UNIVERSITAT DE BARCELONA

Role of Sirtuins in haematopoiesis and leukaemia

Andrés Gámez García-Cervigón

Josep Carreras LEUKAEMIA Research Institute

UNIVERSITAT DE BARCELONA

UNIVERSITAT DE BARCELONA

FACULTAT DE MEDICINA I CIÈNCIES DE LA SALUT

PROGRAMA DE DOCTORAT EN BIOMEDICINA

ROLE OF SIRTUINS IN HAEMATOPOIESIS AND LEUKAEMIA

Doctoral thesis submitted by **Andrés Gámez García-Cervigón** in order to obtain the degree of Doctor of the **Universitat de Barcelona (UB) 2024**.

This thesis has been performed under the co-supervision of **Dr. Alejandro Vaquero García** and **Dr. Berta N. Vazquez Prat** from the Josep Carreras Leukaemia Research Institute. Tutor: **Dr. Manel Esteller Badosa.**

Student:

Andrés Gámez García-Cervigón

Director:	Director:	Tutor:
Dr. Alejandro Vaquero García	Dr. Berta N. Vazquez Prat	Dr. Manel Esteller Badosa

A mi padre

1. ABBREVIATIONS

53BP1	p53 binding protein 1
Ac-Coa	Acetyl-coenzyme A
AcK	Anti-pan-acetyl-lysine antibody
ADPR	ADP-Ribose
ADPRT	ADP-Ribosyltransferase
ALP	All-lymphoid progenitor
AML	Acute myeloid leukaemia
AMPK	(AMP)-activated kinase
ARTC	Clostridial toxin-like ADPRT
ARTD	Diphtheria toxin-like ADPRT
ATM	Ataxia-telangiectasia mutated
ATR	Ataxia-telangiectasia and Rad3-related
BAF	BRG1-associated factor
BAFF	B-cell activating factor of the TNF family
BAFFR	BAFF receptor
B-ALL	B-cell acute lymphoblastic leukaemia
BCL6/11B	B-cell leukaemia/lymphoma 6/11B
BCR	B-cell receptor
BD	Bromodomain
BIR	Baculovirus IAP Repeat domain
BLP	B-lymphoid progenitor
BM	Bone marrow
bp	DNA base pair
C/EBPα	CCAAT/enhancer binding family protein
cBAF	Canonical BAF
CBP	CREB-binding protein
CDK9	Cyclin-dependent kinase 9
CFA	Complete Freund adjuvant
CHD	Chromodomain helicase DNA binding
ChIP-Seq	Chromatin immunoprecipitation coupled to sequencing
CLL	Chronic lymphocytic leukaemia
CLP	Common lymphoid progenitor
CML	Chronic myeloid leukaemia
CMP	Common myeloid progenitor
COG	Children's oncology group
Co-IP	Co-immunoprecipitation
CXCL12	C-X-C motif chemokine 12
DAMP	Damage-associated molecular pattern
DAPI	4',6-diamidino-2-phenylindole
DC	Dendritic cell
DCD	Double chromodomain
DDA	Data-dependent acquisition
DEG	Differentially expressed gene
DGE	Differential gene expression
DN1-4	Double negative 1-4 cell
DNMT	DNA methyltransferase
DOT1L	Disruptor of telomeric silencing 1-like
DP	Double positive cell
DPF	Double PHD finger domain
DSB	DNA double-strand break
DUB	Deubiquitinase
EBF1	Early B-cell factor 1
ELK4	ETS domain-containing protein
ERK1/2	Extracellular signal-regulated kinase 1/2
ETP	Early Thymic progenitor
EV	Empty vector
EZH2	Enhancer of zeste homolog 2

FAD	Flavin adenine nucleotide
FDR	False discovery rate
FL	Foetal liver
FLT3L	FMS-like tyrosine kinase 3 ligand
FO B cell	Follicular B-cell
FOXO	Forkhead box O
GABPβ1	GA binding protein β1
GC B cell	Germinal center B-cell
GCN5	General control non-derepressible 5
GDH	Glutamate dehydrogenase
GEO	Gene expression omnibus
GFI1	Growth factor independent 1
GMP	Granulocyte-monocyte progenitor
GNAT	GCN5-related N- acetyltransferases
GSEA	Gene set enrichment analysis
HAT	Histone acetyltransferase
hCD4	Human CD4
HD	Partial homeodomain
HDAC	Histone deacetylase
HIF1	Hypoxia inducible factor 1
HPLC	High-performance liquid chromatography
HPSC	Haematopoietic and progenitor stem cell
HR	Homologous recombination
HSC	Haematopoietic stem cell
ID	Inhibitory domain
ID2	Inhibitor of differentiation 2
IDR	Intrinsically disordered region
Igh	Immunoglobulin heavy chain locus
IL7	Interleukin 7
IL7R	Interleukin 7 receptor
ILC	Innate lymphocyte cell
IRF4/7/8	Interferon regulatory factor 4/7 /8
ISWI	Imitation Switch
JAK3	Janus kinase 3
JMJD3	Jumonji domain-containing 3
KMT	Lysine methyltransferase
LC-qMS/MS	Low-input liquid chromatography coupled to quantitative mass spectrometry
LLPS	Liquid-liquid phase separated condensates
LMMP	Lymphoid-primed multipotent progenitor
LSD	Lysine-specific demethylases
LT-HSC	Long-term HSC
Ly6d	Lymphocyte antigen 6D
MACS	Magnetic-activated cell separation
MBT	Malignant brain tumor domain
M-CSF	Macrophage colony-stimulating factor
MEP	Megakaryocyte–erythrocyte progenitor
MLL	Mixed-lineage leukaemia
MPP	Multipotent progenitor
NAM	Nicotinamide
NBS1	Nijmegen breakage syndrome 1
ncBAF	Non-canonical BAF
NCOR1	Nuclear receptor corepressor 1
NF-κB	Nuclear factor Kappa B
NHEJ	Non-homologous end joining

NK cells	Natural killer cells
NP-HEL	NPP hapten-conjugated hen egg lysozyme
NuRD	Nucleosome remodelling and deacetylase complex
mADPRT	mono-ADPRT
OP	Octapeptide
PAMP	Pathogen-associated molecular pattern
PARP	Poly-ADPRT
PAX5	Paired-box 5
PC	Plasma cell
PCA	Principal component analysis
PCAF	Polybromo-associated BAF
pDC	Plasmacytoid dendritic cell
Ph	Philadelphia chromosome
PHD	Plant homeodomain
PI3K	Phosphatidylinositol 3-kinase
PLZF	Promyelocytic leukaemia zing finger
PP	Protein phosphatase
PPARγ	Peroxisome proliferator-activated receptor γ
PPI	Protein-protein interaction
PRC1	Polycomb repressive complex 1
PRMT	Protein arginine methyltransferases
PSM	Peptide-spectrum matches
PWWP	Pro-Trp-Trp-Pro domain
RT-qPCR	Real-time quantitative polymerase chain reaction
RAG1/2	Recombination activation genes 1/2
RNA Pol	RNA polymerase
RPA2	Replication protein 2
rRNA	Ribosomal RNA
SAM	S-adenosyl-L-methionine
SCF	Stem cell factor
scRNA-Seq	Single-cell RNA sequencing
sgRNA	Short-guide RNA
Sir2	Silent information regulator 2
SIRT	Sirtuin
SP	Single positive T-cell
STAT5	Signal transducer and activator of transcription 5
ST-HSC	Short-term HSC
SWI/SNF	Switch/sucrose non-fermentable
TAD	Transactivation domain
TARGET	Therapeutically applicable research to generate effective treatments
TCF3	Transcription factor 3
TCR	T-cell receptor
TET2	Ten-eleven translocation 2
T_{FH} cell	Follicular helper $CD4^+$ T-cell
T_H cells	Helper $CD4^+$ T-cells
TSA	Trichostatin A
TSP	Thymus seeding progenitor
TSS	Transcription start site
TTD	Tandem tudor domain
UTR	Untranslated region
UTX	Ubiquitously transcribed tetratricopeptide repeat, X chromosome)
V(D)J	Variable, Diversity and Joining gene segments
WRN	Werner syndrome protein
zf-CW	zinc-finger CW domain

II. ABSTRACT

Haematopoiesis is the process by which haematopoietic stem cells replenish all blood cell types, including myeloid cells and lymphocytes, throughout life in mammals. In particular, B lymphocytes or B-cells are the key mediators of humoral immunity by secreting antigen-specific antibodies that identify and neutralise noxious agents. B-cells develop in the bone marrow through sequential differentiation into several progenitor stages that progressively acquire a specialised identity and lose their lineage potential to become committed to the B-cell lineage. This is orchestrated by a number of lineage-specific transcription factors, including PAX5, which controls virtually all aspects of B-lymphopoiesis. Importantly, defects in haematopoietic progenitors can lead to the development of haematological malignancies, such as B-cell acute lymphoblastic leukaemia (B-ALL), the most common cancer in children. In B-ALL, PAX5 is a potent tumour suppressor that is frequently targeted by monoallelic inactivating mutations, and the reconstitution of PAX5 physiological levels is known to induce leukaemic cell death. However, the mechanisms controlling PAX5 protein dynamics are poorly understood, which has prevented its exploitation as a pharmacological target.

Haematopoiesis provides a unique model to study the epigenetic basis of cell identity and malignancy. Furthermore, epigenetic regulators are frequently targeted by deleterious mutations in haematological malignancies. One of such regulators are sirtuins, a family of NAD^+-dependent deacetylases that coordinate cellular adaptations to stress, play critical functions in organismal physiology and ageing, and have emerged as attractive anti-cancer targets. Surprisingly, while the roles of sirtuins in immunity have been extensively studied, very little is known about their relevance in haematopoietic differentiation and leukaemia, which prompted us to address this gap of knowledge.

Here, we identified a molecular mechanism linking the sirtuin family member SIRT7 with B-cell development and malignancy through the regulation of PAX5. We describe that SIRT7 is upregulated in developing B-cells, where it collaborates with PAX5 to transcriptionally repress lineage potential and sustain B-cell development and commitment. SIRT7 imposes a dynamic deacetylation switch on a single PAX5 residue that controls its protein stability and functions. Acetylated PAX5 poorly binds to chromatin and fails to regulate most of its target genes and to drive both development and commitment. In contrast, deacetylated PAX5 stringently controls transcription and ultimately restores B-cell commitment but not differentiation, which depends on the coordinated action of both proteoforms. Importantly, the SIRT7/PAX5 interplay is conserved in B-ALL cell lines and patients, where SIRT7 exerts a strong tumour suppressor function, presumably through this mechanism. Our findings define an important function of SIRT7 in B-cell identity and may provide a rationale to therapeutically target PAX5 in B-ALL.

III. INTRODUCTION

1. Haematopoiesis

Haematopoiesis is the developmental process by which all blood cell types are produced throughout mammalian life. The classical model of haematopoiesis entails the sequential differentiation of haematopoietic stem cells (HSCs) into discrete progenitor populations which hierarchically develop to generate the diverse blood cell types. In adults, this process primarily occurs in the bone marrow (BM), where HSCs reside within a microenvironment that sustains their long-term self-renewal and controls their proliferation, differentiation and migration[1]. HSCs are thought to undergo a first round of asymmetric division that produces two cells with different fates: a long-term (LT)-HSC that retains long-term self-renewal ability, and a primed short-term (ST)-HSC with reduced self-renewal potential, that further develops into a multipotent progenitor (MPP). MPPs split haematopoietic differentiation into two different branches by differentiating into oligopotent common lymphoid progenitors (CLPs) and common myeloid progenitors (CMPs), which lose multipotency and generate the precursors that give rise to all the cell mature cell types of the four blood lineages: erythrocytes, megakaryocytes, lymphocytes and myeloid cells[2] (**Fig. I1**).

In the classical model, successive fate decisions progressively branch the haematopoietic system from oligopotent progenitors to unipotent precursors, ultimately generating various specialised cell types. Thus, classical haematopoietic differentiation conceptually takes the form of a tree and is defined by two key features: hierarchy and irreversibility. First, each blood cell type follows a well-defined differentiation pathway involving several progenitor stages that gradually build their identity and functions. Second, each differentiation pathway is unidirectional, i.e. after each fate decision, cell plasticity is reduced and lineage commitment is increased, while downstream cell states cannot dedifferentiate into their more precursors. However, this rigid model has been challenged over the past 20 years by new conceptual and technological advances. For example, lineage tracing techniques and loss-of-function mouse models have demonstrated lineage conversion of committed B-lymphocyte progenitors, and even dedifferentiation of mature B-cells *in vivo*[3-6]. In cancer, clinical cases of lineage switching have been described and are thought to provide a mechanism for immune evasion. This has been observed for example in some B-ALL relapses, in which leukaemic blasts lose their B-cell markers and acquire a distinct myeloid phenotype[7-9]. In addition, it has recently been proposed that haematopoiesis works as a continuum of cellular states rather than a stepwise system of discrete populations making divergent fate decisions[10]. It is now clear that early

progenitors with identical immunophenotypes display impressive functional heterogeneity, and it has been suggested that HSCs can already be biased towards specific lineages prior to differentiation[11-14]. Thus, our view of haematopoietic differentiation is evolving to a more fluid model.

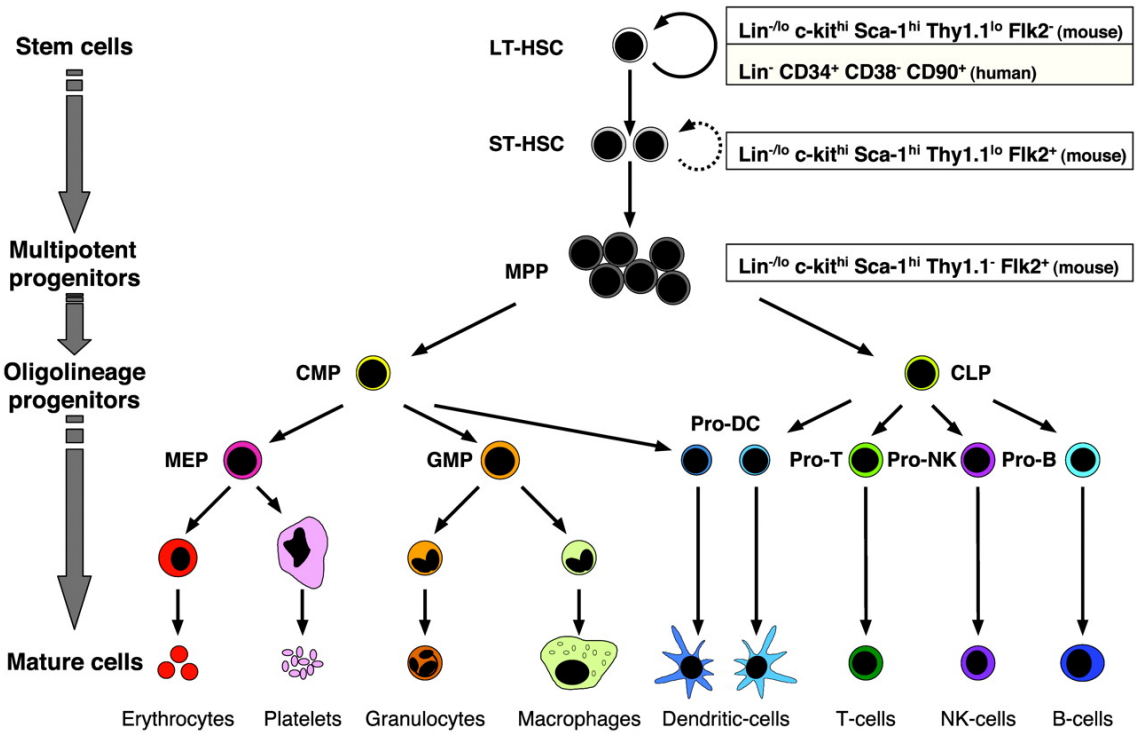


Fig. I1 | Summary of the hierarchical differentiation from HSCs to mature blood cell types
HSCs self-renew and differentiate into MPPs, that branch off haematopoietic differentiation into myeloid (CMP) and lymphoid (CLP) oligopotent progenitors. CMPs and CLPs subsequently differentiate into intermediate precursors that ultimately generate all blood cell types. From Passegue *et al.* (2003)[2].

Either hierarchical or continuum, the haematopoietic system replenishes the mammalian blood daily with 2-4x10[11] cells of four major cellular lineages[15]: (i) erythrocytes, which carry haemoglobin to transport oxygen to tissues; (ii) megakaryocytes, which produce platelets, the key mediators of coagulation during wound healing; (iii) myeloid cells, the central players in innate immunity; and (iv) lymphocytes, the most recently evolved blood cell type, which provide antigen-specific adaptive immunity and memory. As depicted in **Fig. I1**, erythrocytes, platelets and myeloid cells are derived from a common CMP origin. CMPs branch into megakaryocyte–erythrocyte progenitors (MEPs), which further bifurcate into erythroid and megakaryocytic unipotent precursors; and granulocyte-monocyte progenitors (GMPs), which are responsible for the generation of all myeloid cell types (mast cells, basophils, eosinophils, dendritic cells, monocytes and neutrophils). On the

other hand, CLPs give rise to all lymphoid lineages (B-cells, innate lymphocyte cells (ILCs), NK cells and T-cells), and, interestingly, are also an important source of dendritic cells[2].

1.1. Myelopoiesis

Myeloid cells are the main component of innate immunity, the first layer of active immunity. They play critical functions in the organismal defense against pathogens and exogenous molecules and participate in other physiological processes, including tissue repair and angiogenesis[16]. In the context of the immune response, myeloid cells mediate immediate and generic reactions upon the recognition of molecules that are common to many pathogens (pathogen-associated molecular patterns, PAMPs). Similarly, damaged host cells release damage-associated molecular patterns (DAMPs) that are recognised by pattern recognition receptors (such as Toll-like receptors or C-type lectin receptors). In both cases, receptor activation leads to the secretion of cytokines and other molecules that create an inflammatory environment aimed at eliminating the detrimental agent and recruiting more immune cells[17]. In fact, myeloid cells are also responsible for priming lymphocyte activation through antigen presentation, so they represent the first trigger of adaptive immunity.

In the classical model of haematopoietic differentiation, each fate decision is determined by the timely expression of lineage-instructive transcription factors and by the sensing of extracellular cues. From the perspective of transcriptional regulation, myelopoiesis is orchestrated by PU.1 (or SPI1), CCAAT/enhancer binding family proteins (C/EBPα, C/EBPβ and C/EBPδ), and Interferon regulatory factor 8 (IRF8), among others (**Fig. I2**). Intriguingly, although these transcription factors are decisive elements of myeloid differentiation, none of them is myeloid-specific. Instead, early myeloid specification appears to be determined by their balanced and stage-specific expression. For example, PU.1 is expressed at varying levels by many blood cells and their precursors, and its abundance is a major determinant of the early determination of myeloid versus lymphoid fate[18-20]. PU.1 (encoded by the *Spi1* gene) is required for both myeloid and lymphoid differentiation, as *Spi1*$^{-/-}$ mice display a complete loss of monocytes, granulocytes and lymphocytes[21]. Furthermore, conditional knock-out models have shown a complete absence of CMPs and CLPs in PU.1-deficient mice[22,23]. Shortly after the generation of CMPs, C/EBPα mediates their differentiation into GMPs, which are completely absent in *Cebpa*$^{-/-}$ mice, and participates in the granulocyte versus monocyte lineage decision[24,25]. Indeed, the relative expression of C/EBPα and PU.1

has been proposed to determine the macrophage-versus-neutrophil lineage choice, although IRF8 also takes an active part of this decision[26,27]. However, transcription factor-guided fate decisions are highly stage-specific across haematopoiesis, as illustrated by an elegant report by Heinz *et al*[28]. This article investigated the genomic distribution of PU.1 in mature macrophages and B-lymphocytes, and that of C/EBPα in macrophages, using chromatin immunoprecipitation coupled to sequencing (ChIP-Seq). They found that more than half of the PU.1 peaks were cell type-specific, and that C/EBPα colocalised with PU.1 in macrophage-specific regulatory elements. Importantly, the differential binding of PU.1 strongly correlated with the specific expression programmes of both cell types, and loss of PU.1 led to a sharp reduction in the occupancy of C/EBPα, highlighting how two factors can oppose or exploit each other depending on the cellular context[28]. Thus, in the absence of stage-specific transcription factors, myeloid specification and commitment rely on the balanced and time-resolved expression of these factors, which perform cell type-specific functions depending on the repertoire of available collaborators and antagonists.

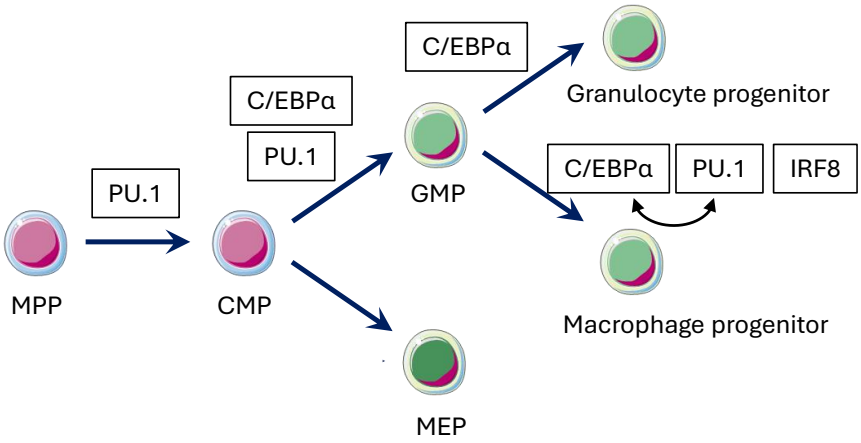


Fig. I2 | Overview of how transcription factors drive myelopoiesis
PU.1 facilitates differentiation MPP differentiation into CMPs, that further are further specified into GMPs (by C/EBPα and PU.1) or MEPs. In GMPs, C/EBPα promotes granulocyte differentiation and collaborates with PU.1 and IRF8 to facilitate macrophage differentiation. Adapted from Stavast *et al.* (2018)[29].

1.2. Lymphopoiesis

Lymphocytes are one of the most outstanding innovations of recent evolution. They provide adaptive immunity, which can be raised on demand against virtually any threatening agent, and have the unique ability to preserve the acquired immunity for extremely long periods of time. The cornerstone of adaptive immunity is the expression of

genetically encoded monoclonal B-cell and T-cell receptors (BCRs and TCRs) that recognise cognate antigens with great specificity. Thus, a repertoire of B- and T-lymphocytes with enormous functional and genetic diversity constitutes adaptive immunity. B-cells play a leading role in humoral immunity by expressing antigen-specific immunoglobulins (BCRs) in their membranes and secreting them massively into the blood stream upon activation by their antigen. In contrast, two major types of T-cells can be distinguished according to their functional features, both of which rely on antigen presentation to their TCRs for activation: $CD8^+$ T-cells, which play cytotoxic functions against tumour cells or cells infected with intracellular pathogens, and the heterogeneous group of $CD4^+$ T-cells or helper T-cells (T_H cells), which orchestrate immune responses by regulating the functions of innate immune cells and B-cells.

In addition, a third group of lymphocytes, the innate lymphocyte cells (ILCs), do not express antigen-specific receptors and are involved in innate immunity:. ILCs are a heterogeneous group of lymphocytes that can be divided into class I ILCs (natural killer (NK) cells and ILC1), class II ILCs (ILC2) and class III ILC3s. NK cells are well-known for their ability to kill tumour cells and cells infected with viruses or intracellular bacteria, whereas ILC1-3 have been identified more recently and play regulatory functions based on the secretion of specific cytokines[30]. Hence, there is an evident functional analogy between NK cells with $CD8^+$ T-cells and ILCs with $CD4^+$ T-cells. Indeed, the transcriptional programmes that regulate the functions of several T_H cell and ILC populations, as well as those of their cytotoxic NK and T-cell counterparts, overlap substantially[31]. Intriguingly, genetic deletion of the transcription factor GATA3 in mice impairs the development of several ILC and T_H populations but has been proposed to be dispensable for that of cytotoxic NK or $CD8^+$ T-cells, suggesting that, after diverging into their different lineages, T-cells and NK cells share a common lineage instructive network[32,33]. Earlier in development, the DNA binding inhibitor ID2 (Inhibitor of differentiation 2) seems to be the central driver of ILC specification in CLPs, by inhibiting the B- and T-cell fate. However, $ID2^+$ ILC precursors do not differentiate into conventional NK cells, suggesting that, despite their functional similarity to ILC1, NK cells develop via an independent differentiation pathway that remains to be characterised in detail[34] (**Fig. I3**).

Another key aspect of early fate determination is the need to repress specific transcriptional regulators at precise developmental stages, and this is achieved by self-reinforcing networks that progressively emerge during differentiation. For example, E-proteins, that drive B- and T-cell specification, are strongly antagonised by ID family

proteins, such as ID2, which promotes the ILC fate[35,36]. The E-protein TCF3 (Transcription factor 3, or E2A) initiates the transcriptional networks that precede B-lymphopoiesis, by inducing the master regulator of B-cell specification, EBF1 (Early B-cell factor 1)[37]. Subsequently, one of the most important functions of EBF1 is to strongly repress ID2 expression to abolish ILC potential. Indeed, deletion of EBF1 in B-cell progenitors leads to ID2 derepression and facilitates their conversion into ILCs[38]. Again, this should not be interpreted as a general rule for lymphoid development, as, for example, ID2 and ID3 are upregulated later in mature CD8⁺ T-cells, which rely on these proteins for their survival and terminal differentiation[39].

Given the recent and common evolutionary origin of B- and T-cells, it is not surprising that they develop through parallel differentiation pathways, that share many of their key features. The following sections discuss the transcriptional regulation of B-cell and T-cell development.

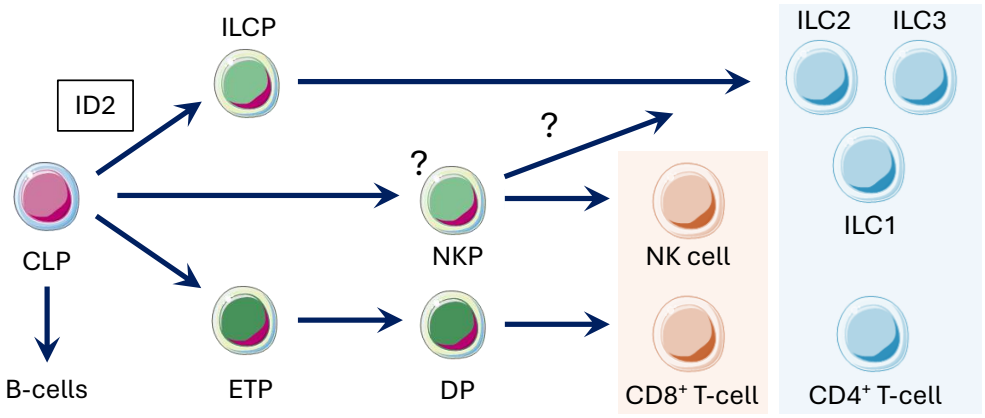


Fig. I3 | Schematic representation of the analogy between T-cell and NK cell development
Helper-like ILC1-3 differentiate from a common ID2⁺ ILCP, whereas a putative NKP precedes the development of cytotoxic NK cells, and potentially that of ILC1-3. ETPs differentiate into several T-cell progenitor stages, including the DP stage, to give rise to helper CD4⁺ and cytotoxic CD8⁺ T-cells. Cytotoxic NK and CD8⁺ T-cells and helper-like ILC1-3 and CD4⁺ T-cells are grouped in shaded blue and orange squares, respectively.

