]. Moreover, the phenotype of another *Ikzf1* mutant mouse strain with a missense mutation of the third ZF (H191R mutation, referred to as *Ikzf1*$^{Plstc}$) was similar to that of *Ikzf1-DN* mutant mice[30]. Although these three mutations should result in a lack of functional Ikaros, the mutation that eliminates the middle ZFs responsible for DNA binding capacity (*Ikzf1-DN*) and the missense mutation in the third ZF (*Ikzf1*$^{Plstc}$) retain the ability to form heterodimers with other IKZF family proteins. Therefore, the more severe phenotype caused by the *Ikzf1-DN* and *Ikzf1*$^{Plstc}$ mutation may reflect the consequence that heterodimerization of mutant Ikaros with other family member proteins interferes with the functions of the binding partner. However, the mechanisms underlying the regulation of heterodimer and homodimer formation of IKZF proteins remain unclear. Thus, to gain a better understanding of the disorders involving IKZF family proteins, it is necessary to determine whether, to what extent, and in which context the functions of the AIOLOS–IKAROS heterodimers and IKAROS–IKAROS or AIOLOS–AIOLOS differ.

The patients developed B cell malignancy in adulthood besides profound B cell developmental defects. Somatic variants and the deletion of *IKZF1* and *IKZF3* have been associated with B-ALL[19,20]. Additionally, germline *IKZF1* variants could increase susceptibility to B-ALL[21]. A possible explanation for the development of lymphoid malignancy in the patient's family is the loss of functional AIOLOS by the dominant negative effect of homodimer formation, as indicated by the spontaneous development of lymphoma in aged Aiolos knock-out mice[22]. As an alternative hypothesis, heteromeric interference by mutant AIOLOS contributes to lymphomagenesis. In pre-pro-B and pro-B cells, the expression of IKAROS predominates that of AIOLOS and the level of AIOLOS expression reaches that of IKAROS in pre-B cells and later stages of B cells. Hence, the functional abrogation of IKAROS in mature B cells may have contributed to the tumorigenesis of these cells. Although further investigations are required to determine the

precise mechanism underlying the pathogenesis of lymphomagenesis in patients carrying the AIOLOS mutant, the results of the present study implicate heteromeric interference in a wide range of biological phenomena, including tumorigenesis.

The B cell developmental disorder and T cell abnormalities in *Ikzf3*$^{+/G158R}$ mutant mice were restored by the elimination of the C-terminal domain of the Aiolos$^{G158R}$ mutant. This observation sheds light on a potential therapeutic approach since the deletion of the dimerization domain of the *IKZF3*$^{G159R}$ allele in hematopoietic stem cells could rescue the immunodeficiency phenotype and possibly susceptibility to the development of hematologic malignancies observed in the patients.

The results of the present study revealed heterodimeric interference among IKZF family proteins. Various transcription factors, such as Myc, NF-κB, and STAT, function in heterodimer formation. Hence, it is conceivable that similar heteromeric interference by transcription factors would be involved in disease pathogenesis. Although the classical gene knock-out approach is useful to discern the outcomes of loss of protein function, the precise mechanism regulating the phenotype cannot be always elucidated. Besides the functional redundancy within family members, it is difficult to evaluate to what extent the loss of heterodimer function contributes to the phenotype. Particularly, the pathogenesis of heterozygous missense variations in dimer-forming proteins has been mostly interpreted as DN or haploinsufficiency effects, thereby underestimating the functions of heterodimers. This study highlights the importance of the functions of protein complexes rather than individual proteins. The pathogenic mechanism underlying heteromeric interference is currently limited to heterodimer formation. However, future studies might discover that this concept also applies to interference in the function of multi-protein complexes in other human diseases.

## Methods

### Study approval

Informed consent was obtained from patients and their family members. The study was conducted in accordance with the Helsinki Declaration and approved by the ethics boards of Tokyo Medical and Dental University (G2019–010A, G2019–052A, and A2021–156A) and by the genetic recombination board of RIKEN Yokohama Branch (2020–3).

### Mouse strains

*Ikzf3*$^{G158R}$ mutant allele was generated by CRISPR-Cas9 mediated genome editing. Custom sgRNA was designed to target exon 4 of the *Ikzf3* gene. The repair template DNA for the homology-directed repair (HDR) was prepared in the form of a single-stranded DNA oligonucleotides (ssODN) by Integrated DNA technology (IDT). The sgRNA fused to tacrRNA, Cas9 mRNA and the repair template ssODN were co-injected into the cytoplasm of C57BL6/J fertilized eggs by the Animal Facility Group at RIKEN, IMS. Founder off-springs that harbor the mutant allele were selected and crossed to wild-type C57BL6/J mice to establish the *Ikzf3*$^{G158R}$ mouse line. The founders that had indel and frameshift mutation at the site of Cas9-mediated double-strand DNA break were selected to establish

the *Ikzf3* knock-out mouse line. C-terminal zinc finger deletion mutant mice were generated by CRISPR-Cas9 genome editing with sgRNA targeting the exon 8 of *Ikzf3* gene using eggs obtained by *in vitro* fertilization between *Ikzf3*$^{+/G158R}$ sperm and C57BL6/J oocytes. The sequences of sgRNA and template ssODN are shown in Supplementary Tables 3 and 4. All mice were maintained in the specific pathogen free animal facility at the RIKEN, IMS, and all animal procedures were in accordance with institutional guidelines for animal care and with the protocol (28–017) approved by IACUC of RIKEN Yokohama Branch. All mice were housed in social groups, provided with food and water and maintained on a 12 h light/12 h dark cycle, at ~22 °C ambient temperature with ~40% humidity.

## Plasmids

The entire coding region of human *IKZF1* and *IKZF3* were PCR-amplified from cDNA of healthy individual peripheral blood, and cloned into pCR4 plasmid using TOPO TA Cloning Kit for Sequencing (Thermo Fisher Scientific). The missense mutant *IKZF3* plasmid was generated by site-directed mutagenesis and ΔZF5/6 mutant *IKZF3* plasmid for Co-IP experiment was generated by PCR-based mutagenesis. cDNA fragment encoding wild-type or mutant AIOLOS was subcloned into pcDNA3 expression vector, and was used for electrophoretic mobility shift assay. For co-localization assay, GFP and mRFP coding sequences were fused to the 3' end of wild-type or mutant *IKZF3* and *IKZF1* cDNA to make AILOS-GFP and IKAROS-RFP fusion protein, respectively, by fusion PCR. Each cDNA fragment encoding such fusion protein was cloned into pcDNA3 expression vector. Primers sequences are shown in Supplementary Table 5.

## Cells

COS-7 fibroblast cell line was purchased from JCRB cell bank; HEK293T cell line was purchased from JCRB cell bank and ATCC; NIH3T3 cell line was purchased from ATCC. These cell lines were maintained in Dulbecco's Modified Eagle Medium (DMEM, High-glucose, Thermo Fisher Scientific) supplemented with 10 % heat inactivated fetal bovine serum (FBS, Gibco) and 1 % Penicillin-Streptomycin (Thermo Fisher Scientific).

NALM-6 human pre-B cell line was purchased from RIKEN BRC. NALM-6 cells were maintained in RPMI 1640 medium (Thermo Fisher Scientific) supplemented with 10 % heat inactivated FBS and 1 % Penicillin-Streptomycin. In order to generate NALM-6 lacking AIOLOS expression, frameshift mutation was induced in the *IKZF3* gene by the CRISPR-Cas9 genome editing. Custom sgRNA was designed to target exon 2 of the *IKZF3* gene. The sequences of sgRNA is shown in Supplementary Table 3. The expression vectors of GFP (pCMV-EGFP, Addgene), sgRNA fused to tacrRNA (pU6 gRNA, Addgene) and Cas9 (hCas9, Addgene) were transiently transfected to NALM-6 by electroporation using NEPA-GENE21(NEPAGENE). GFP positive cells were single cell sorted by flow cytometry after 24 to 48 hours of transfection. Clones harboring for biallelic indel mutations were selected by sequencing of genome, and loss of AIOLOS protein expression was confirmed by immunoblotting.

For the reconstitution of FLAG-tagged *IKZF3* in *IKZF3*-deficient NALM-6 cell line, cDNA fragment encoding wild-type AIOLOS or mutant AILOS$^{G159R}$ with FLAG-tag sequences at

the N-terminus was cloned into pRetroX-TetOne-Puro vector (Clontech). *IKZF3* deficient NALM-6 cells were retrovirally transduced with one of these vectors, and were selected with puromycin to establish NALM-6 cells for inducible FLAG-tagged AIOLOS protein. Expression of FLAG-tagged AIOLOS protein was induced by adding 1 ng/ml doxycycline to the medium.

## Antibodies

The following monoclonal antibodies were used for flow cytometry for human subjects: anti-CD8 (BW135/80), anti-CD19 (LT19), anti-CD45RO (UCHL1), anti-CD4 (VIT4), anti-CD3 (BW264/56), anti-TCRγδ (11F2), anti-TCRαβ (BW242/412), anti-CD127 (MB15–18C9), anti-CD38 (IB6), anti-CD123 (AC145), anti-CD45 (5B1) (all from Miltenyi Biotec, all used at 1:20); anti-CD16 (3G8), anti-CD45RA (2H4LDH11LDB9), anti-CD56 (N901 (HLDA6)), anti-TCRVα24 (C15), anti-TCRVβ11 (C21), anti-CD8 (SFCI21Thy2D3), anti-CD25 (B1/49.9), anti-CD10 (ALB1), anti-HLA-DR (Immu-357), anti-CD11c (BU15) (All from Beckman Coulter, all used at 1:20); anti-CCR6 (11A9), anti-CXCR3 (1C6/CXCR3), anti-IgM (G20–127), anti-Lineage cocktail (SK7 (CD3), MφP9 (CD14), 3G8 (CD16), SJ25C1 (CD19), L27 (CD20), NCAM16.2 (CD56)) (all from BD Biosciences); anti-CCR4 (2G12), anti-CD161 (HP-3G10), anti-HLA-DR (LN3) (all from BioLegend, all used at 1:20)

The following monoclonal antibodies were used for flow cytometry for mouse subjects: anti-CD43 (S7, used at 1:100), anti-IgM (R6–60, used at 1:100), anti-B220 (RA3–6B2), anti-CD24 (M1/69, used at 1:2000), anti-CD93 (AA4.1), anti-CD21 (7G6, used at 1:100), anti-CD95 (Jo2), anti-GL7 (GL7), anti-CD19 (1D3), anti-TCRβ (H57–597), anti-CD4 (RM4–5), anti-CD8a (53–6.7), anti-CD44 (IM7), anti-CD3ε (145–2C11, used at 1:100), anti-CD62L (MEL-14), anti-CD11b (M1/70), anti-CD335 (29A1.4) (all from BD Bioscience, all used at 1:200 unless otherwise specified); anti-CD23 (B3B4), anti-CD45 (30-F11), anti-CD4 (GK1.5) (all from eBioscience, all used at 1:200); anti-IgD (11–26c.2a, Biolegend, used at 1:200).

The following monoclonal antibodies were used for immunoblot analysis, cellular localization assay, and ChIP-assay: anti-Aiolos (EPR9342(B), abcam, used at 1:1000); anti-Aiolos (9D10) and anti-Ikaros (4E9 and Ik14) (all from EMD Millipore); anti-Aiolos (D1C1E, Cell signaling technology, used at 1:1000); anti-Actb (AC-15, used at 1:10000), and anti-FLAG (M2) (all from Sigma-Aldrich); anti-Rabbit IgG Alexa Fluor 488, anti-Mouse IgG Alexa Fluor 568 (all from Thermo Fisher Scientific, all used at 1:250); anti-GAPDH (2275-PC-100, R&D Systems, used at 1:1000); anti-HA (TANA2, MBL, used at 1:1000).

## Whole-exome sequencing

Genomic DNA was extracted from the peripheral blood of the patients and healthy family members. Exon capture was performed with SureSelect Human All Exon V4 (Agilent technologies). Massively parallel sequencing was performed on a HiSeq2000 (Illumina), generating 100 bp paired-end reads. Burrows-Wheeler Aligner was used to align sequences to the human genome reference sequence (GRCh37/hg19). Variant allele numbers were counted by Variant Caller with Multinomial probabilistic Model (VCMM) and genes were

annotated by ANNOVAR. Identified variants with minor allele frequency of <0.0001 were selected, according to National Center for Biotechnology Information's dbSNP (https://www.ncbi.nlm.nih.gov/snp/). Scores of the variants were obtained from the CADD web server (https://cadd.gs.washington.edu). Selected results identified by whole-exome sequencing were confirmed by Sanger sequencing.

### Template selection and homology modeling

Since the three-dimensional structure of human AIOLOS is not yet experimentally determined, a model of the four N-terminal zinc fingers of human AIOLOS was constructed using homology modeling. First, a template search was conducted using BLASTP against RCSB protein data bank (PDB) using the human AIOLOS sequence as a query and possible template sequences were identified. A six-finger fragment of human PR/SET domain 9 (PRDM9c) with (PDB ID: 5V3G)[24,31,32] was selected as a template considering its sequence identity of 38%, 0% gap, 55% positives and BLAST bit score of 224.

Homology models were generated using MODELLER (version 9.17)[33]. Initially, the target and template sequences were aligned using the align2d command of MODELLER that uses global dynamic programming, with linear gap penalty function for aligning the two profiles. Then, 20 three-dimensional models of human AIOLOS with four N-terminal ZFs and DNA were built with MODELLER[33] and one model was selected based on the lowest DOPE (Discrete optimized protein energy) score and visual inspection[34]. The selected model of human AIOLOS with four N-terminal ZFs in complex with DNA and zinc ions was used for molecular dynamics simulations and analyses.

