

Chromatin Reader Dido3 Regulates the Genetic Network of B Cell Differentiation

Fernando Gutiérrez del Burgo

Centro Nacional de Biotecnología: Centro Nacional de Biotecnología <https://orcid.org/0000-0001-6659-5762>

Tirso Pons

Centro Nacional de Biotecnología <https://orcid.org/0000-0002-7616-7750>

Enrique Vázquez de Luis

Fundación Centro Nacional de Investigaciones Cardiovasculares Carlos III: Fundacion Centro Nacional de Investigaciones Cardiovasculares Carlos III <https://orcid.org/0000-0003-1157-2592>

Carlos Martínez-A

Centro Nacional de Biotecnología <https://orcid.org/0000-0002-2121-189X>

Ricardo Villares (✉ rvillares@cnb.csic.es)

Centro Nacional de Biotecnología <https://orcid.org/0000-0001-7562-6700>

Research

Keywords: Dido1, H3K4me3, PRC2, B cells, hematopoiesis, bone marrow.

Posted Date: November 11th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-1031396/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

BACKGROUND

The development of hematopoietic lineages is based on a complex balance of transcription factors whose expression depends on the epigenetic signatures that characterize each differentiation step. The B cell lineage arises from hematopoietic stem cells through the stepwise silencing of stemness genes and balanced expression of mutually regulated transcription factors, as well as DNA rearrangement, in a complex process involving epigenetic remodeling.

RESULTS

Here we report the impact on B cell differentiation of the lack of DIDO3, a reader of histone post-translational modifications, in the mouse hematopoietic compartment. We found reduced DNA accessibility in hematopoietic precursors, leading to a severe deficiency specifically in the generation of successive stages of B-cell differentiation. The expression of essential transcription factors and differentiation markers is impaired, as is the process of somatic recombination. DIDO3-deficient cells show transcriptional alterations of a number of polycomb repressive complex 2, suggesting the involvement of DIDO3 in determining the specific activity of PRC2 in the B cell lineage, including VH-DJH rearrangement.

CONCLUSIONS

Taken together, our data suggest that DIDO3 is an epigenetic reader involved in the specific differentiation of B cell precursors in the hematopoietic compartment of mice.

Background

B cells originate from hematopoietic stem cells (HSC) through a process defined by the activation or silencing of key transcription factors and parallel DNA recombination events. HSC can develop into multipotent progenitors (MPP) that are able to give rise to B cells by passing through lymphoid progenitor cell stages known as lymphoid-primed multi-potential progenitor (LMPP), common lymphoid progenitor (CLP), prepro-B, pro-B, pre-B, and immature B cells¹. It is currently considered that B cell differentiation is initiated by interaction of a PU.1-, IKAROS-, and IL7R-expressing progenitor cell with stromal cells and IL7². IKAROS complexes might initially suppress myeloid gene expression in lymphoid progenitors and then participate in B and T cell specification by respectively silencing T and B cell-specific loci³. Progenitor progression to the B cell lineage is then controlled by the transcription factors E2A (E12/47), EBF1, FOXO1, and PAX5⁴. E2A and EBF1 specify the B cell lineage by activating B lymphoid gene expression in prepro-B cells⁵. PAX5 subsequently controls B cell commitment at the transition to the pro-B cell stage, and is necessarily maintained to sustain B-lineage identity⁶.

In the CLP subpopulation, somatic rearrangements are first detected between the diversity (DH) and joining (JH) gene segments of the heavy chain (HC) locus⁷. CLP can mature into four different cell types: B lymphocytes, T lymphocytes, NK, and dendritic cells^{8,9}. In mice, upregulation of B220 expression in a CLP subpopulation allows differentiation to prepro-B cells. Prepro-B cells determined for B cell differentiation are characterized by tyrosine kinase AA4.1 (CD93) expression¹⁰. Entry into the B lineage is marked by CD19 expression and the end of DH-JH rearrangement in the early pro-B state. IL7R signaling contributes to early pro-B cell survival and transient proliferation¹¹ and sequential reorganization of Ig genes¹². Heavy chain recombination culminates in the late pro-B stage with the productive rearrangement of VH heavy chain variable segments with DJH (VH-DJH). Productive rearrangement of *Igh* is essential for pre-BCR signaling, a crucial checkpoint for survival and further differentiation¹³.

Post-translational histone modifications constitute one of the epigenetic mechanisms that enables control of chromatin structure, DNA accessibility, and gene expression in a dynamic, cell type-specific manner. Proteins that specifically “read” these modifications must play important roles in development, differentiation, and tumorigenesis¹⁴.

Dido1 is a gene highly conserved throughout evolution; its orthologs have been identified in chordates and insects. The shortest coded protein isoform, DIDO1, includes nuclear location and nuclear export signals (NLS, NES), and a plant Zn-finger homeodomain (PHD); it interacts with chromatin mainly through di/trimethylated lysine of histone H3 (H3K4me2/3)^{15,16}, and works as a reader of the chromatin state. DIDO1 nonetheless has yet to be assigned to known complexes. Both PHD and NLS, but not NES, are present in all three natural isoforms. DIDO2 is the least abundant and includes a transcription elongation factor S-II central domain (TFIIS2M), necessary for the nucleolytic activity of RNA polymerase II, and a *Spn* paralog and ortholog C-terminal (SPOC) domain of unknown function, but which is usually found in proteins with transcriptional co-repressor activity. DIDO3 is the main isoform and includes all domains found in DIDO2 as well as a coiled-coil (CC) region that probably has a tertiary or quaternary structural function. There are also proline-, alanine-, and arginine-rich low complexity regions in DIDO3.

All attempts to date to delete *Dido1* have resulted in cell lethality. N-terminal truncation results in a shorter protein, expressed by an internal promoter, that is associated with aneuploidy, centrosome amplification, and DNA damage in the form of centromere-localized breaks¹⁷. Mice with the *Dido1*ΔNT mutation show hematological myeloid neoplasms, and alterations in *DIDO1* are associated with the myelodysplastic syndrome in humans¹⁸. Deletion of the *Dido* 3' terminal exon (E16) renders a truncated form of DIDO3. This deficiency leads to embryonic lethality; embryos die by gestation day 8.5¹⁹. Embryonic stem cells (ESC) derived from these DIDO3 C-terminal mutant embryos do not differentiate correctly, but retain their capacity for self-renewal.

Due to the embryonic lethality of *Dido* exon 16, we studied the effect of DIDO3 deficiency in the hematopoietic lineage using mice bearing a floxed exon 16 of *Dido1*²⁰ and a Cre recombinase gene under the control of the *Vav1* promoter. In these mice, the B cell population is severely reduced by a primary

defect in the pro-B cell stage progression. By analyzing the transcription levels of B cell-characteristic genes at distinct differentiation stages in these DIDO3-deficient mice, we detected a reduction in expression of B cell markers and of transcription factors specifically in prepro-B samples; there was no evident alteration in master regulators such as *Spi1* (PU.1) and *Ikzf1* (Ikaros). This finding suggests that DIDO3 is involved in the chromatin restructuring needed for the B cell differentiation pathway, and thus controls developmental cell plasticity and lineage fate decisions.

Results

To characterize the *Dido1* expression pattern in different human and mouse tissues, cell lines, and cancer types, we retrieved publicly available data from web bio-sources. Publicly accessible expression data using the Genevisible tool (<https://genevisible.com/search>). The Affymetrix microarray data (Mouse Genome 430 2.0 and Human Genome U133 Plus 2.0) explored with Genevisible web bio-source (<https://genevisible.com>) indicated the highest *Dido1* mRNA expression in the nervous and hematopoietic systems (Suppl Fig. 1).

Loss of the DIDO3 protein results in anemia and severe peripheral lymphopenia. To examine the *in vivo* role of DIDO3 in hematopoiesis, we deleted exon 16 of *Dido1* in murine HSC early in development. Mice homozygous for the floxed allele (*Dido1*^{flE16/flE16}) were crossed with mice carrying one functional and one mutated allele of *Dido1* (*Dido1*^{ΔE16}, which lacks the DIDO3 C terminus) and the *Vav1:iCre* transgene. The resulting *Dido1*^{flE16/ΔE16};R26R-EYFP;iCre^{Vav1} mice (hereafter *Dido1*ΔE16) were used as experimental subjects, and *Dido1*^{flE16/+};R26R-EYFP;iCre^{Vav1} littermates (hereafter wt) were used as controls, except where otherwise specified. Ablation of *Dido1* exon 16 in early development was confirmed by qPCR on DNA samples from sorted lineage (lin)⁻Kit⁺Sca1⁺(LSK) cells derived from bone marrow, which showed >99% reduction in *Dido1*ΔE16 DNA content (Suppl. Fig. 2A). Hematologic analysis showed a significant decrease in cellular components when *Dido1*ΔE16 mice were compared with heterozygous littermate controls. White blood cell numbers dropped to 33% of wt (1.71 ± 1.32 vs. $5.11 \pm 1.71 \times 10^3$ cells/ml, n = 8, p<0.001) and red cell numbers to 92% of wt values (9.65 ± 0.36 vs. $10.54 \pm 0.51 \times 10^6$ cells/ml, n = 8, p<0.01). Within the leukocyte fraction, the lymphocyte subpopulation was responsible for the defect observed, with a reduction in cell numbers to 27% of wt (1.02 ± 1.20 vs. $3.84 \pm 1.61 \times 10^3$ cells/ml, n = 8, p<0.0001)(Suppl. Fig. 2B). Peripheral blood samples from wt and *Dido1*ΔE16 mice were analyzed by flow cytometry for EYFP reporter expression; >97% of circulating cells were EYFP-positive in both lymphoid and myeloid gates, as determined by forward scatter (FS) vs. side scatter (SS) gating (Suppl. Fig. 2B).

We calculated total numbers of various cell subsets in peripheral blood from complete blood count and flow cytometry data. *Dido1*ΔE16 mice had significantly reduced lymphocyte subset counts, including B220 (58 ± 26 vs 2675 ± 423 cells/μl in control samples, n = 6, p<0.0001), CD4 (301 ± 123 vs. 749 ± 232 cells/μl, p<0.01), CD8 (109 ± 48 vs. 423 ± 87 cells/μl, p<0.001), and NK1.1 (50 ± 15 vs. 163 ± 69 cells/μl, p<0.05), while monocyte (90 ± 8 vs 90 ± 18 cells/μl) and granulocyte numbers (590 ± 348 vs. 550 ± 161 cells/μl) were unaffected (Suppl. Fig. 2C).

Bone marrow analysis

As *Dido1* was previously linked more closely to development and stem cell differentiation functions, we analyzed the B cell lineage in bone marrow. We used flow cytometry to identify $\text{lin}^- \text{Sca1}^+ \text{Kit}^+ \text{IL7R}^-$ (LSK) and common lymphoid progenitor $\text{lin}^- \text{Sca1}^+ \text{Kit}^+ \text{IL7R}^+$ (CLP) cell populations (Fig. 1A, B), and found no significant differences in the relative proportions of each. *Dido1* Δ E16 mice nonetheless showed reduced total bone marrow cellularity and significantly lower absolute numbers of LSK cells (Fig. 1C). From *Dido1* Δ E16 mice, we obtained $2.85 \pm 0.83 \times 10^4$ lin^- cells, vs. $3.93 \pm 0.45 \times 10^4$ from wt mice ($n = 6$, t-test ** $p < 0.01$). The numbers for LSK ($2.09 \pm 0.64 \times 10^4$ vs. $3.34 \pm 1.0 \times 10^4$, t-test * $p < 0.05$) and CLP cells ($1.13 \pm 0.50 \times 10^3$ vs. $1.92 \pm 1.17 \times 10^3$, t-test $p > 0.05$) were also reduced.

We analyzed the distribution of the distinct B cell precursor subpopulations, namely prepro-B, pro-B, pre-B, and immature B cells (Fig. 2A). The number of B220⁺ cells was substantially lower in *Dido1* Δ E16 mice than in the controls (Fig. 2B). The populations most affected were pre-B and immature B cells, although the number of prepro-B and pro-B cells was similar in mice of both genotypes (Fig. 2C).

Functional analysis of bone marrow precursors

Signaling from the pre-BCR complex promotes the survival and proliferation of cells that have successfully rearranged the *Igh* locus. This is a first checkpoint in which those rearrangements that give rise to a non-functional heavy chain lead to death by apoptosis. In a second checkpoint, immature B cells whose BCR is activated by an autoantigen in the bone marrow can undergo receptor editing to modify its light chain, or they can be deleted by apoptosis. To detect cell death in distinct precursor populations, we used flow cytometry to analyze bone marrow samples from *Dido1* Δ E16 and wt mice, labeled with appropriate antibodies and with annexin V and DAPI. Apoptosis levels were slightly higher, although non-significant, in *Dido1* Δ E16 mice. In the immature B cell population ($\text{EYFP}^+ \text{B220}^+ \text{CD19}^+ \text{IgM}^{\text{low}}$), both the number of apoptotic cells and the annexin V mean fluorescence intensity (MFI) in bone marrow of *Dido1* Δ E16 mice showed a significant increase relative to controls (Fig. 3A, B). This result suggests that DIDO3 affects the size of the immature B cell population in *Dido1* Δ E16 mice by increasing the number of cells that undergo apoptotic processes at this stage. Our data support the idea that DIDO3 is necessary to overcome the BCR signaling-dependent checkpoint.