1.2.1. Transcriptional regulation of T-cell development

T-cell progenitors develop from CLPs in the bone marrow but soon migrate to a specialised organ, the thymus, to complete their maturation. It should be noted that the classical model separating myeloid versus lymphoid differentiation early in CMP and CLP cells was later challenged by the identification of lymphoid-primed multipotent progenitors

(LMPPs), which exhibit granulocytic, monocytic and lymphoid potential and precede CLPs in lymphoid development[40]. However, for simplicity, CLPs will be referred to here as the common B- and T-cell precursors. In CLPs, the first branching point is the expression of the surface marker Ly6d (Lymphocyte antigen 6D), which segregates these progenitors into Ly6d$^-$ ALPs (All-lymphoid progenitors) and Ly6d$^+$ BLPs (B-lymphoid progenitors), with the former preceding the latter in development[41] (**Fig. I4**). After the ALP stage, an uncharacterised thymus-seeding progenitor (TSP) migrates to the thymus and gives rise to the first T-cell committed precursor, the ETP (Early Thymic progenitor), in a process that is critically dependent on Notch signalling in the thymus. Indeed, ectopic expression of Notch1 in the bone marrow inhibits B-cell development and induces T-cell development *in situ*, and Notch1 inactivation completely depletes T-cells in mice[42,43]. Intriguingly, Notch1 deletion was later reported to allow intrathymic progenitors to develop into B-cells, highlighting the fundamental role of Notch signalling in early T-cell specification[44].

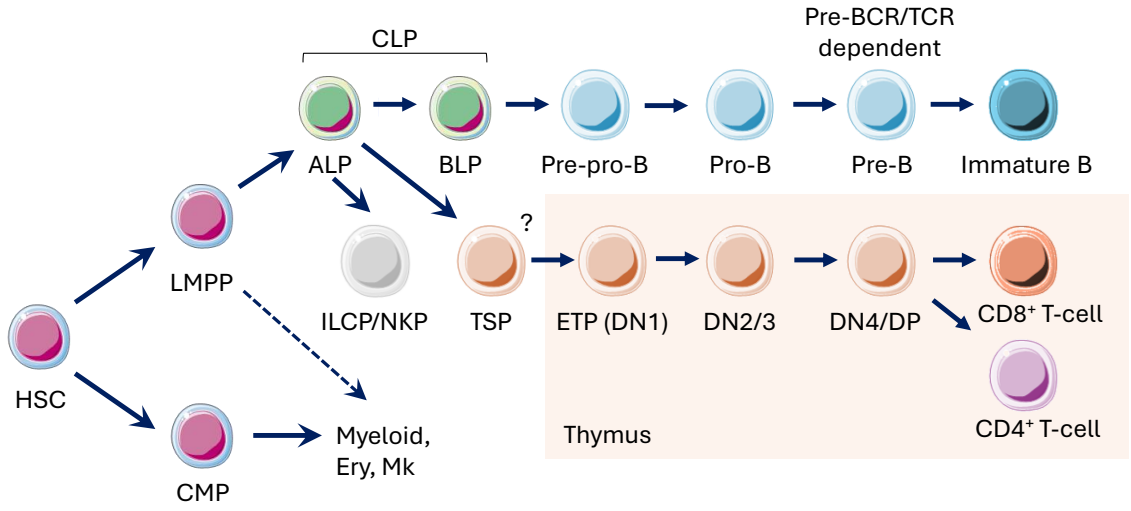


Fig. I4 | Overview of the analogy between T-cell and B-cell development
LMPPs differentiate into CLPs (consisting of ALPs and B-cell biased BLPs) but preserve some myeloid potential. CLPs hierarchically differentiate into pre-pro-B cells, pro-B, pre-BCR dependent pre-B cells and finally into immature B-cells. A CLP-derived TSP migrates into the thymus and differentiates into DN1, DN2/3 (pre-TCR-independent, analogous to pro-B cells) and pre-TCR-dependent DP cells, and ultimately differentiate into CD8$^+$ or CD4$^+$ T-cells. Ery, erythrocytes; Mk, megakaryocytes.

Once in the thymus, ETPs undergo a sequential differentiation pathway that parallels that of B-cell progenitors, with the generation of functional and safe TCRs and BCRs being the cornerstone of both processes. This is accomplished through the stepwise rearrangement of variable, diversity and joining (V(D)J) gene segments by RAG1 and RAG2 (Recombination activation genes 1 and 2), which underlies the extraordinary diversity of antigen-specific

receptors (described in more detail in the next section). Thus, ETPs (or DN1) initially differentiate into $CD4^-CD8^-$ double negative 2 (DN2) cells, which undergo the first round of V(D)J recombination, become fully committed to the T-cell lineage and further differentiate into DN3 cells. In analogy to B-cell progenitors, DN2-3 cells are also referred to as pro-T cells due to their independence from TCR signals. At the DN3 stage, two key events take place: first, DN3 cells exit the cell cycle, and then they undergo the second round of V(D)J recombination. This constitutes a major checkpoint of T-cell development, as only those clones that make in-frame rearrangements will express a functional pre-TCR that provides the signals necessary for survival. These clones finally complete their maturation through the downstream DN4 and $CD4^+CD8^+$ double-positive (DP) stages, which undergo positive and negative selection and eventually segregate into $CD4^+$ or $CD8^+$ single positive (SP) T-cells[45].

From the perspective of transcriptional regulation, T-cell development cannot be attributed to any T-cell-specific transcription factor. Instead, the generation of a T lineage identity is guided by Notch signalling and the coordinated action of several transcription factors[46]. Notch signalling in the thymus is the starting point of T-cell differentiation. In ETPs, Notch rapidly induces the expression of the pioneer transcription factors TCF1 and GATA3[47,48]. Although neither TCF1 nor GATA3 are T-cell specific, they are the first instructive factors to drive T-cell specification into DN2 progenitors[49]. TCF1 activates a primitive T-cell program by promoting a first wave of accessible chromatin in T-cell regulatory elements[50]. Thereby, TCF1 induces the early expression of T-cell genes, including GATA3[47]. Interestingly, ectopically expressed TCF1 in fibroblasts binds to T-cell regulatory elements and induces the *de novo* formation of open chromatin in these regions, leading to the derepression of T-cell genes in these cells. These observations highlight the pioneering ability of TCF1[50]. In parallel, GATA3 plays a primordial role in the early commitment of T-cell progenitors, by abolishing the B-cell potential in DN2 cells[51]. Indeed, GATA3 can bypass TCF1 deficiency in early T-cell progenitors to abrogate their lineage promiscuity[52]. Of note, genome-wide analyses revealed the stage-specificity of GATA3 functions, as it binds to largely different regions in differentiating and mature T-cell subsets, thereby regulating distinct transcriptional programmes[53].

During the transition to the DN2 stage, the expression of TCF1 and GATA3 keeps increasing[46]. At this stage, they collaborate to induce another critical transcription factor, BCL11B (B-cell leukaemia/lymphoma 11B), whose expression is fairly restricted to T-cells

within the haematopoietic system[54]. BCL11B is the determinant of T-cell commitment, as T-cell progenitors lacking BCL11B fail to downregulate stem cell-related genes and display strong NK and even myeloid cell potential[54-56]. Indeed, BCL11B interacts with transcriptional repressor complexes and directly suppresses the expression of ID2 and PLZF (Promyelocytic leukaemia zinc finger), an important regulator of innate immune responses, among other alternative lineage genes[57]. Interestingly, BCL11B recruits some of its interaction partners to its target loci, as ChIP-Seq experiments of these factors revealed that they relocated or lost their genomic occupancy upon BCL11B deletion[57]. Furthermore, *Bcl11b*$^{-/-}$ DN2 cells are largely impaired to further proceed in development, indicating that BCL11B is also involved in differentiation[55]. In fact, BCL11B not only behaves as a repressor but can also activate the expression of its target genes[58,59]. Hence, T-cell programs are established by the combinatorial action of transcription factors that create self-reinforcing feedforward networks to activate T-cell genes and repress lineage-inappropriate genes (**Fig. I5**).

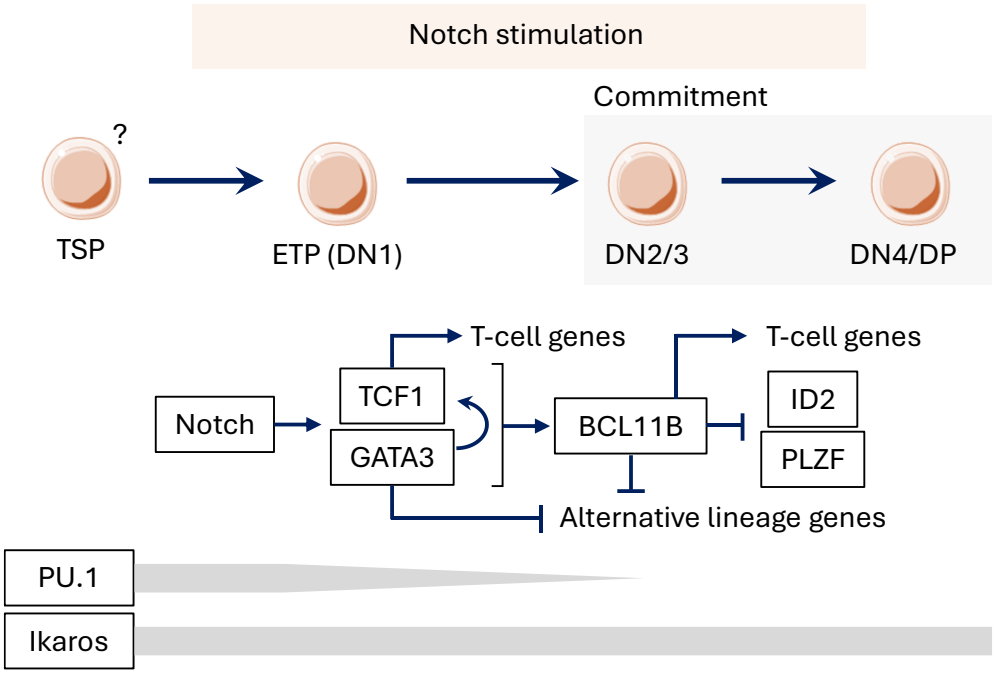


Fig. I5 | Extrinsic and transcriptional control of T-cell development
Notch signals in the thymus instruct early T-cell development in collaboration with transcription factors. Notch induces TCF1 and GATA3, that induce T-cell genes and repress alternative lineage genes, respectively. GATA3 further induces TCF1 expression and, together, both proteins activate BCL11B expression, that further activates T-cell genes and repress alternative lineage genes, including ID2 and PLZF. This process is assisted by Ikaros throughout T-cell development and by decreasing expression of PU.1 in its earliest stages.

However, so far only the appearance of new players along the T-cell differentiation pathway has been considered. Although these factors are critical for initiating and maintaining the T-cell programmes, they are clearly not sufficient to drive the T-cell fate. While suppression of many HSC/MPP factors is required to preclude stemness in early T-cell progenitors, others are also required for later stages of lymphoid development. This is the case of the pan-haematopoietic factors PU.1 and Ikaros (IKZF1), which are already expressed in HSCs and MPPs but play vital functions in early and late T-cell development. In HSCs, Ikaros deficiency results in the loss of lymphoid potential, as neither lymphocytes nor their progenitors can be detected in *Ikzf1* null mice[60-63]. However, unlike other stemness genes, Ikaros expression is observed throughout T-cell (and B-cell) development, where it plays several stage-specific functions. For example, Ikaros participates in gene repression during the DN3-to-DN4 transition and at the DP stage[64,65]. Indeed, the introduction of a common T-ALL haploinsufficient mutation affecting the DNA-binding domain of Ikaros in mice leads to a T-lymphopoiesis arrest at the DN3 stage and to T-ALL development with full penetrance[66]. Afterwards, Ikaros is also required for negative selection and for the CD4 versus CD8 fate decision[67]. Even in mature T-cells, the polarisation of $CD4^+$ T-cells into several T_H subsets has been shown to be transcriptionally regulated by Ikaros[68,69].

In the case of PU.1, although it is required for T-cell development, it also appears to oppose T-cell commitment, and its expression is progressively reduced, but readily detectable, until the DN2 stage[70]. Yet, PU.1 binds to more than 30,000 sites in ETPs and still binds to 5,000 sites in DN2 cells. Along this trajectory, the accessibility of regions that lose PU.1 binding is lost, and the expression of their nearby genes is consequently reduced, suggesting that PU.1 participates in the early specification of ETPs rather than in later stages. In addition, ectopically expressed PU.1 in DN3-like cells binds to condensed chromatin regions and induces accessibility in these regions, leading to transcriptional activation of the associated genes[71]. But if PU.1 opposes T-cell development after the ETP stage, why should it be expressed in DN2 cells? First, PU.1 has been shown to be required for the optimal cell cycle progression of ETP and DN2 cells. Second, PU.1 *slows down* the activation of the T-cell developmental programmes in DN2 cells, suggesting that it regulates the timing of genes that are required in later stages[72]. Not surprisingly, PU.1 is commonly mutated in T-cell leukaemias, and T-cell progenitors that fail to downregulate PU.1 are more prone to malignant transformation[73,74].

Thus, the combined action of pre-existing and newly expressed factors, that play stage-specific functions and must be turned on and off with an appropriate timing, drives the establishment of the T-cell identity.

1.3. B-cell development

B-cell differentiation follows the same rules as its T-cell counterpart, with one fundamental exception: the expression of B-cell-specific factors, which is not observed in T-cells. As anticipated above, several profound analogies exist between B- and T-cell lymphopoiesis, in terms of their developmental trajectories and the general rules governing their transcriptional regulation (**Fig. I4**). Similar to thymocytes, B-lymphoid progenitors undergo a series of cellular stages that first specify and then commit them to the B-cell lineage, in a process that is driven by the time-resolved expression of a toolkit of transcription factors.

B-cell development begins when a CLP differentiates into the precursor of progenitor B-cells, or pre-pro-B cell, the first cell to express B-cell markers. After that, an iterative process of mutually exclusive proliferation and differentiation (i.e. V(DJ) recombination) stages starts. This process is driven from the inside by lineage-instructive transcription factors and from the outside by cytokines and chemokines that provide the cues that guide the entire developmental process. The two most important of these extrinsic factors are IL7 (Interleukin 7) and CXCL12 (C-X-C motif chemokine 12), which are expressed by stromal cells at discrete niches in the bone marrow and enforce the migration of B-cell progenitors to reciprocally exclude the signals of each other. In IL7-rich niches, the transcription factor STAT5 (Signal transducer and activator of transcription 5) is activated by phosphorylation and induces Myc-dependent proliferation while inhibiting differentiation in pre-B cells. Subsequently, migration to $CXCL12^{high}IL7^{low}$ niches leads to cell cycle exit and facilitates V(D)J recombination, in which the CXCL12/CXCR4 axis itself also plays an active role beyond migration[75]. Consistently, the knockouts of *Il7, Il7ra* (that encodes a subunit of the IL7 receptor (IL7R)), *Stat5a* and *Stat5b* arrest B-cell development at the pre-pro-B cell stage, although *Cxcl12* deficiency arrests B-cell development at a later stage, suggesting that it is dispensable in the earliest progenitors[76-79].

Upon CLP specification into pre-pro-B cells, the B-cell identity starts to be shaped. Pre-pro-B cells initiate the first round of V(D)J recombination and, once completed, differentiate into pro-B cells, that soon become committed to the B-cell lineage. Pro-B

cells undergo several cycles of proliferation and then exit the cell cycle to complete the recombination of the immunoglobulin heavy chain (*Igh*) locus, leading to the first critical checkpoint: the production of an in-frame *Igh* rearrangement and the subsequent expression of a functional IgH chain. In those clones that succeed, the IgH chain is transported to the cell surface and assembles the pre-BCR, a protein complex that provides the signals that enable survival and that defines the next stage in the B-cell developmental pathway, the pre-B cell. As pro-B cells do, pre-B cells first undergo a strong clonal proliferative burst, which is mediated by IL7/IL7R/STAT5 signalling, and then exit the cell cycle, instructed by the pre-BCR and by the escape of IL7-expressing niches[80]. Thus, pre-BCR signalling reactivates the V(D)J recombination program to perform the last step required to generate a B-cell: the rearrangement of the immunoglobulin light chain genes (*Igk* or *Igl*), another critical checkpoint. Upon succeed, B-cell progenitors express all the components of a functional B-cell and differentiate into immature B-cells, that express fully functional BCRs containing IgM molecules, and eventually migrate to the spleen to complete their maturation and give rise to the different B-cell subsets.

1.3.1. V(D)J recombination

Mature B-cells express an antigen-specific BCR on their surface, consisting of a functional antibody or immunoglobulin (Ig) and two signalling adaptors, Igα and Igβ, which mediate downstream signal transduction. Upon activation of the BCR by its cognate antigen, B-cells can differentiate into antibody-secreting cells and start to secrete massive amounts of their encoded antibody. Antibodies are formed by two heavy chains (IgH) and two light chains (Igκ or Igλ) containing several conserved Ig domains and an antigen-specific region. Antibody diversity is achieved through random V(D)J recombination of the Igh and *Igk* (or *Igl*) loci, which encode for the heavy and light chains, respectively, and this process is mediated by RAG1 and 2 (Recombination activation genes 1 and 2) proteins, which are exclusively expressed in B-cells and T-cells.

In mice, the germline *Igh* locus spans 2.8Mb and consists of 113 V_H (variable), 13 D_H (diversity) and 4 J_H (joining) gene segments[81]. In pre-pro-B cells, RAG proteins randomly select a D_H and a J_H element and perform the first *Igh* somatic rearrangement (D_H-to-J_H) though a DNA cut-and-paste mechanism that brings these exons together (**Fig. I6**). Later in pro-B cells, a second rearrangement, V_H-to-D_HJ_H, occurs, in a process that critically depends on loop extrusion of the entire *Igh* locus by the B-cell master regulator PAX5

(Paired box 5)[81,82]. This second rearrangement leads to the assembly of a functional *Igh* locus and facilitates differentiation into pre-B cells. Thus, pre-B cells express a clonally recombined IgH composed of a particular combination of V, D, and J elements, and use this IgH to ensemble their pre-BCR, allowing the recombination of the loci encoding the light chain proteins. The *Igk* (or *Igl*) locus lacks any diversity elements, and therefore its recombination involves only the rearrangement of variable and joining elements, which has recently been proposed to take place through different mechanisms than those driving *Igh* recombination[83]. Thus, a substantial part of the enormous diversity of antibody specificities (and similarly, of TCR specificities) is explained by the combinatorial variety of V, D, and J elements in the *Igh* and *Igk* loci. However, other mechanisms also contribute to this diversity, including class-switching recombination, which switches the original IgM isotype to more specialised IgD, IgG, IgA, and IgE antibody isotypes, and somatic hypermutation, which enhances antibody affinity by introducing point mutations in the antigen recognition sites. Both processes take place in activated mature B-cells and critically depend on T_H cell-mediated stimulation[84-86].

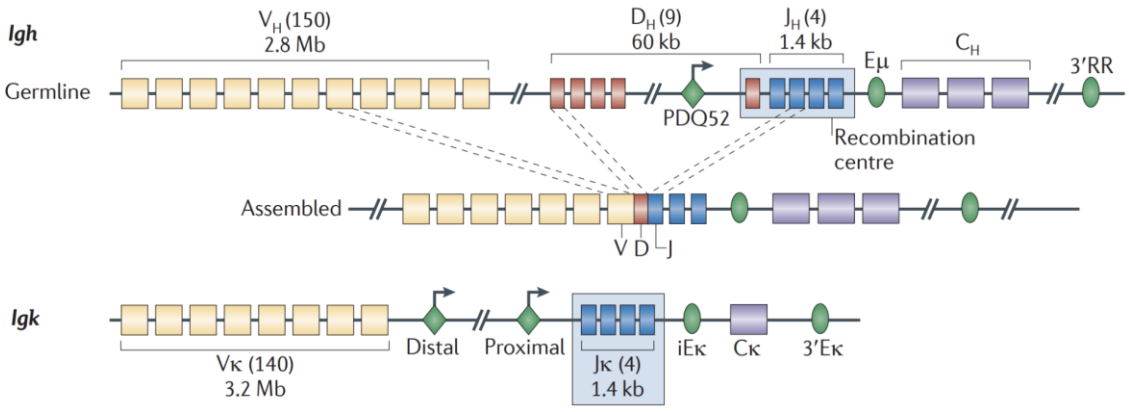


Fig. I6 | Schematics of V(D)J recombination of the *Igh* and *Igk* loci
Germline V, D and J elements (*Igh*) or V and J (*Igk*) elements are stochastically rearranged by RAG1 and Rag2 recombinases to assemble functional *Igh* and *Igk* loci. Retrieved from Schatz *et al.* (2011)[87].

In pre-B cells, the proliferation versus recombination programmes are orchestrated by the interplay of the pre-BCR and IL7R. As discussed above, the latter promotes proliferation and survival by activating the Janus kinase 3 (JAK3)/STAT5 and Phosphatidylinositol 3-kinase (PI3K)/AKT pathways. Concomitantly, IL7/IL7R/PI3K suppresses V(D)J recombination through AKT, which phosphorylates FOXO1, a transcription factor that induces RAG1/2 expression, and mediates its proteasomal degradation[80]. In contrast, pre-

BCR and CXCL12/CXCR4 signalling collaborate to activate ERK1/2 (Extracellular signal-regulated kinases 1/2), which mediates cell cycle exit, and induce the expression of IRF4 (Interferon regulatory factor 4), which facilitates, together with TCF3 and PU.1, the opening of the *Igk* locus, making it accessible to RAG1/2[75,88].

1.3.2. Transcriptional regulation of B-cell development

The first step in the generation of a B-cell is taken by Ikaros, the trigger of lymphoid potential. Ikaros is a zinc finger protein that primes MPPs to the lymphoid fate and critically contributes to their development into CLPs through both transcriptional activation and repression[62,89,90]. Ikaros induces the expression of the transcriptional repressor GFI1 (Growth factor independent 1), and they subsequently collaborate to repress PU.1 and, thereby, the myeloid potential, through two different mechanisms. First, GFI1 directly interacts with PU.1 to inhibit its transactivation activity[91,92]. Second, GFI1 binds to multiple regulatory regions of the *Spi1* gene (encoding for PU.1), including its promoter, and directly interferes with the PU.1 autoregulatory loop that controls its expression[92]. Conversely, increased PU.1 expression displaces GFI1 from these regions to promote the myeloid fate, highlighting how the early lymphoid versus myeloid bias relies on the balance between the activities of Ikaros/GFI1 and PU.1 (see **Fig. I7**).

Downstream of lymphoid priming by Ikaros and GFI1, E-proteins (especially TCF3 (or E2A)) collaborate with these factors to strengthen the lymphoid potential. While the expression of Ikaros, GFI1 and PU.1 is unaffected in the CLPs of TCF3-deficient mice, many of their canonical targets (including *Rag1* and *Dntt*) are not properly expressed in these cells, indicating that the contribution of TCF3 is essential for this gene regulatory network[93]. Next in the cascade, TCF3 induces the transcription of FOXO1, and it in turn activates the expression of EBF1, the first B-cell-specific factor, although TCF3 is also able to directly activate EBF1 itself[37,94]. Later, FOXO1 critically participates in the cell cycle exit of committed progenitors and in the induction of the programmes that facilitate V(DJ) recombination. Thus, FOXO1 not only guides early development by inducing EBF1 but also coordinates the mechanisms that control its progression[95]. Indeed, one of the key targets of EBF1 is FOXO1, so the induction of EBF1 by FOXO1 actually creates a self-reinforcing feedback that stabilises the specification of these progenitors to the B-cell pathway[96].

EBF1 expression provokes dramatic changes in the nucleus of the CLP, leading to its differentiation into the first truly unipotent B-cell progenitor: the pre-pro-B cell, which already shows features of B-lineage cells, including B220 expression in its membrane[97]. *Ebf1*$^{-/-}$ mice display a complete absence of peripheral B-cells due to an arrest at the pre-pro-B cell stage. In these mice, both CLPs and pre-pro-B cells show severely impaired expression of multiple early B-cell genes and are completely unable to perform D_H-to-J_H recombination[97,98]. Indeed, ectopic EBF1 expression in HSCs has been shown to enforce B-cell differentiation at the expense of other fates, and can partially overcome the B-cell deficiencies caused by knocking out other key regulators of early B-cell development, such as IL7Rα, Ikaros and PU.1[99-103]. However, EBF1 operates in the context of a complex transcription factor network that requires the cooperation of all its core members. For example, although EBF1 expression partially rescues B-cell development in Ikaros-null mice, these progenitors fail to become committed to the B-cell lineage and to properly undergo V(D)J recombination[101]. Furthermore, EBF1 requires the collaboration of E-proteins, among others, to regulate the expression of its target genes[104,105].