### Molecular dynamics simulations and analyses

The homology modeled structures with G159 and R159 (generated by mutating G159 to R159 *in silico* keeping the secondary structure intact) in complex with DNA and zinc ions were subsequently used as the starting structures for molecular dynamics simulations. Hydrogen atoms were added using the Xleap tool of AMBER 14[35]. The wild-type and G159R mutant structures were subsequently solvated in an octahedral box of three-point transferable intermolecular potential (TIP3P) waters with ~10 Å distance between the protein surface and the box boundary[36]. Each system was then electrostatically neutralized by the addition of $Na^+$ counter ions. The SANDER module of AMBER14 package with ff14SB force field and ZAFF (Zinc Amber Force Field) was used for parametrizing zinc ion and subsequent simulations[37,38]. Each system was initially subjected to 2000 steps of steepest descent followed by 2000 steps of conjugate gradient minimization to eliminate close van der Waals contacts. Subsequently, each system was gradually heated from 0 to 300 K in 500 ps followed by a constant temperature equilibration at 300 K for 3000 ps. Following this, 100 ns production molecular dynamics runs were carried out with periodic boundary conditions in the NPT ensemble with Berendsen temperature coupling and isotropic molecule-based scaling at a constant pressure of 1 atm[39]. The SHAKE algorithm was used to constrain bond lengths involving bonds to hydrogen(s) and the particle-mesh Ewald (PME) method was used for the calculation of long range electrostatic forces[40,41]. The analyses of simulations were carried out by CPPPTRAJ module of AMBER 14 and

Visual Molecular Dynamics (VMD) software[42]. Figures were generated using PyMOL (http://www.pymol.org).

### Electrophoretic mobility shift assay

HEK293T cells were transfected with pcDNA3/WT or G159R mutant AIOLOS (Fig. 2d) or pcDNA3.1/C-DDK-WT or G159R mutant AIOLOS (Fig. 2e), using Lipophectamine LTX Reagents (Thermo Fisher Scientific) or Effectene (Qiagen), respectively. 24–48h hours after transfection, nuclear extract was prepared using NE-PER nuclear and cytoplasmic extraction reagents (Thermo Fisher Scientific). For Fig. 2d, nuclear extracts were incubated with $^{32}$P-labeled IK-BS4 probe (5'-TGACAGGGAATACACATTCCCAAAAGC-3')[25]. Anti-AIOLOS monoclonal antibody (EPR9342(B), abcam) was used for a super-shift assay. For Fig. 2e, Nuclear extracts were pre-incubated with 0.5 μl of anti-AIOLOS (D1C1E, Cell signal technology) for 10 min and then incubated with biotinylated IK-BS1 (5′-BIOTIN-TCAGCTTTTGGGAATACCCTGTCA-3')[23]. Gel mobility shift assays were performed with LightShift Chemiluminescent EMSA kit (Thermo Fisher Scientific) according to manufacturer's instruction.

### Cellular localization assay

NIH3T3 cells were transfected with indicated plasmids using Effectene (Qiagen) according to the manufacturer's instruction. 24 hours after transfection, cells were washed, fixed and permeabilized as previously described[18]. The cells were stained using primary anti-HA antibodies (Biolegend), and/or anti-Aiolos antibodies (D1C1E, Cell signaling technology). The cells were washed and incubated with Alexa Fluore 488 (green color) and/or Alexa Fluore 568 (red color)-conjugated secondary antibodies (Thermo Fisher Scientific). Images were collected on a ZOE fluorescent cell imager (Bio-Rad, original magnification 175x).

For fusion protein assays, COS-7 cells were transfected with pcDNA3 expression vector encoding IKAROS-mRFP together with pcDNA3 vector encoding either AIOLOS-GFP or AILOS$^{G159R}$-GFP fusion protein. One μg of each plasmid were mixed with 100 μl of Opti-MEM (Thermo Fisher Scientific) and 12 μl of 0.25 mg/ml polyethylenimine (PEI). Mixtures were incubated at room temperature for 15 minutes and were added to culture medium. Forty-eight to seventy-two hours after transfection, cells were washed with PBS and were fixed with 4 % paraformaldehyde for 10 minutes. The cells were stained with DAPI and were examined under FV10i confocal laser scanning microscope (Olympus).

### Immunoblot

Thymocytes were washed with ice-cold PBS, and lysed with RIPA Lysis and Extraction Buffer (Thermo Fisher Scientific) supplemented with Halt™ Protease Inhibitor Cocktail (Thermo Fisher Scientific) by incubating on ice for 10 min. Cell debris were removed by centrifugation at 15,000 x *g* for 10 min. Lysate was mixed with 1x volume of 2x Laemmli Sample Buffer (Bio-Rad), and boiled at 95 °C for 10 min followed by a brief incubation on ice. The electrophoresis was performed with 10 % polyacrylamide gel (e-PAGEL, ATTO), and proteins were transferred onto a polyvinylidene difuoride membrane (EMD Millipore). After the blocking with 1x TBS-T buffer (nakarai tesque) with 5 %(w/v) of skim-milk (Wako), the membrane was incubated with primary antibodies in 1x TBS-T over-night at

4 °C. Membrane was washed with 1x TBS-T, and incubated with horseradish peroxidase-conjugated sheep anti-mouse IgG or donkey anti-rabbit IgG polyclonal IgG (GE Healthcare) for 30 min at room temperature. Target proteins were developed by use of Amersham ECL Select™ Western Blotting Detection Reagent and Amersham Imager 680 (GE Healthcare). Densitometry was done with Image J software.

### Co-immunoprecipitation

HA-tagged wild-type *IKZF1* was cloned into pcDNA3 expression vector. FLAG-tagged wild-type *IKZF3*, G159R mutant *IKZF3*, or ΔZF5/6 mutant *IKZF3* were cloned into pcDNA3 expression vector. HEK293T ($1 \times 10^6$ cells) cells were co-transfected with 1 μg of HA-tagged *IKZF1* and FLAG-tagged *IKZF3* (wild-type, G159R, or ΔZF5/6) and cultured for 24 h. Cells were lysed in lysis buffer (25 mM Tris-HCl (pH 7.5), 150 mM NaCl, 2 mM $MgCl_2$, 1 % NP-40, 5 % Glycerol) supplemented with protease inhibitors and DNase. Lysates were incubated with anti-FLAG antibody (M2, Sigma-Aldrich) or Mouse IgG conjugated to Dynabeads Protein G (Thermo Fisher Scientific) overnight at 4 °C. The beads were washed with wash buffer (25 mM Tris-HCl (pH 7.5), 150 mM NaCl, 0.1 % NP-40) for three times, followed by wash with 50 mM NH4HCO3 solution. Eluted protein samples were separated by SDS-PAGE, and immunoprecipitated with anti-HA antibody (TANA2, MBL) or anti-AIOLOS antibody (EPR9342(B), abcam).

### Luciferase reporter gene assay

The IK-BS1 Luciferase report construct containing four copies of high affinity IKAROS binding site (IK-BS1, 5'-TCAGCTTTTGGGAATACCCTGTCA-3') has been previously described[18]. Briefly, HEK293T cells were co-transfected with pGL4.11-IKBS1, pRL-TK (Renilla luciferase), and FLAG tagged WT and/or G159R mutant AIOLOS with or without pcDNA3-HA-IKAROS using Effectene (Qiagen). The next day, cells were lysed, and luciferase activity was measured using the Dual-Luciferase reporter assay system (Promega, following manufacturer's instructions) on a Synergy4 Microplate reader (BioTek). The firefly luciferase experimental reporter activity was normalized to the Renilla luciferase activity and then the results were normalized to the empty vector control. Each experiment was carried out in duplicate.

### RNA-seq

Bone marrows harvested from *Ikzf3*$^{+/+}$, *Ikzf3*$^{+/G158R}$, and *Ikzf3*$^{G158R/G158}$ mice were used for cell sorting to isolate pre-B cells ($B220^+AA4.1^+IgM^-CD43^-CD24^+$). Total RNA was extracted by TRIzol reagent (Thermo Fisher Scientific), followed by removal of DNA with RNA Clean & Concentrator-5 (ZYMO RESEARCH). mRNA was obtained by poly (A) selection with mRNA DIRECT kit (Thermo Fisher Scientific). RNA-seq libraries were prepared using NEBNext Ultra RNA Library Prep Kit for Illumina (New England Biolabs) following manufacturer's protocol. Briefly, first and second strand of cDNA was synthesized from mRNA obtained by poly (A) selection. cDNA was end-repaired, dA-tailed and were ligated with adaptor. The adapter ligated DNA fragments were size-selected by dual bead-based size selection using AMPure XP Beads (Beckman Coulter). The size-selected DNA was amplified by PCR for 12 to 15 cycles. The purified DNA was diluted to 2 nM and sequenced with a HiSeq1500 sequencer (Illumina).

For analysis of differentially expressed genes, RNA-seq reads were mapped to murine genome and read counts were calculated with STAR using refGene as genome indexes files[43]. Generated read counts were analyzed with the R package edgeR[44].

For gene ontology (GO) analysis, genes with more than $\log_2$ 0.5 of fold-change between *Ikzf3*$^{+/+}$ and *Ikzf3*$^{+/G158R}$ pre-B cells, with average logCPM of >3 and FDR of <0.05 were extracted. GO enrichment analysis was performed with the R package[45].

## ChIP

Thymocytes or NALM-6 cells following 48 h of doxycycline treatment were washed twice with PBS. Cells were cross-linked for 10 to 15 minutes with 1 % formaldehyde, followed by quenching with glycine. Cells were washed 3 times with ice-cold PBS. Cells were lysed in lysis buffer 1 (50 mM HEPES-KOH (pH 7.5), 140 mM NaCl, 1 mM EDTA, 10 % glycerol, 0.5 % NP-40, 0.25 % Triton X-100, 1x Complete Mini Protease inhibitor (Roche), 1 mM PMSF) for 10 minutes at 4 °C and pelleted. Cells were further lysed with lysis buffer 2 (10 mM Tris-HCl (pH 8.0), 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 1x Complete Mini Protease inhibitor (Roche), 1 mM PMSF) for 10 minutes at 4 °C and pelleted. The pellet was resuspended with lysis buffer 3 (10 mM Tris-HCl (pH 8.0), 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1 % sodium deoxycholate, 1x Complete Mini Protease inhibitor (Roche), 1 mM PMSF), and sonicated with a Microson Sonicator XL2000 for 6 cycles of 15-second sonication at level 6. Chromatin was cleared by centrifugation at 13500 rpm for 10 minutes. Supernatants were incubated with 5 to 10 μg of antibodies that are prebound to Dynabeads M-280 Sheep anti-mouse IgG (Thermo Fisher Scientific). The mixtures were rotated at 4 °C overnight. The beads were washed twice with low salt buffer (20 mM Tris-HCl (pH 8.0), 150 mM NaCl, 2 mM EDTA, 0.1 % SDS, 1 % Triton X-100), twice with high salt buffer (20 mM Tris-HCl (pH 8.0), 400 mM NaCl, 2 mM EDTA, 0.1 % SDS, 1 % Triton X-100), and finally twice with RIPA buffer (50 mM HEPES (pH 7.6), 500 mM LiCl, 1 mM EDTA, 1 % NP-40, 0.7 % Sodium deoxycholete). Chromatin-protein complexes were eluted with elution buffer (50 mM Tris-HCl (pH 8.0), 10 mM EDTA, 1 % SDS) and reverse cross-linked overnight at 65 °C. The elution was treated with 0.2 mg/ml of RNaseA at 37 °C for 1 hour, followed by proteinase K treatment at a concentration of 0.2 mg/ml at 55 °C for 1 hour. DNA was extracted by phenol-chloroform precipitation and resuspended in TE buffer. Chromatin from 1 to 2 × $10^8$ cells/IP was used for ChIP-seq.

## ChIP-seq

ChIP-seq libraries prepared from ChIP samples of thymocytes or NALM-6 cells following doxycycline induction were prepared according to NEBNext ChIP-Seq Library Prep Master Mix Set for Illumina (New England Biolabs). Up to 10 ng of chromatin-immunoprecipicated DNA and input DNA were end-repaired, dA-tailed and ligated with adaptor. The adapter ligated DNA fragments were size-selected by dual bead-based size selection using AMPure XP Beads (Beckman Coulter). The size-selected DNA was amplified by PCR for 12 to 15 cycles. The purified DNA was diluted to 2 nM and sequenced with a HiSeq1500 sequencer (Illumina), or NextSeq sequencer (Illumina). Hisat2 was used to align sequences to the mouse genome reference sequences (mm9) or to the human genome reference sequence (GRCh38)[46]. ChIP peaks were identified using MACS2[47]. Differential bindings, correlation

heatmap, and binding affinity heatmap by transcription factors were analyzed using Diffbind package in R[48]. Binding sequences for the transcription factors were analyzed using HOMER[49].