We used BrdU (5-bromo-2'-deoxyuridine) incorporation in DNA *in vivo* to identify and examine proliferating cells, and DAPI to determine DNA content and cell cycle stages. The percentage of cells that incorporated BrdU after 18 h post-intraperitoneal inoculation was quantified for pro-B/pre-B and immature/mature cells. Cells from *Dido1* Δ E16 mice in pro-B/pre-B states showed BrdU incorporation similar to that of control mice (Fig. 3C). In the immature cell population, there was a non-significant increase in the percentage and the MFI of cells that incorporated BrdU (Fig. 3D) (MFI: pro-B/pre-B: *Dido1* Δ E16 = 4.25 ± 1.232 vs. wt = 4.68 ± 0.852 ; t test: $p = 0.52$; $n = 3$; immature/mature B cells, *Dido1* Δ E16 = 3.92 ± 1.144 vs. wt = 2.05 ± 0.443 ; t test: $p = 0.09$; $n = 3$). The percentage of S phase cells of the pro-B/pre-B populations of both mice was similar, although an increase in the percentage of

*Dido1*ΔE16 mature/immature B cells was observed in this phase (Fig. 3D). The increase in the percentage of immature B cells in the S phase in *Dido1*ΔE-16 compared to control mice and the increase in cells that undergo apoptosis suggest that immature B cells from *Dido1*ΔE16 mice undergo an S phase block that eventually leads to apoptosis.

In vitro differentiation assays were carried out to determine whether the *Dido1*ΔE16 mouse B cell precursors showed alterations in their capacity to differentiate towards the B lineage. We purified cells lacking lineage-specific markers (*lin*⁻) by negative selection and plated them in a semi-solid matrix and medium optimized for B cell development, supplemented with IL7. In cultures from wild type mice, between 25 and 35 colonies were counted in plates seeded with 2 x 10⁶ *lin*⁻ cells/ml; between 5 and 10 smaller cell groups/plate were also registered. No colonies were found in plates seeded with *lin*⁻ precursors from *Dido1*ΔE16 mice, although we detected some clusters consisting of a small number of disaggregated and undifferentiated cells (Suppl. Fig 3). B lineage cells (B220⁺) from the conditional mutant mice thus showed differentiation defects when stimulated *in vitro*, with blocking of proliferation and maturation towards B precursors in these conditions.

Hematopoietic precursors from *Dido1*ΔE16 mice failed to repopulate the bone marrow of irradiated recipient mice

The small numbers of early hematopoietic precursors in *Dido1*ΔE16 mice and their failure in *in vitro* differentiation assays led us to test whether their function was affected *in vivo* in the absence of DIDO3. As transplant recipients, we used nine lethally irradiated B6-CD45.1 mice (10 Gy); at day 1 post-irradiation, the mice received a retro-orbital injection of 150 μL of a mixture of total bone marrow cells (3.75 x 10⁶) from CD45.2⁺EYFP⁺ *Dido1*ΔE16 and CD45.2⁺EYFP⁻ *Dido1*⁺ donor mice, at a 1:1 ratio (Suppl. Fig. 4).

Peripheral blood was monitored at day 14 post-inoculation; we observed a lower CD45.2⁺EYFP⁺ *Dido1*ΔE16 cell contribution to circulating blood populations; of the blood cells, 12.78 ± 1.97% were CD45.2⁺EYFP⁺ *Dido1*ΔE16 and 80.13 ± 3.43% were CD45.2⁺EYFP⁻ *Dido1*⁺. A further 6.73 ± 3.45% were endogenous surviving cells (CD45.1⁺), mainly of the CD3⁺ phenotype, reported to be thymus-derived radioresistant cells²¹ (Fig. 4A). At day 21 post-inoculation, ~25% of circulating leukocytes were B cells, most of which were CD45.2⁺EYFP⁻ *Dido1*⁺, with only 0.05% CD45.2⁺EYFP⁺ *Dido1*ΔE16 and 0.01% CD45.1 B cells. This imbalance was partially anticipated, given the severely biased composition of the transplanted bone marrow, with higher percentages of pre-B and immature B progenitors in wt mice.

At 28 days post-transplant, four mice were sacrificed and their bone marrow analyzed by flow cytometry. The highest percentage of hematopoietic cells observed was CD45.2 (5.91 ± 6.26% CD45.2⁺EYFP⁺ *Dido1*ΔE16; 55.64 ± 9.42% CD45.2⁺EYFP⁻ *Dido1*⁺), with a residual percentage of CD45.1 cells (0.99 ± 0.49%). We also verified that only 3.02 ± 1.96 x 10⁴ cells were CD45.2⁺EYFP⁺ *Dido1*ΔE16 vs. 2.24 ± 1.21 x 10⁶ CD45.2⁺EYFP⁻ *Dido1*⁺ cells; 8.45 ± 6.49 x 10³ CD45.1 cells were also found. This result suggests a lower level of nesting and/or expansion capacity in DIDO3-deficient cells.

At 40 days post-transplant, the five remaining mice were sacrificed and bone marrow populations were analyzed by flow cytometry (Fig. 4B). From $6.04 \pm 1.33 \times 10^5$ lin^- cells/femur, only $1.68 \pm 1.13 \times 10^4$ cells/femur were $\text{CD45.2}^+\text{EYFP}^+$ *Dido1* Δ E16, and $1.64 \pm 0.41 \times 10^4$ cells/femur were CD45.1^+ .

The bone marrow analysis yielded means of $2.25 \pm 1.87 \times 10^3$ LSK and $1.31 \pm 1.25 \times 10^2$ CLP $\text{CD45.2}^+\text{EYFP}^+$ *Dido1* Δ E16 cells, significantly lower than the $1.91 \pm 0.80 \times 10^4$ LSK and $8.58 \pm 2.34 \times 10^2$ CLP $\text{CD45.2}^+\text{EYFP}^-$ *Dido1* $^+$ cells. B cell precursors were identified by B220 expression and encompass cell populations from the prepro-B phase to immature and mature B cells. Approximately 99% of B220^+ cells were CD45.2EYFP^- *Dido1* $^+$ ($2.15 \pm 1.01 \times 10^6$ cells/femur), and only $2.96 \pm 1.90 \times 10^4$ cells/femur were $\text{CD45.2}^+\text{EYFP}^+$ *Dido1* Δ E16 (Fig. 4C). These data indicate that, relative to LSK-wt cells, the LSK-*Dido1* Δ E16 cells and the remainder of the B lineage progenitor cells had a cell-autonomous disadvantage both in nesting and in repopulating the bone marrow of irradiated mice.

Identification of essential regulators of stem cell function and the B cell pathway by ATAC-seq analysis of bone marrow LSK cells

Although the proportion of LSK cells within the lin^- population is similar in wt and *Dido1* Δ E16 mice, adoptive transfer experiments showed a lower repopulation capacity, especially considering that this undifferentiated population was overrepresented in the mutant fraction of the inoculum. As DIDO1 was demonstrated to be a chromatin interactor that recognizes accessible regions of chromatin with H3K4me3 histone marks, we analyzed this population by Omni-ATAC-seq²². This technique detects differences in chromatin accessibility, as assessed by sensitivity to DNA digestion by transposase Tn5 in a process termed “tagmentation”. The tagmented fragments were subjected to massive sequencing, and two biological replicas with accessible regions in the range of 10^4 were used for differential analysis between wt and *Dido1* Δ E16 LSK cells.

Accessible regions were annotated with the ChIPseeker program²³ by associating the position of the accessible regions in each replica with the structure of the genes (promoter, 5'-UTR, exon, intron, 3'-UTR, or intergenic) in the mouse reference genome assembly GRCm38/UCSC mm10. After filtering low quality reads, an average of 52 million high quality reads were obtained, and 96.1% clean reads were mapped to the GRCm38/UCSC mm10 assembly using Bowtie2 (bowtie-bio.sourceforge.net/bowtie2/) and in parallel BWA-MEM 0.7.15 (<http://bio-bwa.sourceforge.net>). The largest percentage of accessible regions was concentrated in gene promoters and introns and in intergenic regions, in both LSK-wt and LSK-*Dido1* Δ E16 cells (Fig. 5A). In the *Dido1* Δ E16 samples, however, there was a relative reduction in the number of introns and intergenic regions relative to promoters, which was related to the preferential association of DIDO1 to enhancer sequences over promoters²⁴.

A Gene Set Enrichment Analysis (GSEA, <http://www.gsea-msigdb.org>) of the ATAC-seq data was performed. GSEA is a computerized method that determines whether gene sets defined *a priori* show concordant and significant differences between two biological states²⁵. In seeking overlaps with

genesets in the C2 collection (“curated genesets”) of MSigDB, we found that the promoters with the least accessibility in LSK-*Dido1*ΔE16 coincided with (1) genes with mixed epigenetic marks (H3K4me2 and H3K27me3) in mouse embryonic fibroblast (MEF) HCP genes (high-CpG-density promoters), $p < 1.79 \times 10^{-49}$; (2) silencing marks (H3K27me3) in embryonic cells, $p < 4.88 \times 10^{-31}$; (3) human EED targets (a core subunit of PRC2), $p < 3.72 \times 10^{-27}$, and (4) human SUZ12 targets (also a core subunit of PRC2), $p < 3.61 \times 10^{-25}$ (Supplementary Table 1). In accordance with these data, our previous microarray data²⁶ analyzed here by GSEA, indicate that genes silenced in DIDO3-deficient ESC correlate significantly with those silenced in mouse *Suz12*-deficient ESC; the genes overexpressed in these cells also coincided with those overexpressed in the *Suz12*^{-/-} cells. There thus appears to be a relationship between epigenetic regulation of gene expression by PRC2 and the presence of DIDO3. An online tool (<http://bioinformatics.psb.ugent.be/webtools/Venn/>) was used to construct a Venn diagram that represents differentially-expressed genes (DEG) associated with distinct genesets by GSEA (Fig. 5B). There was also correlation with (5) genes expressed after p53 activation ($p < 1.48 \times 10^{-24}$), in agreement with the *Cdkn2a* alterations observed in *Dido1*ΔE16 mouse samples at later developmental stages (see below).

When analysis was limited to non-promoter regions (intra- or intergenic), using the MSigDB C3 collection (“regulatory target gene sets”), there was significant overlap with (1) the targets of the putative histone-lysine N-methyltransferase PRDM6 ($p < 7.5 \times 10^{-24}$), and (2) the set of genes with at least one occurrence of the highly conserved motif M17 AACTTT, a potential binding site for an unknown transcription factor in the region spanning up to 4 kb around their transcription start sites ($p < 5.91 \times 10^{-69}$) (Fig. 5C). Within the 79 genes common to the three genesets, we found a number of essential regulators for stem cell function and B cell pathway entrance, including *Akt3*, *Runx1*, *Ebf1*, *Igfr1*, *Tcf4*, *Cdk6*, and *Etv6*. GSEA analysis of the few identified sites that show higher accessibility in *Dido1*ΔE16 LSK cells yielded no relevant results.

To validate the expression data in additional samples by RT-qPCR, we selected several genes in the list of regions with different Tn5 accessibility in LSK-wt and LSK-*Dido1*ΔE16 cells: genes related to the Polycomb group (PcG) proteins (*Bmi1*, *Ezh2*), maintenance and differentiation of stem cells (*Runx1*, *Nanog*, *Ikzf2* (Helios), *Tcf12* (Heb)), differentiation towards lineage B (*Klf3*, *Il7r*, *Ebf1*, *Tcf3* (E2A, E12/E47), *Spi1* (PU.1), *Foxo1*) or towards lineage T (*Gata3*), and cell proliferation (*Mycn* (N-myc), *Myc*, *Cdk6*). The analysis showed that expression of genes related to differentiation towards the B lineage was non-significantly lower in LSK-*Dido1*ΔE16 than in LSK-wt cells. Other genes such as *Gata3* or *Cdk6* were expressed more significantly in LSK-*Dido1*ΔE16 cells, whereas *Nanog* levels were 16 times higher than in LSK-wt cells (Fig. 5D). *Dido1*ΔE16 CLP samples showed reduced *Ebf1* expression, whereas *Nanog* levels were normal.

TLDA analysis of gene expression in B cell precursors.