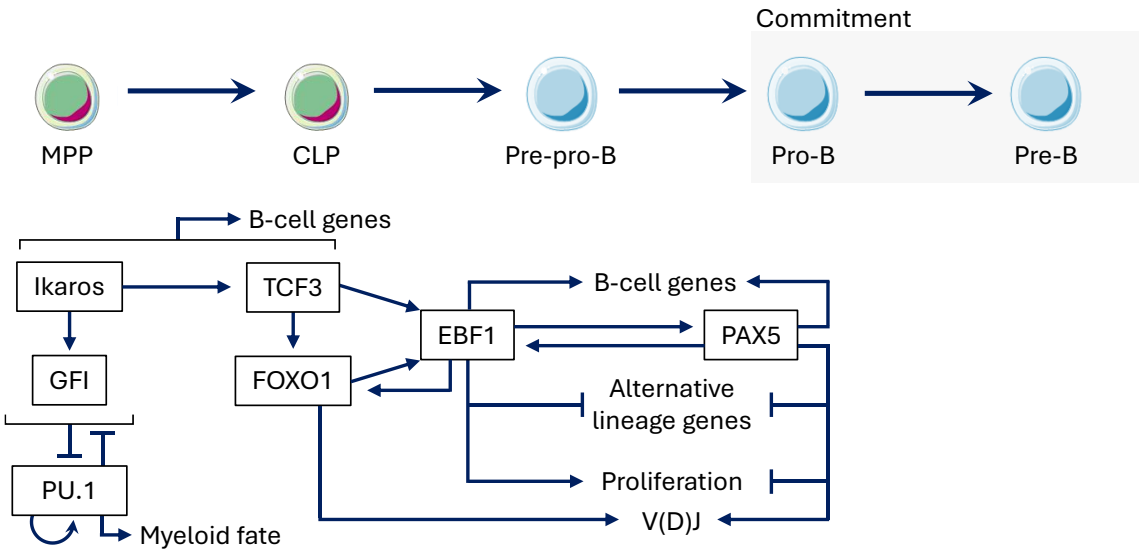


Fig. I7 | Transcriptional control of early B-cell development
In MPPs, Ikaros induces the expression of the transcriptional repressor GFI, which represses PU.1 expression and thereby the myeloid fate. The balance between Ikaros-GFI and PU.1 expression determines lymphoid-versus-myeloid differentiation. In lymphoid-biased MPPs, Ikaros induces TCF3 expression, and they collaborate to induce a primitive B-lymphoid gene expression programme. TCF3 induces FOXO1 and EBF1, which is also induced by FOXO1. EBF1 specifies CLPs to pre-pro-B cells by inducing early B-lymphoid genes, including PAX5, and repressing lineage-inappropriate genes. Upon PAX5 expression in pro-B cells, Ikaros, TCF3, EBF1 and PAX5 collaborate to drive lineage commitment and the B-cell developmental programs, including the coordinated regulation of proliferation and V(D)J recombination.

EBF1 paves the way for unipotency by repressing lineage-inappropriate genes, thus creating a scenario of increasingly B-cell specific gene expression, with genes associated with alternative fates being progressively silenced and genes required for differentiation progressively activated[102,106]. In fact, EBF1 behaves as a pioneer transcription factor because it can bind to regions of closed chromatin and poise them for activation, even when ectopically expressed in T-cells[106-108]. Two elegant reports described that EBF1 performs its pioneering functions in two steps: first, it transiently binds to genomic regions of inaccessible chromatin (as well as to accessible regions that need to be repressed)[109]. Secondly, EBF1 leverages its highly unstructured C-terminal domain to form the so-called liquid-liquid phase-separated (LLPS) condensates, which behave as transient, membrane-less *organelles* with key functions in chromatin regulation. In the context of these LLPS condensates, EBF1 recruits to its target loci BRG1, the catalytic subunit of a crucial chromatin remodelling complex (the BRG1-associated factor (BAF) complex), which catalyses the repositioning of the surrounding nucleosomes and thus enables the accessibility of other transcription factors to these loci[110]. Not surprisingly, EBF1 has recently been described to contribute to the three-dimensional organisation of chromatin in pro-B cells[111].

Thus, specification to the B-cell lineage culminates in differentiation into pre-pro-B cells when EBF1 is expressed. In pre-pro-B cells, EBF1 mediates a decisive function from which there is no return: the activation of the PAX5 gene[112]. PAX5 is the master regulator of B-cell differentiation and commitment, and it governs virtually all the processes that occur during B-cell development[113]. Its expression starts at the pro-B cell stage and is restricted to B-cells within the haematopoietic system. The first function that PAX5 plays in pro-B cells is to further increase EBF1 expression, thus creating a strong positive feedback loop[37]. Together, PAX5 and EBF1 then mediate the completion of B-cell lineage commitment. Indeed, while *Pax5*$^{+/-}$ pro-B cells display normal commitment, the combined heterozygosity of PAX5 and EBF1 leads to lineage promiscuity and partial developmental arrest[114]. In *Pax5*$^{-/-}$ mice, B-cell development is completely arrested in pro-B cells, and these pro-B cells are not committed to the B-cell fate, as they can differentiate into several alternative lineages, including T-cells, NK cells and macrophages[115-117]. Further, conditional deletion of PAX5 abolishes lineage commitment even in mature B-cells[3].

PAX5 coordinates B-cell development and prevents malignant transformation of B-cell progenitors through different mechanisms, all of them based on genome regulation. It

interacts with epigenetic regulators (such as the BAF remodelling complex, the histone acetyltransferase CBP (CREB-binding protein) and the nuclear receptor corepressor NCOR1, among others) and guides them to shape the chromatin landscape of B-cell progenitors, meaning that it promotes the deposition or the elimination of active and repressive histone marks by recruiting histone modifiers to its target loci[118]. Thus, PAX5 behaves as both a transcriptional repressor and activator. Importantly, PAX5 regulates distinct transcriptional programmes during early and late differentiation[119]. Indeed, it has been shown to have an impressively stage-specific genomic occupancy, with less than half of its binding sites overlapping between early progenitors and mature B-cells[120]. In addition, PAX5 plays a key role in the establishment and dynamics of the three-dimensional chromatin organization of B-cell progenitors[81,120]. Hence, PAX5 binding along differentiation mediates the dynamic changes in genome organisation and gene expression that take place during differentiation.

Many examples of how genome regulation by PAX5 drives B-cell development have been described. For instance, PAX5 regulates pro-B cell expansion by repressing the expression of Myc, a central mediator of proliferation in pro-B cells, and this is opposed by EBF1, that binds to the same regulatory regions to promote *Myc* expression and pro-B cell proliferation[121]. PAX5 also participates critically in V_H-D_HJ_H recombination by facilitating the contraction of the whole *Igh* locus through the repression of a single gene: the cohesin release factor *Wapl* (Wings Apart-Like Protein Homolog). While the cohesin complex is responsible for the looping of the *Igh* locus, Wapl evicts this complex from chromatin to terminate its functions. In PAX5-deficient pro-B cells, Wapl impairs cohesin-mediated locus contraction, whereas Wapl repression by PAX5 allows cohesin to contract the *Igh* locus, leading to proper V_H-D_HJ_H recombination and to structural changes in the whole pro-B cell genome[81,82,122]. After the transition to pre-B cells, PAX5 is still at the forefront of B-cell development regulation by directly inducing the expression of *Blnk*, one of the key signalling proteins that regulate pre-BCR signalling[123]. Finally, PAX5 has been reported to prevent malignant transformation of B-cell progenitors, together with Ikaros and EBF1, by directly controlling glucose uptake and ATP production, thereby restricting their metabolic demands[124,125].

In summary, from the perspective of transcriptional regulation, B-cell development involves the prevalence of a lymphoid bias in MPPs, which facilitates the emergence of a primitive and labile B-lymphoid transcriptional programme. Following specification by

EBF1, the genomic landscape becomes increasingly refined, and, ultimately, the expression of PAX5 dynamically orchestrates the architectural and transcriptional changes that ultimately give rise to a mature B-cell.

2. Leukaemia

Haematological malignancies are a heterogeneous group of cancers that arise from defects in haematopoietic progenitors or mature blood cells. They include acute leukaemias, chronic leukaemias, lymphomas and myelodysplastic syndromes, among others. Depending on the cell of origin, these malignancies can be classified as lymphoid or myeloid neoplasms. For instance, defects at different stages of B-cell differentiation lead to B-cell leukaemia, lymphoma or myeloma, whereas T-cell leukaemia or lymphoma and monocytic leukaemia result from the malignant transformation of T-cells or monocytes, respectively. Lymphoid and myeloid leukaemias are further categorized in acute (lymphoid, ALL; myeloid, AML), when they emerge from an undifferentiated progenitor, and chronic (CLL and CML), that result from the malignant transformation of a mature cell[126].

Contrary to other cancers, haematological malignancies are prevalent in children, where they account for the 40% of all cancers, although some of them are also common in adults. In children, B-ALL comprises about 70% of all leukaemia cases, followed by acute myeloid leukaemia (AML), which accounts for 15-20% of the cases[127]. These two types of leukaemia originate from the uncontrolled proliferation of poorly differentiated lymphoid (in B-ALL, mainly pro-B or pre-B cells) or myeloid progenitors, in which the interaction of germline and somatic mutations produces a differentiation block that leads the accumulation and transformation of these progenitors. Indeed, many of the most prevalent mutations in AML and B-ALL disrupt or alter the activity of the key transcription factors and epigenetic modifiers that control their differentiation such as *IKZF1* (the gene encoding for Ikaros), *PAX5*, *CEBPA* or *TET2* (Ten-eleven translocation 2)[127,128]. Furthermore, haematological malignancies often carry large chromosomal rearrangements that lead to the creation of fusion proteins (such as ETV6::RUNX1, BCR::ABL1 or TCF3::PBX1), as well as to complex karyotypes, including hyperploidy and hypoploidy[127]. However, these mutations do not appear to be sufficient for the development of overt leukaemia, as clonal oncogenic mutations in haematopoietic progenitors are observed in about 5% of newborns, but most of them never develop leukaemia[129]. Yet, these mutations induce a pre-cancerous state that predisposes to the development of leukaemia upon the acquisition of secondary mutations, which can be triggered by exposure to infections, among other factors[130,131].

From a translational perspective, haematological malignancies have been at the forefront of the therapeutic discovery for several decades. An outstanding example is the development of ABL1 kinase inhibitors, which have greatly improved the outcome of several types of leukaemia carrying the BCR::ABL1 mutation. In these patients, the BCR::ABL1 fusion, also known as the Philadelphia (Ph) chromosome, leads to constitutive activation of ABL1, which can be specifically targeted using these specific inhibitors[132]. The treatment of several haematological malignancies has also improved significantly with the advent of immunotherapies, which harness the host immune system to fight cancer using a variety of strategies. For instance, one of the major advances in anti-cancer therapies in recent years has been the development of chimeric antigen receptor T-cells, or CAR T-cells. In the case of B-ALL, anti-CD19 CAR T-cell-based therapies have been shown to significantly improve patient outcomes[133]. This therapy is based on the collection of T-cells from the patient and the *ex vivo* expansion and genetic modification of these cells, which are transduced with a chimeric TCR-based construct. This construct contains the variable region of an antibody against CD19, a surface marker expressed by all B-cells, including B-ALL blasts. Thus, when these genetically modified T-cells are reintroduced into the patient, they target cancer cells and mount a cytotoxic response against them[134]. Thanks to these advances, the prognosis for many haematological malignancies that were incurable 50 years ago is becoming increasingly favourable. However, due to the enormous diversity of these malignancies, some subtypes still have intermediate or poor prognoses, so extensive research is still required to improve the survival of these patients.

2.1. B-cell acute lymphoblastic leukaemia

B-ALL is the most common cancer in children and the second most common acute leukaemia in adults[135,136]. Although there have been great advances in the treatment of B-ALL over the last decades, many patients still relapse after an initial remission, making overall long-term survival an ongoing challenge, especially in the case of adults[135]. In addition, patients over the age of 60 often fail to tolerate standard chemotherapy regimens, which forces to reduce their doses and leads to suboptimal responses and to a very poor long-term survival of around 10-15%[137].

B-ALL comprises a vast number of different molecular subtypes, which are dramatically different in terms of long-term survival and in their frequencies in children and adults[138].

Most of these subtypes are defined by the presence of recurrent chromosomal aberrations, in particular gene translocations (such as KMT2A translocations and BCR::ABL1, TCF3::PBX1 or ETV6::RUNX1 fusions) and chromosome number alterations (including low hypodiploidy and hyperploidy B-ALL)[139]. The presence of such genetic aberrations allows for an increasingly specialised stratification of these patients. However, about a 30% of B-ALL cases, the so-called B-other cases, do not carry any of these aberrations, so they cannot be stratified according to this criterion[140]. Fortunately, the advent of deep sequencing techniques and massive gene expression profiling has more recently led to the identification of novel molecular subtypes based on specific mutations or gene expression patterns. One such case is the Ph-like class, one of the worst prognostic B-ALL subtypes, accounting for approximately 10% and 20% of childhood and adult cases, respectively[141,142]. This subtype is characterised by a gene expression pattern that closely resembles that of the BCR::ABL1 (or Philadelphia chromosome-positive, Ph^+) subtype. Even though Ph-like ALL lacks the BCR::ABL1 translocation, this subtype comprises cases that display a variety of mutations that lead to constitutive activation of the ABL kinase and thus phenocopy the BCR::ABL1 translocation. Indeed, BCR::ABL1 inhibitors seem to work well in preclinical models and are currently in clinical trials for the treatment of Ph-like cases[143].

As anticipated above, B-ALL is a *dual-hit* disease. It is well documented that driver oncogenic mutations in B cell progenitors are not sufficient to generate overt leukaemia. Instead, these mutations initiate a clonal, pre-malignant state that requires secondary mutations to develop into disease[144]. In B-ALL, the most common secondary mutations occur in B-lymphoid transcription factors. For example, mutations in the B-lymphoid transcription factor PAX5 occur in 30% of all B-ALL cases and are even more prevalent in BCR::ABL^+ B-ALL, where more than half of the cases show PAX5 alterations[145,146]. Similarly, mutations affecting other central B-cell development regulators, including EBF1, TCF3 or Ikaros have been detected in B-ALL, and more than 80% of BCR::ABL1 cases have been shown to display Ikaros mutations, suggesting a close link between these alterations[146]. It is therefore not surprising that the defective functions of these key genes lead to a developmental arrest that facilitates the malignant transformation of pre-leukemic undifferentiated blasts. Indeed, much evidence has shown that these transcription factors actually act as potent tumour suppressors in different ways. In mice, compound heterozygous loss of PAX5 and EBF1 produces B-ALL, and PAX5 heterozygosity cooperates with oncogenic mutations (such as BCR::ABL1 expression or constitutive STAT5 activation)

to induce B-ALL with almost complete penetrance[124,147-149]. Conversely, reconstitution of PAX5 physiological levels induces leukaemia regression in human and mouse B-ALL cells and restores the differentiation capacity of malignant B-ALL cells in a mouse model of B-ALL[125,150]. From a functional perspective, some of the roles that PAX5 normally plays in B-cell progenitors serve themselves to prevent malignancy, such as its ability to repress the *Myc* locus and thereby limit proliferation, or its participation in V(D)J recombination[81,121]. In addition, PAX5 and Ikaros prevent oncogene-driven malignant transformation of B-cell progenitors by imposing a chronic energetic constraint, but inactivating mutations in these genes in B-ALL relieve this metabolic constraint and critically contribute to B-ALL malignancy[125]. Mechanistically, PAX5 and Ikaros bind to the promoters of several genes controlling glucose uptake and metabolism to repress them. In B-ALL, dominant-negative mutations of PAX5 and Ikaros, as well as heterozygous deletion of *Pax5*, revert the repression of these genes and thereby facilitate a metabolic permissiveness that sustains oncogenic progression and that can be abrogated by restoring the expression of wild-type PAX5 and Ikaros[125]. Thus, PAX5 and other B-lymphoid transcription factors may be attractive targets for novel therapeutic strategies to fight B-ALL.

Importantly, unlike many tumour suppressors that that typically undergo biallelic deletions, the vast majority of *PAX5* inactivating mutations are monoallelic[145,151], which opens a compelling window for reinforcing the functions of the *Wt PAX5* allele and thus circumvent its haploinsufficiency. However, although many studies have addressed how PAX5 gene expression is regulated and how it drives B-cell development and relates to human leukaemia, almost nothing is known about how the PAX5 protein is regulated, and no upstream regulators have been described. Therefore, there is a lack of rationale for modulating PAX5 functions in B-ALL, which has prevented its therapeutic exploitation.

3. Epigenetics

A long-standing question that has been and still is a matter of extensive research is how the complexity of living beings can be explained from their genomes; how a single genome is able to give rise to billions of extraordinarily specialised cells, with hundreds of different identities and functions, from a single cell. One of the prevailing answers points to genome regulation, which allows every single cell type to express a particular set of genes that defines its identity and functions. Thanks to the development of high-throughput DNA sequencing techniques, as well as increasingly sophisticated imaging, molecular biology and biochemical techniques, in the last two decades we have witnessed extraordinary advances in our understanding of genome regulation. Thus, our vision of genome regulation nowadays includes very diverse mechanisms: DNA and RNA modifications, complex combinations of histone modifications, histone replacements, nucleosome sliding, three-dimensional chromatin organization, liquid-liquid phase separation (LLPS)-mediated transcriptional control, etc. All these phenomena are the object of study of epigenetics, the research field that seeks to understand how chromatin modifications affect cellular and organismal function in health and disease. The term epigenetics was initially coined by Waddington, who referred to "the branch of biology which studies the causal interactions between genes and their products, which bring the phenotype into being"[152]. Currently, epigenetics is defined as the study of inheritable modifications of the chromatin that do not involve changes in the DNA sequence. These modifications are highly cell-type specific and play a central role in physiology and pathology by regulating nearly all aspects of genome biology, including chromatin accessibility, gene silencing, transcription or RNA splicing and folding, which ultimately controls the functions and identity of the cell. Indeed, epigenetic drugs are emerging as promising targets for anti-cancer therapies, and mutations in epigenetic regulators are among the leading causes of malignant transformation and cancer progression[153].

3.1. Chromatin and the nucleosome

Eukaryotic DNA is packed within the cellular nucleus in the context of a macromolecular structure named chromatin, which is composed of genomic DNA, histones and accessory proteins, and whose fundamental and repetitive unit is the nucleosome. Nucleosomes consist of a protein octamer containing two copies of each core histone (H2A, H2B, H3 and H4), that is wrapped by 145-147 base pairs (bp) of DNA and followed by a linker DNA

of variable length[154]. In the context of nucleosomes, there is an extensive interaction interphase between DNA and histones, in part because histones are highly basic proteins (rich in lysine and arginine residues) with abundant positive charges, allowing for the establishment of electrostatic interactions with the negatively charged DNA[154]. In addition, histones have highly unstructured N- and C-terminal tails that protrude out from the nucleosome core but play, as described below, a crucial role in its regulation (**Fig. I8**).

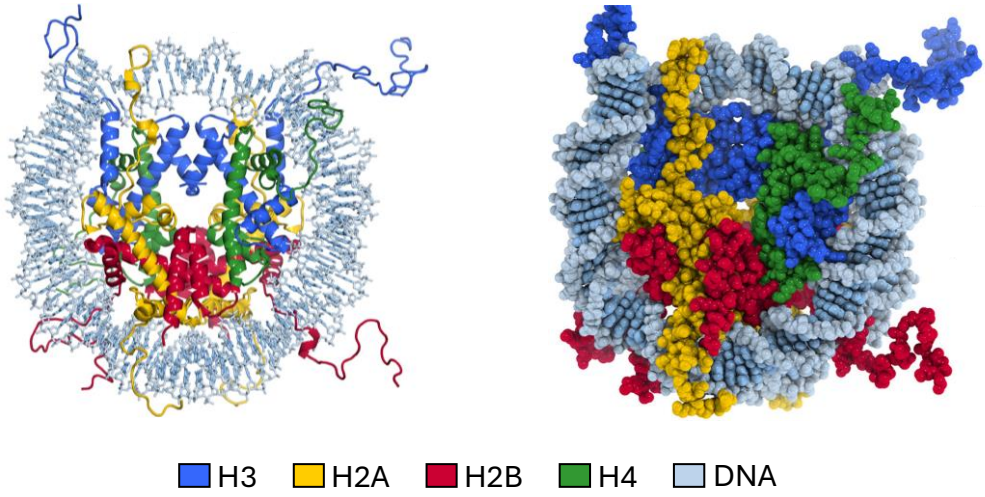


Fig. I8 | Nucleosome structure
Cartoon and sticks (left) and molecular surface (right) representations of the nucleosome core crystal structure. Adapted from McGinty *et al.* (2015)[155].

Chromatin is organised into several levels of compaction. First, nucleosome arrays adapt a conformation known as the 11 nm fiber or *beads on a string* structure[156]. Secondly, a 30 nm fiber has been proposed as the next level of compaction, which could be stabilised by linker H1/H5 histones and has been shown *in vitro* to achieve a 6-fold compaction relative to the 11 nm fiber[157,158]. Finally, metaphasic cells condense their chromatin by folding it into 300-700 nm fibers, which eventually compact into chromosomes. In addition to these hierarchical levels of chromatin compaction, chromatin is structurally and functionally organised into two different types: euchromatin and heterochromatin. Euchromatin comprises those regions of the genome that are relaxed or *open*, making these regions more accessible to their regulators and thereby transcriptionally active. Conversely, heterochromatin consists of highly condensed chromatin, which is therefore inaccessible and transcriptionally silent, and can be further classified as constitutive or facultative heterochromatin[159]. The former is mostly found in telomeres and pericentric chromosomal regions and plays an important architectural role, so it is conserved across cell types,

whereas the latter is highly dynamic and cell type specific. Thus, facultative heterochromatin can be dynamically decondensed in different cell cycle or developmental stages to facilitate the timely and cell-type-dependent expression of specific sets of genes[160,161].

3.2. Epigenetic regulation of gene expression

Chromatin is a highly dynamic structure that undergoes extensive epigenetic modifications to regulate the gene expression patterns that define cell identities and allow cells to respond to environmental stimuli. Although many different forms of epigenetic regulation have been identified, the two most prevailing mechanisms of chromatin regulation are DNA methylation and post-translational modification of histones. The first of these mechanisms involves the covalent addition of a methyl group into a cytosine nucleotide in the DNA, which usually leads to gene silencing[162]. This process is catalysed by the family of DNA methyltransferases (DNMTs) and removed by TET family enzymes, and defects in their functions have been linked to disease[163]. For example, global hypomethylation and locus-specific hypermethylation is a common hallmark of cancer cells, and TET2 mutations are among the most common genetic alterations in AML and other haematological neoplasms[164,165].

In contrast, the group of histone post-translational modifications comprises a plethora of different modifications with diverse effects in chromatin biology. Histone modification involves the covalent addition of chemical groups (acetyl or other acyl groups, methyl, phosphate, ubiquitin, SUMO, ADP-Ribose, etc) into specific residues of core histones, usually in the N- and C-terminal tails of H3 and H4. The deposition of these modifications (or histone marks) is catalysed by *writer* enzymes (histone acetyltransferases (HATs), methyltransferases, kinases, ubiquitin ligases, etc.), removed by *eraser* enzymes (e.g. histone deacetylases (HDACs), demethylases, phosphatases, deubiquitinases) and interpreted by *readers* (proteins containing bromodomains, chromodomains, etc.)[166]. The concerted action of these enzymes results in complex combinations of histone marks that have profound effects in the structure and function of the chromatin. Indeed, many of these modifications are used to define functional regions of the genome, since they tend to be specifically enriched in particular regions[167] (**Fig. I9**).

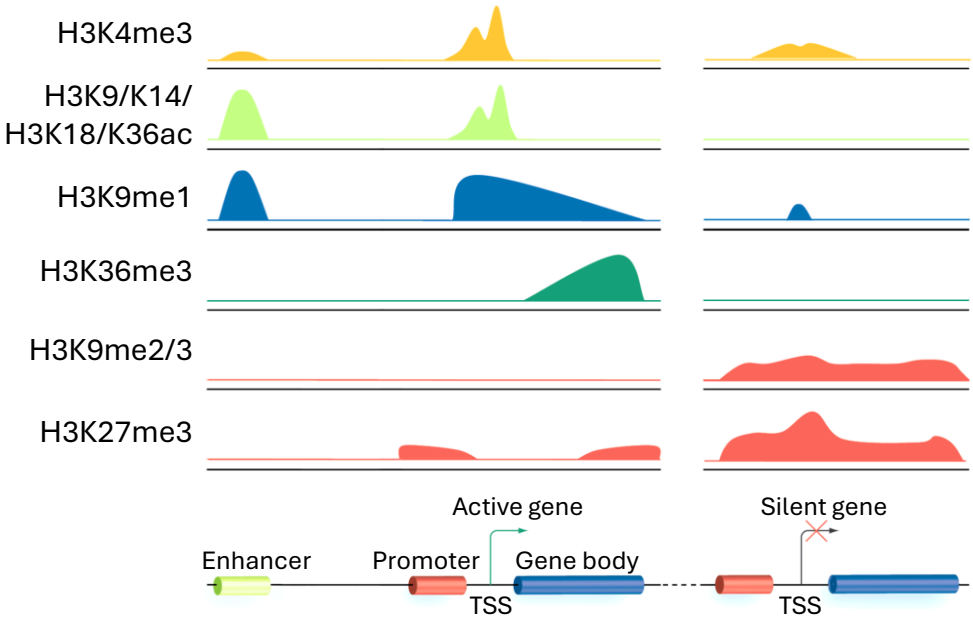


Fig. I9 | Genomic distribution of histone modifications
Schematic representation of the distribution of some histone modifications in active enhancers, promoters, TSSs and gene bodies, as well as in the promoters, TSSs and gene bodies of repressed genes. Adapted from Lim *et al.* (2010)[168].

3.2.1. Histone acetylation

Histone acetylation takes place at specific lysine residues and consists of the covalent addition of an acetyl group from an acetyl-coenzyme A (Ac-CoA) into the ε-amino group of the lysine side chain. As a consequence, the positive charge of the lysine is neutralised, and the histone-DNA interaction is weakened, which increases chromatin accessibility and typically provides a docking site for transcription factors containing bromodomains[169]. Thus, histone acetylation tends to be enriched in the regulatory regions of active genes, such as their promoters and enhancers, and is associated with active transcription, although it is also actively involved in DNA repair, mitosis, and other processes[170].

Acetylation, and more generally, the modification of each histone lysine residue is collectively regulated by the set of writers and erasers that catalyse their addition or removal, although different histone marks can colocalise in the same regions and even influence the regulation of nearby marks. For instance, active enhancers display high levels of acetylated H3-Lys27 (H3K27ac) but also contain other marks such as H3K9me1 (monomethylated H3-Lys9), whereas H3K14ac is more enriched in promoters and gene bodies[171,172]. Other examples of histone acetylation marks that are commonly enriched in

Table I1 | Summary of HATs and HDACs. Some of the main HATs and HDACs, as well as their histone substrates and effects on chromatin regulation are reported. Adapted from Kouzarides et al. (2007)[173], Hyndman *et al.* (2017)[174] and Seto *et al.* (2014)[175].