## Statistical analyses

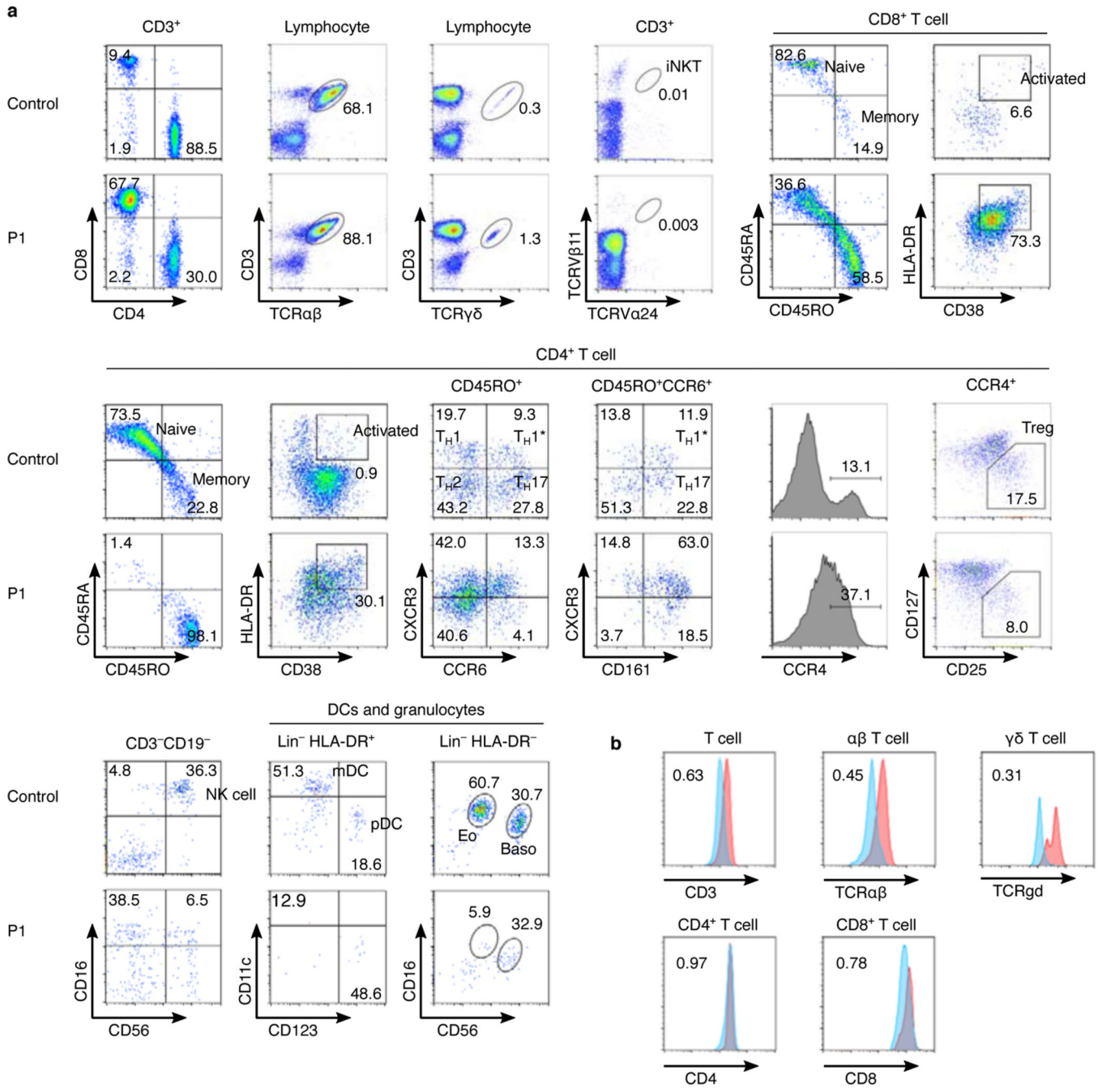
GraphPad Prism 7 was used for all statistical analyses, except for RNA-seq and ChIP-seq data. Significance of differences in *IKZF* gene expressions and cellular frequencies were calculated using one-way ANOVA. Difference of *Ikzf3* allele frequency in *Ikzf3*$^{+/G158R}$ pre-B cells and cellular frequencies in bone marrow chimera experiment were calculated using Student's *t*-test.

For detection of differentially expressed genes in pre-B cells, generalized linear models were deployed with the cutoff of FDR as <0.05. MACS2 was used for the detection of significant binding, using default setting. The signal values of binding enrichment (fold-enrichment) represent the ratio between ChIP-seq read counts and input read counts and were calculated by MACS2. Overlaps and differential bindings of the ChIP-seq peaks were identified using Bedtools, and ChIPPeakAnno or Diffbind package in R, using default parameters. Top 1000-enriched, all differentially bound, and top 1000 differentially bound peaks were subjected for motif analyses by HOMER, with region size set to 50.

## Data availability

Genome-wide reads were deposited at GEO. They are accessible for analysis at GSE167487.

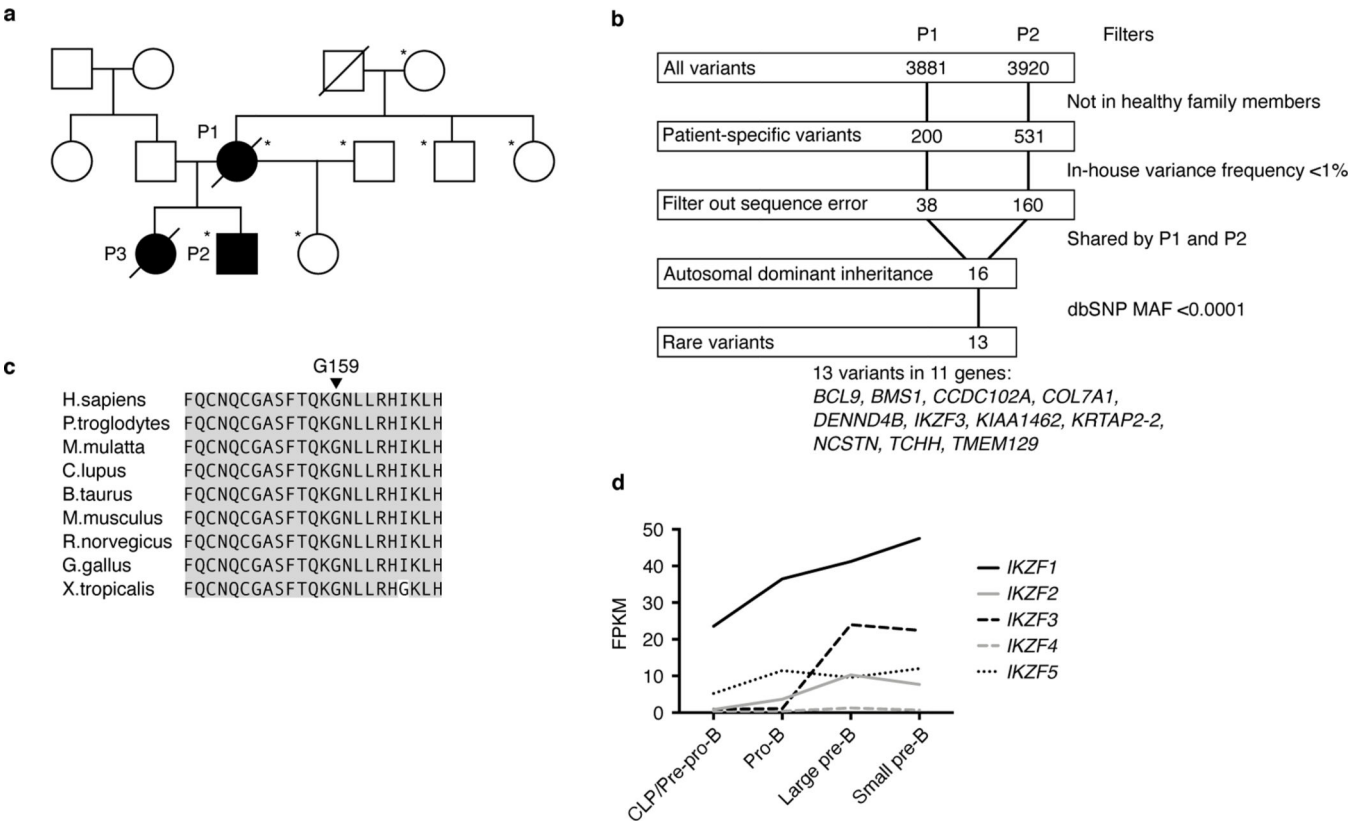
## Extended Data




**Extended Data Fig. 1. Lymphocyte and dendritic cell subset analyses of the P1.**

**(a)** Flow cytometric analysis of peripheral blood lymphocytes and monocytes in the P1 and a control subject. Gating strategies are shown above each plot. Numbers within the plots represent percentage of the defined populations. Briefly, the ratio of CD4/CD8 cells was inverted and T cells had an activated phenotype (HLA−DR$^+$CD38$^+$). CD4$^+$ T cells were skewed to memory phenotype (CD45RA$^-$CD45RO$^+$) and $T_H1$* cells (CD4$^+$CD45RO$^+$CD161$^+$CXCR3$^+$CCR6$^+$) were increased. iNKT cells were decreased. NK

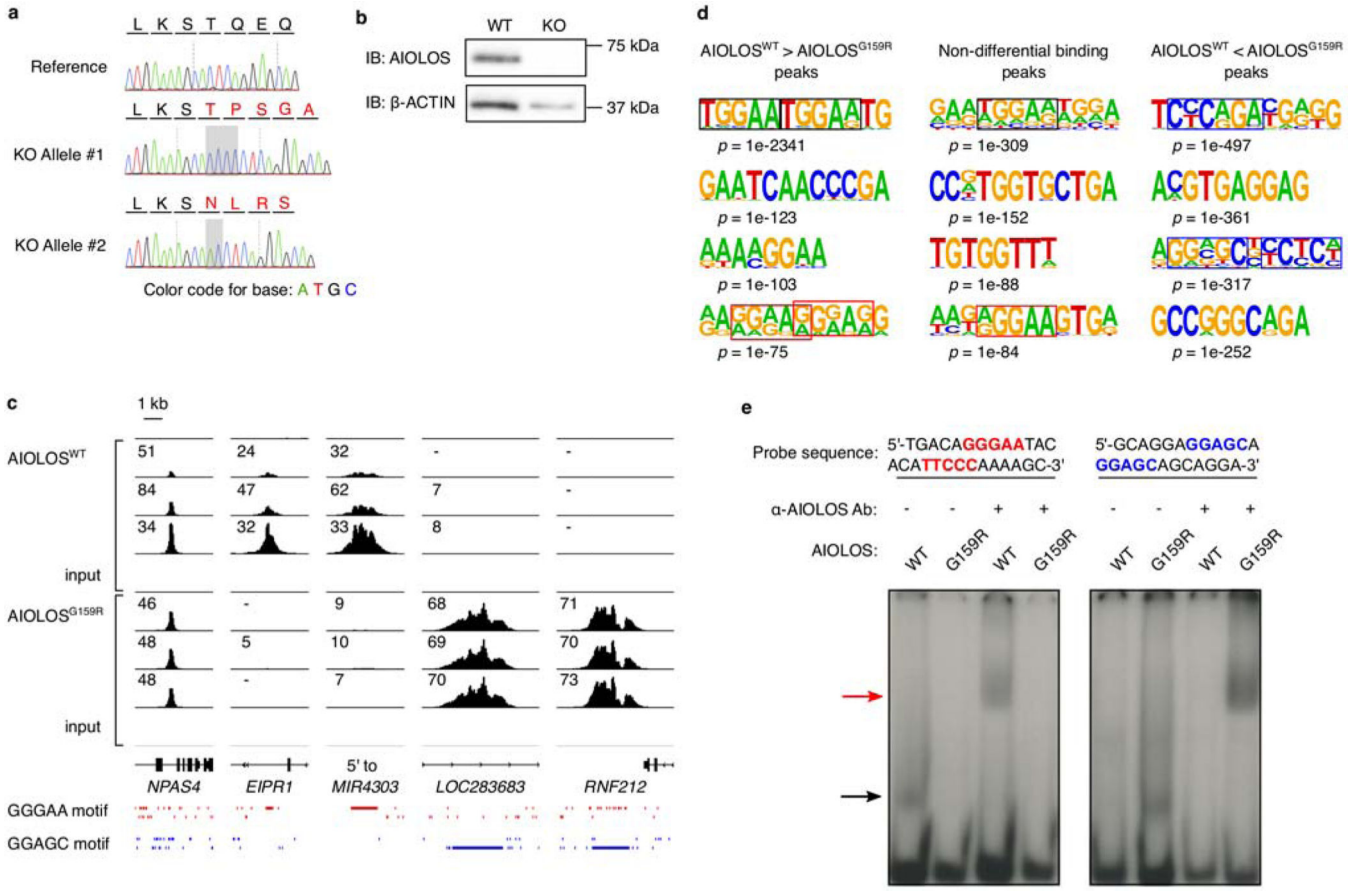
cells, mDCs, and eosinophils were also decreased. (**b)** CD3 and TCR expression levels in T cells of the P1 (blue) and control (red). Numbers represent MFI of P1's T cell subsets relative to a control subject. Surface expression levels of TCRβ, CD3, and CD8 were decreased in T cells of the P1, despite comparable expression of CD4. iNKT: invariant NK T cell, mDC: myeloid dendritic cell, pDC: plasmacytoid dendritic cell, Eo: Eosinophil, Baso: Basophil, MFI: Mean fluorescence intensity




**Extended Data Fig. 2. Whole exome analysis of the patients.**

**(a)** Patients and healthy family members who were analyzed by whole-exome sequencing are indicated with an asterisk. (**b)** Filtering strategy for whole-exome analysis. Patient-specific variants were selected by familial segregation. Variants resulting from sequencing errors were filtered out by ignoring the variants with high in-house frequency. Variants shared by the P1 and the P2 were selected and further narrowed down to rare variants by using cut-off as dbSNP minor allele frequency (MAF) of 0.0001. Candidate variants other than *IKZF3* were *BCL9* (NM_004326.2:c.3934delG, NP_004317.2:p.Gly1312Alafs, and NM_004326.2:c.3936_3937insTTT, NP_004317.2:p.Gly1311_His1312insGly), *BMS1* (NM_014753.3:c.3557C>T, NP_055568.3:p.Ala1186Val), *CCDC102A* (NM_033212.3:c.1135C>T, NP_149989.2:p.Arg379Trp), *DENND4B* (NM_014856.2:c.307G>A, NP_055671.2:p.Val103Ile), *KIAA1462* (NM_020848.2:c.2215G>A, NP_065899.1:p.Gly739Ser), *KRTAP2–2* (NM_033032.2:c.333_334delGA, NP_149021.2:p.Thr112Profs), *NCSTN* (NM_015331.2:c.464A>G, NP_056146.1:p.Glu155Gly), *TCHH* (NM_007113.2:c.1072_1074delGAG,