To analyze expression of several genes related to B cell development in the different precursor subpopulations, we used customized TaqMan Low Density Arrays, tested against preamplified samples

of cDNA synthesized from total RNA from sorted subpopulations. The largest number of differences was detected in the prepro-B subpopulation (Fig. 6A), which suggests that a smaller number of cells is able to follow the B cell pathway at this less-differentiated stage. mRNA levels of transcription factors such as *Tcf3*, *Foxo1*, and *Ebf1* (labeled (1) in Fig. 6A) were decreased in *Dido1* Δ E16 compared to wild type mice. These factors are necessary for precursor differentiation towards the B lineage and are crucial for expression of Pax5, an essential factor for maintenance of B line identity. We also observed lower levels of mRNA encoded by genes related to B cell development and maturation (labeled (2) in Fig. 6A), including *Pax5*, *Igll1*(Igl5), *S1pr1*, *Btk*, *Blk*, *Cd79a* (MB-1, Iga), *Cd79b* (Igb), *Irf4*, *Cr2* (Cd21), *Pou2af1* (OBF1), and *Cxcr4*. In *Dido1* Δ E16 cells, we detected upregulation of genes (labeled (3) in Fig. 6A) such as *Cebpb*, whose relationship with cancer-driven myelopoiesis is observed in humans²⁷, and of *Spn* (Cd43), which is downregulated in the wt cells at this stage²⁸, when V(D)J recombination begins.

As in almost every developmental process, the three “E-proteins” play a fundamental role in hematopoiesis. These transcription factors are characterized by a basic helix-loop-helix (bHLH) domain with capacity to homo- and hetero-dimerize with themselves and other HLH proteins, and to bind to DNA through E-boxes (CACGTG sequences). *Tcf3* (which encodes E2A isoforms E12 and E47) and *Tcf12* (which encodes HEB isoforms HEBcan and HEBalt) have been implicated in B cell development. To test the expression of the distinct isoforms, we designed specific primers for conventional RT-qPCR assays. As already detected in the TLDA assays, *Tcf3* expression was lower in *Dido1* Δ E16 than in wt prepro-B cells; this reduction was due equally to E12 and E47 isoforms (Fig. 6B). However, while the expression of the HEBcan isoform of *Tcf12* (assessed by exon 1/2 levels) was not affected by *DIDO3* absence, HEBalt isoform expression (exon X9 levels) dropped to 10% of the wt in *Dido1* Δ E16 prepro-B cells (Fig. 6C).

Tcf3 levels remained low in *Dido1* Δ E16 pro-B cells (Fig. 6D), and we also found low *Bcl2* and high levels of the cell cycle regulator *Cdkn2a* (p16, cyclin-dependent kinase inhibitor 2A). E2A is necessary to maintain *Ebf1* and *Pax5* expression as well as the B cell genetic program, whereas BCL2 is an antiapoptotic factor, and CDKN2A inhibits the kinase-dependent kinase cyclin 4 (CDK4) and therefore blocks the cell cycle in G1-

In the pre-B *Dido1* Δ E16 population, we again detected a pro-apoptotic environment, with low *Bcl2* and high levels of cyclin D2 (*Ccnd2*), which codes for the regulatory subunit of CDK4 and CDK6, necessary for cell cycle G1/S transition. Our data also showed no silencing of genes that encode pre-BCR components such as *Igll1*, *CD79a*, and *CD79b*, or genes that inhibit pro-B cell differentiation such as *Erg* (Fig. 6E).

In immature B-*Dido1* Δ E16 cells, we detected an increase in the mRNA levels of stemness-related transcription factors (*Pou5f1* (Oct-4), *Nanog*, *Kit*, and *Ly6a* (Sca1)), in addition to high levels of pre-BCR surrogate light chain components such as *Vpreb2* and *Vpreb1*, and of cell cycle regulatory genes such as cyclin-dependent kinase inhibitor *Cdkn2b* (p15INK4b). We also observed overexpression of *Id3*, which forms heterodimers with HLH proteins that lead to inhibition of DNA binding and thus, regulation of gene transcription (Fig. 6F). All these results suggested that *Dido1* Δ E16 affects differentiation processes both in mouse hematopoietic precursors and in embryonic stem cells.

RNA-seq analysis of gene expression in pre-B cells

As the first checkpoint in B cell development is dependent on pre-BCR activation and takes place at the pre-B stage²⁹, where we found the greatest effect of the absence of *DIDO3* (seen as lack of progenitor cell progression), we sequenced the transcriptome (RNA-Seq) of cells in this phase.

Using pooled samples (2-3 mice/pool) to obtain sufficient starting material, we sorted pre-B cells by flow cytometry and purified RNA. Two biological replicates were used to construct libraries for massive sequencing. After filtering low quality reads, an average of 24 million clean reads were obtained and 94.91% clean reads were mapped to the mouse genome. A total of 7846 genes with more than three fragments per kilobase of transcript per million mapped reads (FPKM) was used for subsequent analyses. DEG were defined based on the absolute value of their fold change $|FC| \geq 1.5$ ($\log_2 FC > 0.6$) associated to a false discovery rate (FDR) ≤ 0.05 . We found 120 genes with higher expression in wt pre-B cells, and 346 genes upregulated in *Dido1* Δ E16 cells (Supplementary table 2).

The biological significance of DEG was explored using the network tool STRING v.11.0 (<https://string-db.org>). Gene Ontology (GO) terms with FDR < 0.25 included positive regulation of leukocyte differentiation, cellular response to stimulus, positive regulation of lymphocyte activation, positive regulation of lymphocyte differentiation, positive regulation of T cell differentiation, positive regulation of B-cell differentiation, signal transduction, regulation of signaling, immune response, and regulation of leukocyte cell-cell adhesion. After gene clustering, the highlighted signaling pathways were MTOR, GPCR, Rho-GTPase, interferon, b-catenin, cytokine signaling in the immune system, and antigen processing by ubiquitination and proteasome degradation (Suppl. Fig. 5).

For a second analysis we applied the GSEA tool, using a ranked list of genes with base mean > 1 , ordered by FC. By analysis against the group of genesets defined as targets of specific transcription factors, we identified the genesets of targets for PRC2 components JARID2 and SUZ12 ($p < 0.001$) (Fig. 7). *Aebp2*, a substoichiometric PRC2 subunit that stimulates the K9 and K27 methylation of histone H3³⁰, was upregulated ($\log_2 FC = 2.61$ FDR = 0.067) in *Dido1* Δ E16 pre-B cells.

The relevance of pre-BCR in the pre-B developmental checkpoint and the PRC2 participation in V(D)J recombination led us to study the effect of *DIDO3* deficiency on the transcription of *Igh* and *Igk* loci. Joining (J) and variable (V) regions of the *Igk* locus were not notably affected. There was over-representation of proximal *Igh* V regions (Q52 and 7183 families), whereas intermediate and distal V fragments (mainly J558 and 3609 families) were under-represented in *Dido1* Δ E16 pre-B cells (Fig. 8A and Supplementary table 3). We detected unexpectedly high transcription levels of DNA located between J and D fragments (Fig. 8B). In wt cells, this DNA segment is eliminated by recombination in the previous pro-B stage, or even in prepro-B, by RAG1/2 activity. Finally, we detected expression of the *Igg2b* (*g2b*) gene, which suggested that isotype-switched IgG2b⁺ memory B cells co-purified with pre-B cells (B220⁺CD19⁺CD43⁻IgM⁻) in numbers substantially higher than in wt samples (Fig. 8C).

Discussion

Although DDO3, which is expressed ubiquitously in the organism, presumably has a more general function in determining chromatin structure, its profound influence on the B cell population led us to focus our study on this cell lineage. Speculations on the molecular mechanism involved have been treated elsewhere^{31,32,20,33}. Here we demonstrate that DDO3 plays a role in HSC homing and differentiation, as well as in the later B cell differentiation process.

In lymphopoiesis B, Ebf1 is the key transcription factor for B cell specification and compromise⁵. Ectopic expression of Ebf1 in HSC directs differentiation towards the B lineage at the expense of other lineages³⁴. *Ebf1*^{-/-} mice maintain a developmental blockade of the B lineage in the prepro-B stage and lack mature B cells³⁵. Our mouse model shows partial arrest in B cell development in the prepro- and pro-B populations (B220⁺CD43⁺CD19⁻ and B220⁺CD43⁺CD19⁺, respectively), comparable to the phenotype of *Ebf1*^{+/-} mice. Our RT qPCR assays show that *Ebf1* expression in purified *Dido1*ΔE16 CLP is approximately one-third of that observed in wt CLP cells, where the specification process towards the lymphoid lineage begins. Heterozygotic deletion of *Ebf1* resulted in an impaired response to IL7 *in vitro*³⁶, which provides a possible explanation for the observed stage-specific reduction in cellular expansion. The loss of DDO3 in the LSK state activates *Nanog* transcription, which negatively modulates *Ebf1*. Decreasing EBF1 levels would reduce the potential of CLP-*Dido1*ΔE16 cells for differentiation into the B lineage.

All the canonical E proteins are important for lymphoid development³⁷. HEB (*Tcf12*) and E2 2 (*Tcf4*) are thought to contribute to the total E protein dose needed for B cell development, since the absence of either factor led to a decrease in pro-B cell numbers, but no obvious developmental block³⁸. Of the E proteins, E2A is necessary to maintain *Ebf1*, *Foxo1*, and *Pax5* expression, and hence the B cell genetic program. B cell development in *Tcf3*^{+/-}; *Tcf12*^{-/-} mice is completely blocked at the CLP stage³⁹. It is not known, however, whether this arrest is due to the lack of HEBcan, HEBalt, or both isoforms. In our case, neither is completely absent, but whereas HEBcan levels were normal, HEBalt levels dropped in *Dido1*ΔE16 prepro-B cells to 10% of that in wt, which suggests an undescribed function for this isoform. Lack of E2A function results in an inability to rearrange *Igh* gene segments, which can be rescued with a single dose of E2A⁴⁰. Although it remains to be determined whether this effect is dose-dependent, it could explain our results with *Dido1*ΔE16 pre-B cells, in which a substantial number of RNA reads correspond to the segment between D and J fragments, which was deleted efficiently in wt samples.

We also detected abnormal *Igg2b* expression in *Dido1*ΔE16 pre-B cells. CpG DNA stimulates TLR9, with intense expression of activation-induced cytidine deaminase (AICDA) and production of *Igg2b* germ line transcripts in B cell precursors⁴¹, but we did not find *Aicda* or the TLR9-pathway components overexpression. We thus support the possibility of relatively large numbers of isotype-switched memory B cells⁴² in *Dido1*ΔE16 bone marrow, as were co-purified in our pre-B cell sorting strategy.

In *Dido1*ΔE16 pro-B cells, we found high levels of *Cdkn2a*, which acts as an inhibitor of the kinase-dependent cyclin 4 (CDK4), and low levels of *Bcl2*, involved in blocking cell death (anti-apoptotic). *Cdkn2a* is a known target of EZH2 in several cell types⁴³. *Cdkn2a* codes for p19ARF, which stabilizes p53, as well as p16INK4a, a CDK4 inhibitor⁴⁴, both of which could affect the survival and expansion of adaptive lymphoid cells. CDK4 regulates cell progress in G1 and would be inhibited by excess CDKN2A.

EZH2, a core PRC2 component, is needed to silence *Cdkn2a* in developing B lymphocytes prior to the preBCR checkpoint⁴⁵. The *Cdkn2a*/p53 pathway must be repressed until preBCR selection, but also needs to be accessible to protect developing B cells from oncogenic transformation. Indeed, many early B lineage acute leukemias harbor inactivating mutations at the *Cdkn2a* locus. In contrast, germinal center B cell lymphomas have activating *Ezh2* mutations; in these cells, EZH2 might contribute to tumorigenesis, in part through *Cdkn2a* repression. EZH2-dependent repression of *Cdkn2a* is therefore a critical control point for preBCR selection that could be manipulated by lymphoid cells to promote transformation⁴⁶.

We found significant association of the transcriptional alterations in *Dido1*ΔE16 cells with a number of PRC2 targets and genes bearing H3K27me3 epigenetic marks, which suggests that DIDO3 is involved in determining specific PRC2 activity in the B cell lineage. Moreover, loss of EZH2 is known to cause impaired distal VH–DJH rearrangement and *Igh* locus contraction⁴⁷. There is thus a preference to use proximal V fragments as in our model, which reinforces the view that DIDO3 acts in coordination with PRC2.