Family	HAT/HDAC	Histone substrates	Remarks
GNAT	KAT2A/GCN5	H2B, H3K9/K14/K18/K27/K36, H4K8/K12/K14	Transcriptional activation
	KAT2B/PCAF	H2B, H3K9/K14/K18, H4K8	
P300/CBP	KAT3A/CBP KAT3B/p300	H2AK5, H2BK12/K15, H3K14/K18/K27/K56 H4K5/K8/K12/K16	Transcriptional activation
MYST	KAT5/Tip60	H2AX, H2AZ, H3K14, H4K5/K8/K12	Transcriptional activation, DNA repair
	KAT6A/MOZ	H3K9/K14, H4	Transcriptional activation
	KAT6B/MORF	H4K5/K8/K12/K15	
	KAT7/HBO1	H3, H4K5/K8/K12	Transcriptional activation, DNA repair
	KAT8/MOF	H4K16	Transcriptional activation, DNA repair, cell cycle
Transcription factors	KAT4/TAF1	H3K14, H4	Transcriptional activation
	KAT12/TFIIIC90/GTF3C4	H3K9/K14/K18	RNA-Pol III-mediated transcription
Nuclear receptor co-activators	KAT13A/SRC1/NCOA1	H3K9/K14, H4	Transcriptional activation
	KAT13B/SCR3/NCOA3	H3, H4	
	KAT13C/p600		
	KAT13D/CLOCK	H3K14	
Cytoplasmic HATs	KAT1/HAT1	H2A, H4K5/K12	DNA repair, histone deposition
	HAT4/KAA60	H4	Mitosis
Class I	HDAC1	H3ac, H4K16ac	Transcriptional repression
	HDAC2	H3K9/K56, H4K16ac	
	HDAC3	H2A, H3K9/K14ac, H4K5/K12ac	Transcriptional repression, heterochromatin, DNA repair, cell cycle
	HDAC8	H2Aac, H2Bac, H3ac, H4K16/K20ac	Transcriptional repression, cell cycle
Classs IIa	HDAC4	H3ac	Transcriptional repression, DNA repair
	HDAC5	-	Transcriptional repression, heterochromatin, DNA replication
	HDAC7		Transcriptional repression
	HDAC9	H3ac, H4ac	
Class IIb	HDAC10	H4ac	Transcriptional repression
	HDAC6	-	DNA repair, cell migration
Class IV	HDAC11	H3ac, H4ac	Transcriptional repression, DNA replication
Class III	SIRT1	H3K9/K14/K56ac, H4K16ac, H1K26ac	Transcriptional repression, DNA repair, heterochromatin
	SIRT2	H3K56ac, H4K16ac	Chromosome condensation, DNA repair
	SIRT3	H3K9ac, H4K16ac	Transcriptional repression, DNA repair
	SIRT4	-	-
	SIRT5	-	-
	SIRT6	H3K9/K18/K56ac	Transcriptional repression, telomeric and pericentric chromatin, DNA repair
	SIRT7	H3K18/K36ac	Transcriptional repression,heterochromatin, rRNA regulation

accessible chromatin regions include H3K4ac, H3K9ac, H3K18ac and H3K36ac, as well as H4 modifications, such as H4K12ac and H4K16ac[172]. Similarly, an increasing number of other histone lysine acylations have been reported, including butyrylation, crotonylation and succinylation. Like acetylation, these modifications have been shown to promote chromatin opening[176-178].

Contrary to other histone marks, acetylation is highly dynamic, and its levels at each locus ultimately depend on the balanced action of the responsible HATs and HDACs. In many cases, both HATs and HDACs work in the context of high-molecular-weight multi-protein complexes, such as the Nucleosome remodelling and deacetylase (NuRD) complex, that regulate their catalytic activity and direct their specificity towards specific loci[169]. In the human genome, there are 21 putative HATs that can be classified in three major families based on the homology of their catalytic domains: GNATs (General control non-derepressible 5 (Gcn5)-related N-acetyltransferases), p300/CBP and MYST. Furthermore, some nuclear receptor co-activators have also been shown to have histone acetyltransferase activity, despite lacking consensus HAT domains, and some transcription factors carrying HAT domains belong to an "orphan" HAT family[179,180] (summarised in **Table I1**).

The human proteome contains 18 HDACs, and these are catalogued in four classes, also based on their homology[181]. Class I HDACs (HDAC1-3 and 8) are ubiquitously expressed in the nucleus and display strong histone deacetylase activity, so they generally act as transcriptional repressors, although they also target a few non-histone substrates[182]. Class II HDACs (IIa, HDAC4, 5, 7 and 9; IIb, HDAC6 and 10) are tissue-specific, primarily cytoplasmic and display very weak catalytic activity, especially class IIa HDACs, for which very few substrates have been identified. However, the knockout of several class II HDACs produces severe phenotypes in mice, indicating that they may work independently of their enzymatic activity, play dominant negative functions or be activated in response to specific stimuli[183]. HDAC11 is the only member of class IV HDACs. Although its functions are not well understood, it displays tissue-specific expression and has been reported to be important for immunity and cancer, but its targets remain to be identified[184]. Whereas the catalytic domain of class I, II and IV HDACs shares a similar enzymatic mechanism that that requires zinc as a cofactor and that can be inhibited using Trichostatin A (TSA), class III HDACs or sirtuins have a completely different HDAC domain that requires NAD^+

for its catalysis. Because of the objectives proposed for this PhD thesis, the sirtuin family of HDACs is introduced in detail in section 3.3.

Not surprisingly, several HATs and HDACs have been shown to perform important functions in haematopoietic differentiation and malignancy. For instance, MOZ sustains HSC quiescence and long-term survival, whereas CBP deficiency leads to HSC exhaustion and decreased repopulation capacity[185,186]. Interestingly, while loss of MOF, the main H4K16ac acetyltransferase, strongly impairs HSC maintenance in adults, it does not affect foetal haematopoiesis, despite a global loss of H4K16ac[187]. As discussed below, sirtuins play important roles in HSCs as well. During B-cell development, one of the major epigenetic regulators controlling this process is the histone deacetylase HDAC7. HDAC7 is a direct target of TCF3, EBF1 and FOXO1 and critically participates in transcriptional repression in B-cell progenitors. Consequently, loss of HDAC7 produces increased apoptosis and a severe differentiation block in pro-B cells[188]. Furthermore, HDAC7 expression was recently reported to be a good prognosis predictor in infant pro-B ALL carrying *MLL* translocations, a B-ALL subtype associated with poor clinical outcomes[189].

3.2.2. Histone methylation

This modification occurs in basic residues (lysine, arginine and histidine), mostly lysine and arginine residues. Contrary to other modifications, methylation does not affect the charges of these residues. Instead, it creates docking sites that facilitate the recruitment of specific readers containing chromo, tudor, WD40 or other methylation recognition motifs, although this mechanism is common to other histone modifications[190] (**Fig. I10**). Lysine residues can be mono-, di- or trimethylated (me1, me2 and me3, respectively), whereas arginine residues can be monomethylated or dimethylated symmetrically or asymmetrically (me2s or me2a), and histidine methylation remains poorly studied. Contrary to histone acetylation, methylation can promote active transcription or repress it, depending on the residue and its methylation status[191]. Thus, H3K4, H3K36 and H3K79 methylations are generally associated with active transcription, while H3K9, H3K27 and H4K20 methylations are found in regions of condensed and inactive chromatin[191]. For instance, in the case of H3K4, H3K4me3 is generally found in active promoters, whereas H3K4me1 is enriched in enhancers[192].

Lysine and arginine methylation are catalysed by KMTs (Lysine methyltransferases) and PRMTs (Protein arginine methyltransferases), respectively, that transfer the methyl group from S-adenosyl-L-methionine (SAM) to the ε-amino group of the lysine side chain or to the guanidine group of arginine. There are two classes or PRMTs: type I (PRTM1-4, 6 and 8), that catalyse arginine me1 and me2a; and type II, that catalyse the deposition of me1 and me2s[190]. KMTs are also classified in two families: DOT1L (Disruptor of telomeric silencing 1-like) and SET (Su(var)3–9, Enhancer-of-zeste and Trithorax) domain-containing proteins, that encompass 8 subfamilies: SUV39, SET1, SET2, EZ, RIZ, SMYD, SUV4-20, and the orphan members SET7/9 and SET8. Interestingly, translocations in the SET1-family gene MLL1/KMT2A (Mixed-lineage leukaemia 1) are among of the most common lesions in B-ALL and AML[193]. Another common target of inactivating mutations in haematological malignancies and other cancers is EZH2 (Enhancer of zeste homolog 2), a Polycomb-group KTM that plays a fundamental role in gene repression by catalysing the deposition of H3K27me2/3[194,195]. Indeed, *Ezh2* deletion in mice predisposes for the development of several haematological malignancies, including T-ALL and MDS[196,197].

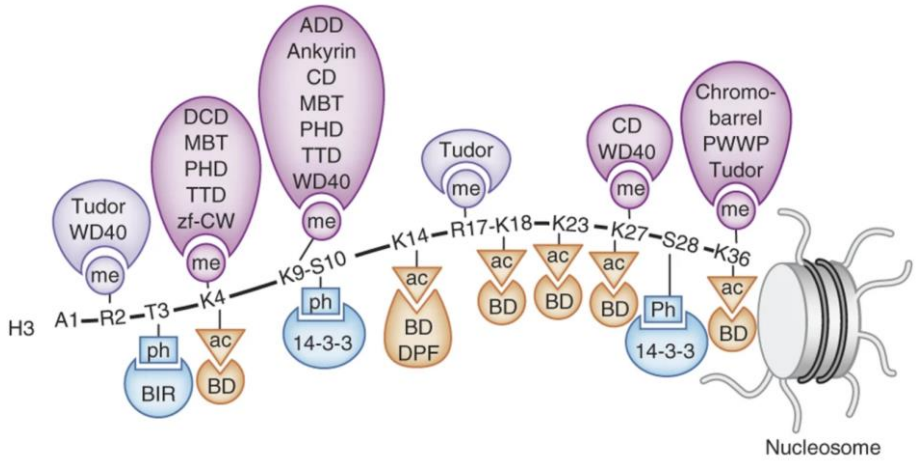


Fig. I10 | Reader domains of histone modifications
Examples of how the presence of histone modifications in specific residues drives the recruitment of proteins containing reader domains. Histone methylation is recognized by Tudor, tandem tudor (TTD), WD40, double chromodomains (DCD), malignant brain tumor (MBT), plant homeodomains (PHD), zinc-finger CW (zf-CW), ADD, Ankyrin, CD, MBT, Chromo-barrel or Pro-Trp-Trp-Pro (PWWP) domains; acetylation (ac) is recognized by bromodomains (BD) and double PHD finger (DPF) domains; and phosphorylation (ph) is recognized by 14-3-3 and Baculovirus IAP Repeat (BIR) domains, among others. Retrieved from Musselman *et al.* (2012)[198].

The removal of these modifications is catalysed by lysine-specific demethylases (LSD1 and LSD2) and Jumonji family proteins, being the Jumonji protein JMJ6 the only known arginine demethylase[199]. In the case of LSD1/2, both proteins use flavin adenine

nucleotide (FAD) as a cofactor to catalyse demethylation, and LSD1 targets H3 mono- and dimethylated lysine modifications; whereas LSD2 specifically demethylates H3K4me1 and H3K4me2. Jumonji family proteins display varying specificities. For example, JMJD3 (Jumonji domain-containing 3) and UTX (Ubiquitously transcribed tetratricopeptide repeat, X chromosome) target H3K27me2 and H3K27me3, whereas JMJD5 specifically targets H3K36me2[200-202]. In haematopoiesis, some of these enzymes seem to compensate each other. However, the combined knock out of *Kdm4a*, *Kdm4b* and *Kdm4c* in HSCs has been shown to impair the long-term repopulation capacity of HSCs, since *Kdm4a/b/c*-deficient HSCs generate significantly reduced numbers of myeloid and lymphoid cells[203].

3.2.3. Histone phosphorylation

Histone phosphorylation consists in the transfer of a phosphate from ATP to the hydroxyl group of the side chain of serine, threonine and tyrosine residues, which significantly increases the negative charge of the residue and thereby affects chromatin structure. Phosphorylation is catalysed by protein kinases and reversed by protein phosphatases (PPs), which are generally subjected to upstream signalling pathways that control their activation. This is a highly dynamic modification with a very rapid turnover, allowing fast responses to environmental stimuli and facilitating the coordination of complex cellular processes such as DNA replication, mitosis, meiosis and DNA repair, although it is also associated with transcription[204-206]. For instance, the repair of DNA double-strand breaks (DSBs) involves the coordination of multiple proteins to sense the break, repair it and terminate the signal. This is regulated at multiple levels by several enzymes, including kinases and phosphatases. Following DNA damage, the H2AX histone variant is phosphorylated at S139 (producing the so-called γH2AX form) by several kinases, including ATR (Ataxia telangiectasia and Rad3-related) and ATM (Ataxia-telangiectasia mutated), to delimitate the damaged region[207]. γH2AX then increases the accessibility of the region and recruits the DNA repair machinery, which culminates with the dephosphorylation of γH2AX by PP2A and PP4C and with the inactivation of upstream kinases to terminate the signal[208,209]. Another example of how histone phosphorylation contributes to orchestrate complex cellular processes is H3S10 phosphorylation by Aurora kinases, which is required for chromosome condensation during mitosis and helps to coordinate the regulation of gene expression in a dynamic manner along the cell cycle[210,211].

3.2.4. Histone ubiquitination and ubiquitin-like modifications

Ubiquitination is associated with almost all cellular processes in eukaryotes. It involves the addition of one or more ubiquitin molecules to a lysine residue of a protein. The catalytic cascade of ubiquitination involves the hierarchical activities of ubiquitin-activating enzymes (E1), ubiquitin-conjugating enzymes (E2), and ubiquitin-ligases (E3), that covalently attach ubiquitin, a 76-amino acid protein, to the ε-amino group of a lysine residue[212]. Conversely, deubiquitinases (DUBs) remove this modification. E3 ligases can catalyse the formation of complex combinations of ubiquitination in a single protein, including mono-, multi-ubiquitination, and a broad range of poly-ubiquitin chains (branched, mixed, etc.). Furthermore, ubiquitination can be linked to any of the 7 lysine residues of the ubiquitin, creating a whole *ubiquitin code* with a variety of effects on the regulation of the target protein[213]. Canonically, polyubiquitination has been linked with proteasomal degradation and selective autophagic degradation of the ubiquitinated protein, whereas mono-ubiquitination has been associated with chromatin regulation and protein trafficking, although the complexity of the ubiquitin code is increasingly defying this classical vision[214].

In the context of chromatin, ubiquitination is most abundant in H2A and H2B, with 10% of all H2A proteins being ubiquitinated in the cell[215]. H2A mono-ubiquitination (H2AK119ub) is catalysed by Polycomb repressive complex 1 (PRC1)[216]. It usually colocalises with H3K27me3, a repressive mark, so it is mainly associated with chromatin compaction and gene repression, although it is also present, to some extent, in regions that lack H3K27me3[217]. Interestingly, loss of PRC1 function in HSCs leads to premature expression of the B-cell transcription factors EBF1 and PAX5 due to epigenetic derepression of their loci, which accelerates B-cell specification but negatively affects HSC self-renewal[218]. Mono-ubiquitination of H2BK123 is important for transcriptional activation and RNA Polymerase (RNA Pol) II elongation, although it can also repress gene expression when it is found in gene promoters[219,220]. Furthermore, H2KB123ub plays a critical role in the deposition of H3K4 and H3K79 methylation marks, which further links this mark with transcriptional activity[221]. H2A ubiquitination is also important for DNA damage repair signalling, since H2AK15ub acts a scaffold for the recruitment of the recruitment of 53BP1 (p53 Binding Protein 1), an essential non-homologous end joining (NHEJ) DNA repair factor[222].

Ubiquitin is not the only protein that can be covalently attached to another protein to regulate its functions, but there is a whole family of ubiquitin-like proteins that also perform this regulatory function and follow similar rules to those of ubiquitin, in terms of their catalytic cycle and their cellular functions. Among them, SUMO1-4 (Small-ubiquitin-related modifier) proteins are the best studied ubiquitin-like proteins[223]. Histone SUMOylation was originally described to lead to transcriptional repression. For instance, SUMOylated H4 was reported to associate with HDAC1and facilitate its recruitment to promote to gene silencing[224]. However, recent advances in the field have uncovered a more complex role of this modification, that has been shown to be involved in chromatin remodelling and transcriptional elongation, among other processes[225].

3.2.5. Histone ADP-Ribosylation

Mono- and poly-ADP-Ribosylation are the result of the transfer of the ADP-Ribose (ADPR) unit from a NAD^+ molecule to specific lysine, arginine, glutamate, aspartate, cysteine or serine residues[226]. Interestingly, mono-ADP-Ribosylation is mostly found in cytoplasmic proteins, whereas poly-ADP-Ribosylation is mainly observed in nuclear proteins. In the case of histones, ADP-Ribosylation is a relatively rare modification that is usually deposited in response to specific stimuli, such as the induction of DNA damage, although all core histones and the linker H1 have been described to undergo this modification[227,228]. Histone ADP-Ribosylation dramatically affects chromatin structure, leading to chromatin opening, and creates a binding site for its readers, such as proteins containing macrodomains, that specifically bind to this modification[229].

Three families of proteins with ADP-Ribosyltransferase (ADPRT) activity have been described: diphtheria toxin-like ADP-Ribosyltransferases (ARTDs, also known as PAR-polymerases or PARPs), which includes 18 members with both mono- and poly-transferase activities; clostridial toxin-like ADP-Ribosyltransferases (ARTCs); and some Sirtuin family members, that besides intrinsic HDAC activity display mono-ART activity[230].

PARP1 was the first member of the family to be identified and is by far the best characterised ADPRT, in part because of its potential as pharmacological target for cancer treatment. PARP1 plays a complex role in genome integrity by stabilizing DNA replication forks, participating in chromatin remodelling and, specially, by recognizing DNA damage sites[231]. In the latter case, PARP1 recruitment to damaged sites is one of the earliest

events in DNA damage repair[232]. Upon recruitment, PARP1 poly-ADP-Ribosylates itself and both histone and non-histone targets, which promotes the *in-situ* assembly of the DNA repair machinery proteins though interactions with their ADPR-binding domains, and in some cases also activates their functions[233]. Indeed, PARP1 inhibitors have shown efficacy in DNA repair-deficient cancers, and some of them are currently approved for the treatment of several cancers[234-236].

3.2.6. Chromatin remodelling complexes

ATP-dependent chromatin (or nucleosome) remodelling complexes are the drivers of another key mechanism of chromatin regulation, that does not involve chemical modification of either DNA or histones. Instead, these complexes regulate chromatin accessibility by specifically targeting the nucleosomes to remove them, slide them across the DNA (nucleosome sliding), or even remove core histones from the nucleosome and replace them with histone variants (nucleosome editing)[237]. While the majority of the nucleosomes are assembled during DNA replication, the activity of nucleosome remodelling complexes allows fine-tuning of the composition and spacing of the nucleosomes to adjust chromatin density on demand. These features make these complexes essential regulators of virtually all aspects of chromatin biology[238]. Not surprisingly, components of these complexes are frequently mutated in human cancers[239].

There are four major families of remodelling complexes, that share some characteristic properties, including the presence of a single ATPase enzyme; one or more other proteins or protein domains that regulate its activity; and interaction domains that facilitate their binding to chromatin factors or site-specific transcription and thus provide target specificity[238]. Each of the four families is specialised in a single remodelling function: ISWI (Imitation SWItch) and CHD (Chromodomain-Helicase-DNA binding) remodellers facilitate the assembly and organization of nucleosomes after DNA replication, although some of their members also play other key functions beyond nucleosome assembly in interphasic cells[240]. The SWI/SNF (SWItch/Sucrose Non-Fermentable) family members increase chromatin accessibility through nucleosome sliding or by removing whole nucleosomes, which exposes the binding sites of transcription and epigenetic factors, thereby facilitating their recruitment[241]. Finally, remodellers of the INO80 family are responsible for nucleosome editing, by removing specific histones from the nucleosome and replacing them with other canonical core histones or with histone variants[242,243].

The SWI/SNF family is further subdivided into three categories, depending on the subunits with which they are assembled: canonical BAF (cBAF, BRG1-or BRM-associated factors), PBAF (Polybromo-associated BAF) and non-canonical BAF (ncBAF) complexes. All of them contain a single catalytic subunit (either SMARCA4 (also known as BRG1) or SMARCA2 (also known as BRM)), that consumes ATP to catalyse nucleosome sliding and eviction[244]. In general, the activity of SWI/SNF complexes promotes gene expression, but they can also promote gene repression in several ways (e.g., by facilitating the binding of repressive transcription factors or by transcriptional interference)[245,246].

Another major mediator of chromatin remodelling in mammals is the NuRD complex, a member of the CHD family that is involved in nucleosome assembly, DNA damage repair, gene expression and genome integrity[247]. NuRD contains seven core components and a high degree of subunit variability that differs among tissues. Indeed, one of the key components of the NuRD complex in lymphocytes is Ikaros, a transcription factor that was described in the first section of this thesis. In lymphocytes, this protein is mostly found in complex with NuRD, and it critically contributes to targeting the complex to specific genomic regions. In fact, loss of Ikaros, which has a dramatic effect on lymphocyte development, leads to a complete redistribution of NuRD in the chromatin, causing the reactivation of transcriptionally inactive regions[248,249]. Besides these regulatory subunits, NuRD also contains an ATPase catalytic subunit (CHD3/4, also known as Mi-2α/β), responsible for its remodeller functions, and importantly, a second catalytic subunit: HDAC1/2, which makes NuRD the only known complex with both remodeller and deacetylase activity[247].

3.3. Sirtuins

Sirtuins are NAD^+-deacetylases that play pivotal functions in the cellular response to various stresses, and that have been implicated in a myriad of cellular processes and in organismal ageing. They have been conserved across evolution, and virtually all species have at least one sirtuin homolog. In prokaryotes and archaea, there are usually one or two sirtuin homologues, whereas yeast and mammals have five and seven copies of sirtuin genes, respectively[250].

Sirtuins are defined according to their homology with the *S. cerevisiae* gene *Sir2* (Silent information regulator 2), the first family member that was identified, and have been phylogenetically classified into five classes[252] (**Fig. I11a**). Classes I and IV are exclusive of eukaryotes and are the only sirtuins that play epigenetic functions. Class I includes the mammalian SIRT1, SIRT2 and SIRT3, as well as several yeast-specific sirtuins, and Class IV encompasses mammalian SIRT6 and SIRT7. Finally, classes II and III contain mammalian SIRT4 and SIRT5, whereas class U includes a group of bacterial-specific sirtuin homologues[253]. Thus, sirtuins have acquired increasing specialization over evolution. One of the ways in which such diversity is reflected in mammals is the fact that the seven mammalian sirtuins (SIRT1-7) display different subcellular locations: SIRT1, 6 and 7 are fundamentally found in the nucleus. In cell lines, SIRT7 is mostly found in the nucleolus, although the diversity of its functions suggests a more complex distribution in primary cells, as we will see. SIRT2 is primarily cytoplasmic but binds to chromosomes in mitotic cells, specifically during the G_2/M transition, whereas SIRT3, 4 and 5 are essentially mitochondrial proteins[250] (**Fig. I11b**). This said, it should be mentioned that an increasing number of reports have found mammalian sirtuins localised in non-canonical compartments under specific contexts[254-256].

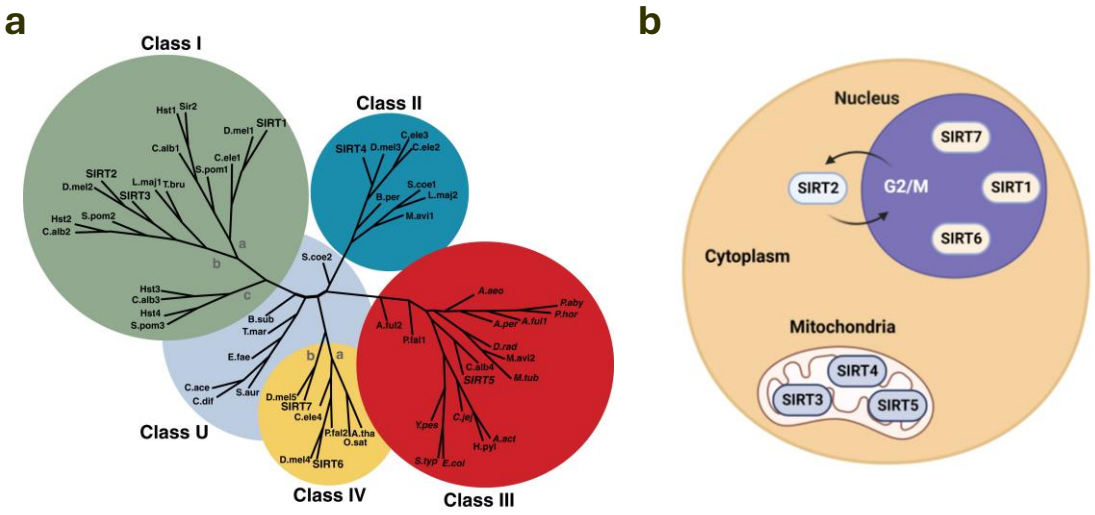


Fig. I11 | Sirtuin classes and subcellular distribution
a, Phylogeny of sirtuin classes throughout evolution. Adapted from Hirschey *et al.* (2011)[251]. **b,** Canonical subcellular distribution of mammalian sirtuins. Created using Biorender.

As anticipated, the HDAC domain of Sirtuins requires NAD^+ as an enzymatic cofactor, which closely links sirtuin activity with the metabolic state of the cell[257]. The deacetylation reaction catalysed by sirtuins consumes one molecule of NAD^+ per acetylated lysine and releases the deacetylated lysine together with a nicotinamide (NAM) molecule and a 2'-O-acetyl-ADP-Ribose molecule, as a consequence of NAD^+ cleavage during the catalysis (**Fig. I12**). Of note, 2'-O-acetyl-ADP-Ribose is a metabolite exclusively generated by sirtuins, whereas NAM is a strong pan-sirtuin inhibitor, which serves as an intrinsic regulatory mechanism of sirtuin activity[258]. Enzymatically, sirtuin-mediated deacetylation consists of two steps. First, sirtuins break the NAD^+ molecule and transfer its ADP-Ribose moiety to the target acetylated lysine, creating an ADP-Ribosylated intermediate and leading to the release of the NAM molecule from the NAD^+. Sirtuins then catalyse the cleavage of this intermediate, resulting in the release of the deacylated protein or peptide and a 2'-O-acetyl-ADP-Ribose molecule. Thus, unlike the other HDACs, which release free acetate during their reaction, sirtuins incorporate the acetyl group into the APD-Ribose[259].