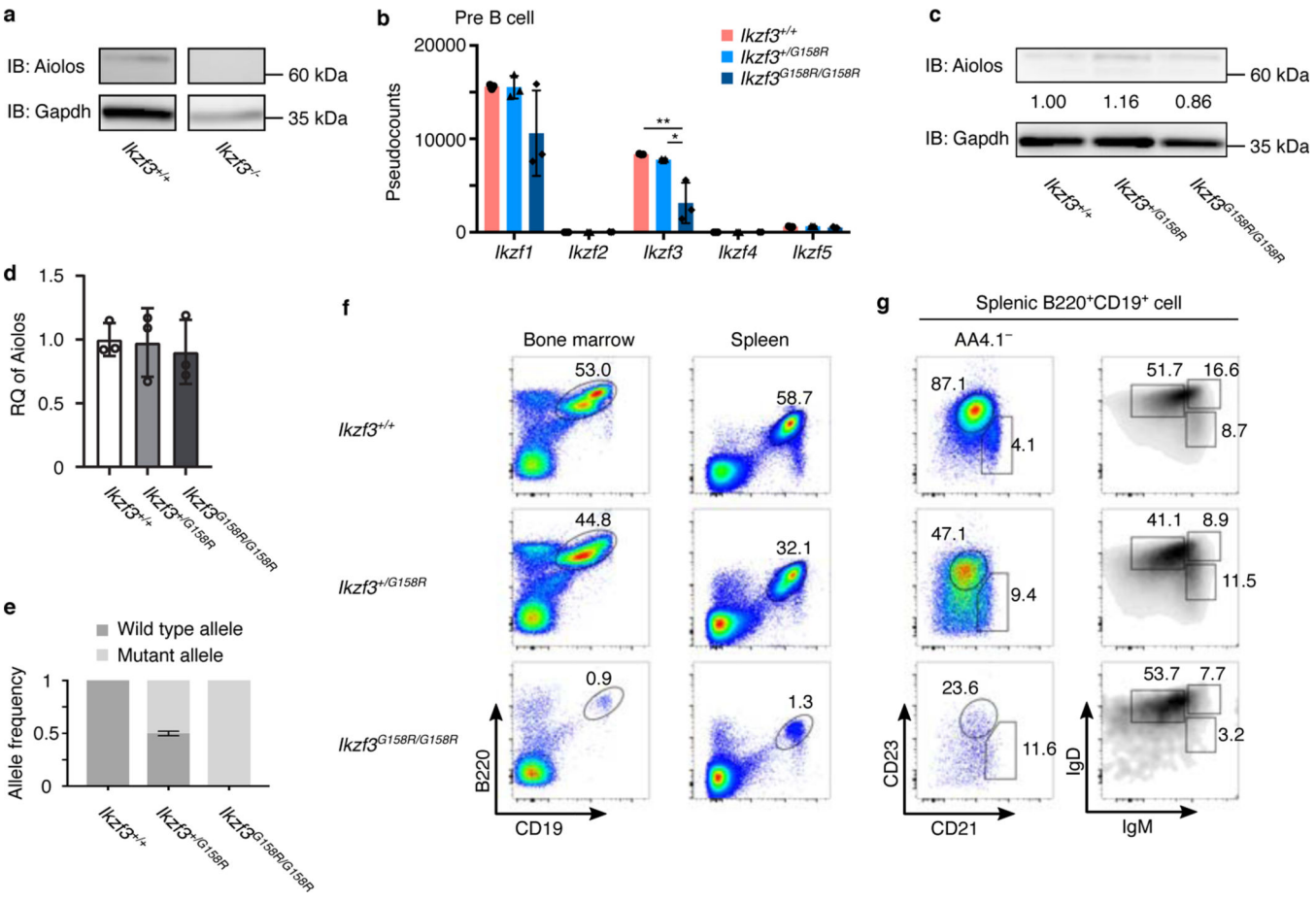
NP_009044.2:p.Glu358del), and *TMEM129* (NM_138385.3:c.40G>C, NP_612394.1:p.Val14Leu). (**c**) Alignment of amino acid sequences of the second zinc finger domain of AIOLOS orthologues from several species. Gray-shaded letters indicate identical amino acid in relation to human AIOLOS. Glycine residue at 159 position in human AILOS is well conserved beyond species. (**d**) Expression pattern of IKZF family genes during human B cell development. CLP/Pre-pro-B cell ($CD34^+CD10^+CD19^-$), pro-B cells ($CD34^+CD10^+CD19^+$), large pre-B cells ($CD34^-CD10^+CD19^+CD79B^+IgM^-$) and small pre-B cells ($CD34^-CD10^+CD19^+CD79B^-IgM^-$) were isolated by FACS sorting from the bone marrow aspirate of a healthy donor. RNA was extracted and subjected to RNA-seq analysis. FPKM of *IKZF* genes in the indicated populations are shown.




**Extended Data Fig. 3. Wild-type AIOLOS and AIOLOS$^{G159R}$ ChIP-seq in NALM-6 human pre-B cell line.**

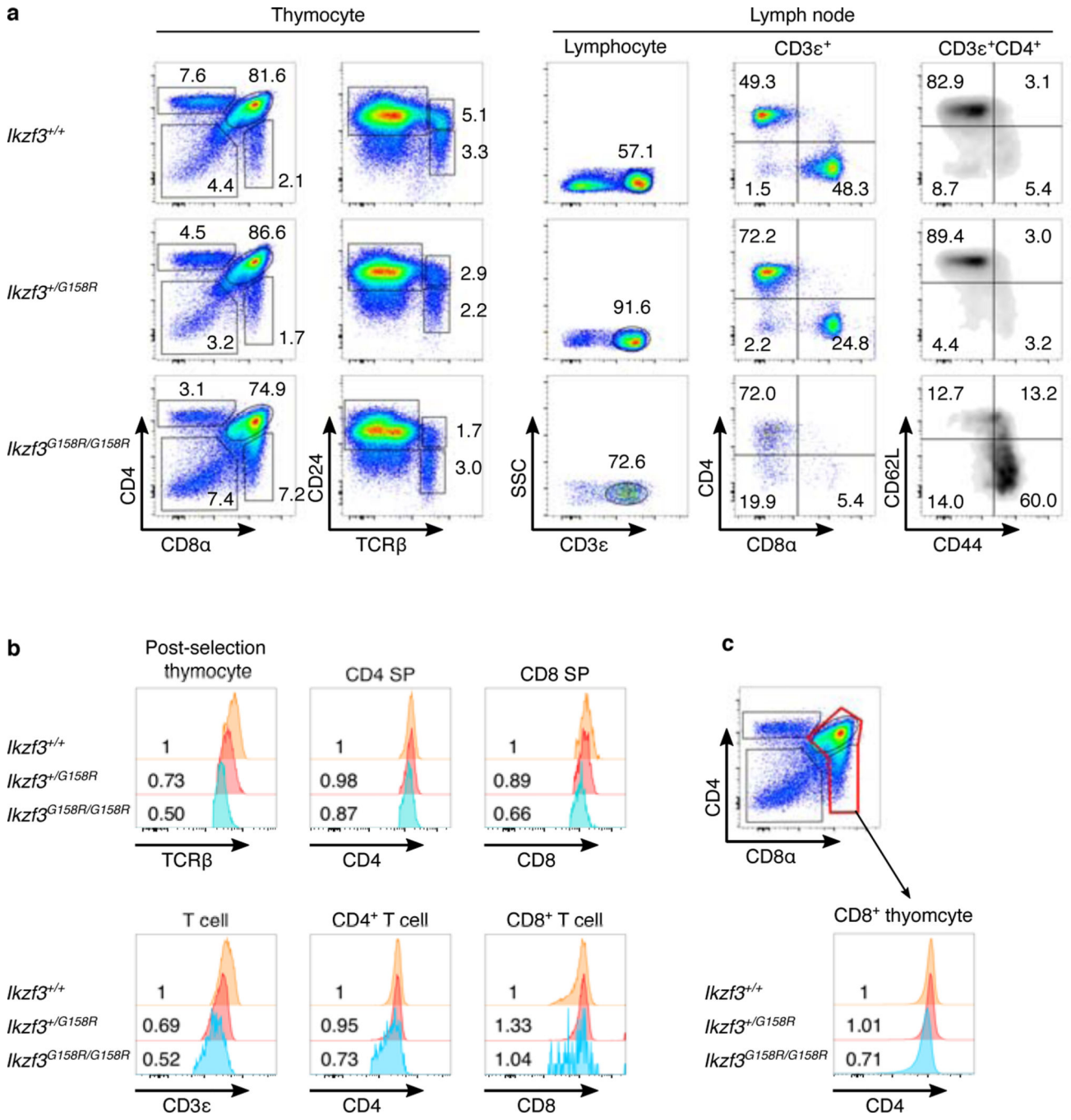
**(a)** Genomic sequence of the *IKZF3* knock-out (KO) NALM-6 cell line. Exon 2 of *IKZF3* was targeted by CRISPR-Cas9. Each allele of *IKZF3* was cloned and sequenced. The knockout clone had an indel in exon 2, resulting in a frameshift and premature termination of the protein. Grey shading indicates inserted nucleotides. Amino acid in red were changed by the frameshift. **(b)** Western blotting of AIOLOS in wild-type (WT) and *IKZF3*-KO NALM-6 cell lines. Representative of three independent experiments. **(c)** Triplicates of ChIP-seq tracks showing five representative loci with unique and common binding by AIOLOS$^{WT}$ and AIOLOS$^{G159R}$ in the *IKZF3*-KO NALM-6 cell line reconstituted

with FLAG-tagged AIOLOS$^{WT}$ or AIOLOS$^{G159R}$. Numbers represent the signal values of binding enrichment of the detected peaks. Structure of the genes are shown at the bottom. Locations of binding motifs (GGGAA and GGAGC) within the ChIP-seq track regions are indicated at the bottom. **(d)** The top significant DNA binding motifs with *p*-values for AIOLOS$^{WT}$ and AIOLOS$^{G159R}$ abstracted from the peaks with all statistically different bindings and non-differential bindings between quadruplicate ChIP-seq samples. The AIOLOS consensus binding sequence (GGGAA) is delineated by the red square and TGGAA motif is delineated by the black square, whereas binding motifs specific to the AIOLOS$^{G159R}$ peaks (GGAGC, GGAGG, and GCAGG) are delineated by the blue square. GGGAA and TGGAA motifs were consistently associated with AIOLOS$^{WT}$, whereas GGAGC, GGAGG, GCAGG, and CCCAGA motifs were repeatedly shown association with AIOLOS$^{G159R}$. Peaks with nondifferential binding between AIOLOS$^{WT}$ and AIOLOS$^{G159R}$ were enriched with relatively low accumulation of AIOLOS canonical binding motifs. **(e)** EMSA showing binding of AIOLOS$^{WT}$ and AIOLOS$^{G159R}$ binding to AIOLOS consensus sequence (indicated in red font, IK-BS4 probe) or AIOLOS$^{G159R}$ specific motif (GGAGC, indicated in blue font) containing probe designed from genome regions with high AIOLOS$^{G159R}$ peaks. Direct binding of AIOLOS$^{G159R}$ to GGAGC motif was observed. Representative of three independent experiments.



**Extended Data Fig. 4. Expression of Aiolos in thymocyte of *Ikzf3*$^{+/+}$, *Ikzf3*$^{+/G158R}$ and *Ikzf3* $^{G158R/G158R}$ mice, and supplementary flowcytometric analysis.**

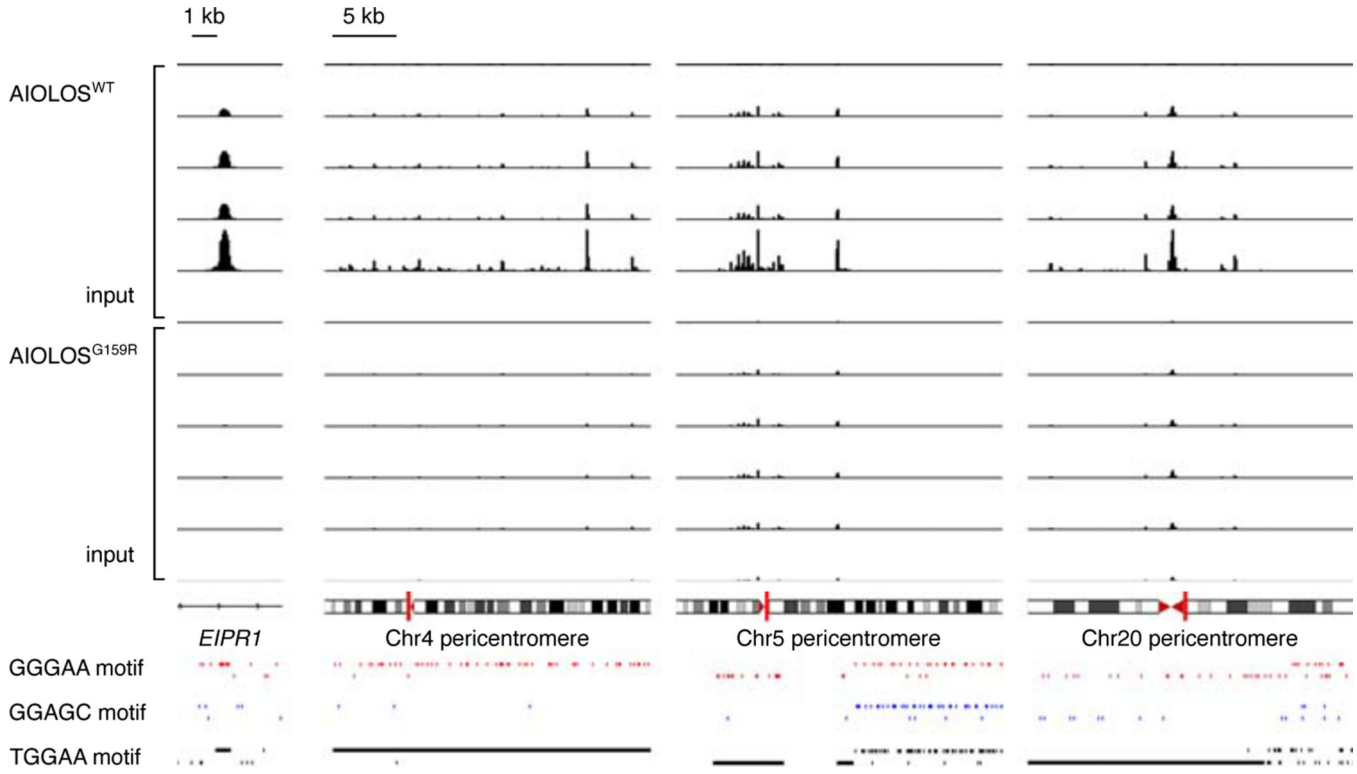
**(a)** Total cell lysates from the thymus of *Ikzf3*$^{+/+}$ and *Ikzf3*$^{-/-}$ mice were subjected for immunoblot using anti-Aiolos antibody. Representative of three independent experiments. (**b**) Expression levels of IKZF family genes determined by RNA-seq in pre-B cells of mice with the indicated genotype ($n$ = 3 for each genotypes). Bar graphs show mean with SD. * $p$ <0.0094, ** $p$ <0.0054, determined by one-way ANOVA. **(c, d)** Total cell lysates from the thymus of *Ikzf3*$^{+/+}$, *Ikzf3*$^{+/G158R}$ and *Ikzf3*$^{G158R/G158R}$ mice were subjected for immunoblot using anti-Aiolos antibody. Expression levels of Aiolos were normalized by Gapdh protein. Numbers indicate relative intensity of Aiolos of indicated genotypes to *Ikzf3*$^{+/+}$ sample(**c**). Graphs showing summary of relative quantity (RQ) of three independent experiments **(d)**. The expression levels of Aiolos were comparable between the genotypes ($n$ = 3 for each genotypes). Bar graphs show mean with SD. (**e**) Relative expression of wild-type and mutant *Ikzf3* alleles in pre-B cells calculated from RNA-seq data ($n$ = 3 for each genotypes). Bar graphs show mean with SD. (**f**) Flowcytometric analysis of CD19 and B220 staining in bone marrow and splenic cells in *Ikzf3*$^{+/+}$, *Ikzf3*$^{+/G158R}$ and *Ikzf3*$^{G158R/G158R}$ mice. (**g**) Flowcytometric analysis of follicular B cell and marginal zone B cells in indicated cell subsets. IgM and IgD expression in B220$^{+}$CD19$^{+}$ cells were also shown.