We showed differential expression of the DIDO isoforms in MPD¹⁸, and others reported that BCR-ABL1 translocation (Philadelphia chromosome; Ph) in myelocytic leukemia (CML) might be involved in modulating *DIDO1* expression in a JAK2V617F-independent manner^{48,49}. Disorders in multiple signaling pathways and genetic abnormalities combined with the Ph are essential for the evolution of different types of leukemia; however, why MPD evolve specifically into CML, acute myeloid leukemia, acute leukemia leukocytosis, or mixed-phenotype acute leukemia is currently unclear. BCR-ABL1 kinase hyperactivity leads to the activation of signaling pathways and deregulation of cellular processes^{50,51}, and hyperactivation of most of these pathways have been confirmed in CML and ALL mouse models. BCR-ABL1 activity channels mainly through JAK/STAT and PI3K/AKT/mTOR pathways, and through CCAAT/enhancer-binding proteins (C/EBP, which regulate normal myelopoiesis as well as myeloid disorders)⁵². Our Gene Ontology analysis of RNAseq data showed that, in cells lacking DIDO3, the expression of PI3K/mTOR and JAK/STAT pathway components, as well as *Cebpb*, are affected at the transcriptional level. Statistical analysis of DNA methylation showed that target genes of NANOG⁵³, Polycomb-group target genes, and/or genes with an AACTTT promoter motif can have a methylation profile that seems to indicate predisposition to development of cancer, such as breast cancer⁵⁴. Such aberrant DNA methylation profiles can be seen in advance of any signs or clinical symptoms of cancer.

Conclusions

In all, our data provide evidence of the involvement of DDO3, a reader of histone post-translational modifications (PTM), in several processes of bone marrow hematopoiesis, including precursor homing, cell cycle progression, apoptosis, and differentiation. DDO3 deficiency is associated with alterations in DNA accessibility and transcriptional defects of major B cell differentiation master regulators. This association seems to be related, at least in part, to the relationship of DDO3 with PRC2. Remaining open questions include the DNA methylation and the general histone PTM landscape in DDO3-deficient cells, as well as the consequences of DDO3 deficiency in peripheral B cell differentiation and hematopoietic cancer development.

Materials And Methods

Mice

In brief, a mouse strain bearing an exon 16 flanked by *loxP* sites was generated and interbred with Cre^{VAV1}-producing mice heterozygous for exon 16 deletion⁵⁵. Hematopoietic populations were isolated from the offspring. Mice were handled according to national and European Union guidelines, and experiments were approved by the Comité Ético de Experimentación Animal, Centro Nacional de Biotecnología, Consejo Superior de Investigaciones Científicas and by the Comunidad de Madrid (PROEX 055-14).

Flow cytometry

Cells from each sample were diluted in a solution of 0.5% PBS-BSA, 0.065% NaN₃. To avoid nonspecific antibody binding, samples were blocked with Fc-Block (CD16/32; BD Pharmingen). Antibodies used were from eBioscience (B220-eF450, CD3-biotin, CD5-percpCy5.5, CD19-AlexaFluor700, CD93-APC, CD127 (IL7R)-PeCy7, Sca1-APC), Beckman Coulter (B220-biotin, B220-PE, CD4-biotin, CD8-PE, CD8-biotin, cKit-APC, Gr1-biotin, Ly6C-biotin), BioLegend (CD4-PeCy7, CD11b-PeCy7, CD45.2-AlexaFluor700), and BD Pharmingen (CD11b-biotin, CD21-APC, CD23-PE, CD43-biotin, CD45.1-PE, IgG1-PE, IgG3-PE, Ly6G-PE, Nk1.1-APC). Biotinylated antibodies were visualized with streptavidin-eF450 (eBioscience), streptavidin-PE (Southern Biotech) or streptavidin-PE-CF594 (eBioscience). Cell death was monitored using annexin V-APC (Immunostep) and DAPI. In *in vivo* assays, the cell cycle was analyzed by measuring BrdU incorporation into cell DNA 18 h after intraperitoneal injection of 200 mg BrdU (Invitrogen). Cells were labeled following the protocol for the BrdU Staining Kit for flow cytometry (Invitrogen). Analytical cytometric data was acquired using a Gallios cytometer (Beckman Coulter).

Competitive bone marrow transplantation

B57.SJL-Ptprca Pepcb/BoyJ (B6-CD45.1) mice were used as transplant recipients. Recipient cells can be distinguished from hematopoietic cells from donor mice that express the CD45.2 alloantigen (present in the C57BL/6 strain). A group of six CD45.1⁺ mice were irradiated (10 Gy) to inhibit cell division and provoke apoptosis of hematopoietic cells. On day 1 post-irradiation, mice were inoculated retro-orbitally with 150 μ l (3.75 x 10⁶ cells) of a 1:1 mixture of bone marrow cells from CD45.2⁺;EYFP⁺; *Dido1* Δ E16 and

CD45.2⁺;EYFP⁻;Dido1^{+/+} mice (Suppl. Fig. 4). Mice were bled at various times to monitor transplant evolution, and were euthanized at 4 weeks. Bone marrow was extracted for flow cytometry analysis.

***In vitro* differentiation of bone marrow cells**

Bone marrow cells from wt and *Dido1*ΔE16 mice (lin⁻: CD3, CD11b, CD19, CD45R (B220), Ly6G/C (Gr1), TER119), were purified with the EasySep Mouse Hematopoietic Progenitor Cell Isolation kit (STEMCELL Technologies) according to the manufacturer's protocol. Isolated cells were grown in a semi-solid methyl cellulose matrix for the Mouse Hematopoietic CFU Assay in Methocult Media 3630 (STEMCELL Technologies), following the manufacturer's instructions. The medium is supplemented with IL7, which promotes growth of B lymphocyte progenitor cells as pre-B colony-forming units (CFU-pre-B).

RT-qPCR

RNA was extracted from sorted cell samples with the RNeasy Plus Mini Kit (50; Qiagen) according to manufacturer's instructions; quality was determined by electropherogram (Agilent 2100 Bionalyzer). To obtain cDNA, the SuperScriptIII First-Strand Synthesis System for RT-PCR kit (Invitrogen) was used. Up to 1 μg total RNA was used as template in a final reaction of 20 μl. qPCR reactions were carried out 384-well plates in an ABI PRISM 7900HT system (Applied Biosystems) using GoTaq qPCR Master Mix (Promega).

TaqMan Low Density Arrays (TLDA)

Precursor B populations (prepro-B, pro-B, pre-B and immature B cells) were obtained by cell sorting from bone marrow of wt and *Dido1*ΔE16 mice, RNA was extracted from these cell populations as described above, and mRNA was retro-transcribed with SuperScript III (Invitrogen); the resulting cDNA was amplified in TaqMan Universal Master Mix II solution, not UNG (2x) (Applied Biosystems), with a pre-amplification step using TaqMan Custom PreAmp Pool (P/N 4441859; Applied Biosystems). Data were analyzed with ExpressionSuite v. 1.1 (Applied Biosystems). The level of significance for Student's t-test was established at p<0.05.

Assay for transposase-accessible chromatin using sequencing (ATAC-seq)

The LSK cell samples obtained by sorting were processed with the DNA Library Prep Kit (Illumina). The tagged DNA was purified with the DNA Clean ConcentratorTM5 Kit (Zymo Research). The samples were pre-amplified with the Nextera Index kit and the Nextera DNA Library Prep kit (both from Illumina), incorporating index sequences at the ends. DNA from the PCR reaction was purified using the DNA Clean ConcentratorTM5 Kit. To quantify tagged DNA, the 1x dsDNA HS Assay kit (Qubit) was used. Samples were then enriched in DNA sizes between 200 and 600 bp (corresponding to 1 to 3 nucleosomes) using AMPure XP magnetic beads (Beckman-Coulter), following the manufacturer's guidelines. To analyze size distribution, DNA analysis was performed by electropherogram as above. To sequence tagged DNA, a mixture was used of libraries with a balanced proportion of three samples each of LSK-wt cells and LSK-*Dido1*ΔE16 cells, each with ~100,000 pmol/l tagged DNA fragments. After samples were sequenced,

fragment size distribution was determined to have maxima corresponding to the lengths of two nucleosomes (~350 bp), one nucleosome (~180 bp), and <1 nucleosome, the last compatible with that expected for naked DNA. Massive sequencing was performed by the CNIC Genomics Service on a HiSeq 2500 platform using a Hi-Seq-Rapid PE Flowcell 2x50 kit.

Sequencing data processing: FASTQ files with the paired-end sequences were aligned to the UCSC mm10 mouse reference genome (genome.ucsc.edu) using Bowtie2 (bowtie-bio.sourceforge.net/bowtie2/) and BWA-MEM 0.7.15 (bio-bwa.sourceforge.net) methods. Sequences aligned in blacklisted regions (sites.google.com/site/anshulkundaje/projects/blacklists) were removed with the Intersect tool from BEDTools (github.com/ark5xfbedtools2) before calculating accessible regions of chromatin using the MACS2 program (github.com/taoliu/MACS/tree/master/MACS2) used in the ENCODE project pipeline (www.encodeproject.org). In parallel, the HOMER tool (homer.ucsd.edu/homer/) was also used to calculate accessible regions of chromatin and to find differentially accessible regions between LSK-wt and LSK-*Dido1* Δ E16 cells (see below). The overlap between accessible regions identified by MACS2 and HOMER was 80-90% (see Supplementary Table 4).

Genomic annotations: The data were obtained, processed, and annotated using tools from R (www.r-project.org/) and Bioconductor (bioconductor.org). To determine the equivalent accessible regions in the replicas, the Intersect tool and the files produced by MACS2 were used. Genes, promoters, UTRs, introns, and exons were allocated to the accessible regions with the CHIPseeker program (github.com/GuangchuangYu/CHIPseeker).

Differential accessibility analysis of chromatin: The sequences (readings) in the accessible and equivalent regions of all replicates were counted and normalized with the StringTie 1.3.3 method (ccb.jhu.edu/software/stringtie). To calculate the differential accessibility of chromatin, the accessible and equivalent regions were analyzed with at least one reading in all replicas, with the DESeq2 (bioconductor.org/packages/DESeq2) and edgeR methods (bioconductor.org/packages/edgeR). Accessible regions with a fold change value >2 and a false discovery rate (FDR) <0.05 were considered statistically significant. We also used the HOMER tool in the differential accessibility analysis of chromatin. Gene Set Enrichment Analysis (GSEA, www.gsea-msigdb.org/) was used for the ATAC-seq data.

RNA-Seq

Pre-B cells extracted from the bone marrow of wt and *Dido1* Δ E16 mice were isolated by sorting. RNA was purified using the RNeasy mini plus kit (Qiagen) and massive sequencing was performed at BGI Genomics (Hong Kong) using the HiSeq 2000 platform.

RNA-seq data analysis. Sequenced reads were processed in parallel using the Galaxy RNA-seq pipeline (usegalaxy.org) and a combined approach with BWA-MEM 0.7.15, StringTie (ccb.jhu.edu/software/stringtie/) and edgeR. The combined approach was previously used to process RNA-seq data of *Dido1* Δ E16 in embryonic stem cells (GEO: GSE152346) Differentially expressed genes

(DEG) were defined according to the absolute value of their log₂ (fold change) ≥ 0.6 and FDR < 0.05 , unless otherwise stated. Gene ontology (GO), pathway annotation, and enrichment analyses were carried out using the Molecular Signatures Database (MSigDB) (go-basic.obo, downloaded Jan 15, 2020 by gsea-msigdb.org). A Venn diagram was constructed to obtain co-expressed DEG among samples using the webtool at <http://bioinformatics.psb.ugent.be/webtools/Venn/>, provided by Bioinformatics & Evolutionary Genomics (Gent, BE). The biological significance of DEG was explored using STRING v. 11.0 (string-db.org).

Declarations

Ethics approval

All animal experiments were designed in compliance with European Union legislation and approved by the Committee for Ethics in Animal Experimentation of the Centro Nacional de Biotecnología (CNB/CSIC) (Proex 322/15).

Consent to participate

Not applicable.

Availability of data and materials

ATAC-seq and RNA-seq data are deposited in the Gene Expression Omnibus under accession code GSE157228. Private access token: gpunwiqctudnib

Competing interests

The authors declare no competing interests.

Funding

This work was supported by grants from the Spanish Ministerio de Ciencia e Innovación (PID2019-110574RB-I00 and SAF2016-75456-R) and the Fundación Ramón Areces (XVIII Concurso Nacional para la Adjudicación de Ayudas a la Investigación en Ciencias de la Vida y de la Materia).

Authors' contributions

Conceptualization: CMA, RV

Writing: FGB, CMA, RV

Funding acquisition: CMA, RV

Formal analysis: FGB, TP, EVdL, RV

Data curation: TP, EVdL

Methodology: FGB, RV

Supervision: RV

Acknowledgements

We thank the Genomics Unit and the Service of Cardiovascular Physiology of the Animal Facility of the Spanish National Center for Cardiovascular Research (CNIC) for ATAC-seq and complete blood count processing, respectively, Dr María López-Bravo for B6-CD45.1 mice, the CNB facility for flow cytometry, and Catherine Mark for editorial assistance.