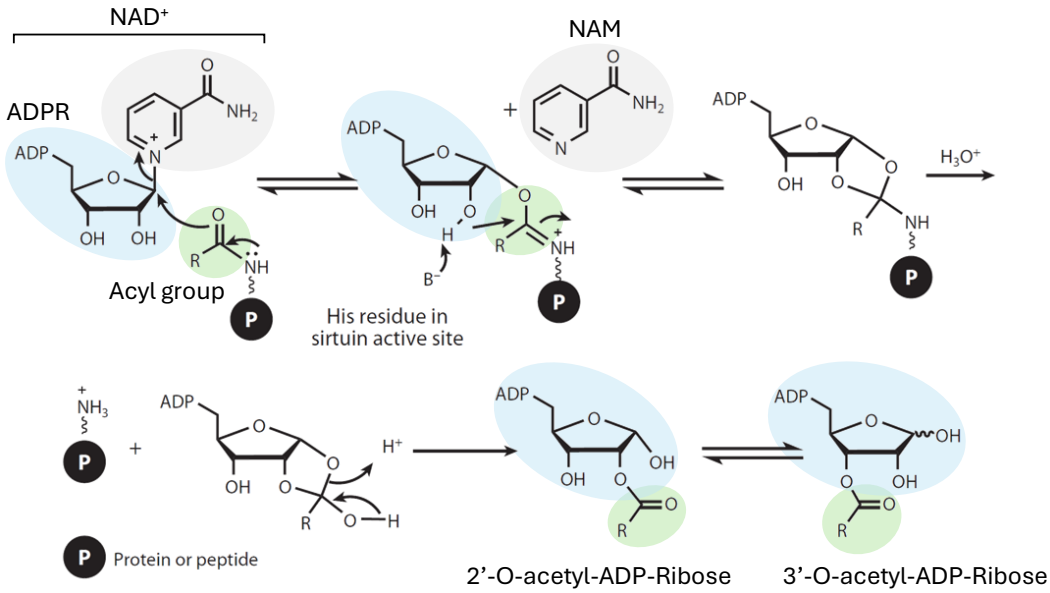


Fig. I12 | Protein deacetylation reaction catalysed by sirtuins
Adapted from Bheda *et al.* (2016)[259].

In addition to their canonical deacetylation reaction, some sirtuins have been reported to use the same mechanism to catalyse the removal of other acyl groups from lysine residues, so the family can be more broadly referred to as NAD^+-dependent deacylases[176,260]. Furthermore, some sirtuins also display a less-characterised mono-ADP-Ribosyltransferase (mADPRT) activity. For instance, SIRT4, which is localised in mitochondria, ADP-Ribosylates glutamate dehydrogenase (GDH) to decrease its activity,

whereas SIRT6 catalyses the ADP-Ribosylation of several non-histone protein substrates, including itself, to promote genome integrity[261,262]. More recently, our lab described that SIRT7 also catalyses its own ADP-Ribosylation upon glucose starvation, which directly contributes to the cellular response against this metabolic stress[263].

3.3.1. Histone sirtuin targets

In general, nuclear sirtuins (SIRT1, 2, 6 and 7) exert their functions by deacetylating either histone marks or non-histone protein substrates, in particular transcription factors or other epigenetic regulators. In the first case, they regulate chromatin structure with various results, including transcriptional repression and DNA damage repair. Otherwise, deacetylation of non-histone substrates typically serves as a mechanism to regulate their subcellular location, interaction partners, catalytic activity or protein stability.

In chromatin, sirtuin activity is closely related to the deacetylation of H3K9ac and H4K16ac, two highly conserved histone marks with well-defined roles in chromatin structure[250]. In mammals, H4K16ac is deacetylated by SIRT1, SIRT2 and nuclear SIRT3. This histone mark interferes with the formation of higher order chromatin compaction and is tightly linked to the progression of the cell cycle, but also to the regulation of other processes, such as DNA replication and repair, transcription and autophagy[250]. Importantly, loss of H4K16ac is a hallmark of human cancer, consistent with the oncogenic roles of sirtuins in several cancers, and hyperacetylation of H4K16ac is observed during ageing, also consistent with the decline in sirtuin levels or activity with age[264-266]. Of interest, H4K16ac deacetylation by SIRT1 is related to gene silencing and heterochromatin formation, whereas SIRT2-mediated deacetylation of this residue is tightly related to its ability to control cell cycle progression but is not associated with heterochromatin formation[267,268].

In the case of H3K9ac, this mark is targeted by SIRT1, SIRT2, SIRT3 and SIRT6. H3K9ac deacetylation precedes the formation of facultative heterochromatin by facilitating the di- or trimethylation of this residue, so deacetylation of this marks represents a chief example of transcriptional repression by sirtuins[269]. Furthermore, H3K9ac deacetylation by SIRT6 in telomeres is required for proper telomeric function and plays an important role during DNA damage repair[270,271]. Interestingly, NF-κB (Nuclear factor Kappa B), a central mediator of inflammatory responses, recruits SIRT6 to proinflammatory genes to facilitate the

deacetylation of H3K9ac at their promoters. In SIRT6-deficient mice, enhanced H3K9ac at these regions leads to hyperactive NF-κB-mediated inflammatory programs, which has been associated with cellular senescence and premature aging in these mice[272].

Other histone marks targeted by sirtuins are H3K56ac, H3K18ac and H3K36ac. The first of them is deacetylated by SIRT1, 2 and 6 and has mainly been associated with DNA damage repair and heterochromatin maintenance, although it has also been linked to transcriptional activation during the cell cycle in yeast[273-275]. H3K18ac was originally thought to be the main histone target of SIRT7, although a more recent report, as well as unpublished data by our laboratory, have identified H3K36ac as the major histone target of SIRT7[276]. H3K18ac is associated with transcriptional activation, so deacetylation this protein is one of the mechanisms by which SIRT7 promotes gene silencing. In cancer, SIRT7-mediated H3K18ac deacetylation at promoters contributes to oncogenic maintenance, as it helps to stabilise the transformed phenotype of cancer cells[277]. Much less is known about H3K36ac, a histone mark that seems to be specifically deacetylated by SIRT7. In yeast, H3K36ac is enriched in the promoters of actively transcribed genes and plays a role in promoting homologous recombination (HR)-mediated DNA repair[278,279]. More recently, H3K36ac was described to be a docking site for the chromodomain of yeast Chd1p, a member of the CHD family of chromatin remodellers, and to promote nucleosome sliding at the transcription start site (TSS)[280].

3.3.2. Non-histone sirtuin targets

Beyond histone mark deacetylation, sirtuins also control several biological processes through the deacetylation of non-histone substrates. This includes not only chromatin regulators and transcription factors, but also proteins that are not related to chromatin, although only the former are discussed here. One of the first mechanisms of sirtuin-mediated heterochromatin formation independently of histone deacetylation was the identification of the interplay between SIRT1 and Suv39h1, the main histone methyltransferase responsible for the deposition of H3K9me3 in heterochromatin regions. SIRT1 was found to interact with Suv39h1 and deacetylate it at K266, in its catalytic SET domain, leading to enhanced activity of Suv39h1, that was largely impaired in the absence of SIRT1[281]. Interestingly, SIRT1 and SIRT2 not only control histone acetylation marks directly but also by regulating some of the main HATs that catalyse for the deposition of these marks: p300 and MOF. In the case of p300, SIRT1 deacetylates Lys1020 and

Lys1024, leading to strong repression of its activity[282]; whereas SIRT2 deacetylates several lysines of its catalytic domain and seems to positively modulate its function[283]. Contrarily, SIRT1-mediated deacetylation of MOF, the main H4K16 HAT, at Lys274 enhances its recruitment to chromatin and its ability to catalyse H4K16 acetylation[284].

In the context of DNA repair, sirtuins also help to control several aspects of this pathway through deacetylation of its players. Indeed, loss of SIRT1, SIRT2, SIRT6 and SIRT7 is associated with increased genome instability stress through different mechanisms[285-288]. For instance, SIRT1 deacetylates NBS1 (Nijmegen breakage syndrome 1), a DNA damage sensor, to facilitate its phosphorylation and activation[289]. SIRT1 also deacetylates core members of the DNA repair machinery, including the NHEJ mediator Ku70 and the HR mediator WRN (Werner syndrome protein), with opposing effects: in the case of Ku70, SIRT1-mediated deacetylation promotes its DNA repair capacity, whereas WRN deacetylation is required for its relocation to nucleoplasm upon repair[290,291]. Another crucial target of sirtuins is PARP1, one of the key players of the DNA damage response. Upon DNA damage, PARP1 recruitment to the lesion is one of the earliest events, as described above. PARP1 recognises the DSB and catalyses the transfer of PAR chains to itself and to other targets, leading to the recruitment of the DNA repair machinery to the lesion[233]. To avoid the activation of PARP1 in the absence of DNA lesions, SIRT1 inhibits its activity by deacetylating it[262]. Conversely, SIRT6 promotes PARP1 functions during DNA repair, although the mechanism does not involve deacetylation but PARP1 mono-ADP-Ribosylation by SIRT6[262]. Recently, our lab identified another mechanism implicating SIRT1 in the timing of the DNA damage response. We found that SIRT1 deacetylates the regulatory subunits PP4R3α/β of the PP4 phosphatase complex, leading to a reduction of its activity and to increased phosphorylation of H2AX and RPA2 (Replication Protein A2), two key events in DNA damage signalling, which ultimately controlled the timing of the response[292].

Sirtuins also deacetylate many transcription factors, including p53, Myc, FOXO (Forkhead box O) proteins, NF-κB, E2F1, BCL6 (B-cell lymphoma 6), HIF1 (Hypoxia inducible factor) and Notch, with various impacts in cell biology. This is especially observed under stress insults, although some of these factors are also regulated by sirtuins in basal conditions[252]. Perhaps the most well-known targets of sirtuins are p53 and FOXO proteins. In response to genotoxic stress, p53 is acetylated at many different residues, which generally leads to its stabilization and its recruitment to its target genes and ultimately

promotes cell cycle exit and apoptosis[293]. Of these residues, SIRT1 and SIRT6 deacetylate K382, thereby reducing the transcriptional activity of p53 and antagonizing cellular senescence[294-296]. *Tp53* haploinsufficiency substantially rescues the dramatic reduction in lifespan displayed by *Sirt6*$^{-/-}$ mice, indicating that premature ageing in these mice is partly due to p53 hyperacetylation, although an increase in cancer incidence is also observed in these compound *Sirt6*$^{-/-}$*Tp53*$^{+/-}$ mutant mice compared to *Wt* mice[297]. Surprisingly, however, homozygous or heterozygous loss of *Tp53* do not rescue embryonic lethality in *Sirt1*$^{-/-}$ mice, whereas the combined heterozygosity of *Sirt1* and *Tp53* leads to a dramatic increase in tumour incidence, reaching 76% in *Sirt1*$^{-/-}$*Tp53*$^{-/-}$ mice at 20 months of age[298]. SIRT2 and nuclear SIRT3 also deacetylate and inhibit p53 to promote cell survival, although the specific residues remain to be identified[299,300]. Finally, SIRT7 has also been reported to reduce p53 transcriptional activity through deacetylation of K320 and K373 in hepatocellular carcinoma cells[301]. However, it seems to be a context-dependent event, since SIRT7 did not deacetylate p53 in other cancer cells[277,302]. Our lab recently reported that loss of *Tp53* does not rescue the embryonic lethality displayed by *Sirt7*$^{-/-}$ mice but rather aggravates it, and that SIRT7 haploinsufficiency in a *Tp53*$^{-/-}$ background increases tumour incidence[303]. The observations that SIRT1, SIRT6 and SIRT7 deficiencies cooperate with that of p53 to drive tumour formation were expected given their tumour suppressor functions. In contrast, the interplay between sirtuins and p53 in life expectancy is puzzling, as it would be expectable that p53 hyperactivation would at least partly explain the shortened longevity of Sirt1$^{-/-}$, *Sirt6*$^{-/-}$ and *Sirt7*$^{-/-}$ mice, SIRT6 and SIRT7, but this is only observed in *Sirt6*$^{-/-}$ mice.

Regarding FOXO family transcription factors, sirtuins target these proteins in the context of stress responses and in immune cells[304-307]. For instance, FOXO3A deacetylation by SIRT2 upon nutrient starvation or oxidative stress enhances its transcriptional activity to facilitate the expression of its target genes, whereas SIRT1 prevents T-cell anergy and terminal differentiation by enhancing FOXO1 protein stability through deacetylation[308]. Another example of stress-dependent transcription factor regulation by sirtuins is the case of HIF1α, the key mediator of cellular adaptation to hypoxic conditions. In resting cells, HIF1α undergoes fast degradation and inactivation, which is assisted by SIRT1-mediated deacetylation. Hypoxia induces SIRT1 downregulation and HIF1α acetylation, which contributes to its stabilization and the activation of its target genes[309].

Sirtuins also play extensive functions in the immune system and in haematopoiesis, mainly related with inflammation and T-cell polarization. In haematopoiesis, all nuclear sirtuins have been found to regulate the physiology of HSCs, although their functions in other haematopoietic progenitors remain largely unexplored[304]. During inflammation, SIRT1, SIRT2 and SIRT6 regulate macrophage polarization by repressing the transcriptional activity of another canonical sirtuin target, among other mechanisms[272,310,311]. A more recent study linked SIRT1 to transcriptional regulation through LLPS for the first time. In the context of viral infections, SIRT1 was shown to deacetylate IRF3 and IRF7, two major regulators of the anti-viral response, which was critical for their ability to undergo phase separation and stimulate interferon expression to reduce the viral load[312].

3.3.3. SIRT7

Although SIRT7 has been considered the least studied mammalian sirtuin for years, this conception is now changing thanks to a growing interest in the biology of this protein. SIRT7 is now recognised as a in important regulator of many physiological processes and as an attractive pharmacological target in disease. Much of what we have learned about SIRT7 comes from the characterization of *Sirt7*$^{-/-}$ mice. SIRT7 deficiency in mice has been primarily linked with ageing, but also with several diseases, including cardiac hypertrophy and fatty liver disease[288,313,314]. These mice display increased genome instability due to impaired DNA damage repair, which leads to the development of a premature progeroid syndrome and to reduced lifespan[288]. Indeed, like other sirtuins, the expression and activity of SIRT7 are known to decline with age in several tissues and cell types[315-318].

SIRT7 is expressed in most tissues, although its expression is particularly high in metabolically active tissues such as liver, spleen and testis, where it has been related with cell proliferation and growth[319]. Originally, SIRT7 was described to be primarily enriched in the nucleolus in cancer cell lines, and several studies reported later that it shuttles into the nucleoplasm in response to various cellular stresses[254,301,319,320]. However, there is increasing evidence suggesting that it may not be a general rule. For instance, this idea was first supported by the finding that SIRT7 deacetylates H3K18ac in the promoters of various target genes in the absence of overt stress[277]. Unpublished data obtained during this thesis further support this idea, since H3K36ac and H3K18ac ChIP-Seq experiments indicated that SIRT7 globally deacetylates both marks in B-cell progenitors. Furthermore, functional proteomics experiments that systematically looked for SIRT7 interacting

proteins and substrates identified a myriad of candidate SIRT7 interactors that were related to various cellular processes, including chromatin remodellers and proteins related to metabolism, transcription, translation or ubiquitination[321-323]. Together, these evidence strongly suggest that SIRT7 may constitutively function out of the nucleolus in resting cells.

One of the most ubiquitous functions of SIRT7 is transcriptional regulation, which it broadly exerts through different mechanisms. In general, SIRT7 behaves as a transcriptional repressor through the deacetylation of H3K18ac, and, presumably, H3K36ac[276,277]. On the contrary, SIRT7 promotes ribosomal RNA (rRNA) expression in several ways. First, it directly interacts with RNA Pol I, the polymerase responsible for rRNA transcription, and deacetylates one of its subunits, PAF53. This facilitates the binding of RNA Pol I to rDNA and enhances its transcriptional activity, which is reversed upon stress[324]. Second, SIRT7 itself is associated with transcriptionally active rDNA regions, and it stimulates RNA Pol I recruitment to these regions[319]. Finally, SIRT7 further promotes RNA Pol I recruitment to rDNA by deacetylating fibrillarin, a pivotal factor in the epigenetic activation of these regions[325]. Surprisingly, SIRT7 is also involved in the formation of heterochromatin at rDNA regions, by deacetylating H3K18ac and recruiting DNMT1 to these regions, which is fundamental for the maintenance of nucleolar structure[325].

Beyond rRNA synthesis and H3K18ac deacetylation, SIRT7 also regulates RNA Pol II- and III-dependent transcription. In the first case, it directly controls RNA Pol II transcriptional elongation through deacetylation of CDK9 (Cyclin-dependent kinase 9). Upon deacetylation by SIRT7, CDK9 becomes activated and phosphorylates RNA Pol II, thereby releasing it from promoter pausing and allowing productive transcriptional elongation[323]. The relationship between SIRT7 and RNA Pol III is less clear. SIRT7 was found to interact with the RNA Pol III-dependent transcription factor complex TFIIIC2 and to bind regions transcribed by RNA Pol III, including tRNA-encoding genes. Interestingly, SIRT7 knockdown reduced the expression of several tRNAs, suggesting a link with RNA Pol-III-mediated transcription, but the mechanisms mediating these observations remain unclear[326].

As other sirtuins, SIRT7 acquires special importance in the context of cellular stress. For instance, our laboratory described that under several stresses, especially glucose starvation, SIRT7 auto-ADP-Ribosylates itself, which promotes its interaction with the histone variant macro H2A.1, a reader of this modification. This mediates the recruitment

of SIRT7 to specific chromatin regions, in which SIRT7 coordinates the response to this stress[263]. SIRT7 also negatively regulates the response to hypoxia independently of its HDAC activity by directly interacting with HIF1α and HIF1β, the two subunits of HIF1, and thereby reducing its protein stability and transcriptional activity. Interestingly, this effect was unaffected by treatment with the pan-sirtuin inhibitor NAM and in cells expressing the deacetylation-inactive SIRT7^{H187} mutant, raising the question of whether SIRT7 regulates HIF1 through ADP-Ribosylation or through steric effects[327]. In the nucleolus, that is a central structure for stress sensing and for the coordination of the subsequent response, SIRT7 participates in the reduction of rRNA expression associated with stress. For example, several cellular stresses eject SIRT7 from the nucleolus, preventing SIRT7-mediated deacetylation of PAF53 and thereby interfering with rRNA transcription[324]. Once out of the nucleolus, nucleoplasmic SIRT7 mediates the stress response through different mechanisms, that have been mainly associated with DNA damage repair[176,288,328]. Finally, SIRT7 contributes to relieve endoplasmic reticulum stress through a mechanism that depends on its HDAC activity and on SIRT7-mediated repression of the *Myc* locus[314].

SIRT7 is actively involved in cellular and organismal metabolism, especially under stress. The closest link between sirtuins and cellular metabolism comes from their NAD^+ dependency. NAD^+ levels mirror the metabolic state of the cell and strongly influence sirtuin activity. Furthermore, sirtuins produce NAM as a byproduct of their deacetylation reaction. NAM acts as a potent sirtuin inhibitor, which represents an intimate crosstalk between the activity of the different sirtuins[329]. A more specific example of the crosstalk between sirtuins in metabolism was identified in the context of adipogenesis. SIRT1 inhibits adipocyte differentiation by recruiting the nuclear corepressor NCOR to the promoter of PPARγ (Peroxisome proliferator-activated receptor γ)[330]. On the contrary, SIRT7 promotes adipogenesis by directly binding to SIRT1 and reducing its recruitment to the PPARγ promoter[331]. Beyond NAD^+ availability, SIRT7 activity is targeted by metabolic cues in several other ways. For instance, PRMT6 methylates SIRT7 at R388 to inhibit its deacetylase activity in conditions of high glucose availability[332]. In addition, SIRT7 is one of the multiple targets of AMPK (AMP-activated protein kinase), an energy sensor of the metabolic status of the cell. Upon starvation, AMPK-mediated SIRT7 phosphorylation evicts SIRT7 from the nucleolus and stimulates its degradation, resulting in reduced rDNA transcription[333].

SIRT7 controls cellular metabolism mainly through mechanisms involving transcriptional regulation. From the epigenetic perspective, SIRT7 regulates mitochondrial biogenesis by

deacetylating H3K18ac at the promoters of genes associated with this process[332]. Similarly, H3K18ac deacetylation by SIRT7 suppresses the expression of genes of the mitochondrial translation machinery, which promotes mitochondrial homeostasis[315]. SIRT7 further regulates mitochondrial function by deacetylating and promoting the transcriptional activity of GABPβ1 (GA binding protein β1), a transcription factor that controls mitochondrial genes[334]. As described above, SIRT7 auto-ADP-Ribosylation coordinates the response to glucose starvation by promoting the enrichment of the histone variant macro-H2A at genes involved in glucose homeostasis[263]. Finally, SIRT7 also collaborates with several transcription factors to regulate metabolism. For instance, ELK4 (ETS domain-containing protein) recruits SIRT7 to chromatin to repress the expression of the gluconeogenic gene *G6pc*, whereas Myc targets SIRT7 to the promoters of ribosomal proteins during endoplasmic reticulum stress[314,335].

Not surprisingly, the central relevance of SIRT7 in genome integrity and cellular stress has led to the identification of diverse functions of this protein in malignancy. Perhaps the most evident tumour suppressive function of SIRT7 is played during malignant transformation, due to its functions in the protection genome integrity. It should be noted, however, that SIRT7 plays an intricate and highly context-dependent role in cancer, up to the point that it can play either oncogenic or tumour suppressive functions in different cancers. Consistently, the expression levels of SIRT7, as well as those of other sirtuins, are associated with different clinical outcomes in different malignancies[336]. From the oncogenic side, SIRT7 overexpression has been described in thyroid, hepatic, bladder, colorectal and lung cancers, and it is a predictor of poor prognosis in some of them[264]. On the contrary, SIRT7 was reported to repress metastasis and to be downregulated in breast cancer metastases[337]. Similarly, SIRT7 strongly inhibits metastasis in oral squamous cell carcinoma through deacetylation of SMAD4, and *Sirt7* haploinsufficiency cooperates with *Tp53* deficiency to enhance cancer metastasis in mice[303,338]. In haematological malignancies, which are one of the focuses of this thesis, only one report has explored the role of SIRT7. In this paper, a significant reduction in the levels of SIRT7 in AML and CML (chronic myeloid leukaemia), compared to healthy donors, was reported. SIRT7 levels were positively correlated with a good response to therapy in these patients, and pharmacological inhibition of oncogenic drivers further increased SIRT7 levels, clearly pointing to a promising tumour suppressor role of SIRT7 in these malignancies[318]. However, the mechanisms implicating SIRT7 in haematological malignancies remain

fundamentally unexplored, which motivated us to explore its potential as a novel player of their biology.

As it happens with haematological malignancies, the roles of SIRT7 in haematopoietic progenitors and immune cells remain largely uncharacterised, although a few papers have explored its functions in these contexts. A role of SIRT7 in the immune system was originally suggested by the observation that one of the tissues with the highest expression of SIRT7 is the spleen, which could underlie important functions of this protein in immune responses[319]. In mouse HSCs, SIRT7 is downregulated with age, and the loss of SIRT7 recapitulates several aspects of HSC ageing, including myeloid bias, decreased reconstitution capacity, and loss of quiescence[288,315]. Age-dependent SIRT7 downregulation has also been reported in human macrophages, and it seems to be related with monocyte to macrophage differentiation, which may link SIRT7 with age-related inflammation. Further implicating SIRT7 in inflammation, a significant infiltration of inflammatory macrophages was observed in the myocardium of *Sirt7*$^{-/-}$ mice, that also displayed increased levels of inflammatory cytokines in the hearts of these mice[313]. However, the mechanisms underlying these observations are not known. Finally, the critical function of SIRT7 in NHEJ DNA repair was found to be important to some extent for IgG class switching in splenic B-cells, and *Sirt7*$^{-/-}$ mice displayed reduced cellularity in the bone marrow, spleen in thymus, strongly suggesting a potential role in haematopoiesis[288]. Hence, although a few functions of SIRT7 in immunity and haematopoietic differentiation have been identified, the influence of SIRT7, as well as that of the other sirtuins, remains poorly understood in these processes.

Importantly, unpublished studies from our laboratory have found compelling evidence strongly suggesting that SIRT7 may be required for B-cell differentiation. In *Sirt7*$^{-/-}$ mice, the total number of B-cells, but not those of T-cells, were found to be significantly reduced, suggesting that SIRT7 may be specifically involved in B-cell development. Importantly, this phenotype appeared to be due to an impairment at the pre-B cell stage, since these mice displayed normal numbers of pro-B cells along with a reduction of pre-B cells and downstream stages. Furthermore, these studies also found a potential candidate to mediate the functions of SIRT7 in B-cell progenitors: the B-cell master regulator PAX5. According to our previous data, PAX5 protein but not RNA levels are significantly reduced in B-cell progenitors, suggesting that SIRT7 may directly regulate PAX5 protein levels, and that this phenomenon may underlie an important function of SIRT7 in B-cell development.

These data motivated us to explore more broadly the potential functions of SIRT7 in haematopoiesis and to investigate in detail the specific functions of SIRT7 during B-cell development and its potential link with malignancy.

IV. OBJECTIVES

The roles of SIRT7 and other sirtuins in haematopoietic differentiation and malignancy remain largely unexplored, and unpublished data from our laboratory has strongly suggested that SIRT7 is required for B-cell development and for the functions of its master regulator PAX5. Based on these data, we hypothesise that SIRT7 plays an important role in the establishment of B-cell identity through the regulation of PAX5. On this basis, the global objective of this project is to explore novel functions of SIRT7 and other sirtuins in haematopoiesis and leukaemia, paying special attention to the characterization of the biochemical interplay between SIRT7 and PAX5 and its repercussion on B lymphopoiesis and leukemic transformation. To this end, we have established the following specific objectives:

1. To explore the expression levels of nuclear sirtuins in mouse and human haematopoietic progenitors.
2. To characterise the relevance of SIRT7 in B-cell development and commitment.
3. To identify the mechanisms by which SIRT7 regulates PAX5 functions and whether their interplay mediates the functions of SIRT7 in B-cell development.
4. To interrogate whether the SIRT7/PAX5 axis is functionally relevant in B-ALL.

V. METHODS

1. Mice

Germline *Sirt7*$^{-/-}$ mice (CD45.2) in the 129Sv genetic background were previously described[288]. IgHEL[339] mice (CD45.2) were obtained from Jackson Laboratories (002595) in the C57BL/6 background and were crossed with 129Sv *Wt* and *Sirt7*$^{-/-}$ mice, and further backcrossed for 10 generations to get a pure 129Sv background. *Sirt7*^{N198Q} mice (CD45.2, C57BL/6) were generated by our laboratory and are currently unpublished. *Pax5*$^{-/-}$ (Ref [116]) (CD45.2) mice and CD45.1 (Jackson Laboratories, 002014) mice were on the C57BL/6 background. Samples from germline C57BL/6 *Sirt7*$^{-/-}$ mice (described previously in Ref [313]) were kindly provided by Dr. Alessandro Ianni (IJC, Barcelona, Spain, and Max Planck Institute for Heart and Lung Research, Bad Nauheim, Germany), whereas samples from *Pax5*$^{-/-}$ mice were kindly provided by Dr. Mikael Sigvardsson (Lund University, Sweden). Heterozygous CD45.1CD45.2 mice were generated by crossing C57BL/6 CD45.1 and *Wt* 129Sv CD45.2 mice for one generation to avoid rejection upon transplantation of 129Sv cells. All mice were bred in Comparative Medicine and Bioimage Centre of Catalonia (CMCiB) animal facility of the Germans Trias i Pujol Research Institute (IGTP). Animal studies were conducted in accordance with appropriate institutional ethics committees and national authorities.