**Extended Data Fig. 5. T cell phenotypes in *Ikzf3*$^{+/G158R}$ and *Ikzf3*$^{G158R/G158R}$ mice.**

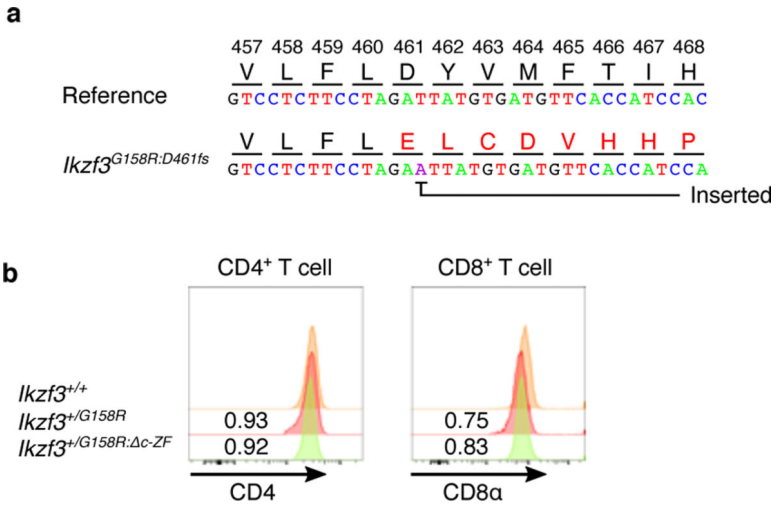
**(a)** Flow cytometric analysis of thymocyte and lymph node T cells in *Ikzf3*$^{+/+}$, *Ikzf3*$^{+/G158R}$, and *Ikzf3*$^{G158R/G158R}$ mice. Expression of indicated surface markers in total thymocytes, lymphocyte gated lymph node cells and CD3e$^{+}$ lymph node cells are shown. Numbers indicate the percentage of cells in each gate or each quadrant. Mature T cells in lymph node of *Ikzf3*$^{G158R/G158R}$ mice showed decrease of CD8$^{+}$ T cells and increase of CD4$^{-}$CD8$^{-}$ T cells. Similar but milder phenotypes were observed in *Ikzf3*$^{+/G158R}$ mice. CD4$^{+}$ T cells in lymph node of *Ikzf3*$^{G158R/G158R}$ mice showed skewing to CD44$^{+}$ memory phenotype,

which also recapitulated the patient's phenotype. (**b)** TCRβ, CD3ε, CD4, and CD8a expression levels in thymocyte and lymph node T cell subsets of *Ikzf3*$^{+/+}$, *Ikzf3*$^{+/G158R}$, and *Ikzf3*$^{G158R/G158R}$ mice. Numbers represent relative MFI against *Ikzf3*$^{+/+}$ mice. Similar to the human patients, *Ikzf3*$^{+/G158R}$ and *Ikzf3*$^{G158R/G158R}$ mice showed decreased expression of TCRβ and CD3ε expressions in thymocytes and lymph node T cells, respectively. (**c**) Emergence of CD4$^{lo}$CD8$^{+}$ cells in thymus of *Ikzf3*$^{G158R/G158R}$ mice. CD4 expression in CD8α$^{+}$ thymocytes (delineated by red line) is shown in the histogram. Numbers represent relative MFI against *Ikzf3*$^{+/+}$ mice.




**Extended Data Fig. 6. Accumulation of Wild-type AIOLOS in pericentromeres.**

FLAG-ChIP-seq was performed in *IKZF3*-KO NALM-6 cell line reconstituted with FLAGtagged AIOLOS$^{WT}$ or AIOLOS$^{G159R}$. Quadruplicates of ChIP-seq tracks showing the pericentromeric regions with TGGAA repeats, and a non-pericentromeric region containing TGGAA repeats. The locations of TGGAA motif are indicated at the bottom. As indicated by the binding motif analyses, AIOLOS$^{WT}$ predominantly bound to TGGAA-rich regions.




**Extended Data Fig. 7. C-terminal sequence of *Ikzf3*$^{G158R:\Delta c\text{-}ZF}$ (*Ikzf3*$^{G158R:D461fs}$) allele and T cell phenotypes of *Ikzf3*$^{+/G158R:\Delta c\text{-}ZF}$ mouse.**

**(a)** Direct sequencing of *Ikzf3* mRNA by *Ikzf3*$^{G158R}$ allele-specific PCR amplification of cDNA amplified using an *Ikzf3*$^{G158R}$ mutant allele-specific 5' primer and universal C-terminal 3' primer. Complementary DNA was synthesized from total RNA extracted from peripheral blood of F0 founder mice generated by CRISPR-Cas9 genome editing. The C-terminal sequence of the *Ikzf3*$^{G158R}$ mutant allele confirmed the single nucleotide insertion (indicated as a purple letter) which results in the frame-shift and disruption of ZF5–6 structure. **(b)** Expressions of CD4 and CD8 in lymph node T cells of *Ikzf3*$^{+/G158R}$, *Ikzf3*$^{+/+}$, and *Ikzf3*$^{+/G158R:\Delta c\text{-}ZF}$ mice. Numbers represent relative MFI against *Ikzf3*$^{+/+}$ mice.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

We thank members of Taniuchi's laboratory for technical advice and discussions; Y. Iizuka for microinjection of RNA for genome editing in mice; H. Asahara and T. Chiba for the advice on genome editing in human cell lines; P. Burrows for critical reading of the manuscript. We apologize to colleagues whose work we could not cite because of space restrictions. This work was supported by RIKEN IMS PID project (I.T.), MEXT KAKENHI JP18H02778 and Health Labour Sciences Research Grant 20FC1053 (T.M.), JSPS Grant-in-Aid for Young Scientists JP20K16884 (M.Y.), JSPS Fund for the Promotion of Joint International Research JP18KK0228 (S.O.), AMED The Practical Research Project for Rare/Intractable Diseases (S.O.), JSPS KAKENHI JP18H02395 (K.Y.J.Z.), and JSPS research fellowship (A.K.P).

## References

1. Tangye SG et al. Human Inborn Errors of Immunity: 2019 Update on the Classification from the International Union of Immunological Societies Expert Committee. J Clin Immunol 40, 24–64 (2020). [PubMed: 31953710]
2. Angulo I. et al. Phosphoinositide 3-Kinase δ Gene Mutation Predisposes to Respiratory Infection and Airway Damage. Science 342, 866–871 (2013). [PubMed: 24136356]

3. Lucas CL et al. Dominant-activating germline mutations in the gene encoding the PI(3)K catalytic subunit p110δ result in T cell senescence and human immunodeficiency. Nat Immunol 15, 88–97 (2014). [PubMed: 24165795]
4. Liu L. et al. Gain-of-function human STAT1 mutations impair IL-17 immunity and underlie chronic mucocutaneous candidiasisHuman STAT1 activating mutations impair IL-17 immunity. J Exp Medicine 208, 1635–1648 (2011).
5. Veerdonk FL van de et al. STAT1 Mutations in Autosomal Dominant Chronic Mucocutaneous Candidiasis. New Engl J Medicine 365, 54–61 (2011).
6. Boutboul D. et al. Dominant-negative IKZF1 mutations cause a T, B, and myeloid cell combined immunodeficiency. J Clin Invest 128, 3071–3087 (2018). [PubMed: 29889099]
7. Minegishi Y. et al. Dominant-negative mutations in the DNA-binding domain of STAT3 cause hyper-IgE syndrome. Nature 448, 1058–1062 (2007). [PubMed: 17676033]
8. Punwani D. et al. Multisystem Anomalies in Severe Combined Immunodeficiency with Mutant BCL11B. New Engl J Medicine 375, 2165–2176 (2016).
9. Bogaert DJ et al. A novel IKAROS haploinsufficiency kindred with unexpectedly late and variable B-cell maturation defects. J Allergy Clin Immun 141, 432–435.e7 (2018). [PubMed: 28927821]
10. Georgopoulos K. et al. The ikaros gene is required for the development of all lymphoid lineages. Cell 79, 143–156 (1994). [PubMed: 7923373]
11. 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). [PubMed: 8986714]
12. Schwickert TA et al. Stage-specific control of early B cell development by the transcription factor Ikaros. Nat Immunol 15, 283–293 (2014). [PubMed: 24509509]
13. Kuehn HS et al. Loss of B Cells in Patients with Heterozygous Mutations in IKAROS. New Engl J Medicine 374, 1032–1043 (2016).
14. Hoshino A. et al. Abnormal hematopoiesis and autoimmunity in human subjects with germline IKZF1 mutations. J Allergy Clin Immun 140, 223–231 (2017). [PubMed: 27939403]
15. Yoshida N. et al. Germline IKAROS mutation associated with primary immunodeficiency that progressed to T-cell acute lymphoblastic leukemia. Leukemia 31, 1221–1223 (2017). [PubMed: 28096536]
16. Nieuwenhove EV et al. A kindred with Mutant IKAROS and autoimmunity. J Allergy Clin Immun 142, 699–702.e12 (2018). [PubMed: 29705243]
17. Eskandarian Z. et al. Assessing the Functional Relevance of Variants in the IKAROS Family Zinc Finger Protein 1 (IKZF1) in a Cohort of Patients With Primary Immunodeficiency. Front Immunol 10, 568 (2019). [PubMed: 31057532]
18. Kuehn HS et al. Germline IKAROS dimerization haploinsufficiency causes hematologic cytopenias and malignancies. Blood 137, 349–363 (2021). [PubMed: 32845957]
19. Mullighan CG et al. Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. Nature 446, 758–764 (2007). [PubMed: 17344859]
20. Mullighan CG et al. Deletion of IKZF1 and Prognosis in Acute Lymphoblastic Leukemia. New Engl J Medicine 360, 470–480 (2009).
21. Churchman ML et al. Germline Genetic IKZF1 Variation and Predisposition to Childhood Acute Lymphoblastic Leukemia. Cancer Cell 33, 937–948.e8 (2018). [PubMed: 29681510]
22. Wang J-H et al. Aiolos Regulates B Cell Activation and Maturation to Effector State. Immunity 9, 543–553 (1998). [PubMed: 9806640]
23. Morgan B. et al. Aiolos, a lymphoid restricted transcription factor that interacts with Ikaros to regulate lymphocyte differentiation. Embo J 16, 2004–2013 (1997). [PubMed: 9155026]
24. Patel A, Zhang X, Blumenthal RM & Cheng X. Structural basis of human PR/SET domain 9 (PRDM9) allele C–specific recognition of its cognate DNA sequence. J Biol Chem 292, 15994–16002 (2017). [PubMed: 28801461]
25. Cobb BS et al. Targeting of Ikaros to pericentromeric heterochromatin by direct DNA binding. Gene Dev 14, 2146–2160 (2000). [PubMed: 10970879]
26. Bruin L. M. O. de, Volpi S. & Musunuru K. Novel Genome-Editing Tools to Model and Correct Primary Immunodeficiencies. Front Immunol 6, 250 (2015). [PubMed: 26052330]

27. Hu Y. et al. Superenhancer reprogramming drives a B-cell–epithelial transition and high-risk leukemia. Gene Dev 30, 1971–1990 (2016). [PubMed: 27664237]
28. Prosser J, Frommer M, Paul C. & Vincent PC Sequence relationships of three human satellite DNAs. J Mol Biol 187, 145–155 (1986). [PubMed: 3701863]
29. Zhang J. et al. Harnessing of the nucleosome-remodeling-deacetylase complex controls lymphocyte development and prevents leukemogenesis. Nat Immunol 13, 86–94 (2012).
30. 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). [PubMed: 12871645]
31. Berman HM et al. The Protein Data Bank. Nucleic Acids Res 28, 235–242 (2000). [PubMed: 10592235]
32. Altschul SF, Gish W, Miller W, Myers EW & Lipman DJ Basic local alignment search tool. J Mol Biol 215, 403–410 (1990). [PubMed: 2231712]
33. Šali A. & Blundell TL Comparative Protein Modelling by Satisfaction of Spatial Restraints. J Mol Biol 234, 779–815 (1993). [PubMed: 8254673]
34. Shen M. & Sali A. Statistical potential for assessment and prediction of protein structures. Protein Sci 15, 2507–2524 (2006). [PubMed: 17075131]
35. Case DA et al. AMBER 14. (http://ambermd.org/doc12/Amber14.pdf, 2014).
36. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW & Klein ML Comparison of simple potential functions for simulating liquid water. J Chem Phys 79, 926–935 (1983).
37. Maier JA et al. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. J Chem Theory Comput 11, 3696–3713 (2015). [PubMed: 26574453]
38. Peters MB et al. Structural Survey of Zinc-Containing Proteins and Development of the Zinc AMBER Force Field (ZAFF). J Chem Theory Comput 6, 2935–2947 (2010). [PubMed: 20856692]
39. Berendsen HJC, Postma JPM, Gunsteren WF van, DiNola, A. & Haak, J. R. Molecular dynamics with coupling to an external bath. J Chem Phys 81, 3684–3690 (1984).
40. Ryckaert J-P, Ciccotti G. & Berendsen HJC Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J Comput Phys 23, 327–341 (1977).
41. Essmann U. et al. A smooth particle mesh Ewald method. J Chem Phys 103, 8577–8593 (1995).
42. Humphrey W, Dalke A. & Schulten K. VMD: Visual molecular dynamics. J Mol Graphics 14, 33–38 (1996).
43. Dobin A. et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21 (2013). [PubMed: 23104886]
44. Robinson MD, McCarthy DJ & Smyth GK edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139–140 (2010). [PubMed: 19910308]
45. Yu G, Wang L-G, Han Y. & He Q-Y clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. Omics J Integr Biology 16, 284–287 (2012).
46. Kim D, Langmead B. & Salzberg SL HISAT: a fast spliced aligner with low memory requirements. Nat Methods 12, 357–360 (2015). [PubMed: 25751142]
47. Zhang Y. et al. Model-based Analysis of ChIP-Seq (MACS). Genome Biol 9, R137 (2008). [PubMed: 18798982]
48. Stark R. & Brown G. DiffBind: differential binding analysis of ChIP-Seq peak data. (2011).
49. 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). [PubMed: 20513432]