References

1. Inlay MA, Bhattacharya D, Sahoo D, Serwold T, Seita J, Karsunky H et al., Ly6d marks the earliest stage of B-cell specification and identifies the branchpoint between B-cell and T-cell development. *Genes and Development* **23**:2376-2381 (2009). doi:10.1101/gad.1836009.
2. Pang SHM, de Graaf CA, Hilton DJ, Huntington ND, Carotta S, Wu L et al., PU.1 is required for the developmental progression of multipotent progenitors to common lymphoid progenitors. *Frontiers in Immunology* **9**:1264 (2018). doi:10.3389/fimmu.2018.01264.
3. Sellars M, Kastner P, Chan S, Ikaros in B cell development and function. *World J. Biol. Chem.* **2**(6):132-139(2011). doi:10.4331/wjbc.v2.i6.132
4. Lin YC, Jhunjhunwala S, Benner C, Heinz S, Welinder E, Mansson R et al. A global network of transcription factors, involving E2A, EBF1 and Foxo1, that orchestrates B cell fate. *Nat. Immunol.* **11**:635-43 (2010). doi: 10.1038/ni.1891.
5. Treiber T, Mandel EM, Pott S, Györy I, Firner S, Liu ET et al., Early B cell factor 1 regulates B cell gene networks by activation, repression, and transcription-independent poising of chromatin. *Immunity.* **32**(5):714-25 (2010). doi: 10.1016/j.immuni.2010.04.013. Epub 2010 May 6. PMID: 20451411.
6. Nutt SL, Heavey B, Rolink AG, Busslinger M, Commitment to the B-lymphoid lineage depends on the transcription factor Pax5. *Nature.* **401**(6753):556-62 (1999). doi: 10.1038/44076. PMID: 10524622
7. Allman D, Sambandam A, Kim S, Miller JP, Pagan A, Well D et al., Thymopoiesis independent of common lymphoid progenitors. *Nat. Immunol.* **4**(2):168-74 (2003). doi: 10.1038/ni878. Epub 2003 Jan 6. PMID: 12514733. .
8. Kondo M, Weissman IL, Akashi K, Identification of clonogenic common lymphoid progenitors in mouse bone marrow. *Cell.* **91**:661-72 (1997). doi: 10.1016/s0092-8674(00)80453-5. PMID: 9393859.
9. Manz MG, Traver D, Miyamoto T, Weissman IL, Akashi K, Dendritic cell potentials of early lymphoid and myeloid progenitors. *Blood.* **97**:3333-41 (2001). doi: 10.1182/blood.v97.11.3333. PMID: 11369621.

10. Rumfelt LL, Zhou Y, Rowley BM, Shinton SA, Hardy RR, Lineage specification and plasticity in CD19-early B cell precursors. *J. Exp. Med.* **203**:675-87 (2006). doi: 10.1084/jem.20052444. Epub 2006 Feb 27. PMID: 16505143; PMCID: PMC2118241.
11. Fleming HE, Paige CJ, Pre-B cell receptor signaling mediates selective response to IL-7 at the pro-B to pre-B cell transition via an ERK/MAP kinase-dependent pathway. *Immunity* **15**:521-31 (2001). doi: 10.1016/s1074-7613(01)00216-3. PMID: 11672535.
12. Bertolino E, Reddy K, Medina KL, Parganas E, Ihle J, Singh H, Regulation of interleukin 7-dependent immunoglobulin heavy-chain variable gene rearrangements by transcription factor STAT5. *Nat. Immunol.* **6**:836-43 (2005). doi: 10.1038/ni1226. Epub 2005 Jul 17. PMID: 16025120.
13. Lutz J, Heideman MR, Roth E, van den Berk P, Müller W, Raman C, et al., Pro-B cells sense productive immunoglobulin heavy chain rearrangement irrespective of polypeptide production. *Proc. Natl. Acad. Sci. U.S.A.* **108** 10644-10649 (2011). doi.org/10.1073/pnas.1019224108
14. Berdasco M, Esteller M, Aberrant epigenetic landscape in cancer: how cellular identity goes awry. *Dev. Cell* **16**;19(5):698-711 (2010). doi: 10.1016/j.devcel.2010.10.005. PMID: 21074720. .
15. Petell CJ, Pham AT, Skela J, Strahl BD, Improved methods for the detection of histone interactions with peptide microarrays. *Sci. Rep.* **9**, 6265 (2019). <https://doi.org/10.1038/s41598-019-42711-y>.
16. Eberl HC, Spruijt CG, Kelstrup CD, Vermeulen M, Mann M, A Map of General and Specialized Chromatin Readers in Mouse Tissues Generated by Label-free Interaction Proteomics. *Mol. Cell* **49**, 368–378 (2013). doi: 10.1016/j.molcel.2012.10.026. PMID: 23201125
17. Guerrero AA, Gamero MC, Trachana V, Fütterer A, Pacios-Bras C, Díaz-Concha NP et al., Centromere-localized breaks indicate the generation of DNA damage by the mitotic spindle. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 4159–4164 (2010). doi.org/10.1073/pnas.0912143106.
18. Fütterer A, Campanero MR, Leonardo E, Criado LM, Flores JM, Hernández JM et al., Dido gene expression alterations are implicated in the induction of hematological myeloid neoplasms. *J Clin Invest* **115**, 2351–2362 (2005). doi: 10.1172/JCI24177. PMID: 16127461; PMCID: PMC1190370.
19. Fütterer A, Raya A, Llorente M, Izpisúa-Belmonte JC, de la Pompa JL, Klatt P et al., Ablation of Dido3 compromises lineage commitment of stem cells in vitro and during early embryonic development. *Cell Death Differ.* **19**, 132–143 (2012). doi: 10.1038/cdd.2011.62.
20. Mora Gallardo C, Sánchez de Diego A, Gutiérrez Hernández J, Talavera-Gutiérrez A, Fischer T, Martínez-A C et al., Dido3-dependent SFPQ recruitment maintains efficiency in mammalian alternative splicing. *Nucleic Acids Research* **47**, 5381-94 (2019) doi:10.1093/nar/gkz235.
21. Bosco N, Swee LK, Bénard A, Ceredig R, Rolink A, Auto-reconstitution of the T-cell compartment by radioresistant hematopoietic cells following lethal irradiation and bone marrow transplantation. *Exp. Hematol.* **38**, 222-232.e2 (2010). doi: 10.1016/j.exphem.2009.12.006. PMID: 20045443.
22. Corces MR, Trevino AE, Hamilton EG, Greenside PG, Sinnott-Armstrong NA, Vesuna S et al., An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nature Methods* **14**, 959–962 (2017). doi:10.1038/nmeth.4396.

23. Yu G, Wang LG, He QY, ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* **31**, 2382–2383 (2015). doi:10.1093/bioinformatics/btv145. PMID: 25765347.
24. Engelen E, Brandsma JH, Moen MJ, Signorile L, Dekkers DH, Demmers J et al., Proteins that bind regulatory regions identified by histone modification chromatin immunoprecipitations and mass spectrometry. *Nature Communications* (2015) doi:10.1038/ncomms8155.
25. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA et al., Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15545–15550 (2005). doi:10.1073/pnas.0506580102.
26. Fütterer A, de Celis J, Navajas R, Almonacid L, Gutiérrez J, Talavera-Gutiérrez A et al., DIDO as a Switchboard that Regulates Self-Renewal and Differentiation in Embryonic Stem Cells. *Stem Cell Reports* **8**, 1062–1075 (2017). doi:10.1016/j.stemcr.2017.02.013.
27. Hirai H, Yokota A, Tamura A, Sato A, Maekawa T, Non-steady-state hematopoiesis regulated by the C/EBP β transcription factor. *Cancer Sci.* **106**, 797–802 (2015). doi:10.1111/cas.12690
28. Rolink A, Grawunder U, Winkler TH, Karasuyama H, Melchers F, IL-2 receptor α chain (CD25JAC) expression defines a crucial stage in pre-B cell development. *Int Immunol* **6**, 1257–1264 (1994). doi:10.1093/intimm/6.8.1257. PMID: 7526894.
29. von Boehmer H, Melchers F, Checkpoints in lymphocyte development and autoimmune disease. *Nature Immunology* **11**, 14–20 (2010). doi: 10.1038/ni.1794
30. Grijzenhout A, Godwin J, Koseki H, Gdula MR, Szumska D, McGouran JF et al., Functional analysis of AEBP2, a PRC2 Polycomb protein, reveals a Trithorax phenotype in embryonic development and in ESCs. *Development* **143**, 2716–2723 (2016). doi: 10.1242/dev.123935. PMID: 27317809. PMCID: PMC5004903.
31. Pacios-Bras C, Pons T, Alonso-Guerrero A, Román-García S, Gutiérrez J, Manrique P et al., DIDO3 regulates microtubule stability at transcriptional and posttranslational levels and is needed for fibroblast adhesion and movement. *SSRN Electronic Journal* (2019) doi:10.2139/ssrn.3428482.
32. Gatchalian J, Fütterer A, Rothbart SB, Tong Q, Rincon-Arano H, Sánchez de Diego A et al., Dido3 PHD modulates cell differentiation and division. *Cell Rep.* **4**, 148–158 (2013). doi: 10.1016/j.celrep.2013.06.014. PMID: 23831028. PMCID: PMC3967714.
33. Sánchez de Diego A, Alonso Guerrero A, Martínez-A C, van Wely KH, Dido3-dependent HDAC6 targeting controls cilium size. *Nat. Commun.* **5**, 3500–3500 (2014). doi: 10.1038/ncomms4500
34. Zhang Z, Cotta CV, Stephan RP, deGuzman CG, Klug CA, Enforced expression of EBF in hematopoietic stem cells restricts lymphopoiesis to the B cell lineage. *EMBO J.* (2003) doi:10.1093/emboj/cdg464.
35. Lin H, Grosschedl R, Failure of B–cell differentiation in mice lacking the transcription factor EBF. *Nature* **376**, 263–267 (1995). doi: 10.1038/376263a0
36. Åhsberg J, Ungerback J, Strid T, Welinder E, Stjernberg J, Larsson M, Early B-cell factor 1 regulates the expansion of B-cell progenitors in a dose-dependent manner. *J. Biol. Chem.* **288**, 33449–33461 (2013). doi: 10.1074/jbc.M113.506261. Epub 2013 Sep 27. PMID: 24078629; PMCID: PMC3829190.

37. Kee BL, E and ID proteins branch out. *Nat. Rev. Immunol.* (2009) doi:10.1038/nri2507.
38. Zhuang Y, Cheng P, Weintraub H, B-lymphocyte development is regulated by the combined dosage of three basic helix-loop-helix genes, E2A, E2-2, and HEB. *Mol. Cell. Biol.* **16**:2898-905 (1996) doi:10.1128/mcb.16.6.2898. PMID: 8649400; PMCID: PMC231283.
39. Welinder E, Mansson R, Mercer EM, Bryder D, Sigvardsson M, Murre C, The transcription factors E2A and HEB act in concert to induce the expression of FOXO1 in the common lymphoid progenitor. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 17402–17407 (2011). doi: 10.1073/pnas.1111766108.
40. Bain G, Maandag EC, Izon DJ, Amsen D, Kruisbeek AM, Weintraub BC et al., E2A proteins are required for proper B cell development and initiation of immunoglobulin gene rearrangements. *Cell* **79**, 885–892 (1994).doi: 10.1016/0092-8674(94)90077-9.
41. Edry E, Azulay-Debby H, Melamed D, TOLL-like receptor ligands stimulate aberrant class switch recombination in early B cell precursors. *Int. Immunol.* **20**:1575-85 (2008) doi:10.1093/intimm/dxn117. Epub 2008 Oct 29. PMID: 18974086.
42. Riedel R, Addo R, Ferreira-Gomes M, Heinz GA, Heinrich F, Kummer J, et al., Discrete populations of isotype-switched memory B lymphocytes are maintained in murine spleen and bone marrow. *Nat. Commun.* **11**, 2570 (2020). doi: 10.1038/s41467-020-16464-6.
43. Velichutina I, Shaknovich R, Geng H, Johnson NA, Gascoyne RD, Melnick AM, et al., EZH2-mediated epigenetic silencing in germinal center B cells contributes to proliferation and lymphomagenesis. *Blood* **116**, 5247–5255 (2010). doi: 10.1182/blood-2010-04-280149.
44. McConnell BB, Gregory FJ, Stott FJ, Hara E, Peters G, Induced Expression of p16INK4a Inhibits Both CDK4- and CDK2-Associated Kinase Activity by Reassortment of Cyclin-CDK-Inhibitor Complexes. *Mol. Cell. Biol.* **19**, 1981–1989 (1999). doi: 10.1128/mcb.19.3.1981. PMCID: PMC83991. PMID: 10022885.
45. Jacobsen JA, Woodard J, Mandal M, Clark MR, Bartom ET, Sigvardsson M et al., EZH2 regulates the developmental timing of effectors of the pre-antigen receptor checkpoints. *J. Immunol.* **198**, 4682–4691 (2017). doi: 10.4049/jimmunol.1700319.
46. Swaminathan S, Duy C, Müschen M, BACH2-BCL6 balance regulates selection at the pre-B cell receptor checkpoint. *Trends Immuno.* **35**, 131–137 (2014). doi: 10.1016/j.it.2013.11.002. PMID: 24332591. PMCID: PMC3943645.
47. Su IH, Basavaraj A, Krutchinsky AN, Hobert O, Ullrich A, Chait BT et al., Ezh2 controls B cell development through histone H3 methylation and Igh rearrangement. *Nat. Immunol.* **4**, 124–131 (2003). doi: 10.1038/ni876. PMID: 12496962
48. Berzoti-Coelho MG, Ferreira AF, de Souza Nunes N, Pinto MT, Júnior MC, Simões BP et al., The expression of Death Inducer-Obliterator (DIDO) variants in Myeloproliferative Neoplasms. *Blood Cells Mol. Dis.* **59**, 25–30 (2016). doi: 10.1016/j.bcmed.2016.03.008. Epub 2016 Apr 8.
49. Berzoti-Coelho MG, Ferreira AF, de Souza Nunes N, Pinto MT, Júnior MC, Simões BP et al., Corrigendum to “The expression of Death Inducer-Obliterator (DIDO) variants in Myeloproliferative