2. Histology

Spleens were collected from *Wt* and *Sirt7*$^{-/-}$ mice before fixation in 10% formalin for 24h. Fixed spleens were washed, dehydrated and embedded in paraffin. Paraffin-embedded spleens were sectioned at 4 µm and stained with haematoxylin and eosin. Visualization and image acquisition of histological sections were performed with an Olympus BX53 microscope.

3. Determination of anti-HEL antibody isotypes

100 µl of an emulsion containing 50 µg of NP-HEL along with Complete Freund Adjuvant (CFA) were intraperitoneally injected into *Wt* and *Sirt7*$^{-/-}$ mice. 14 days after immunization, blood samples were collected from immunized mice or naïve controls, and the serum fraction was separated. For the determination of antibody isotypes, ELISA plates were pre-coated with HEL peptide (6 µg/ml) overnight at 4ºC. Serum dilutions from naïve or immunised mice were added to the plates and incubated at room temperature for 1 hour.

After three washes with a PBS-0.05% Tween solution, rat anti-mouse Ig subclasses (IgM, IgG1, and IgG3) were added, incubated for 1 hour at room temperature and washed three times. A secondary anti-rat Ig was then added, followed by a 1h incubation at room temperature and three additional washes. After washing, a TMB developing solution was added for 15 minutes and stopped with 1N H_2SO_4. Signal intensities were measured at 570 nm (background subtraction) and 450 nm in a Multiskan Sky (ThermoFisher) plate reader.

4. Transplantation experiments

$2x10^6$ *ex vivo* expanded *Wt*, *Sirt7*$^{-/-}$ or *Pax5*$^{-/-}$ pro-B cells or $1.5x10^6$ $Lin^-IgM^+IgD^+$ B cells sorted from the spleens of *Wt* or *Sirt7*$^{-/-}$ mice were washed in PBS containing with 1% heat-inactivated foetal calf serum and resuspended in 200 µL of the same solution. Cells were injected via the tail vein into sublethally irradiated (5 Gy) CD45.2 or CD45.1CD45.2 randomised recipient mice (6 to 10-weeks old). 4 weeks after transplantation, spleens of CD45.1CD45.2 mice or bone marrow and thymus of CD45.2 mice were harvested and analysed by flow cytometry.

5. Flow cytometry and cell sorting

Single-cell suspensions from bone marrow, spleen and thymus samples were obtained by crushing in staining buffer (3% FBS, 2 mM EDTA in PBS). Erythrocytes were lysed in ACK buffer (Gibco), before stopping the reaction by adding 5 volumes of staining buffer. Cells were filtered through 40-µm strainers and incubated with Fc-block (eBioscience) before staining for 30 min (4°C) with specific antibodies. After staining, cells were washed twice, resuspended in staining buffer and analysed with a FACS Canto II (BD Biosciences) or sorted with a FACSAria II (BD Biosciences). Flow cytometry experiments were analysed with FlowJo software. The following flow cytometry antibodies were used: anti-B220 (RA3-6B2, eBioscience), anti-CD11b (M1/70, BD Biosciences), anti-CD127 (eBioSD/199, eBioscience), anti-CD138 (300506, Invitrogen), anti-CD19 (HIB19, eBioscience), anti-CD21 (7G6, BDBiosciences), anti-CD23 (B3B4, BDBiosciences), anti-CD93 (AA4.1, eBioscience), anti-CD38 (90, eBioscience), anti-CD3e (145-2C11, BD Biosciences), anti-CD4 (GSK1.5, eBioscience), anti-CD43 (eBioR2/60 or 1G10, BDBiosciences), anti-CD45.1 (A20, eBioscience), anti-CD45.2 (104, eBioscience), anti-CD8 (53-6.7, eBioscience), anti-Fas (SA367H8, Biolegend), anti-GL7 (GL-7, eBioscience), anti-Gr1 (RB6-8C5, BD Biosciences), anti-hCD4 (RPA-T4, Biolegend), anti-IgD (11-26c, eBioscience), anti-IgG1

(A85-1, BD Pharmigen), anti-IgM (II/41, eBioscience), anti-Ly76 (TER-119, BD Biosciences), anti-NKp46 (29A1.4, Biolegend) and anti-TCRβ (H57-597, BDBiosciences).

The populations analysed were defined as: pre-pro-B ($B220^{+}CD19^{-}$), pro-B ($B220^{+}CD19^{+}IgM^{-}CD43^{+}$), pre-B ($B220^{+}CD19^{+}IgM^{-}CD43^{-}$), large pre-B ($B220^{+}CD19^{+}IgM^{-}CD43^{-}FSC^{high}$), small pre-B ($B220^{+}CD19^{+}IgM^{-}CD43^{-}FSC^{low}$), Immature B ($B220^{+}CD19^{+}IgM^{+}$), BM Mature B ($B220^{high}CD19^{+}IgM^{+}$), marginal zone B ($B220^{+}CD19^{+}CD21^{high}CD23^{-}$), transitional B ($B220^{+}CD19^{+}CD21^{+}CD23^{+}CD93^{+}$), follicular B ($B220^{+}CD19^{+}CD21^{+}CD23^{+}CD93^{-}$), germinal center B ($B220^{+}CD19^{+}IgM^{+}GL7^{+}Fas^{+}$), memory B ($CD19^{+}CD38^{+}CD138^{-}GL7^{-}$), class-switched $IgG1^{+}$ B-cells ($B220^{+}IgG1^{+}$) and plasma cells ($B220^{low}CD138^{+}$).

Apoptosis was measured in pro-B and pre-B cells gated from BM samples previously stained cell surface markers, 7AAD (BD Biosciences) and Annexin V-FITC (Abcam), according to the manufacturer's protocol. Cell cycle distribution was measured in gated BM large and small pre-B cells previously stained with surface markers. Stained cells were fixed and incubated for at least 1h with 1 μg/mL DAPI in permeabilization buffer (eBioscience™ Foxp3 / Transcription Factor Staining Buffer) before FACS analysis.

For intracellular staining of SIRT7 and PAX5, BM-derived single-cell suspensions were stained with cell surface markers and washed. Stained samples were fixed for 30 min and permeabilized with Foxp3/Transcription Factor Fixation/Permeabilization buffers (eBioscience), following the manufacturer's instructions. Cells were then incubated with 2% FBS for 10 min before staining with isotype control, anti-SIRT7 (D3K5A, Cell signaling) or anti-PAX5 (1H9, eBioscience) antibodies for 1h at room temperature. Finally, cells were washed twice with permeabilization buffer supplemented with 2% FBS. For anti-SIRT7 staining, cells were then incubated for 1h at room temperature with a polyclonal anti-IgG (H+L) secondary antibody (Invitrogen) and analysed by flow cytometry.

$pSTAT5^{Y694}$ staining performed in gated BM pre-B cells. Bone marrow cells were previously incubated in RPMI (Gibco) (30 min, 37°C) and subsequently stimulated with IL7 (5 ng/mL) in the same medium (30 min at 37°C). Stimulated cells were fixed with Foxp3/Transcription Factor Fixation buffer, washed and incubated for 1h with ice-cold 100% methanol. Cells were stained with cell surface markers (30 min at room temperature) and an anti-$pSTAT5^{Y694}$ antibody (47/Stat5(pY694), BD Biosciences, 1h at room temperature), washed and analysed by flow cytometry.

6. Isolation of primary B-cell progenitors

Bone marrow *Wt* and *Sirt7*$^{-/-}$ pro-B cells were enriched by anti-CD19 Magnetic-activated cell separation (MACS)-enrichment cells and further purified by cell sorting. Briefly, single-cell suspensions from BM samples were prepared in staining buffer before staining with Fc-block for 20 min and for 30 min with an anti-CD19-Biotin antibody (1D3, BD Biosciences). Stained cells were washed and further incubated with Streptavidin MicroBeads (Miltenyi) before magnetic separation. Purified CD19^{+} cells were cultured overnight in B-cell medium (Opti-MEM, 10% heat-inactivated FBS, 25 mM HEPES, 50 µg/mL Gentamicin, and 50 µM β-mercaptoethanol) supplemented with 10 ng/mL IL7, SCF and FTL3L (Miltenyi) before sorting Lin^{-}CD19^{+}B220^{+}IgM^{-} cells. Foetal liver *Pax5*$^{-/-}$ pro-B cells were kindly provided by Dr. Mikael Sigvardsson (Lund University, Sweden) and obtained as described[121].

7. Cells and reagents

HAFTL, NALM-20, TANOUE, REH, TOM-1, KOPN-8, SD-1 and SEM B-lymphoblastoid cells were grown in RPMI containing 10% heat-inactivated FBS and 100 U/ml penicillin/streptomycin (Gibco). Platinum E, HEK293F, *SIRT7*$^{-/-}$ HEK293F (Ref [263]) cells were maintained in DMEM (Gibco) containing 10% FBS and 100 U/ml penicillin/streptomycin. OP9 cells were cultured in MEM-α (Gibco) supplemented with 20% FBS and 100 U/ml penicillin/streptomycin. Primary pro-B cells obtained from *Wt*, *Sirt7*$^{-/-}$ or *Pax5*$^{-/-}$ mice were co-cultured with mitomycin C-inactivated OP9 cells on B-cell medium supplemented with 10 ng/mL IL7, SCF and FTL3L (Miltenyi). All the cells were maintained in a humidified atmosphere containing 5% CO_2 at 37°C. Transient transfections of HEK293F and *SIRT7*$^{-/-}$ HEK293F were performed with polyethylenimine and the corresponding constructs. For retroviral infection of HAFTL cells, B-ALL cells and primary pro-B cells, Platinum E cells were transiently transfected with a pVSV-G construct (encoding the viral envelop) and pMIG bicistronic vectors (encoding either the selection marker (hCD4 or GFP) alone (empty vector) or with the coding sequences of SIRT7^{Wt}, SIRT7^{H187Y}, PAX5^{Wt}, PAX5^{K198Q} or PAX5^{K198R}). Retroviral supernatants were filtered (0.45 µm) and cells were resuspended in the filtered retroviral supernatants. Transduced cells were centrifuged (1.5h, 1000g, 32°C) for improving transduction efficiency and incubated overnight before sorting hCD4^{+} or GFP^{+} cells.

Lentiviral transduction of HAFTL cells was similarly performed. In this case, lentiviruses were produced in HEK293T cells with pCW-Cas9, psPAX2 and pMDG2 constructs. After transduction, HAFTL cells were selected with 2 µg/mL puromycin (Invivogen) and subsequently transduced with retroviruses containing pLKO5-sgRNA-EFS-GFP constructs with an unspecific short guide RNA (sgRNA) sequence or sgRNAs targeting SIRT7 (pooled 5'-CACCGCACCGCGAGCGGCTCAGACCGCCA-3', 5'-CACCGCGTAGGTGTCACGCATCC TG-3' and 5'-CACCGCGGAGGAAGGTCCGCGAAC-3', previously validated in our laboratory). Transduced cells were treated with 1 µg/mL doxycycline, single GFP^+ clones were sorted and grown in 96 well plates and clones were analysed by anti-SIRT7 immunoblotting.

Treatments were performed with 5 mM NAM (nicotinamide, Sigma) for 48h; 100 µg/mL cycloheximide (Sigma-Aldrich for the 3, 6, 9 or 24h); 2 µM lactacystin (Santa Cruz Biotechnology) for 8h; or 1 µM TSA (Sigma-Aldrich) for 3h.

For *in vitro* competitive growth experiments, TANOUE cells retrovirally transduced with empty pMSCV6-IRES-hCD4, pMSCV-SIRT7-IRES-hCD4 or pMSCV-PAX5-IRES-hCD4 constructs were grown for 3 or 7 days, collected and stained with an anti-hCD4 antibody before FACS analysis of the percentage of $hCD4^+$ / $hCD4^-$ cells.

The following plasmids were used for transfections: pcDNA/4T0-SIRT7-HA, pCMV6-PAX5-Myc-FLAG, pMSCV-Ikaros-FLAG-IRES-GFP (pMIG), pcDNA3.1-EBF1-FLAG, pMSCV-SIRT7-IRES-hCD4 (pMIG), pMSCV-PAX5-IRES-hCD4 (pMIG), pMSCV-PAX5-IRES-GFP (pMIG), pcDNA/4T0-p300-Myc, pcDNA/4T0-PCAF-FLAG, pLX304-GTF3C3-V5, pLX304-NCOA3-V5, pcDNA3-Myc-DTX2, pcDNA3.1-Myc-MDM4. Ikaros-FLAG, EBF1-FLAG, pLKO5-sgRNA-EFS-GFP, pCW-Cas9, pVSV-G, psPAX2 and pMDG2 constructs were from Addgene; the pMIG-derived hCD4 vector was kindly provided by Dr. Jose Luis Sardina (IJC, Spain); pMIG-PAX5-IRES-GFP was kindly provided by Dr. Michael Sigvardsson (Lund University, Sweeden); Myc-MDM4 was a gift from Dr. Stjepan Uldrijan (Masaryk University, Czech Republic); Myc-DTX2 was from Dr. Danny Huang (NYU, USA); and the expression vectors for acetyltransferases were a gift from Dr. Thomas Braun (Max Planck Institute for Heart and Lung Research, Germany).

8. Immunoprecipitation

Immunoprecipitations were performed with cell pellets lysed in RIPA buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.5% sodium deoxycholate, 0.1% SDS, 1% NP40, 2 mM $MgCl_2$) containing cOmplete Protease Inhibitor (Roche). Cell lysates were incubated with benzonase nuclease (Millipore) at 4°C for 8h. Cell lysates were then clarified by centrifugation (17 000 g for 10 min at 4°C) and incubated overnight with conjugated anti-FLAG agarose beads (Millipore) or anti-HA agarose beads (Pierce), or with control (IgG, #2729, Cell signalling) or anti-SIRT7 (C-3, Santa Cruz) antibodies at 4°C with gentle rotation. For endogenous immunoprecipitations, samples were incubated with Protein G agarose beads (Pierce) for a further 2h at 4ºC with gentle rotation. The immunoprecipitated protein complexes were washed 5 times with lysis buffer (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 10% glycerol, 1 mM EDTA) and then eluted with Laemmli buffer supplemented with 10% β-mercaptoethanol. Samples were then boiled at 95°C for 5 min and analysed by immunoblotting.

9. Cell fractionation, gel filtration high performance liquid chromatography (HPLC) and immunoblotting

Cellular fractionation experiments were performed as described previously[263], using the Dignam method. Size exclusion HLPC experiments were performed with nuclear exacts from *Wt* or *Sirt7*$^{-/-}$ HAFTL cells, that were purified and lysed in native conditions following the Dignam method[263]. Nuclear lysates from HAFTL cells were incubated overnight with benzonase nuclease, clarified by centrifugation and concentrated using Amicon Ultra centrifugal filters (Millipore), before molecular weight fractionation on a gel filtration column Superose 6 (Cytiva, USA, fractionation range: $5x10^3$–$5x10^6$ Da). The eluted fractions were denatured with Laemmli buffer containing 10% β-mercaptoethanol and analysed by immunoblot. Densitometric quantification of immunoblotting experiments was performed with ImageJ software.

The following antibodies and dilutions were used for immunoblotting: anti-Acetyl-lysine (9814, Cell signaling, 1:200), anti-Actin (A1978, Merck, 1:5,000), anti-EBF (C-8, Santa Cruz, 1:500), anti-Fibrillarin (B1, Santa Cruz Biotechnology 1:1,000), anti-FLAG (M2, Sigma-Aldrich, 1:10,000), anti-H3 (ab1791, Abcam, 1:30,000), anti-H3K18ac (#9675, Cell signaling, 1:1,000), anti-H3K36ac (D9T5Q, Cell signaling, 1:1,000), anti-H3K36me3 (4909, Cell signaling, 1:1,000), anti-H3K9ac (C5B11, Cell signaling,

1:1,000), anti-HA (6908, Sigma 1:10,000), anti-Myc (9B11, Cell signaling, 1:1,000), anti-PAX5 (D19F8, Cell signaling, 1:500), anti-SIRT7 (D3K5A, Cell signaling, 1:500), anti-STAT5 (3H7, Cell signaling, 1:1,000), anti-V5 (ab9116, Abcam, 1:2,000) and H4K16ac (C15200219, Diagenode, 1:1,000).

10. *In vitro* deacetylation assays

For SIRT7 and PAX5 purification, *SIRT7*$^{-/-}$ HEK293F cells were transfected with vectors encoding PAX5-Myc-FLAG or SIRT7-FLAG for 48h. Cells were treated overnight with 5 mM NAM and for 3h with 1 µM TSA to hyperacetylate PAX5 before harvesting. Cells were lysed and incubated with benzonase nuclease for 8h before clarification and overnight incubation with anti-FLAG beads (Millipore). After immunoprecipitation, samples were washed 5 times at 4ºC using BC500 buffer (20 mM Tris-HCl (pH 8.0), 500 mM NaCl, 10% glycerol, 1 mM EDTA). Washed samples were eluted in native conditions with 0.6 µg/mL FLAG peptide (GenScript) and finally dialysed in BC100 buffer (20 mM Tris-HCl (pH 8.0), 100 mM NaCl, 10% glycerol, 1 mM EDTA) to remove the FLAG peptide. For *in vitro* PAX5 deacetylation assays, purified PAX5 and SIRT7 were incubated in deacetylation buffer (10 mM Tris-HCl pH 8.0, 150 mM NaCl, 1mM DTT, 10% Glycerol) (1h, 37°C), in the presence or absence of 1.25 mM NAD^+. The reaction was stopped by adding 5X Laemmli buffer containing 10% β-mercaptoethanol before anti-PAX5, anti-pan-Acetyl-lysine and anti-SIRT7 immunoblot analysis.

11. Pre-B cell proteome and PAX5 acetylome

For determination of the pre-B cell proteome, proteins were extracted from BM sorted pre-B cells with 6 M urea, 100 mM Tris pH 8.0 and the help of a bioruptor. Proteins were quantified using a nanodrop at 280nm and precipitated with TCA/Acetone. Samples were then reduced with 10mM DTT and alkylated with 55mM CAA. Proteins were resuspended in 6 M urea, 100 mM Tris pH 8.0 and digested with LysC/Trypsin. The Lys-C digestion was performed for 16h at 30ºC, whereas trypsin digestion was performed for 8h at 30ºC. The reactions were stopped with 10% formic acid. The peptides were desalted with a C18 reverse-phase ultramicrospin column and desiccated in a speedvac. Total proteome samples were separated using a C18 analytical column (nanoEaseTM M/Z HSS C18 T3 (75 µm × 25 cm, 100Å, Waters) with a 180 min run, comprising three consecutive steps with linear gradients from 3% to 35 % B in 150 min, from 35 % to 50 % B in 5 min, from 50 % to

85% B in 2min, followed by isocratic elution at 85 % B in 5 min and stabilization to initial. The mass spectrometer was operated in a data-dependent acquisition (DDA) mode, and data were acquired with Xcalibur software 4.0.27.10 (Thermo Scientific).

For identification of PAX5 acetylated residues, *SIRT7*$^{-/-}$ HEK293F cells were transiently transfected with PAX5-Myc-FLAG and either an empty vector or a vector encoding SIRT7-FLAG. After anti-FLAG immunoprecipitation, beads were washed three times with 100 mM Tris pH 8.0 and then resuspended in 6M Urea, 100 mM Tris pH 8.0. Samples were then reduced with 10mM DTT and alkylated with 55mM CAA. Samples were digested with 1µg trypsin (16 hours, 30°C). The reaction was stopped with 10% Formic acid and the peptides desalted with a polyLC C18 pipette tip and dried in a speedvac. The acetylome were separated using an Evosep EV1000 column (150 µm x 150 mm, 1.9 µm) (Evosep) with an 88 min run. The spectrometer was working in positive polarity mode and singly charge state precursors were rejected for fragmentation. Data were acquired with Xcalibur software 4.2.28.14 (Thermo Scientific).

For both total proteome and acetylome analyses the peptides were reconstituted with 3% ACN and 0.1% FA aqueous solution at 100 ng/µL and 800 ng were injected into the mass spectrometer.

12. Semiquantitative PCR, RT-qPCR and RNA-Seq

IgH segments semiquantitative PCR was performed with genomic DNA from IgM+ cells sorted from the spleens of *Wt* or *Sirt7*$^{-/-}$ mice. The PCR was performed as in Ref[340], using degenerate primers. For RNA-Seq and RT-qPCR, RNA was isolated from cell pellets using the Maxwell RSC simplyRNA Tissue Kit (Promega) and according to the manufacturer's instructions. cDNA for RT-qPCR was synthesised with the Transcriptor First Strand cDNA Synthesis Kit (Roche), following the manufacturer's instructions. RT-qPCR reactions were performed in a QuantStudio 5 Real-Time PCR System. Primer sequences are shown in **Methods Table 1.**

RNA for RNA-Seq was similarly extracted from *Wt* and *Sirt7*$^{-/-}$ pro-B cells and pre-B cells (sorted from BM) or *Pax5*$^{-/-}$ *ex vivo* expanded pro-B cells transduced with either empty vectors or with vectors encoding PAX5^{Wt}, PAX5^{K198Q} and PAX5^{K198R}. After library construction, 150 bp paired-end sequencing was performed on a DNBSEQ-G400 (MGI Tech).

Methods Table 1. List of primers used in this study.

RT-qPCR primers	
Rag1-Fw	5'-GTCGCAAGAGAACTCAGGCT-3'
Rag1-Rev	5'-ACGGGATCAGCCAGAATGTG-3'
Trbc1-Fw	5'-ACCTTCTGGCACAATCCTCG-3'
Trbc1-Rev	5'-GGCCTCTGCACTGATGTTCT-3'
Trbc2-Fw	5'-TGGCACAATCCTCGAAACCA-3'
Trbc2-Rev	5'-TGATTCCACAGTCTGCTCGG-3'
Thy1-Fw	5'-GCTCTCCTGCTCTCAGTCTTG-3'
Thy1-Rev	5'-TGTTATTCTCATGGCGGCAGT-3'
Cd3e-Fw	5'-TCACTCTGGGCTTGCTGATG-3'
Cd3e-Rev	5'-TTGCGGATGGGCTCATAGTC-3'
Zap70-Fw	5'-GCTTGAAGGAGGTCTGTCCC-3'
Zap70-Rev	5'-TATCCGTCCGAGTTCAGGGT-3'
Tigit-Fw	5'-CTGTGCTGGGACTCATTTGCT-3'
Tigit-Rev	5'-AGACTCCTCAGGTTCCATTCCT-3'
Pax5-Fw	5'-GGAGGATCCAAACCAAAGGT-3'
Pax5-Rev	5'-TTGTCACAGACTCGCTCTGC-3'
Prf1-Fw	5'-ACCTCCACTCCACCTTGACT-3'
Prf1-Rev	5'-AGGGCTGTAAGGACCGAGAT-3'
Gzmb-Fw	5'-GAAGCCAGGAGATGTGTGCT-3'
Gzmb-Rev	5'-GCACGTTTGGTCTTTGGGTC-3'
Klrb1-Fw	5'-GCTGACGGTCTGGTTGACTG-3'
Klrb1-Rev	5'-TGTTTGTGAGATGAGGGCACA-3'
***Igh* recombination primers**	
D_HL-Fw	5'-GGAATTCGMTTTTTGTSAAGGGATCTACTACTGTG-3'
V_H5558-Fw	5'-CGAGCTCTCCARCACAGCCTWCATGCARCTCARC-3'
V_H-Q52-Fw	5'-CGGTACCAGACTGARCATCASCAAGGACAAYTCC-3'
V_H-7183-Fw	5'-CGGTACCAAGAASAMCCTGTWCCTGCAAATGASC-3'
J_H3-Fw	5'-GTCTAGATTCTCACAAGAGTCCGATAGACCCTGG-3'
Kat8-Fw	5'-TATCTGCCTTTCTCTGTCAATGGG-3'
Kat8-Rev	5'-AGGTGAGCCAGGTTAGGACTTGG-3'

13. ChIP-Seq

10^7 pro-B cells were collected, washed once with PBS and fixed. For anti-PAX5 ChIP-seq, cells were fixed with 1 mg/mL DSG (ThermoFisher) for 30 min followed by an additional fixation with 1% formaldehyde for 10 min. For H3K36ac, H3K36me3 and H3K18ac ChIP-Seq, cells were fixed for 10 min with 1% formaldehyde. The crosslinking was stopped with 1/20 quenching buffer. Fixed cells were then washed twice with ice-cold PBS supplemented with 0.5% BSA (w/v). Nuclei were purified and sonicated with the truChIP Chromatin Shearing kit (Covaris), following the manufacturer's protocol. Nuclear lysates were diluted with one volume of 2X dilution buffer supplemented with 0.1% SDS and protease inhibitors and clarified by centrifugation at 10 000 g (5 min, 4°C). ChIPs were performed overnight at 4°C with 10 µg of an anti-PAX5 antibody (ab183575, Abcam) or 10 µL of anti-H3K36ac (D9T5Q, Cell signaling), anti-H3K18ac (#9675, Cell signaling) or anti-H3K36me3 (4909, Cell signaling) antibodies previously conjugated with 20 µL ChIP-grade

Protein A/G Magnetic Beads (ThermoFisher) for 4h at 4°C. Samples were then washed once with low salt wash buffer (0.1% SDS, 1% Triton, 2 mM EDTA, 20 mM Tris-HCl pH 8.1, 150 mM NaCl), once with high salt wash buffer (0.1% SDS, 1% Triton, 2 mM EDTA, 20 mM Tris-HCl pH 8.1, 500 mM NaCl), once with LiCl immune complex wash buffer (10 mM Tris-HCl pH 8, 250 mM LiCl, 1% NP40, 1% sodium deoxycholate, 1mM EDTA) and twice with modified TE buffer (0.1mM EDTA, 10 mM Tris-HCl pH 8). Low salt, high salt and LiCl wash buffers were previously supplemented with protease inhibitors. After washing, the crosslink of ChIP and input samples was reversed with elution buffer ($NaHCO_3$ 0,1M, 1% SDS, 10 µg RNase A (ThermoFisher)) for 30 min at 37°C, followed by an additional incubation (6h, 65°C) with 50 µg Proteinase K (Apollo Scientific). The eluted DNA was cleaned up with NucleoSpin Gel and PCR Clean-up columns (Macherey-Nagel) and used for library construction and 100 bp paired-end sequencing on a DNBSEQ-G400 instrument

.