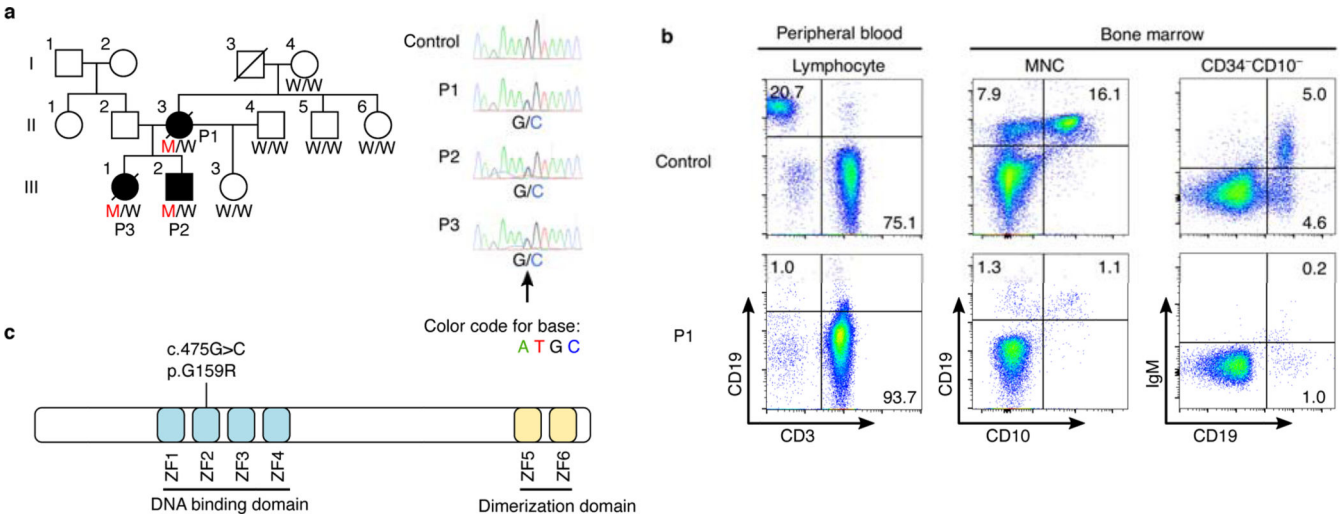



**Figure 1. A heterozygous *IKZF3* variant associated with B cell deficiency.**

**(a)** The pedigree of patients with a heterozygous *IKZF3* variant. Squares and circles represent males and females, respectively. Black symbols represent family members with B cell deficiency and harboring a heterozygous *IKZF3* variant. The right panel shows the *IKZF3* sequences of the patients (P1–P3) and a control subject. The arrow indicates the heterozygous *IKZF3* c.475G>C variant. **(b)** Flow cytometric analysis of peripheral blood and BM of P1 and a healthy control. Numbers indicate the percent of cells in each quadrant. **(c)** A schematic of AIOLOS, which is encoded by *IKZF3*. The blue and yellow boxes represent the N-terminal ZFs that mediate DNA binding and the C-terminal ZFs that mediate dimerization, respectively. The missense variant G159R is located at the second ZF.

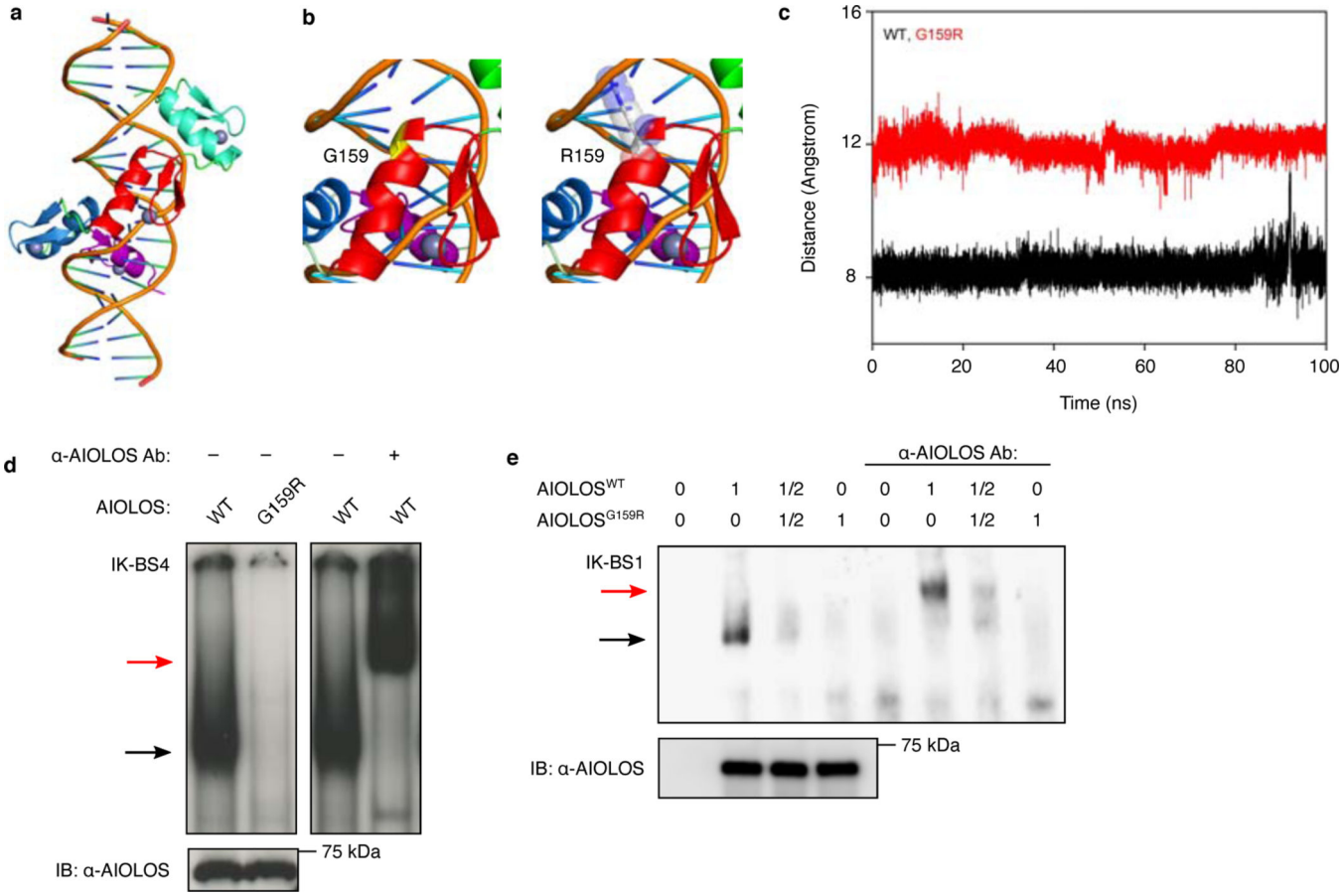



**Figure 2. Impaired binding of AIOLOS$^{G159R}$ to the AIOLOS consensus sequence.**

**(a)** The structure of human AIOLOS with four N-terminal ZFs in a complex with DNA generated by homology modeling using human PR/SET domain 9 (PRDM9) as a template. The N-terminal four ZFs (ZF1–4) of AIOLOS are shown in cyan, red, blue, and purple, respectively. Zinc ions bound to the four ZFs are shown as gray spheres. **(b)** Homology modeled structure of the DNA binding domain of the second ZF of AIOLOS showing that G159 is in contact with DNA and the side chain of R159 in the AIOLOS$^{G159R}$ mutant creates steric clashes with DNA. G159 is shown in yellow and R159 in gray. **(c)** A line plot showing the distance between Cα-G159 with N1-guanine 28 and Cα-R159 with N1-guanine 28 of DNA in WT AIOLOS (black) and the AIOLOS$^{G159R}$ mutant (red), respectively, indicating that the G159 to R159 substitution pushed the DNA away from the second ZF, thereby reducing the binding capacity of AIOLOS$^{G159R}$. The distance data were obtained from the molecular dynamics simulation trajectories of WT AIOLOS and AIOLOS$^{G159R}$ mutant. **(d)** EMSA showing *in vitro* DNA binding of WT AIOLOS and the AIOLOS$^{G159R}$ mutant. Nuclear lysates of HEK293T cells transfected with an expression vector encoding each protein were incubated with a radiolabeled IK-BS4 probe. Anti-AIOLOS antibody (Ab) generated a supershift of the DNA–protein complexes (red arrow). Cell lysates were tested for AIOLOS expression with an anti-AIOLOS Ab (lower panel). Data are representative of three independent experiments. **(e)** Competitive EMSA using an IK-BS1 probe with nuclear lysates of HEK293T cells transfected with WT

AIOLOS and/or the AIOLOS$^{G159R}$ mutant at the indicated ratios. Cell lysates were tested for AIOLOS expression with an anti-AIOLOS Ab (lower panel). Data are representative of three independent experiments.

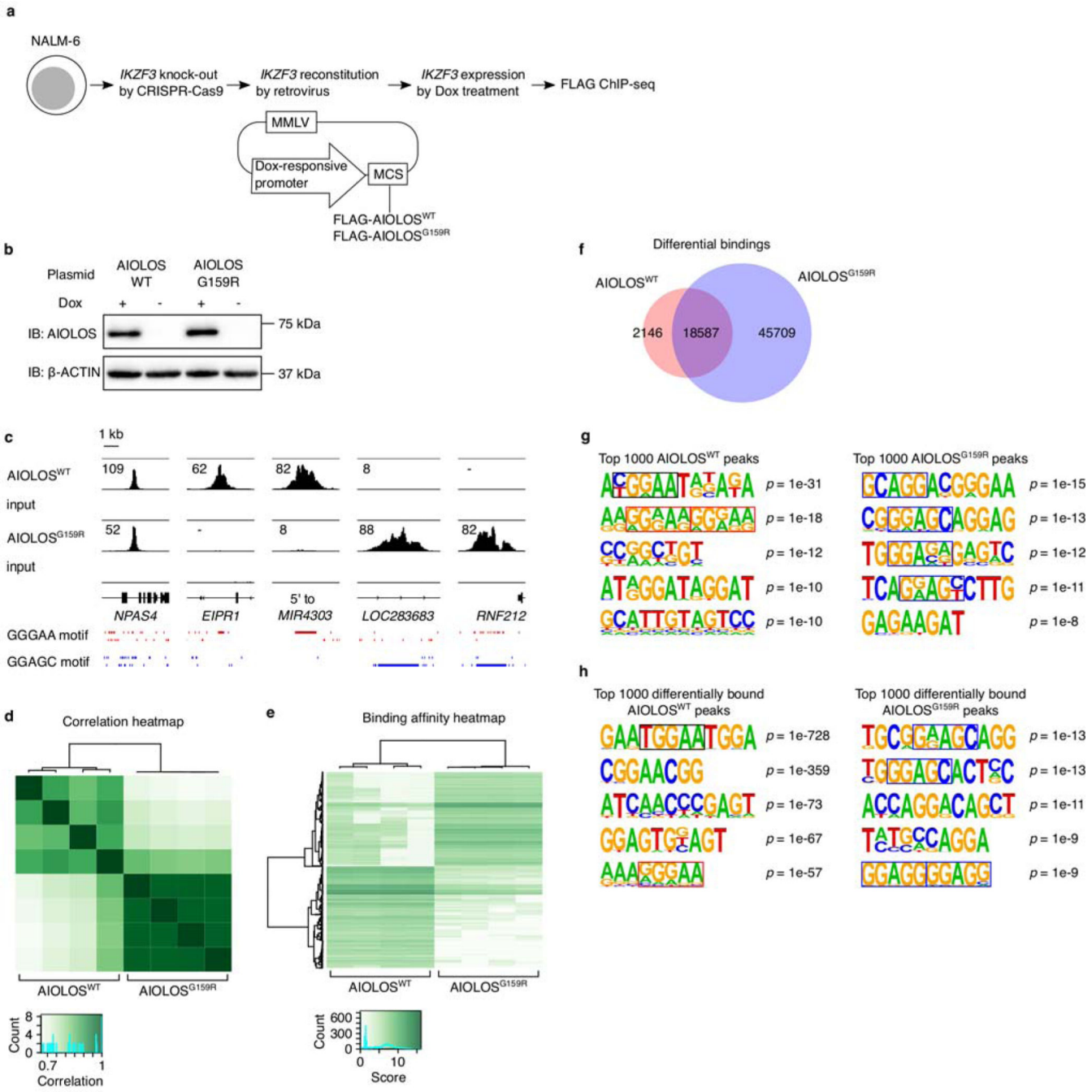



**Figure 3. Altered DNA binding specificity of the AIOLOS$^{G159R}$ variant.**

**(a)** A schematic of AIOLOS ChIP-seq of human NALM-6 pre-B cells. *IKZF3* knocked-out in NALM-6 pre-B cells by targeting exon 2 using CRISPR-Cas9. Dox-inducible, FLAG-tagged AIOLOS$^{WT}$ or AIOLOS$^{G159R}$ was retrovirally transduced into the *IKZF3* knock-out NALM-6 cells. AIOLOS was induced by doxycycline treatment and FLAG ChIP-seq was performed. (**b)** A immunoblot of transduced AIOLOS$^{WT}$ and AIOLOS$^{G159R}$ before and after induction by doxycycline. **(c)** ChIP-seq tracks showing representative loci with unique and common binding by AIOLOS$^{WT}$ and AIOLOS$^{G159R}$. Numbers represent the