- Neoplasms” [Blood Cells Mol. Dis. **59** (2016) 25–30] (S1079979616300225) (10.1016/j.bcmed.2016.03.008)). *Blood Cells Mol. Dis.*(2018) doi:10.1016/j.bcmed.2017.08.013.
50. Kumar S, Wuerffel R, Achour I, Lajoie B, Sen R, Dekker J et al., Flexible ordering of antibody class switch and V(D)J joining during B-cell ontogeny. *Genes Dev.* **27**, 2439–44 (2013). doi: 10.1101/gad.227165.113. PMID: 24240234. PMCID: PMC3841733.
51. Sattler M, Griffin JD, Molecular mechanisms of transformation by the BCR-ABL oncogene. *Semin. Hematol.* **40**, suppl. 2, 4–10 (2003). doi: 10.1053/shem.2003.50034.
52. Kang ZJ, Liu YF, Xu LZ, Long ZJ, Huang D, Yang Y et al., The Philadelphia chromosome in leukemogenesis. *Chin. J. Cancer* **35**, (2016). doi: 10.1186/s40880-016-0108-0. PMID: 27233483; PMCID: PMC4896164.
53. Tommasi S, Zheng A, Yoon JI, Besaratinia A, Epigenetic targeting of the Nanog pathway and signaling networks during chemical carcinogenesis. *Carcinogenesis* **35**, 1726-36 (2014) doi:10.1093/carcin/bgu026.
54. Widschwendter M, Teschendorff A, Jacobs I, Method for predicting risk of developing cancer - Google Patents WO2012104642A1 (2012).
55. Fütterer, A., Talavera-Gutiérrez, A., Pons, T., de Celis, J., Gutiérrez, J., Domínguez Plaza, V., & Martínez-A, C. Impaired stem cell differentiation and somatic cell reprogramming in DIDO3 mutants with altered RNA processing and increased R-loop levels. *Cell death & disease*, 12(7), 637 (2021). doi: 10.1038/s41419-021-03906-2

Figures

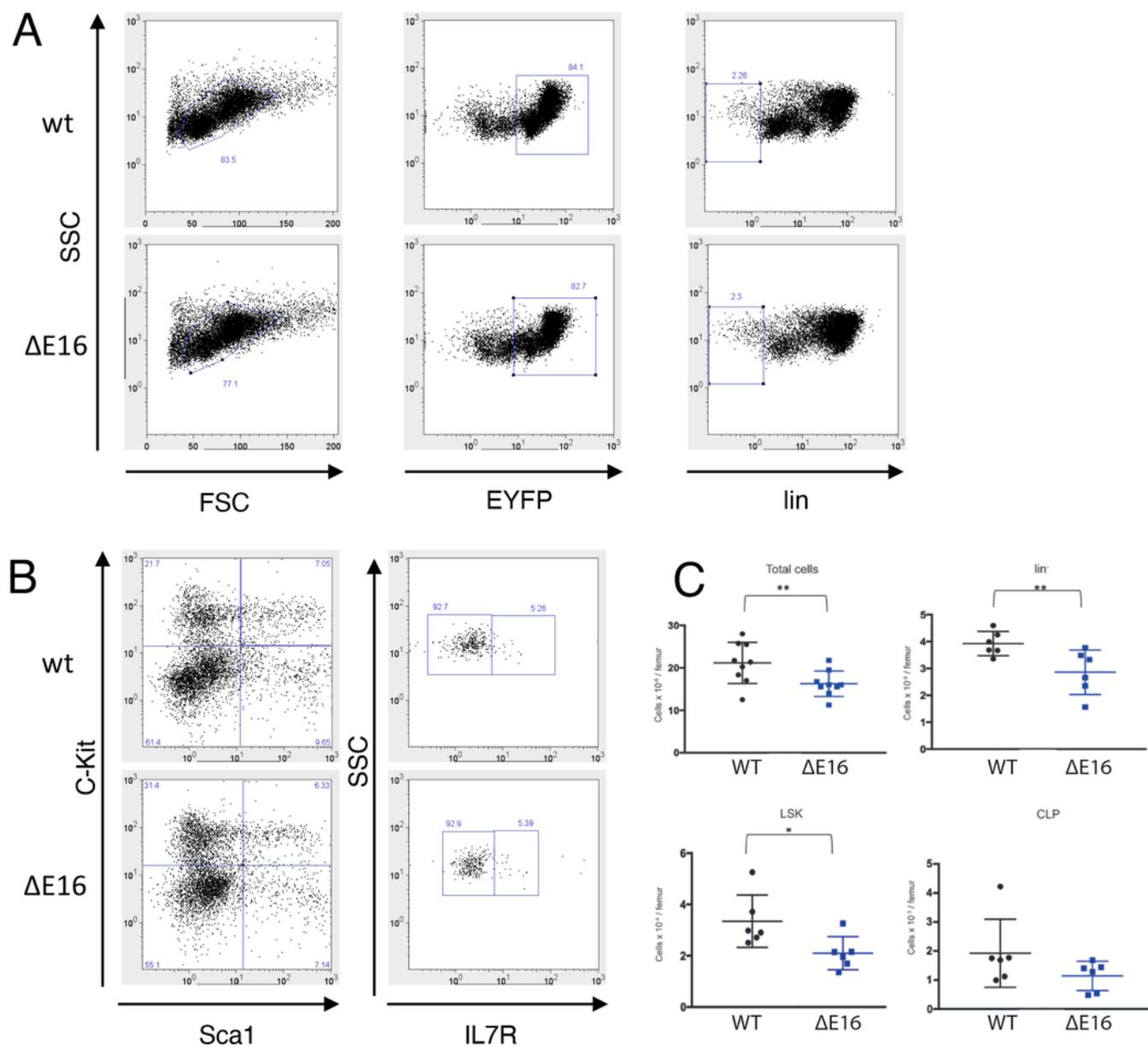


Figure 1

Figure 1

Characterization of hematopoietic progenitor cell populations. A. Total bone marrow cells in wt and *Dido1*ΔE16 mice were selected by EYFP⁺ and *lin*⁻ (Gr1⁻ Ter119⁻CD11b⁻CD3⁻B220⁻) phenotype. B. *Lin*⁻ cells were further analyzed for Sca1, c-Kit, and IL7R expression to characterize LSK (Sca1+cKit+IL7R⁻) and CLP (Sca1+cKit+IL7R⁺) populations. C. Quantification of total (wt n = 9, *Dido1*ΔE16 n = 9), *lin*⁻, LSK and CLP cells (wt n = 6, *Dido1*ΔE16 n = 6) per single femur (t-test ** p<0.01, * p<0.05). Bars, mean ± SD.

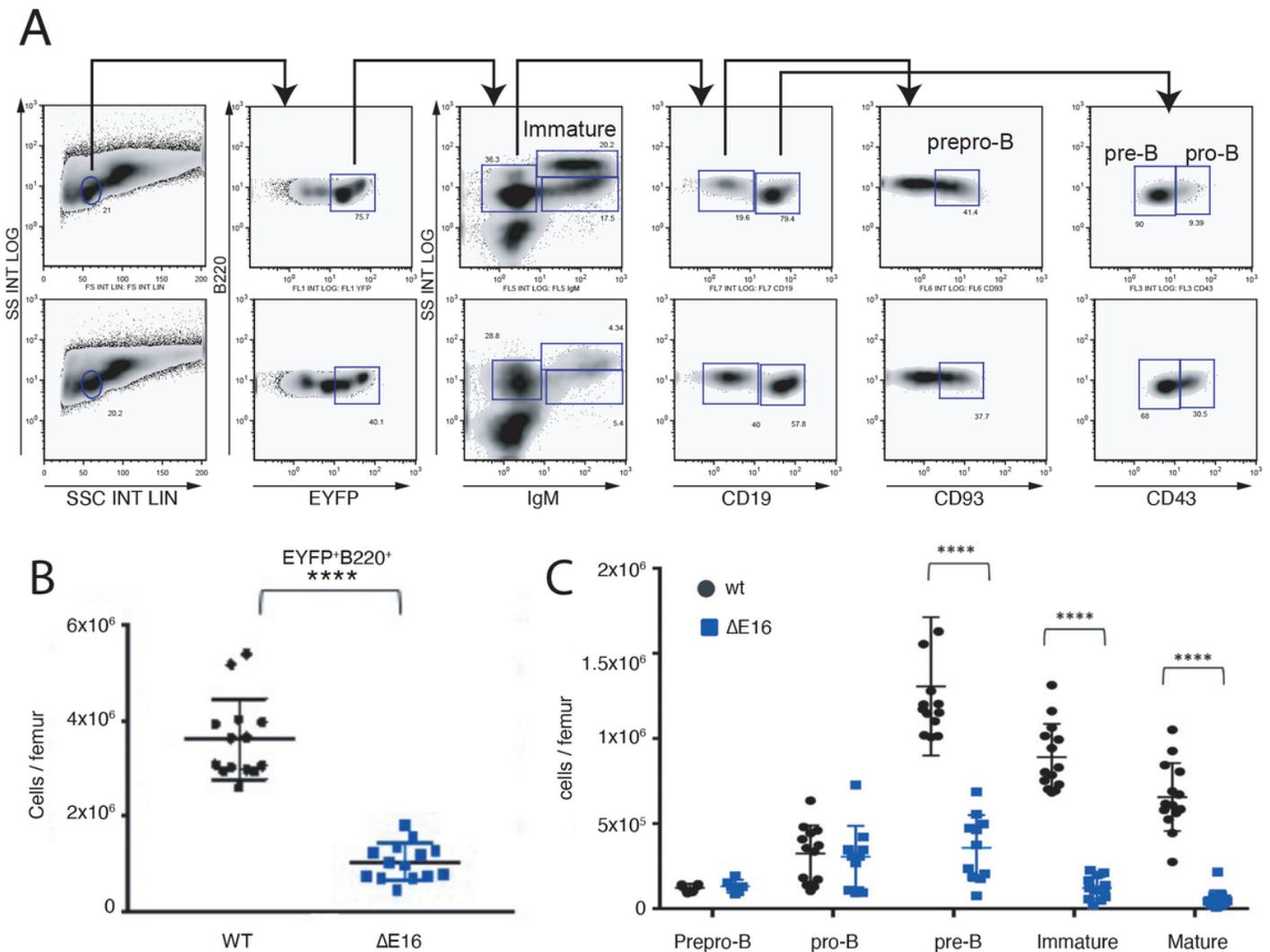


Figure 2

Characterization strategy by flow cytometry for lymphoid progenitor populations in mouse bone marrow. A. Total cells in wt and Dido1 $\Delta E16$ mice cells were selected in FS/SS plot. EYFP⁺ cells were gated by IgM and B220 markers to distinguish populations of mature EYFP⁺B220⁺IgM^{high} cells and immature EYFP⁺B220⁺IgM^{low} cells. Of the EYFP⁺B220⁺IgM^{low} subpopulation, prepro-B cells were identified as CD19⁻CD93⁺, pre-B as CD19⁺CD43⁻, and pro-B as CD19⁺CD43⁺. B. Numbers of EYFP⁺B220⁺ cells per femur in wt compared to Dido1 $\Delta E16$ mice (n = 14; ****, p<0.0001). C. Total bone marrow B precursors and mature recirculating B cells isolated from a single femur. Prepro-B (n = 6); pro-B, pre-B (wt n = 13, Dido1 $\Delta E16$ n = 11); immature and mature B cells (wt n = 14, Dido1 $\Delta E16$ n = 13). B, C. Mean \pm SD are shown, t-test **** p<0.0001.