14. Public data analysis

Mouse and human bone marrow 10X sc-RNA-Seq data were derived from the Broad Institute Single Cell Portal (https://singlecell.broadinstitute.org) (studies SCP978 and SCP101, respectively). For analysis of human sc-RNA-Seq data (study SCP101), the Loom files were downloaded and re-analysed using Scanpy[341] before cell type annotation with DecoupleR (v1.34)[342] and PanglaoDB[343]. Mouse sc-RNA-Seq feature, t-SNA plots and normalised counts were downloaded from the Single Cell Portal (study SCP978). Microarray data were from the Immgen Consortium datasets[344] (GSE15907). Genes regulated by PAX5 genes were obtained from Ref [119] (GSE38046), whereas PAX5 ChIP-Seq data was from Ref [81]. The list of mammalian acetyltransferases was from a previous report[174], and PAX5 protein-protein interactions combined the data from another study[345] with the curated interactions compiled in the BioGRID repository. Proteomics and RNA-Seq data from B-ALL patients were retrieved from Ref [346]. Raw RNA-Seq and proteomics data from this study are available at the Proteome Xchange Consortium (PXD010175) and at the European Genome-phenome Archive (EGA) (EGAS00001003079), respectively. Proteomics data of PAX5 and SIRT7 protein levels in B-CLL cell lines were obtained from Ref [347]. Raw data of this study are deposited at the PRIDE archive (PXD002004). Children's Oncology Group (COG) clinical trial P9906 gene expression (GSE11877) and patient outcome data were obtained from the Genomic Data Commons (https://portal.gdc.cancer.gov) and were generated by the TARGET (Therapeutically Applicable Research to Generate Effective Treatments, https://www.cancer.gov/ccg/ research/genomesequencing

/target) initiative, whereas the presence of PAX5 mutations in these patients were obtained from Ref [348].

15. Proteomics data analysis

For total proteome and acetylome, the RAW thermo files were processed with MaxQuant 1.6.7.0 (mouse database downloaded from https://www.uniprot.org/, December, 2018) or with PEAKS X+ software (mouse database downloaded from https://www.uniprot.org/, November, 2019), respectively. In both cases, only those entries that were reviewed were included. The following parameters were used: trypsin was selected as enzyme; a maximum of two or three missed cleavages were allowed for total proteome and acetylome, respectively. For total proteome, carbamidomethylation was set as fixed modification, whereas oxidation in methionines and acetylation of protein N-terminal were considered variable modifications. Protein quantification was performed using the iBAQ intensity. For the acetylome, carbamidomethylation was set as a fixed modification, while methionine oxidation; lysine and N-terminal acetylation; asparagine and glutamine deamidation; serine, threonine and tyrosine phosphorylation; and aspartic acid, tyrosine, threonine, serine, glutamine and asparagine dehydration were used as variable modifications. The mass tolerance for the parental ion and MS/MS fragments were set to 10 ppm and 0.5 Da. In both analyses, a 1%FDR at peptide and protein level was used to filter the results.

Data processing and statistical analyses were performed with R (https://cran.r-project.org/) and R studio (https://www.rstudio.com/). Statistical analysis for the pre-B cell proteome was performed using the 'limma' package. For the PAX5 acetylome, PSMs (peptide-spectrum matches) were calculated with a homemade algorithm and the PEAKS' output files "peptide.csv" and "DB search psm.csv".

16. RNA-Seq analysis

FastQC was used for quality checking of raw FastQ files. Raw counts were obtained with Salmon[349] "quant" pseudoalignment (mm10 reference mouse genome), with fragment-level GC bias correction (--gcBias), eight threads (-p 8) and selective alignment enabled (--validateMappings). The quant.sf files were subsequently used to import transcript-level quantification data into a Summarised Experiment with Tximeta[350], that was then imported

into an R environment. DESeq2[351] was used for differential gene expression (DGE) analysis. Counts were normalised by DESeq2's median of ratios method and transformed to z-score for data visualization with ggplot2. For obtaining BigWig files, raw FastQ files were mapped onto a reference mm10 mouse genome with Bowtie2[352] to generate SAM files. SAM files were then converted into BAM files with SAMTools[353], skipping alignments with MAPQ smaller than 37 (samtools view -bS -q 37). Sambamba[354] was used to sort and filter BAM files, eliminating unmapped and duplicated reads (-F "[XS] == null and not unmapped and not duplicate"). Index BAI files were created with SAMTools "index". BigWig files were generated using deepTools[355] "bamCoverage" (--binSize 20 --normalizeUsing BPM --ignoreForNormalization chrX --extendReads 150 --centerReads --smoothLength 60). Visualization of BigWig files was performed with IGV[356] software.

GSEA (Gene Set Enrichment Analysis) was performed with the software[357] provided by the Broad Institute (https://www.gsea-msigdb.org/), using RNA-Seq normalised count values and default parameters. Enrichr tool[358] was used to obtain gene ontology terms.

17. ChIP-Seq analysis

FastQC was used for quality checking of raw FastQ files before alignment onto a reference mm10 mouse genome with Bowtie2[352] to generate SAM files. SAM files were converted into BAM files with SAMTools[353], skipping alignments with MAPQ smaller than 37 (samtools view -bS -q 37). BAM files were then sorted and filtered with Sambamba[354], eliminating unmapped and duplicated reads (-F "[XS] == null and not unmapped and not duplicate"). Index BAI files were created using SAMTools "index". BigWig files were generated with deepTools[355] "bamCoverage" (--binSize 20 --normalizeUsing BPM --ignoreForNormalization chrX --extendReads 150 --centerReads --smoothLength 60). Visualization of BigWig files was performed with IGV[356] software. MACS2[359] "callpeak" was used for peak calling (-f BAM --nomodel --extsize 20 -g mm -B), considering both IP (-t) and input (-c) samples for each peak calling. Only significant ($q<0.05$) peaks were considered for downstream analyses. Differential peak analysis was performed with ChIPpeakAnno, whereas heatmaps were created with either DiffBind or plotHeatmap.

18. Data accessibility

All ChIP-Seq and RNA-Seq data have been deposited at the National Center for Biotechnology Information Gene Expression Omnibus (GEO) and are available under the accession number: GSE246370. Total pre-B cell proteome and PAX5 acetylome experiments have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository and are available under the identifier PXD046457. H3K18ac, H3K36ac, and H3K36me3 ChIP-Seq data are not publicly available at the time of submission of this thesis.

VI. RESULTS

1. Expression of nuclear sirtuins in haematopoietic progenitors

Many reports have shown important functions played by nuclear sirtuins in various immune cell types, especially myeloid cells and T-cells, as well as in HSCs. However, the roles of nuclear sirtuins in other haematopoietic progenitors remain largely unexplored. Therefore, we sought to identify novel functions of nuclear sirtuins in haematopoiesis to address this gap of knowledge. To do so, we took an unbiased approach based on the profiling of the expression levels of nuclear sirtuins (SIRT1, SIRT2, SIRT6 and SIRT7), at the single-cell level, in mouse haematopoietic progenitors. We used a publicly available single-cell RNA-Seq (scRNA-Seq) dataset from FACS-sorted lineage negative (Lin^-) progenitors from mouse bone marrow. This dataset identified and clustered haematopoietic progenitor and stem cells (HPSCs), as well as B-cell, T-cell, macrophage and ILC progenitors, thus allowing to comprehensively cover haematopoietic lineages. Overall, *Sirt1* and *Sirt2* displayed a relatively homogeneous expression pattern in all the progenitors analysed, although macrophage and ILC lineage cells, as well as HPSCs, expressed the highest *Sirt1* levels (**Fig. R1a,b**). Contrarily, *Sirt2* high expression was modestly skewed towards B- and T-cells, suggesting that *Sirt2* may be important for lymphocyte biology. *Sirt6* expression was surprisingly low in all cell types, which suggested that it may be dispensable for haematopoietic differentiation, despite its known functions in immunity. Of note, *Sirt7* was by far the nuclear sirtuin displaying the highest expression levels across cell types, and this was particularly remarkable in the B-cell cluster, where *Sirt7* expression was even higher. This suggested that SIRT7 may be specifically required for mouse haematopoiesis in general and for B-lymphopoiesis in particular; that SIRT2 may be also required for lymphoid development; and that SIRT1 may play a role in myelopoiesis and ILC development.

The data from the scRNA-Seq profiling prompted us to further explore the upregulation of SIRT7 in B-cell progenitors. To do so, we first tried to confirm whether this upregulation leads to increased SIRT7 protein levels in B-lineage cells. We isolated bone marrow from three different mice and purified the B-cell fraction using magnetic-assisted cell sorting and an antibody against CD19, a membrane protein expressed by all B-lineage cells. Thus, we separated $CD19^-$ and $CD19^+$ BM cells and analysed the protein levels of SIRT7 in both fractions by immunoblotting. Notably, SIRT7 expression was barely detectable in $CD19^-$ cells. In comparison, $CD19^+$ B-cells displayed strikingly high SIRT7 protein levels, which strongly argued for a role of SIRT7 in B-cell biology (**Fig. R1c**). To track further how the upregulation of SIRT7 takes place during development, we next profiled its RNA expression

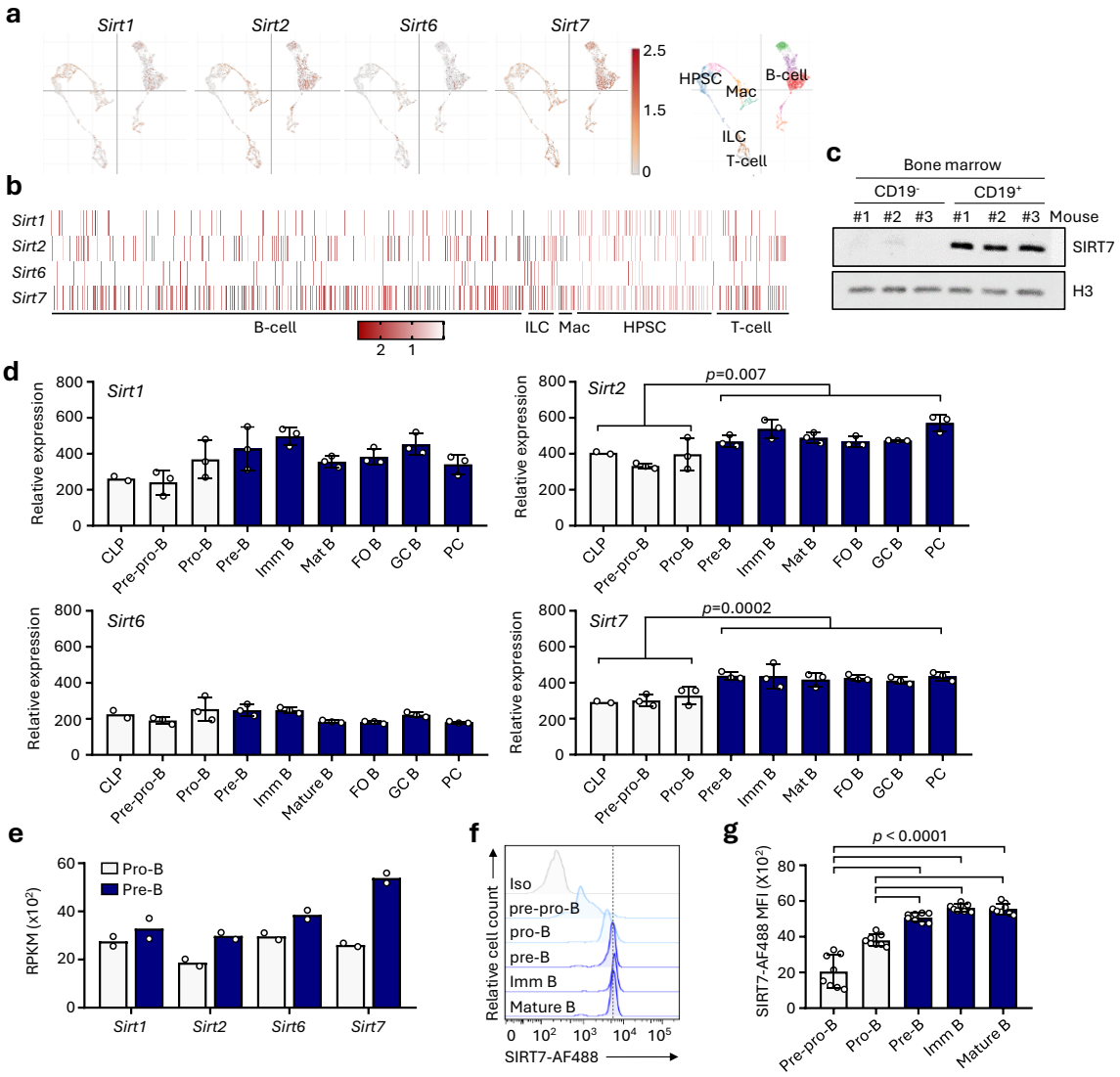


Fig. R1 | SIRT7 is upregulated during early B-cell development in mice

a, tSNE plots displaying the single-cell expression profiles of nuclear Sirtuins (left) and sc-RNA-Seq feature plot (right). **b,** Heatmap displaying the single-cell expression of nuclear Sirtuins in murine bone marrow. **a,b,** Data and tSNE plots were retrieved from the Broad Institute Single Cell Portal (SCP978). **c,** Immunoblot analysis of SIRT7 levels in murine CD19$^-$ and CD19$^+$ bone marrow cells (n = 3). **d,** Microarray expression (GSE15907) of nuclear Sirtuins in B lymphopoiesis stages (n = 3; CLP, n = 2). Data are presented as mean ± s.d. A two-tailed *t*-test was performed comparing pooled CLP–pro-B and pooled pre-B–PC B-cells. **e,** RNA-Seq expression patterns of nuclear Sirtuins in pro-B and pre-B cells. Data are presented as mean (n = 2). **f,g,** Representative FACS plot (**f**) and SIRT7 MFI (**g**) measured by intracellular flow cytometry in the indicated populations (n = 8). In **g**, data are presented as mean ± s.d. Statistical significance was determined by one-way ANOVA with Sidak multiple comparisons. Mac, macrophage; HPSC, haematopoietic and progenitor stem cell; CLP, common lymphoid progenitor; Imm B, immature B-cell; Mat B, mature B-cell; FO B, follicular B-cell; GC B, germinal center B-cell; PC, plasma cell.

levels across B-cell developmental stages, as well as those of SIRT1, SIRT2 and SIRT6, using published microarray data from sorted CLPs, pre-pro-B, pro-B, pre-B, immature (Imm B), mature , follicular (FO B), germinal center (GC B) and plasma cells (PC). Consistent with the scRNA-Seq data, SIRT1 and SIRT6 did not undergo any relevant

changes during development, whereas SIRT2, and specially SIRT7, were upregulated in pre-B cells and downstream stages (**Fig. R1d**). We confirmed these data by sorting pro-B and pre-B cells from mouse bone marrow to perform RNA-Seq, which revealed that only SIRT2 and SIRT7 were upregulated in pre-B cells, with SIRT7 showing the most pronounced upregulation and displaying the highest levels in pre-B cells (**Fig. R1e**). Finally, we decided to corroborate these data at the protein level, so we established a FACS panel to measure the intracellular protein levels of SIRT7 in B-cell progenitor stages. Interestingly, SIRT7 was upregulated in pro-B cells relative to pre-pro-B cells, which was not observed in the microarray data, and its expression continued to increase in pre-B cells, where it reached a plateau (**Fig. F1f,g**). Thus, SIRT7 is upregulated during early B-cell development in mice, suggesting that it may play an important role in this process.

Next, we aimed to determine whether SIRT7 upregulation in B-cells was conserved also in humans. We then took a similar approach and analysed publicly available sc-RNA-Seq data from human bone marrow, although this dataset contained an important limitation: in this case, all the bone marrow mononucleated cells were purified and sequenced, without isolating the progenitor cells. Furthermore, although our analysis effectively clustered several immune cell lineages (B-cells, T-cells, NK cells, plasma cells, classic and plasmacytoid dendritic cells (DC and pDC, respectively), platelets and erythroid cells), it did not have enough resolution to identify specific cell types within each lineage, with two exceptions: it separated B-cells from plasma cells, and DC from pDC (**Fig R2a**). Although this analysis identified different lineages compared to those found in mice, we observed a similar pattern of *SIRT7* expression, with lymphoid cells in general expressing the highest levels (**Fig. R2b**). Surprisingly, *SIRT7* expression collapsed in plasma cells, the

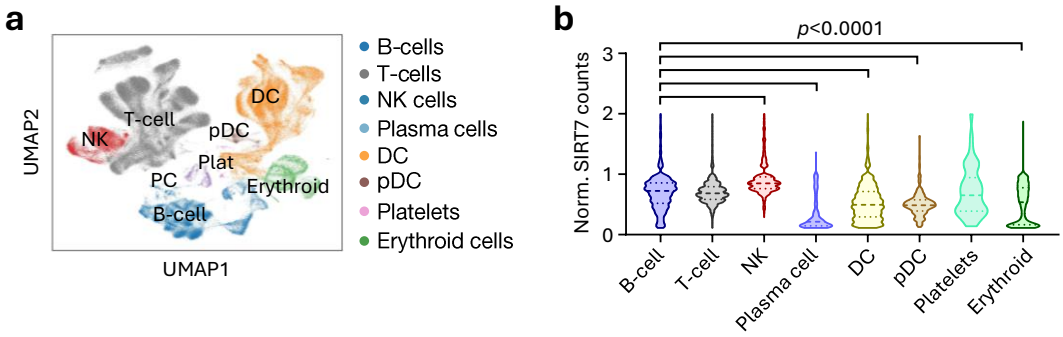


Fig. R2 | Human SIRT7 is preferentially expressed in lymphocytes
a,b, sc-RNA-seq feature plot (**a**) and single-cell expression of SIRT7 (**b**) in the indicated human bone marrow cell types. Data were retrieved from the Broad Institute Single Cell Portal (SCP101). In panel **b**, only those cells with detectable counts are plotted. Dashed lines indicate the median and the first and fourth quartiles. Statistical significance was assessed by one-way ANOVA with Dunnet comparison. DC, dendritic cell; pDC, plasmacytoid dendritic cell; Plat, platelets.

B-cell subset responsible for antibody secretion upon B-cell activation. Furthermore, NK-lineage cells displayed the highest *SIRT7* expression in humans, followed by B-cells and T-cells. Because this analysis included both progenitor and differentiated cells, these data suggested further potential functions of SIRT7 in human NK cell development or functions and supported the idea that SIRT7 may be required for B-cell biology also in humans.

2. SIRT7 is required for early B-cell development

The above data strongly supported the previous findings obtained by our laboratory suggesting that SIRT7 may indeed be involved in B-cell development, which compelled us to validate these observations. To this end, we collected bone marrow samples from *Wt* and *Sirt7*$^{-/-}$ 129/Sv mice and measured by FACS the total numbers of B220^{+}CD19^{+} B-cells in these samples. This clearly showed that the B-cell number was significantly reduced in SIRT7-deficient mice, thus confirming our previous data (**Fig. R3a**). Next, we took a similar approach to explore in more detail how the lack of SIRT7 impacts B-cell development. Thus, we analysed the numbers of pre-pro-B, pro-B, pre-B, immature B and mature B cells in the bone marrow of these mice and observed significant reductions in the numbers of pre-B cells and downstream stages in *Sirt7*$^{-/-}$ mice (**Fig. R3b,c**). In contrast, *Sirt7*$^{-/-}$ mice

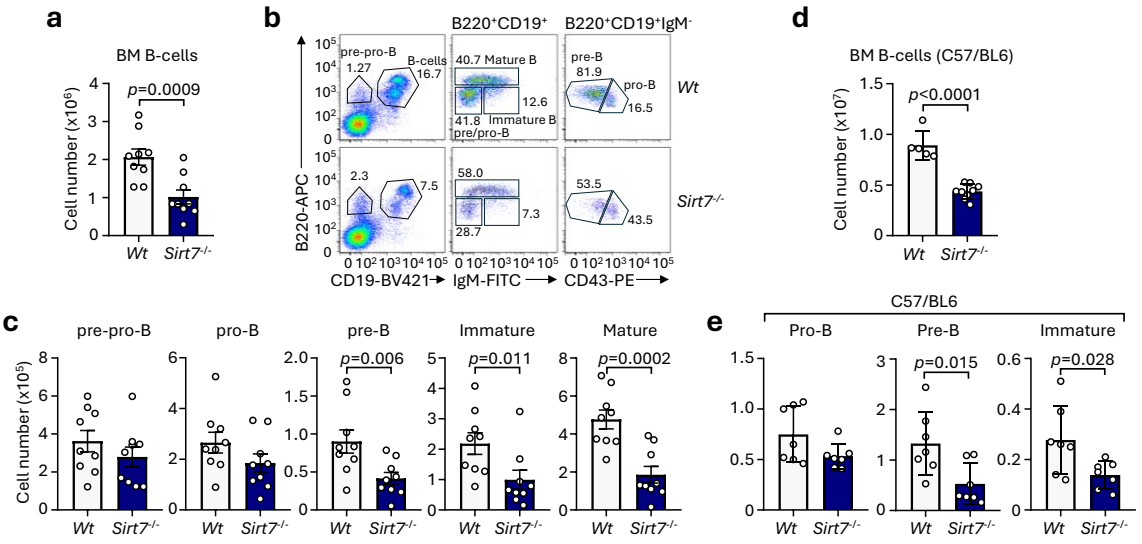


Fig. R3 | SIRT7 is required for early B-cell development

a, Numbers of B220^{+}CD19^{+} B-cells in the bone marrow of 129Sv *Wt* and *Sirt7*$^{-/-}$ mice (n=9). Data are presented as mean ± SEM and were analysed by one-tailed *t*-test. **b,c,** Representative FACS plots (**b**) and number (**c**) of pre-pro-B, pro-B, pre-B, immature and mature B-cell populations in the bone marrow of *Wt* and *Sirt7*$^{-/-}$ mice (n=9). Data in **c** are shown as in panel **a**. **d,e,** Number of B220^{+}CD19^{+} B-cells (**d**) and percentages of pre-pro-B, pro-B, pre-B and Immature B-cell populations (**e**) determined by flow cytometry in the bone marrow of *Wt* and *Sirt7*$^{-/-}$ C57BL/6 mice. Data are presented as mean ± s.d. and significance was assessed by two-tailed *t*-test (*Wt*, n=5 (**d**) and 7 (**e**); *Sirt7*$^{-/-}$, n=9 (**f**) and 7 (**g**)). GC B, germinal center B-cell.

had normal numbers of pre-pro-B and pro-B cells, which indicated a developmental defect at the pro-B-to-pre-B cell transition. To exclude the possibility that this was a strain-specific defect, we repeated these analyses using an alternative *Sirt7*$^{-/-}$ mouse model (reported in Ref [313]), which was generated with a different knock-out strategy and were maintained in the C57BL/6 background. Consistent with our findings in 129/Sv mice, C57BL/6 *Sirt7*$^{-/-}$ mice also displayed reduced numbers of total B-cells due to a defect starting at the pre-B cell stage (**Fig. R3d,e**), which confirmed that SIRT7 promotes pre-B cell development in mice. Because our lab mainly works with the 129/Sv strain, we hereafter used this strain for all studies involving SIRT7-deficient mice.

2.1. SIRT7 promotes B-cell immunity

After completing differentiation, naive B-cells migrate to secondary lymphoid organs, particularly the spleen, where they accumulate in anatomical regions called follicles (follicular B-cells) or in their vicinity (marginal zone B-cells) and wait to encounter their cognate antigens. We next investigated how the reduced output of B-cell progenitors from the bone marrow impacted the biology of peripheral B-cells, using different approaches. First, we evaluated whether SIRT7 deficiency caused any histological changes in the spleen, by collecting the spleens of *Wt* and *Sirt7*$^{-/-}$ mice, fixing them and staining them with haematoxylin and eosin. SIRT7 deficiency led to a loss of the structural architecture of the spleen, with the spleens of *Sirt7*$^{-/-}$ mice showing fewer and less organised follicles than those of *Wt* mice, which suggested compromised B-cell immunity in the absence of SIRT7 (**Fig. R4a**). In line with the histology, FACS analysis of splenic CD19^{+} B-cells showed concomitant reductions in both the percentage and the total numbers of splenic B-cells in mice lacking SIRT7 (**Fig. R4b,c**).