signal values of binding enrichment of the detected peaks. Structures of the corresponding genes are shown below the ChIP-seq tracks. Locations of binding motifs (GGGAA and GGAGC) within the ChIP-seq track regions are indicated at the bottom in red and blue. The upper and lower rows indicate motifs of the plus and minus strands, respectively. (**d)** A correlation heatmap and clustering of the quadruplicate ChIP-seq samples of AIOLOS$^{WT}$ and AIOLOS$^{G159R}$ induced in *IKZF3* knock-out NALM-6 cells. Hierarchical clustering is shown above the heatmap. **(e)** A binding affinity heatmap and clustering of the quadruplicate ChIP-seq samples. Rows represent the binding loci and the color intensity reflects binding strength. Hierarchical clustering of the differentially bound sites is indicated to the left of the heatmap. **(f)** A Venn diagram representing differential and non-differential bindings of AIOLOS$^{WT}$ and AIOLOS$^{G159R}$, calculated from quadruplicate ChIP-seq samples. (**g, h**) The top significant DNA binding motifs with *p*-values for AIOLOS$^{WT}$ and AIOLOS$^{G159R}$. Top 1000 highly enriched peaks bound by representative ChIP-seq samples of AIOLOS$^{WT}$ or AIOLOS$^{G159R}$ (**g**), and top 1000 statistically different bindings between AIOLOS$^{WT}$ and AIOLOS$^{G159R}$ quadruplicate ChIP-seq samples (**h**). The AIOLOS consensus binding sequence (GGGAA) is delineated by the red square, whereas binding motifs specific to the AIOLOS$^{G159R}$ peaks (GGAGC, GGAGG, and GCAGG) are delineated by the blue square. The TGGAA sequence (delineated by black square) is AIOLOS$^{WT}$ enriched peaks. Abbreviations: MCS, multiple cloning site; MMLV, Moloney murine leukemia virus

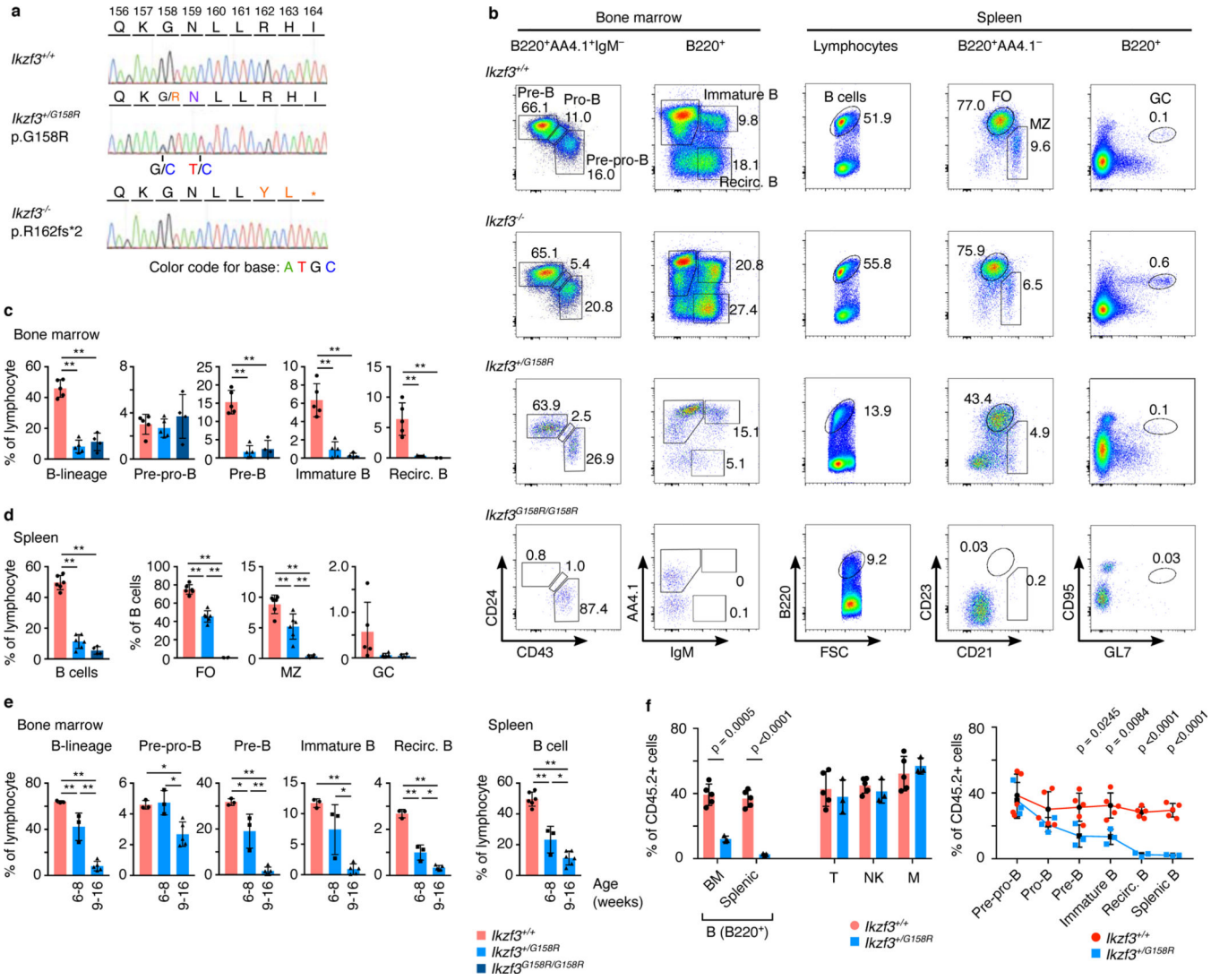



**Figure 4. *Ikzf3*$^{G158R}$ knock-in mice recapitulate impaired B cell differentiation in the patient.**

**(a)** Genomic sequences of the *Ikzf3*$^{G158R}$ and *Ikzf3* knock-out alleles. Nucleotide and amino acid sequences are indicated. Orange font indicates altered amino acids. Purple font indicates the amino acid coded by the synonymous mutation. **(b)** Flow cytometric analysis of BM B cell progenitors and splenic B cell subsets with the indicated genotypes of mice at 9–16 weeks of age. **(c, d)** Frequencies of B cell progenitor subsets in BM (**c**) and splenic B cell subsets (**d**) of mice with the indicated genotypes ($n = 5$ for *Ikzf3*$^{+/+}$ and *Ikzf3*$^{+/G158R}$, $n = 4$ for *Ikzf3* $^{G158R/G158R}$ (**c**); $n = 6$ for *Ikzf3*$^{+/+}$ and *Ikzf3*$^{+/G158R}$, $n = 4$ for *Ikzf3* $^{G158R/G158R}$ (**d**)). Bar graphs show mean with SD. (**e**) Frequencies of B cell progenitor subsets and splenic B cells of *Ikzf3*$^{+/+}$ and *Ikzf3*$^{+/G158R}$ mice at the indicated (BM: $n = 3$ for *Ikzf3*$^{+/+}$ and *Ikzf3*$^{+/G158R}$ (6–8 weeks of age), $n = 5$ for *Ikzf3*$^{+/G158R}$ (9–16 weeks of age), spleen: $n = 6$ for *Ikzf3*$^{+/+}$(6–8 weeks of age), $n = 3$ for *Ikzf3*$^{+/G158R}$ (6–8 weeks of age), $n = 6$ for *Ikzf3*$^{+/G158R}$ (9–16 weeks of age)). Bar graphs show mean with SD. * $p < 0.05$, ** $p < 0.005$, determined by one-way ANOVA (**c–e**). (**f**) BM chimera experiment using BM of *Ikzf3*$^{+/+}$ and *Ikzf3*$^{+/G158R}$ mice. CD45.2$^{+}$ *Ikzf3*$^{+/+}$ or *Ikzf3*$^{+/G158R}$ BM was mixed

with CD45.1$^+$ *Ikzf3*$^{+/+}$ BM at a 1:1 ratio and injected to sublethally irradiated CD45.1$^+$ *Ikzf3*$^{+/+}$ mice. Percentages of CD45.2$^+$ cells within BM B-lineage cells, splenic B cells (B220$^+$), T cells (TCRβ$^+$), NK cells (TCRβ$^-$CD56$^+$), and myeloid cells (CD11b$^+$) at 8 weeks after transplantation are shown. Moreover, the percentages of CD45.2$^+$ cells among B cell progenitors in the BM are indicated ($n$ = 5 for *Ikzf3*$^{+/+}$ and $n$ = 3 *Ikzf3*$^{+/G158R}$). Bar graphs and line graphs show mean with SD. $P$ values were determined by $t$-test.

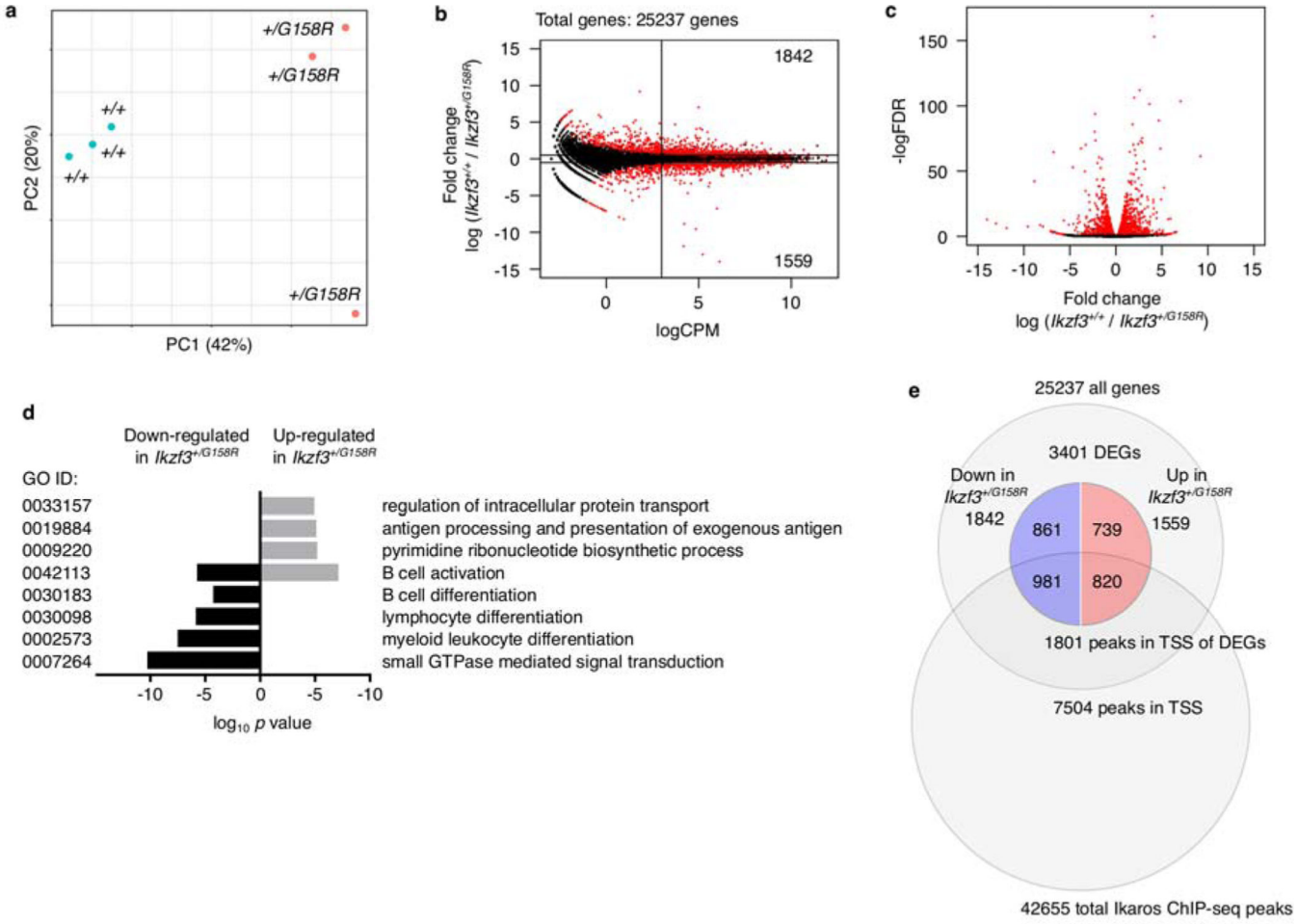



**Figure 5. Genes involved in B cell development are dysregulated in *Ikzf3*$^{+/G158R}$ pre-B cells.**

**(a)** Principal component analysis of the pre-B cell transcriptomes of *Ikzf3*$^{+/+}$ and *Ikzf3*$^{+/G158R}$ mice (n = 3/genotype). **(b)** MA plot (M, log ratio; A, mean average) of gene expression of *Ikzf3*$^{+/+}$ vs. *Ikzf3*$^{+/G158R}$ pre-B cells assessed by RNA-Seq. Each symbol represents one gene and red symbols indicate differentially expressed genes (DEGs) with an FDR of <0.05. The numbers of down- and up-regulated genes are indicated in the upper- and lower-right corners, respectively. The vertical lines represent $\log_2 0.5$ of fold change in expression and the horizontal line represents an average logCPM of 3. These thresholds identified 608 and 778 DEGs up- and down-regulated in *Ikzf3*$^{+/G158R}$ pre-B cells, respectively. **(c)** A volcano plot of gene expression in *Ikzf3*$^{+/+}$ vs. *Ikzf3*$^{+/G158R}$ pre-B cells. Red symbols indicate DEGs with an FDR of <0.05. **(d)** Gene ontology (GO) analysis of DEGs between *Ikzf3*$^{+/+}$ and *Ikzf3*$^{+/G158R}$ pre-B cells. DEGs with $\log_2$ fold change of >0.5 with an average logCPM of >3 were included for analysis. Each row represents a GO term with *p*-values of up- or down-regulated genes in *Ikzf3*$^{+/G158R}$ pre-B cells. **(e)** A Venn diagram showing overlaps of 3401 DEGs between *Ikzf3*$^{+/+}$ and *Ikzf3*$^{+/G158R}$ pre-B cells with known Ikaros binding sites in the WT pre-B cells (42655 total peaks and 7504 peaks in the TSS, GSE86897)[27]. Red and blue semicircles represent up- and down-regulated DEGs in *Ikzf3*$^{+/G158R}$ pre-B cells, respectively. Abbreviations: FDR, false discovery ratio; TSS, transcription start site