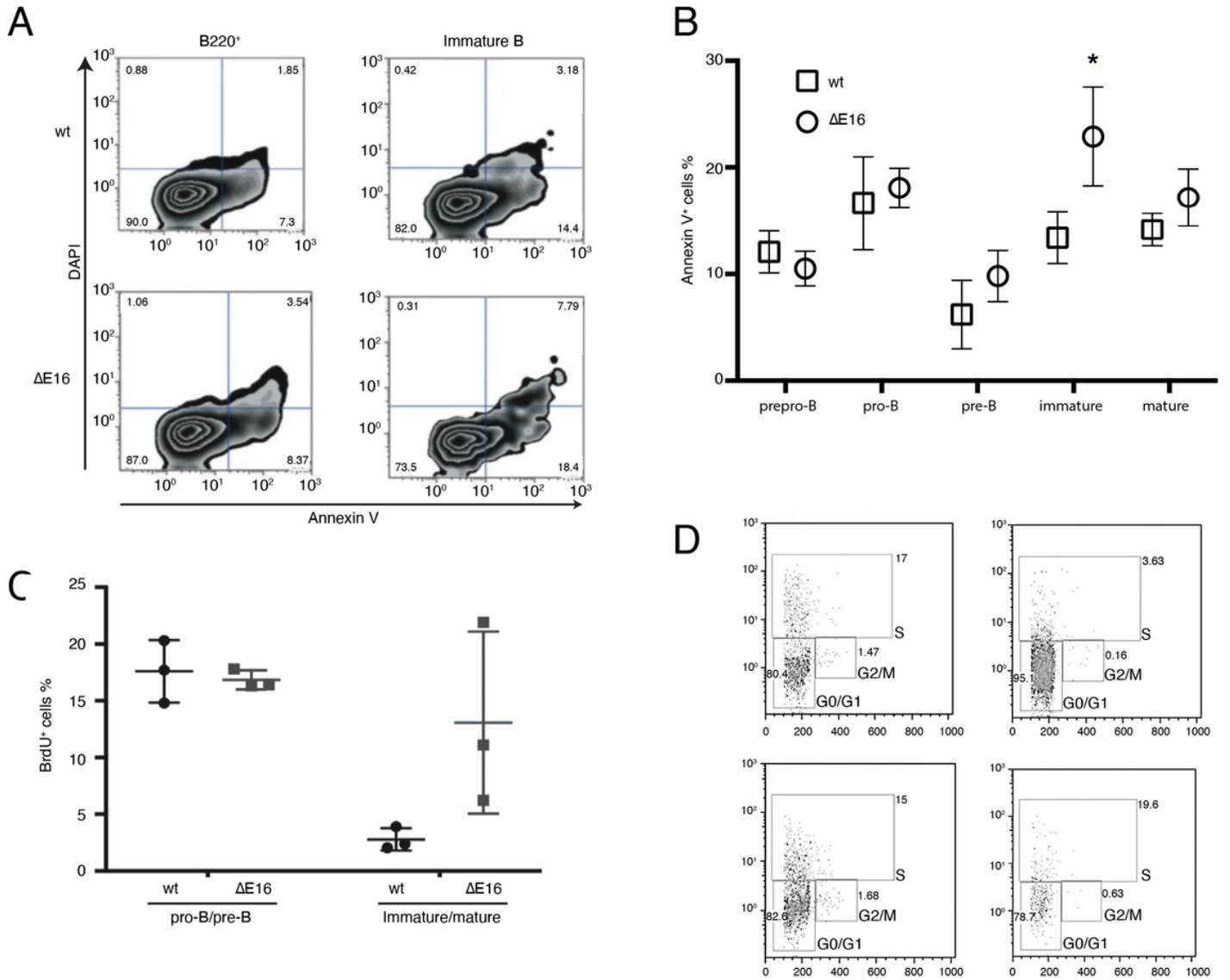


Figure 3

Effect of lack of Dido3 on apoptosis and proliferation. A. Representative graphs of annexin V binding total B220⁺ and immature B cells from bone marrow of wt and Dido1 $\Delta E16$ mice. Lower left, viable cells. Lower right, cells in early apoptosis (which bind annexin V and exclude DAPI). Upper right, cells in late apoptosis (which bind annexin V and incorporate DAPI). Upper left, necrotic cells. B. Mean percentages of annexin V⁺ cells in different precursor subpopulations (mean \pm SD, n = 3, t-test * p<0.05). C. Percentages (mean \pm SD; n = 3) of cells incorporating BrdU in pro-B/pre-B and immature/mature B cell populations, from bone marrow of wt and Dido1 $\Delta E16$ mice 18 h after intraperitoneal BrdU injection. D. Representative flow cytometry analysis of the cell cycle in pro-B/pre-B precursors (EYFP+B220⁺CD19⁺IgM⁺) and immature/mature B cells (EYFP+B220⁺CD19⁺IgM⁺) from bone marrow of BrdU-treated WT and Dido1 $\Delta E16$ mice.

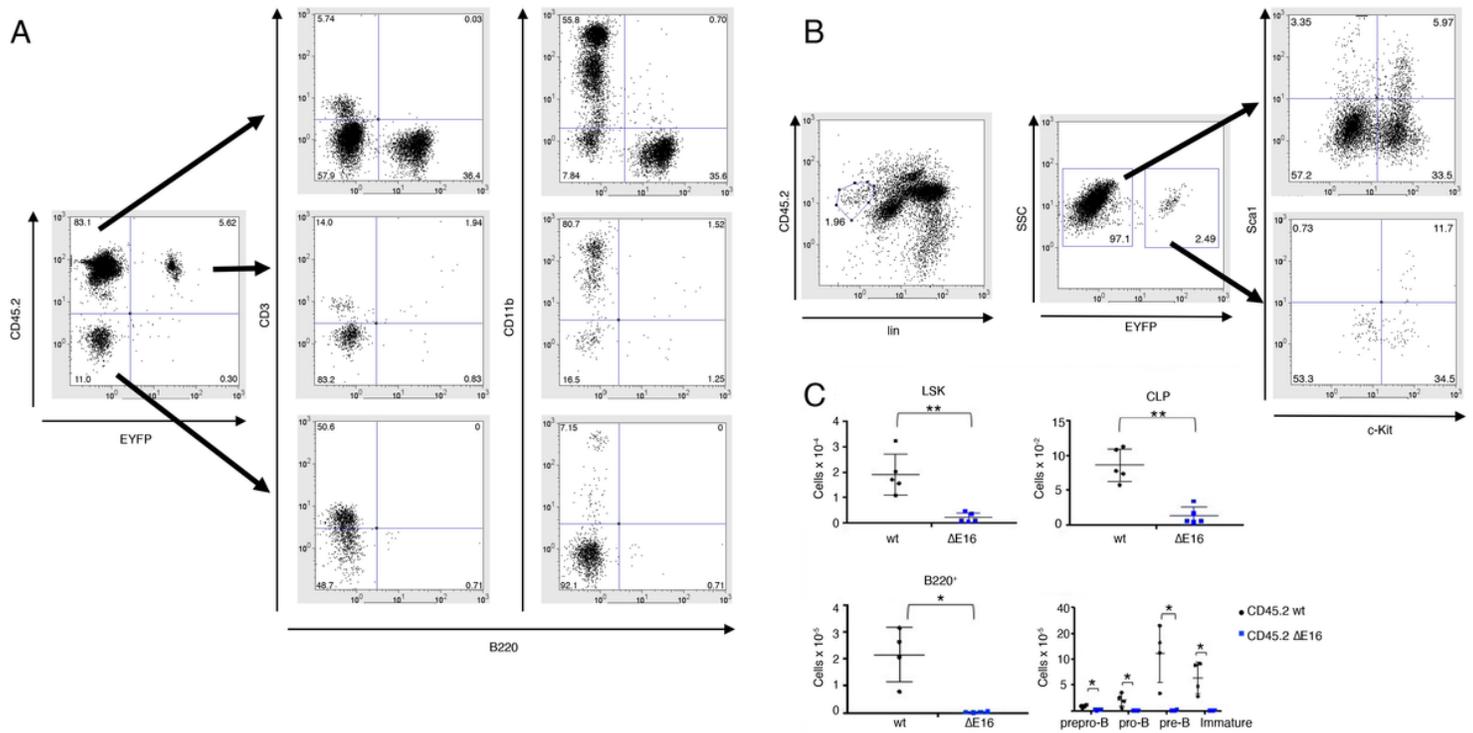


Figure 4

Effect of DIDO3 deficiency on bone marrow transplantation. A. Characterization of peripheral blood white cells in lethally irradiated mice at 21 days post-transplant; total CD45.1+ (host) and CD45.2+ (donor) cells in blood. Dido1 ΔE16 cells were identified by EYFP expression. B. Representative analysis of flow cytometry data to quantify CD45.1+ (host) and CD45.2+ (Dido1 ΔE16 (EYFP+) and wt) populations of *lin*⁺ cells in transplanted mice. C. Total numbers of hematopoietic precursors per femur recovered from transplanted mice 40 days post-transplant. Student's t-tests compare CD45.2 wt and CD45.2⁺ EYFP⁺ Dido1 ΔE16 LSK and CLP cells (** p < 0.01; n = 5), and cells from different B cell precursor populations (*, p < 0.05; n = 4).

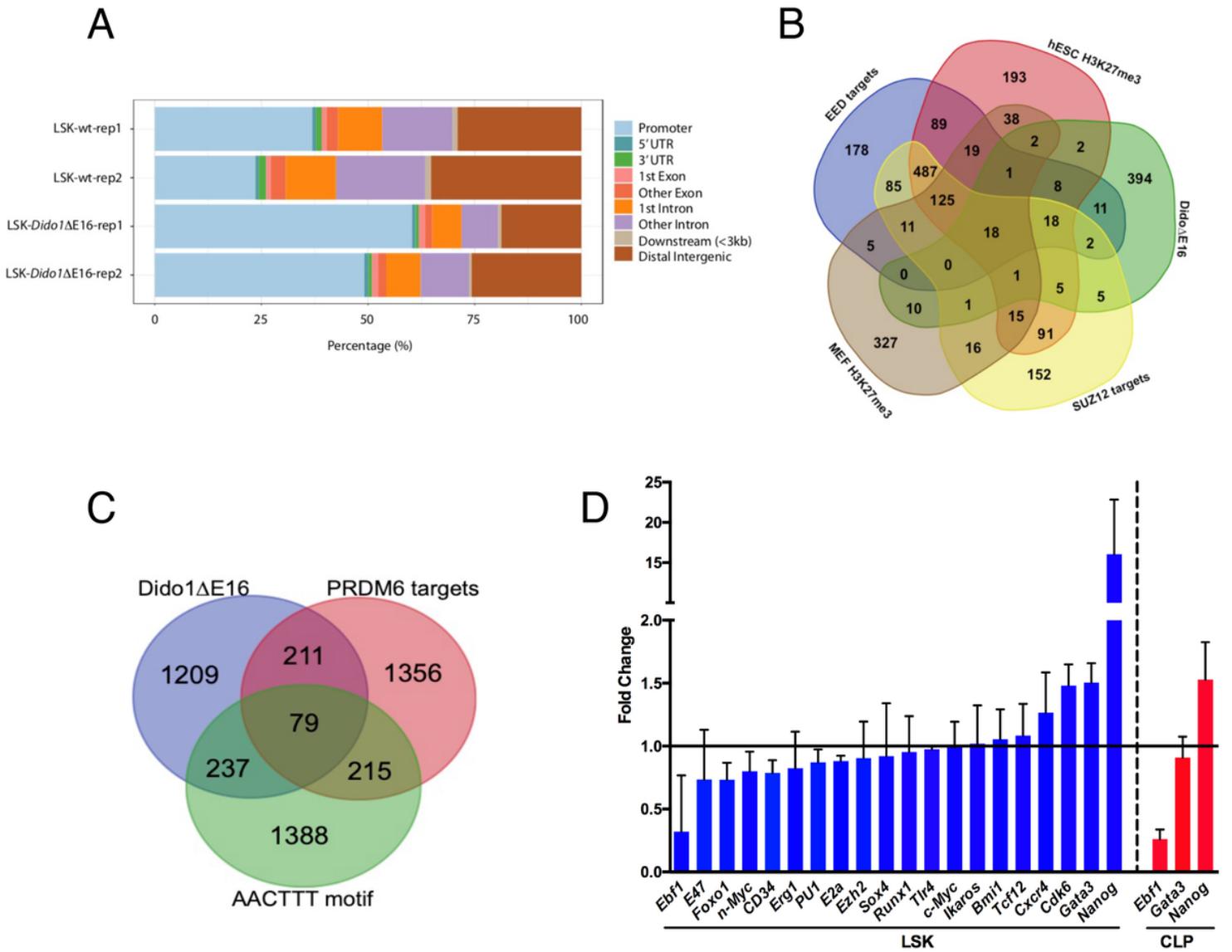


Figure 5

Effect of DIDO3 deficiency on chromatin accessibility. A. Distribution by distance to transcription start sites of chromatin open in wt and closed in *Dido1* Δ E16 LSK cells, as determined by ATAC-seq (mean values, $n = 3$). B. Venn diagrams show overlaps in promoter regions of differentially accessible chromatin with genesets associated with PRC2 activity. C. Overlap of differentially accessible chromatin in intronic and intergenic regions. D. Validation by real-time qPCR of the expression of potentially affected genes.

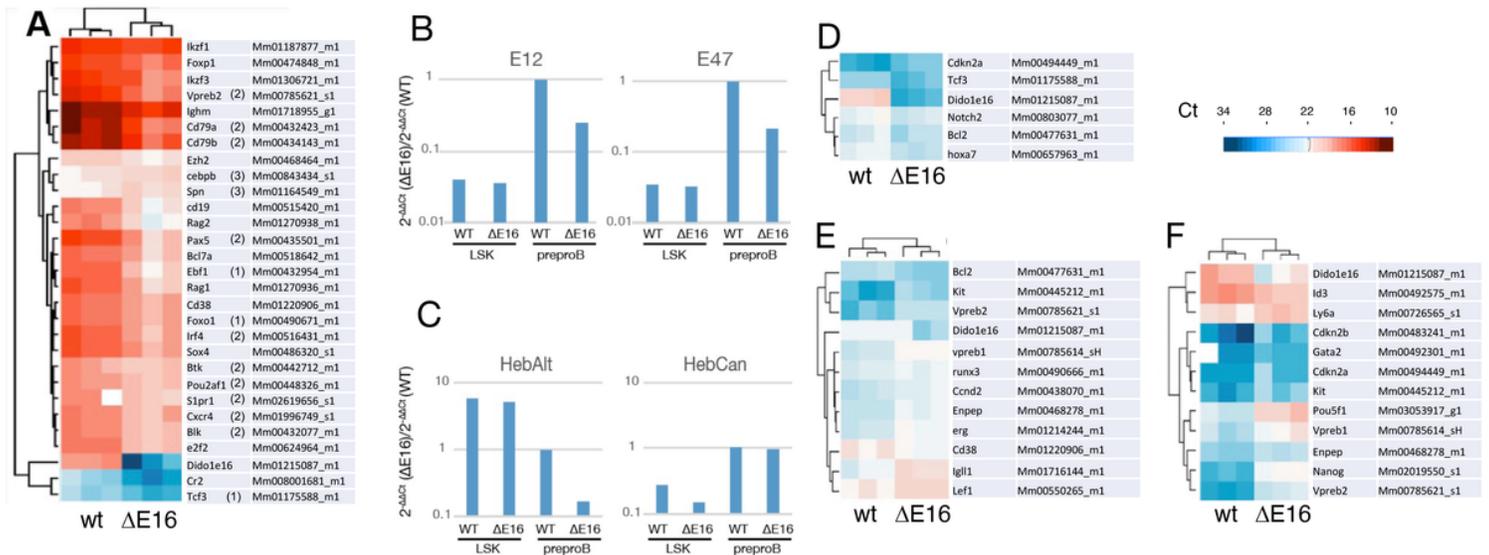


Figure 6

Differences in gene transcription levels in distinct B cell precursor populations from bone marrow. Three biological replicates were analyzed from wt and Dido1 Δ E16 mice. Cells were sorted by flow cytometry. The RNA from each precursor population was extracted, preamplified, and used as template for the TLDA assay. Maximal expression is represented in dark red and minimal expression, dark blue A. Heatmap showing hierarchical clustering and expression of differentially expressed genes ($p < 0.05$) by the prepro-B population. B. Differential expression of Tcf3 and its two isoforms in prepro-B cells and in lin \bar{c} precursors. Values are relative to wt prepro-B levels. C. Expression of various Tcf12 isoforms in Dido1 Δ E16 prepro-B cells, each relative to expression of the same region in wt cells. D-F, heatmaps as in A, for other B progenitor subpopulations. D. Heatmap for the pro-B population. E. Heatmap of the pre-B population. F. Heatmap of the immature B cell population.

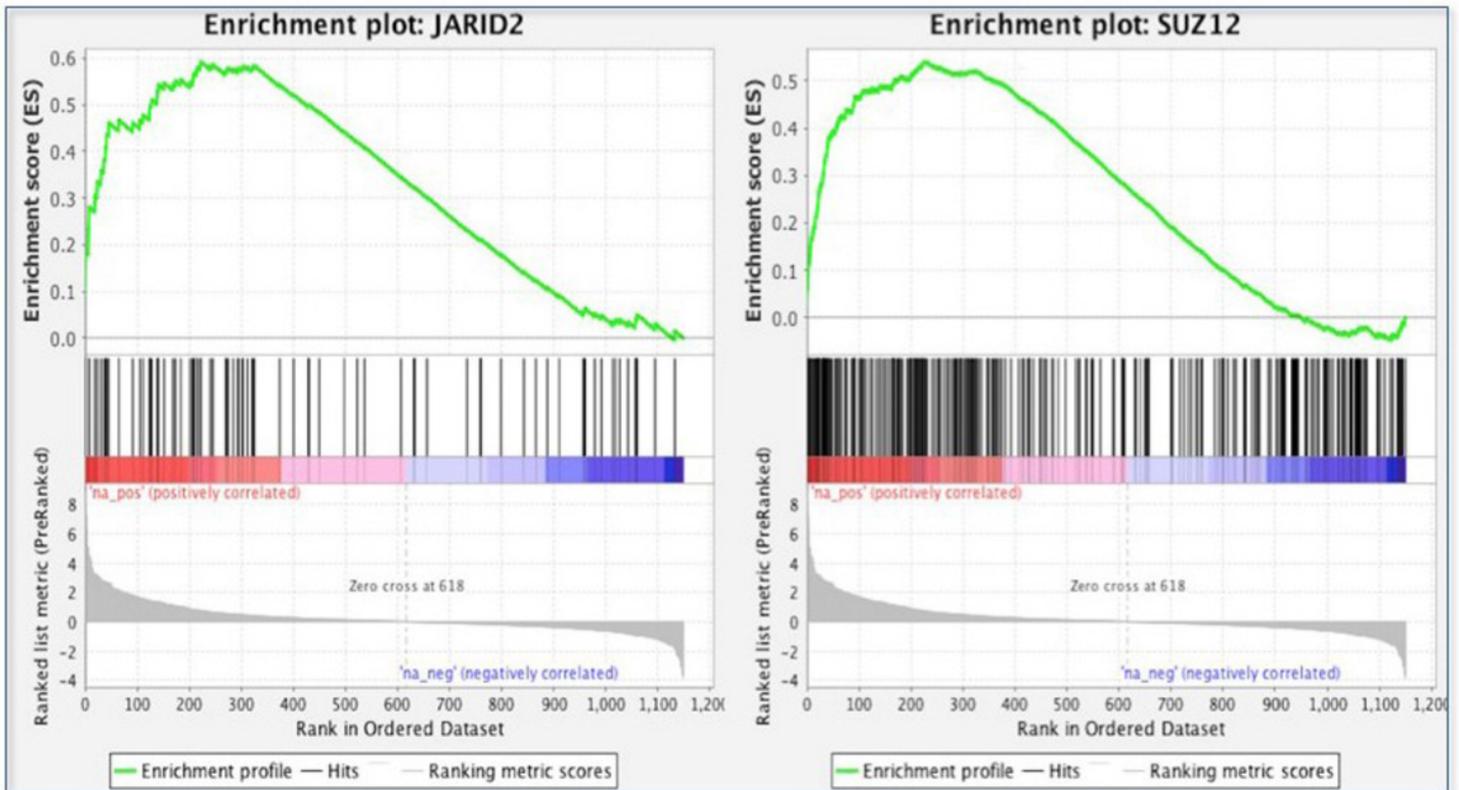


Figure 7

GSEA enrichment graphs of genesets relative to PRC2 activity and differentially expressed genes in pre-B cells. RNA-seq expression data ranked according to fold change in the transcription level in pre-B Dido1 Δ E16 cells, relative to pre-B wt cells. Maximal correlation was observed between the expression of altered genes in pre-B Dido1 Δ E16 cells and genes identified by ChIPseq as targets of PRC2 components JARID2 ($p < 0.001$; FDR = 0.213) and SUZ12 ($p < 0.001$; FDR = 0.217).

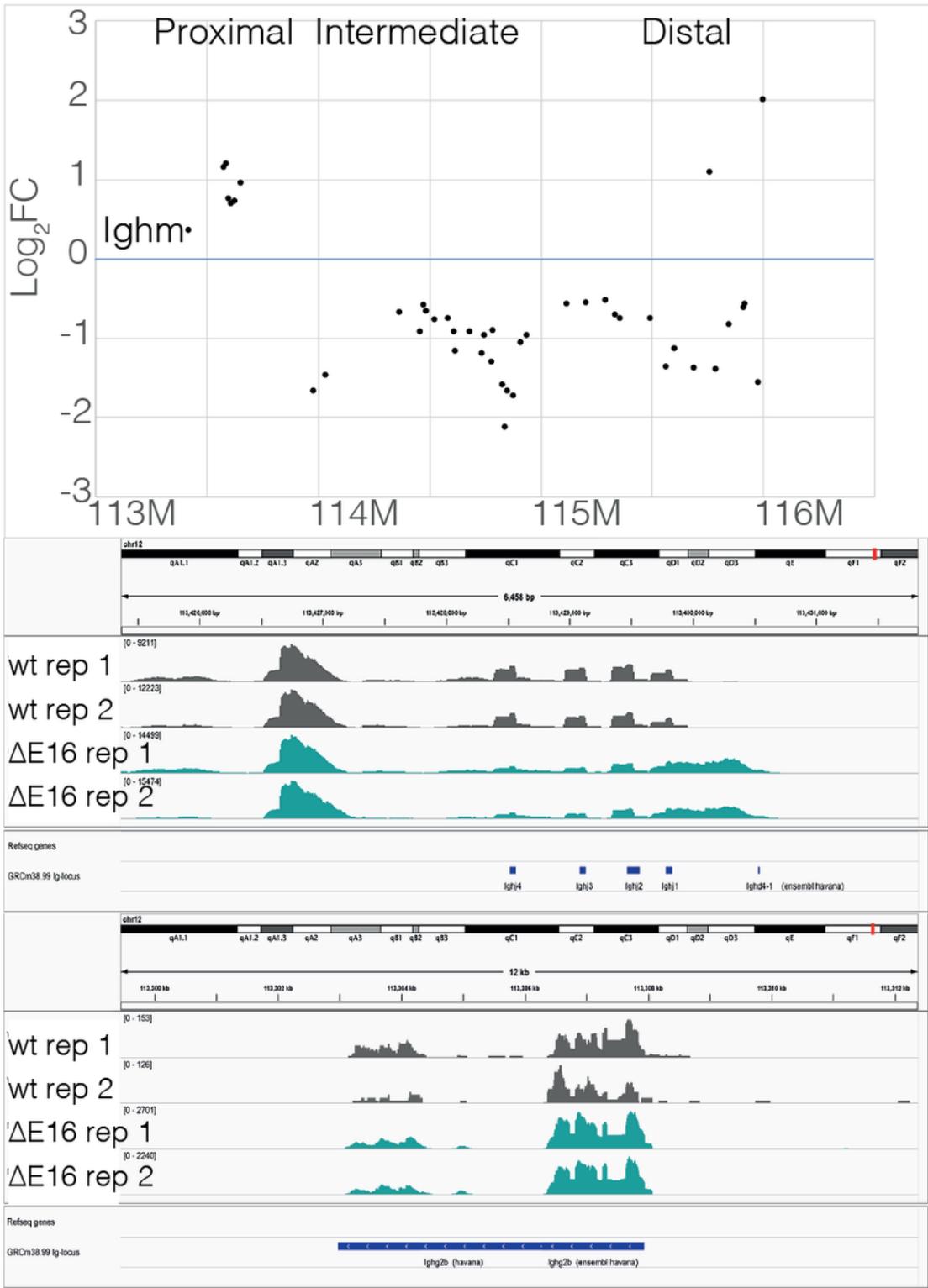


Figure 8

Analysis of the Igh locus in Dido1 ΔE16 pre-B cells. A. Relative abundance of Ighv regions in the transcriptome of Dido1 ΔE16 pre-B cells, shown as logFC of wt levels (*, p<0.05). Bottom, chromosome coordinates; top, the main regions (proximal, intermediate, distal) in which distinct V families map. B. Integrative Genomics Viewer (IGV; <http://www.broadinstitute.org>) plot with RNA-seq data showing transcription from the DNA region between D and J regions; J fragments and the more distal D fragment

(D4) positions are represented. C. IGV plot with RNA-seq data showing transcription of the Igg2b gene in Dido1 Δ E16 pre-B samples.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementaryfigure1.docx](#)
- [Supplementaryfigure2.docx](#)
- [Supplementaryfigure3.docx](#)
- [Supplementaryfigure4.docx](#)
- [Supplementaryfigure5.docx](#)
- [Supplementarytable1.xlsx](#)
- [Supplementarytable2.xlsx](#)
- [Supplementarytable3.xlsx](#)
- [Supplementarytable4.docx](#)