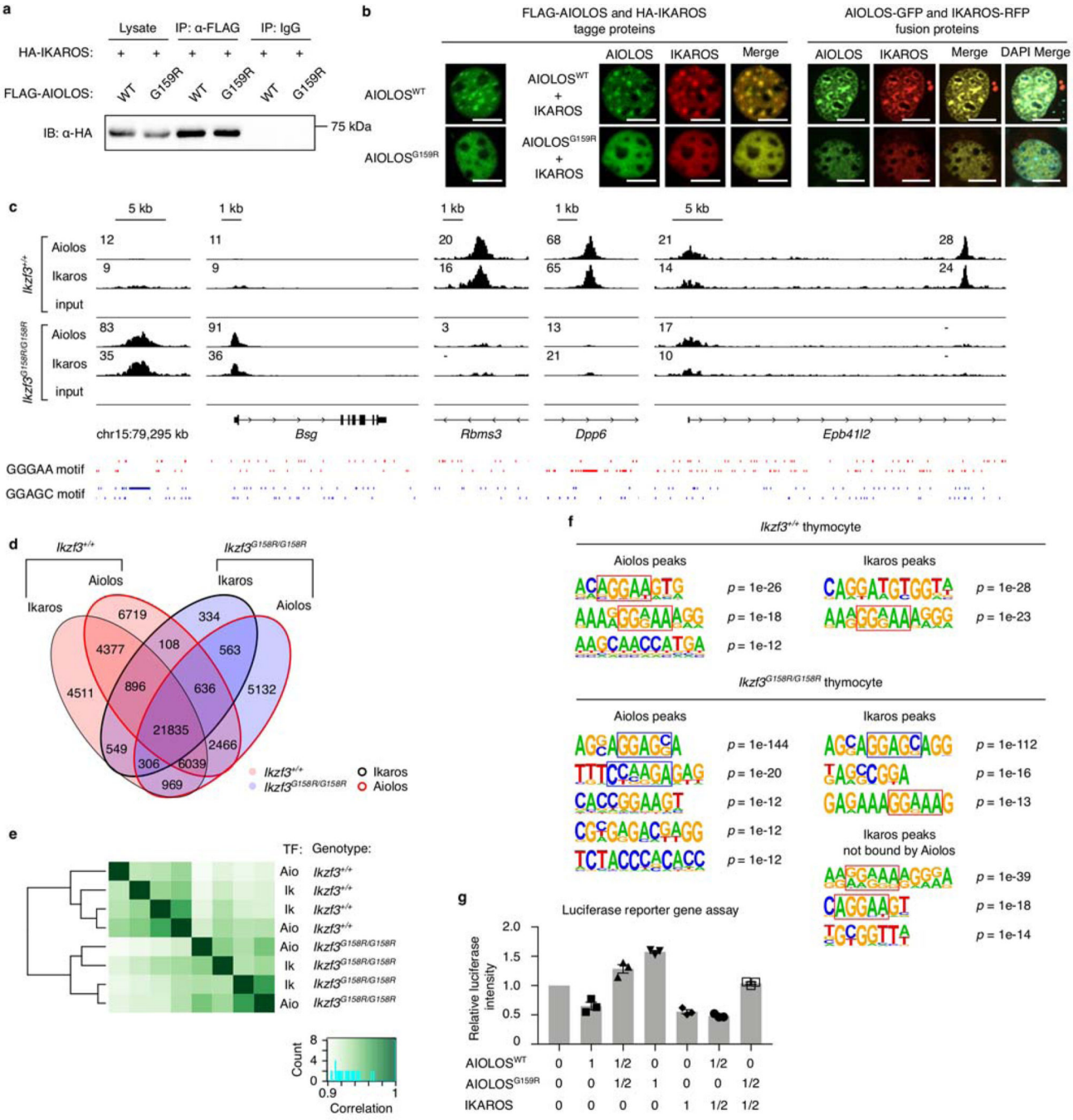



**Figure 6. AIOLOS$^{G159R}$ and Aiolos$^{G158R}$ interfere with Ikaros function.**

**(a)** Co-IP of FLAG-AIOLOS (WT or G159R) and HA-IKAROS. Immunoprecipitation with anti-FLAG Ab was performed following transfection of HEK293T cells with FLAG-AIOLOS (WT or G159R) and HA-IKAROS. A immunoblot with anti-HA Ab is shown. Representative of three independent experiments. **(b)** NIH3T3 cells were transfected with FLAG-tagged AIOLOS (WT or G159R) with or without HA-tagged IKAROS. The cells were stained with anti-HA and/or anti-Aiolos Abs and then incubated with secondary Abs conjugated with Alexa Fluore 488 for AIOLOS or Alexa Fluore 568 for IKAROS

(left). COS-7 cells were transiently transfected with RFP-IKAROS and either GFP-tagged AIOLOS or GFP-AIOLOS$^{G159R}$ (right). The speckled foci represent heterochromatin formation. Representatives of three independent experiments for each condition. Scale bar, 10 μm. **(c-f)** ChIP-seq analysis of Aiolos and Ikaros binding in thymocytes of *Ikzf3*$^{+/+}$ and *Ikzf3*$^{G158R/G158R}$ mice. Summary of two independent experiments. Representatives binding loci are shown along with the corresponding chromosomal positions and gene structures. Numbers represent signal values of binding enrichment of the identified peaks. Locations of binding motifs within the ChIP-seq track regions are indicated at the bottom **(c)**. A Venn diagram showing overlapped binding by Aiolos and Ikaros in thymocytes of *Ikzf3*$^{+/+}$ and *Ikzf3*$^{G158R/G158R}$ mice. Figures represent distinct peak numbers identified in each sample **(d)**. A correlation heatmap and clustering of the genome-wide binding by Aiolos and Ikaros in *Ikzf3*$^{+/+}$ and *Ikzf3*$^{G158R/G158R}$ thymocytes in duplicate **(e).** The top significant DNA binding motifs for Aiolos and Ikaros in *Ikzf3*$^{+/+}$ and *Ikzf3*$^{G158R/G158R}$ thymocytes are shown with *p*-values. Peaks with >20-fold enrichment were subjected to analysis. Binding motifs of the top 1000 enriched peaks of Ikaros not co-bound by Aiolos in *Ikzf3*$^{G158R/G158R}$ thymocytes are also shown. The consensus binding sequence of Ikaros and Aiolos is delineated by the red squares, whereas the binding motifs specific to Aiolos$^{G158R}$ and Ikaros in *Ikzf3*$^{G158R/G158R}$ thymocytes are delineated by the blue squares **(f). (g)** The luciferase reporter gene assay results of HEK293T cells transfected with various combinations of vectors are indicated below the plot. The results were normalized to those of cells transfected with empty vectors. Data are presented as the mean with SEM of three independent experiments. Representative of three independent experiments. Abbreviation: TF, transcription factors

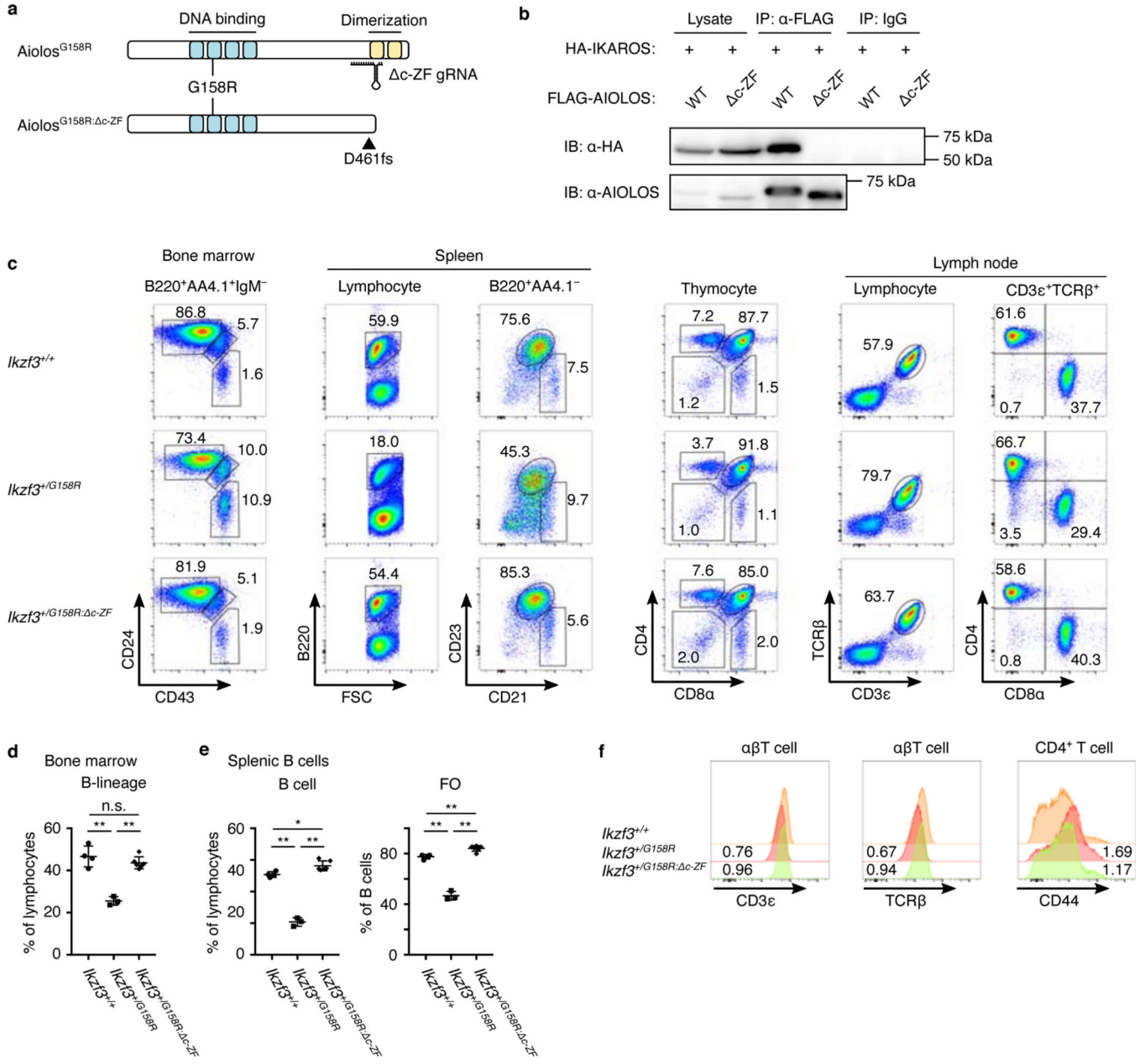



**Figure 7. Restored B cell development and T cell abnormalities by removal of the dimerization domain of the Aiolos$^{G158R}$ mutant.**

**(a)** A schematic of the C-terminal ZF deleting experiments in *Ikzf3*$^{+/G158R}$ mice. Guide RNA targeting exon 8 of *Ikzf3* and Cas9 mRNA were co-injected to the fertilized eggs. Founder offspring harboring a C-terminal deletion on the *Ikzf3*$^{G158R}$ allele (*Ikzf3*$^{+/G158R:\Delta c\text{-}ZF}$ mice) were selected by mouse crossing. **(b)** Co-immunoprecipitation of FLAG-tagged WT or Δc-ZF mutant AIOLOS and HA-IKAROS using total cell lysates of HEK293T cells transfected with FLAG-AIOLOS (WT or Δc-ZF mutant) and HA-IKAROS. Immunoblots with anti-HA Ab and anti-AIOLOS Ab are shown. Representative of three independent experiments. **(c)** Flow cytometric analysis of BM B cell progenitors, splenic B cell subsets, thymocytes, and lymph node T cells of 6-week-old *Ikzf3*$^{+/+}$, *Ikzf3*$^{+/G158R}$,

and *Ikzf3*$^{+/G158R:\Delta c\text{-}ZF}$ mice. Numbers adjacent to the outlined areas indicate the percent of cells in each gate. **(d, e)** Frequency of BM B-lineage cells (**d**), splenic B cells, and FO B cell subsets (**e**) in *Ikzf3*$^{+/+}$, *Ikzf3*$^{+/G158R}$, and *Ikzf3*$^{+/G158R:\Delta c\text{-}ZF}$ mice ($n$ = 4 for *Ikzf3*$^{+/+}$, $n$ = 3 for *Ikzf3* $^{+/G158R}$, and $n$ = 6 for *Ikzf3*$^{+/G158R:\Delta c\text{-}ZF}$). Bar graphs show mean with SD. * $p$ <0.05, ** $p$ <0.005, determined by one-way ANOVA (**d, e**). (**f**) Expression of TCRβ, CD3ε, and CD44 in thymocytes and lymph node T cell subsets of *Ikzf3*$^{+/+}$, *Ikzf3*$^{+/G158R}$, and *Ikzf3*$^{+/G158R:\Delta c\text{-}ZF}$ mice. Numbers represent relative mean fluorescence intensity against *Ikzf3*$^{+/+}$ mice.