We went further in the characterization of the phenotype produced by SIRT7 deficiency in B-cell immunity by analysing the major populations of mature B-cells in the spleen; namely, transitional, marginal zone, follicular, germinal center, memory, plasma and class-switched IgG1^{+} B-cells. As expected, we found significant reductions in the number of most of these populations, supporting the idea that defective B-cell development leads to impaired B-cell immunity in *Sirt7*$^{-/-}$ mice (**Fig. R4d-i**). In particular, the numbers of transitional B-cells, that represent an intermediate stage between BM immature B-cells and peripheral mature B-cells, displayed a strong reduction (**Fig. R4d,e**). Transitional B-cells further develop into (i) T-cell-independent marginal zone B-cells (innate-like B-cells

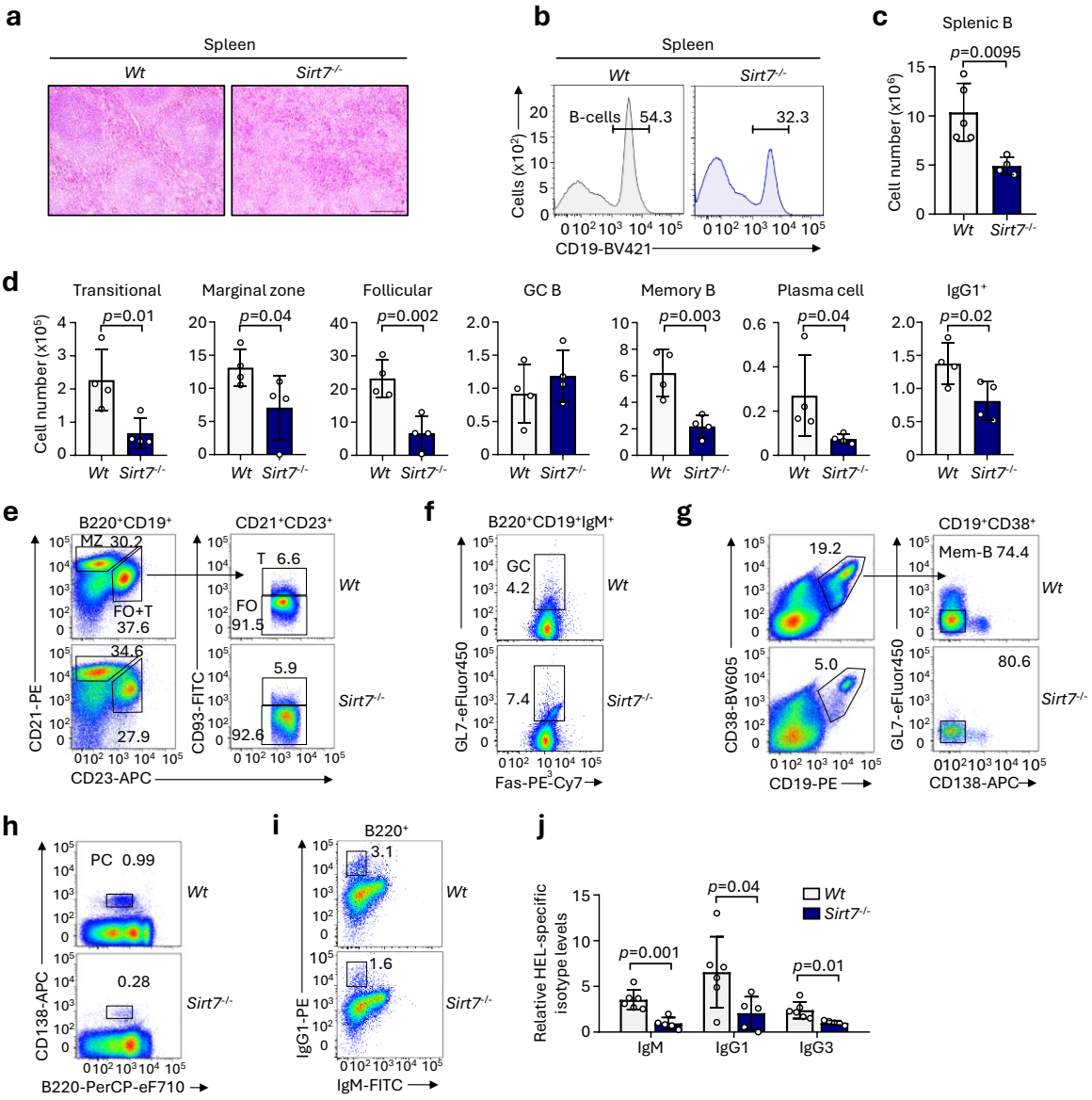


Fig. R4 | SIRT7 promotes B-cell immunity

a, Haematoxylin-eosin staining of histological sections from the spleens of *Wt* and *Sirt7*$^{-/-}$ mice. Images are representative of five spleens per genotype. Scale bar = 500 µm. **b,c,** Representative FACS plots (**b**) and numbers (**c**) of splenic CD19^{+} B-cells. Data in panel **c** are presented as mean ± s.d and statistical significance was determined by one-tailed *t*-test (*Wt*, n = 5; *Sirt7*$^{-/-}$, n = 4). **d,** Number of splenic transitional, marginal zone, follicular, germinal center (GC) and class-switched IgG1^{+} B-cells, as well as bone marrow memory B-cells and plasma cells, from *Wt* and *Sirt7*$^{-/-}$ mice. Data are shown as in panel **c** (n= 4). **e-i,** Representative FACS plots of splenic marginal zone, follicular, transitional (**e**), germinal center (**f**), BM memory B-cells (**g**), BM plasma cells (**h**) and class-switched splenic IgG1^{+} B-cells (**i**) from *Wt* and *Sirt7*$^{-/-}$ mice (n = 4). **j,** Levels of anti-HEL antibody isotypes in the sera of *Wt* and *Sirt7*$^{-/-}$ mice 14 days after immunization, relative to naive mice, as determined by ELISA. Data are shown as mean ± s.d. and significance was assessed by two-tailed *t*-test (*Wt*, n = 6; *Sirt7*$^{-/-}$, n = 5). GC B, germinal center B-cell

that eventually differentiate into plasmablasts (or plasma cells) upon activation); or (ii) follicular B-cells, that differentiate into plasmablasts or memory B-cells upon antigen activation, in a T-cell-dependent manner. As expected, both populations were significantly

reduced in *Sirt7*$^{-/-}$ mice, although this effect was more evident in the case of follicular B-cells, suggesting that marginal zone B-cell differentiation was less dependent on SIRT7 (**Fig. R4d,e**). When follicular B-cells encounter their antigen, they migrate to the center of the follicle and form germinal centers (GCs), where GC B-cells can interact with follicular helper T (T_{FH})-cells, that critically participate in their downstream differentiation into plasma cells or memory B-cells[360]. Unexpectedly, SIRT7 null mice had normal numbers of GC B-cells, due to a marked increase in their percentage compared to *Wt* mice (**Fig. R4d and R4f**). This suggested that SIRT7 deficiency may provoke a second developmental arrest in late B-cell differentiation. Indeed, the numbers of downstream memory B-cells and plasma cells were also significantly reduced in *Sirt7*$^{-/-}$ mice, as were those of class-switched IgG1^{+} B-cells, although their reduction was more modest (**Fig. R4d and R4g-i)**. Collectively, these data indicated that SIRT7 is required for the optimal generation of most mature B-cell subsets and that it may play stage-specific functions during early and late B-cell development.

To directly assess the functional relevance of the above observations in B-cell immunity, we also immunised *Wt* and *Sirt7*$^{-/-}$ mice with the T-cell-dependent antigen NP-HEL (NPP hapten-conjugated hen egg lysozyme) through intraperitoneal injection. 14 days after immunisation, we collected peripheral blood from these mice, isolated the serum fraction, and measured the levels of serum HEL-specific IgM, IgG1 and IgG3 antibodies by ELISA. As expected, SIRT7-deficient mice poorly responded to immunization, since we found significant reductions in the levels of all three isotypes analysed (**Fig. R4j**). This confirmed that defective B-cell development leads to impaired B-cell-mediated immune responses in in *Sirt7*$^{-/-}$ mice.

2.2. SIRT7 plays a pro-B cell-intrinsic role in early B-cell development and commitment that requires its deacetylase activity

Because in this project we used a whole-body knock-out model, at this point we could not rule out the possibility that the B-cell developmental arrest caused by SIRT7 deficiency is due to a defect in the bone marrow microenvironment rather than a B-cell autonomous impairment. To exclude this possibility, we performed *in vivo* transplantation experiments using the CD45.1/CD45.2 system. This system leverages the existence of two different allelic variants of the pan-haematopoietic surface marker CD45 (CD45.1 and CD45.2) that differ only in five amino acids of their extracellular domains, which allowed the

development of monoclonal antibodies that individually recognise each of these variants[361]. Since CD45 is expressed by most haematopoietic and immune cell types and most mouse strains express the CD45.2 variant, this system provides a unique approach to perform lineage tracing experiments by purifying CD45.2 cells and transplanting them into backcrossed CD45.1 recipient mice, which enables to trace the progeny of the transplanted cells. To prevent immune rejection against the donor cells, we crossed *Wt* CD45.2 129/Sv mice with CD45.1 C57BL/6 mice to breed mixed-background CD45.1CD45.2 mice as recipients (since CD45.1 mice are in the C57BL/6 background and our *Sirt7*$^{-/-}$ colony is 129/Sv). Thus, we sorted Lin^{-}B220^{+}CD19^{+}IgM^{-} pro-B cells from *Wt* and *Sirt7*$^{-/-}$ mice, expanded them *ex vivo* in the presence of OP9 feeder cells for a short period, separated the pro-B cells from the coculture and injected intravenously these cells into sublethally irradiated CD45.1CD45.2 recipient mice. Four weeks after transplantation, we collected the spleens of the recipient mice and determined by FACS the number of CD45.1^{-}CD45.2^{+}CD19^{+} cells, containing the progeny of the injected pro-B cells (**Fig. R5a**). As expected, *Wt* CD45.2 pro-B cells efficiently repopulated the splenic B-cell compartment, whereas *Sirt7*$^{-/-}$ CD45.2 pro-B cells largely failed to do so, as deduced by the almost complete absence of CD45.1^{-}CD45.2^{+}CD19^{+} B-cells in mice transplanted with *Sirt7*$^{-/-}$ pro-B cells (**Fig. R5b**). Since *Sirt7*$^{-/-}$ pro-B cells failed to develop into mature B-cells in the context of a *Wt* bone marrow, this experiment confirmed that SIRT7 plays a B-cell-intrinsic role in B-cell development.

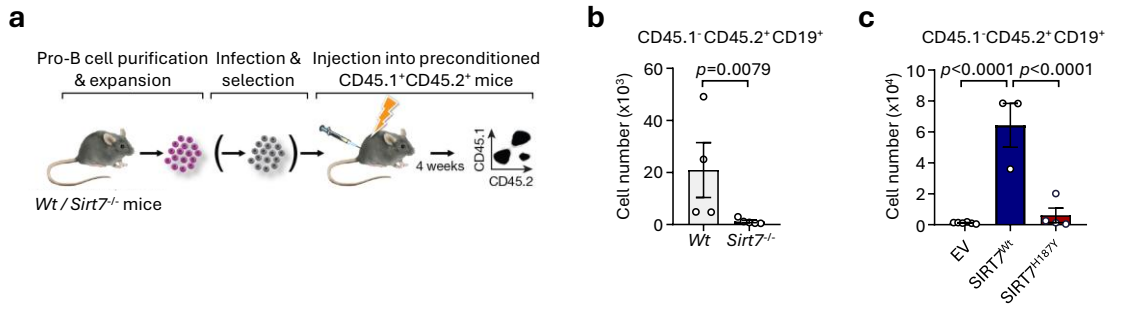


Fig. R5 | SIRT7 plays a pre-B cell-intrinsic role in B-cell lymphopoiesis through its deacetylase activity
a-c, Schematic representation (**a**) and number of CD45.1^{-}CD45.2^{+}CD19^{+} B-cells (**b,c**) in spleens of CD45.1CD45.2 recipient mice 4 weeks after transplantation. Data are presented as mean ± SEM and were analysed by one-tailed *t*-test (*Wt*, n = 4; *Sirt7*$^{-/-}$, n = 5) (**b**) or by one-way ANOVA with Sidak multiple comparisons (EV, n = 6; SIRT7^{Wt}, n = 3; SIRT7^{H187Y}, n = 4) (**c**). In panel **a**, the brackets indicate that the infection and selection steps were only performed for **c**.

As described above, SIRT7 performs many of its functions through deacetylation of either histone marks or transcription factors. Therefore, we leveraged the CD45.1/CD45.2 system to explore (i) whether the developmental arrest induced by SIRT7 in B-cell progenitors is reversible, and (ii) whether the catalytic activity of SIRT7 is required for B-cell development. To this end, we purified and expanded *Sirt7*$^{-/-}$ pro-B cells using the same methodology as before and transduced these cells with retroviruses expressing either an empty vector (EV) or the coding sequences of SIRT7^{Wt} or the deacetylase-dead SIRT7^{H187Y} mutant, in which the mono-ADRT activity of SIRT7 is unaffected[263]. In both cases, the vectors contained a cassette with an internal ribosome entry site (IRES) and the coding sequence of the human CD4 (hCD4) marker. Thus, after transduction, we purified the infected cells by anti-hCD4 cell sorting, expanded them again and injected them into recipient mice (see the diagram in **Fig. R5a**). Four weeks later, *Sirt7*$^{-/-}$ pro-B cells ectopically expressing *Wt* SIRT7 efficiently repopulated the splenic B-cell compartment, whereas cells expressing either the EV or the SIRT7^{H187Y} mutant were largely unable to do it (**Fig. R5c**). Thus, we concluded that SIRT7 deficiency causes a reversible B-cell developmental arrest, and the deacetylase activity of SIRT7 mediates its functions in B-cell development.

One key event that takes place at this stage is the establishment of lineage commitment, so we reasoned that SIRT7 might also play a role in this process. To test this hypothesis, we took a similar approach, based on the purification, *ex vivo* expansion and transplantation of Lin^{-}B220^{+}CD19^{+}IgM^{-} pro-B cells from *Wt* and *Sirt7*$^{-/-}$ mice. After four weeks, we analysed by FACS the presence of pro-B cell-derived CD45.1^{-}CD45.2^{+} TCRβ^{+} T-cells or NKp46^{+} NK cells in the spleens of the recipient mice, which would be indicative of defective commitment. As expected, we found no evidence of lineage conversion in mice injected with *Wt* pro-B cells, since no pro-B cell-derived T-cells nor NK cells were detectable (**Fig. R6a**). Remarkably, this was not the case for *Sirt7*$^{-/-}$ pro-B cells. We found that about 30% of the progeny of the injected *Sirt7*$^{-/-}$ cells lost CD19 expression, indicating that they differentiated into an alternative lineage (**Fig. R6a,b**). In particular, all CD19^{-} cells stained positive for either TCRβ or NKp46, which indicated that *Sirt7*$^{-/-}$ pro-B cells display abnormal T-cell and NK potential and confirmed that SIRT7 is required for B-cell commitment (**Fig. R6c**). Next, we also explored whether impaired commitment in *Sirt7*$^{-/-}$ progenitors was inherited by mature B-cells, as has been described with the B-cell transcription factor PAX5[3], whose conditional knock-out in mature B-cells reverts lineage commitment and leads to lineage conversion. To test this possibility, we sorted Lin^{-}

IgM$^+$IgD$^+$ mature B-cells from the spleens of *Wt* and *Sirt7*$^{-/-}$ mice, injected these cells into sublethally irradiated CD45.1CD45.2 mice, and analysed their spleens 6 weeks after transplantation. Importantly, no pro-B cell-derived TCRβ$^+$ T-cells nor NKp46$^+$ NK cells were detected in this case, suggesting that SIRT7 is dispensable for the maintenance of lineage commitment once it has been established (**Fig. R6d**). Thus, these data strongly supported the possibility that SIRT7 plays different roles during early and late B-lymphopoiesis. Overall, our data demonstrates that SIRT7 is required for early B-cell development, commitment and immunity, through a mechanism that depends on its deacetylase activity.

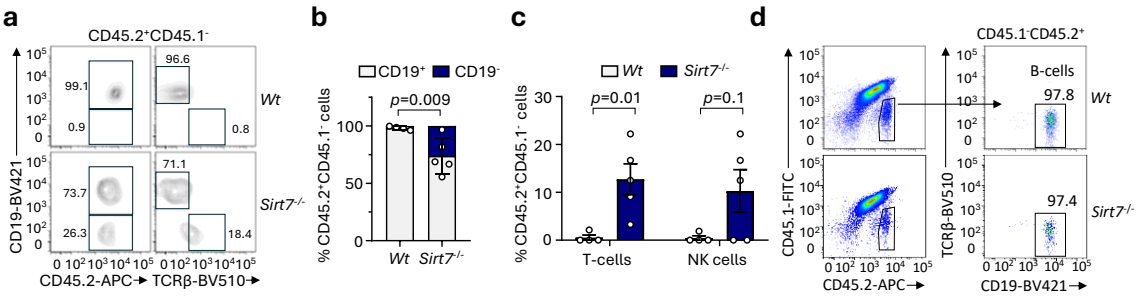


Fig. R6 | SIRT7 is required for B-cell lineage commitment

a-c, Representative FACS plots (**a**) and percentages of donor-derived CD45.1$^-$CD45.2$^+$CD19$^+$ and CD45.1$^-$CD45.2$^+$CD19$^-$ cells (**b**), and of CD45.1$^-$CD45.2$^+$TCRβ$^+$ T-cells and CD45.1$^-$CD45.2$^+$NKp46$^+$ NK cells (**c**) in the spleens of recipient mice injected with *Wt* or *Sirt7*$^{-/-}$ pro-B cells. Data are presented as mean ± s.d. (**b**) or mean ± SEM (**c**) and were analysed by two-way ANOVA with Fisher's LSD test (**b**) or multiple *t*-tests with Holm-Sidak comparison (**c**) (*Wt*, n = 4; *Sirt7*$^{-/-}$, n = 5). **d,** Representative FACS plots of recipient CD45.1CD45.2 mice 6 weeks after injection of *Wt* or *Sirt7*$^{-/-}$ Lin$^-$IgM$^+$IgD$^+$ splenic B-cells (n = 3).

2.3. The mADPRT activity of SIRT7 is dispensable for B-cell development

Our laboratory has shown that, in addition to its deacetylase activity, SIRT7 carries a second catalytic activity that mediates in part its functions in stress response[263]. In that paper, we described that both activities were mediated by two different active sites, and that conservative mutations of two different residues specifically abrogated one activity without affecting the other (HDAC, H187Q; mADPRT, N189Q). More recently, we developed a mouse model carrying a point mutation in the *Sirt7* gene, that leads to the expression of an endogenous SIRT7^{N189Q} mutant, in which the mADPRT, but not the HDAC, activity is abrogated. Although our data showed that the HDAC activity of SIRT7 was indispensable for B-cell development, we used this mouse model to evaluate the requirement of the mADPRT activity for this process. To do so, we collected bone marrow and spleen samples from *Wt* and *Sirt7*^{N189Q} mice and measured the number of total B-cells

and B-cell progenitors in these mice. We found no major differences in the numbers of any of the populations analysed, which indicated that the mADPRT activity of SIRT7 is dispensable for B-cell development and provided further support to the observation that the deacetylase activity of SIRT7 mediates its major functions in B-lymphopoiesis (**Fig. R7a-c**).

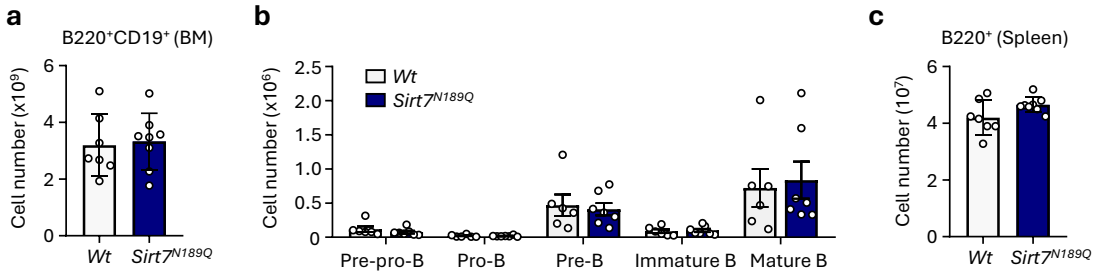


Fig. R7 | The mADPRT activity of SIRT7 is dispensable for B-lymphopoiesis
a, Numbers of B220$^+$CD19$^+$ B-cells in the bone marrow of *Wt* and *Sirt7*^{N189Q}. Data are presented as mean ± s.d. and were analysed by one-tailed *t*-test. **b,** Number of pre-pro-B, pro-B, pre-B and immature B-cell populations in the bone marrow of *Wt* and *Sirt7*^{N189Q} mice. Data are shown as in panel **a**. **c**, Number of B220$^+$ B-cells in the spleens of *Wt* and *Sirt7*^{N189Q} mice. Data are shown as in **a (**Wt, n =7; *Sirt7*^{N189Q}, n=8).

3. SIRT7 promotes pre-B cell survival and expansion through a V(D)J recombination-independent mechanism

The pro-B-to-pre-B cell transition represents a key checkpoint in the B-cell developmental pathway. To pass through this checkpoint, pro-B cells must perform in-frame recombination of V_H, D_H and J_H elements in the *Igh* locus and generate a productive rearrangement of the locus to express IgH molecules, and ultimately assemble a pre-BCR. The pre-BCR complex initiates a signalling cascade that allows early pre-B cells (or large pre-B cells) to evade apoptosis and progress to the next developmental stages[88]. Together with the IL7/IL7R/JAK/STAT5 axis, the pre-BCR coordinates the clonal expansion of large pre-B cells and their subsequent transition to small pre-B cells, that exit the cell cycle and perform the recombination of the *Igk* locus[88]. Since SIRT7 plays a key role in DNA repair and has been linked to cell proliferation in several cell types, we reasoned that it may contribute to the repair of DNA double-strand breaks during V(D)J recombination and/or to the survival and proliferative signalling that follows its completion. Thus, we designed a battery of experiments to address the relevance of SIRT7 in apoptosis, cell cycle and V(D)J recombination in B-cell progenitors.

3.1. SIRT7 promotes pre-B cell expansion

First, we performed Annexin V and 7AAD staining in bone marrow samples previously stained with markers of B-cell progenitors. *Sirt7*$^{-/-}$ pro-B cells displayed a normal percentage of Annexin V^{+}7AAD^{+} apoptotic cells, whereas *Sirt7*$^{-/-}$ pre-B underwent increased apoptosis, suggesting impaired progression through the pre-BCR checkpoint (**Fig. R8a,b**). Next, we stained and permeabilised bone marrow samples and subsequently incubated them with 4',6-diamidino-2-phenylindole (DAPI) to measure their DNA content by FACS. Consistent with an impaired pre-BCR checkpoint progression, *Sirt7*$^{-/-}$ large pre-B cells (FSC^{high}) were prominently arrested at G1, as inferred by the increased percentage of cells in this phase of the cycle and the decreased percentages of cells in S and G_2/M phases (**Fig. R8c,d**). As expected, no cycling small pre-B cells (FSC^{low}) were detected in either *Wt* or *Sirt7*$^{-/-}$ mice. To give further support to this observation, we also analysed the responsiveness of *Wt* and *Sirt7*$^{-/-}$ pre-B cells to IL7 stimulation. To do so, we incubated bone marrow samples in serum-free medium for 30 min at 37ºC to extinguish the signal of the pathway and subsequently added 10 ng/mL IL7 to the medium and incubated the samples for a further 30 min to stimulate it. We then permeabilised the cells and stained them with pre-B cell markers together with an antibody against phosphorylated STAT5^{Y694}, the main mediator of pre-B cell proliferation. FACS measurement of phospho-STAT5^{Y694} in pre-B cells revealed impaired activation of the pathway in *Sirt7*$^{-/-}$ pre-B cells, since phosphorylation of STAT5^{Y694} was significantly reduced in these cells (**Fig. R8e,f**). Furthermore, it is worth mentioning that when we expanded pro-B cells *ex vivo* for the *in vivo* transplantations described above, a procedure that requires IL7 stimulation, we noted that *Sirt7*$^{-/-}$ pro-B cells reproducibly proliferated much more slowly than *Wt* cells did, and they displayed noticeably compromised survival (not shown). Of note, we also measured the protein expression levels of the α chain of IL7R (also known as CD127) in pre-B cells but did not find any major changes, indicating that SIRT7 promoted IL7-mediated STAT5 activation downstream of the receptor (**Fig. R8g**). This raised the possibility that SIRT7 directly regulates STAT5 activation, which prompted us to explore whether SIRT7 and STAT5 were interaction partners. Therefore, we immunoprecipitated endogenous SIRT7 from HAFTL cells and analysed the eluates by immunoblotting. Indeed, SIRT7 and STAT5 co-immunoprecipitated together, suggesting a molecular interplay between these two proteins that may be important for STAT5 activation and pre-B cell proliferation (**Fig. R8h**). Overall, these data strongly suggest that SIRT7 promotes cell survival and IL7-dependent proliferation. However, although these observations were

compelling and might partly explain the phenotype of *Sirt7*$^{-/-}$ progenitors, we did not develop further this issue, which would be very interesting to explore in more detail in the future.

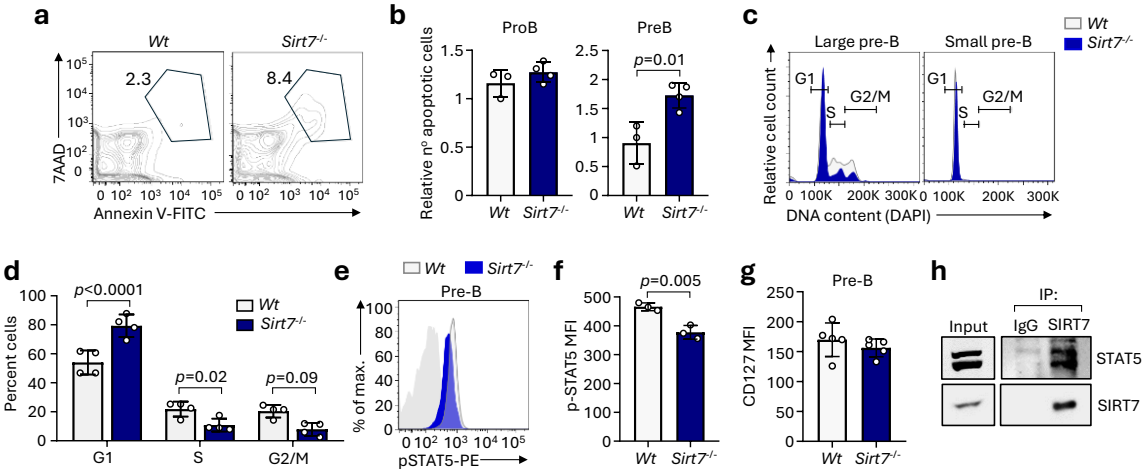


Fig. R8 | SIRT7 sustains B-cell survival and proliferation
a, Representative apoptosis FACS plots in pre-B cells from the bone marrow of *Wt* and *Sirt7*$^{-/-}$ mice (n = 3). **b,** Fold-changes (relative to *Wt*) of 7AAD^{+}Annexin V^{+}-gated pro-B and pre-B cells in the bone marrow of *Wt* and *Sirt7*$^{-/-}$ mice. Data are presented as mean ± s.d. and were analysed by one-tailed *t*-test (n = 3). **c,d,** Representative FACS plots (**c**) and flow cytometric quantification of the percentages of cells (**d**) in the indicated cell cycle stages. Data are shown as mean ± s.d. Statistical significance was determined by two-way ANOVA with Sidak multiple comparison (n = 4). **e,f,** Representative FACS plots (**e**) and intracellular flow cytometric quantification (**f**) of p-STAT5^{Y694} MFI. In panel **e**, the filled grey profile indicates the isotype control. In panel **f**, data are presented as mean ± s.d. and were analysed by two-tailed *t*-test (n = 3). **g,** Flow cytometric quantification of CD127 expression in gated pre-B cells from the bone marrow of *Wt* and *Sirt7*$^{-/-}$ mice. Data are shown as mean ± s.d. (n = 5). **h,** Anti-SIRT7 co-immunoprecipitation followed by anti-STAT5 and anti-SIRT7 immunoblotting. Representative of 3 independent experiments.

3.2. SIRT7 is required for long-range V(D)J recombination but promotes B-cell development through a V(D)J-independent mechanism

SIRT7 plays a crucial role in DNA repair by controlling NHEJ and HR pathways in several cell types. In the case of NHEJ, SIRT7 recruits the protein 53BP1, which is an essential factor of the DNA repair machinery, to double-strand breaks[288]. In lymphocytes, 53BP1 is known to be indispensable for the joining of distal V_H gene segments during V(D)J recombination but not for proximal rearrangements[362]. This is particularly important for the pro-B-to-pre-B cell transition, where V_H-to-D_HJ_H rearrangements often involve long-range recombination, in contrast to D_H-to-J_H and V_K-to-J_K events, which involve only short-range recombination. Thus, defective V_H-to-D_HJ_H could underlie the phenotype caused by SIRT7 deficiency in B-cell progenitors, an idea that was supported by two independent pieces of

evidence. First, real-time quantitative PCR (RT-qPCR) performed with RNA extracted from pro-B cells revealed reduced expression of the *Rag1* recombinase in *Sirt7*$^{-/-}$ pro-B cells (**Fig R9a**). Second, we directly analysed the usage of some of the most frequent V_H, D_H, and J_H gene families by sorting splenic IgM$^+$ cells from *Wt* and *Sirt7*$^{-/-}$ mice, extracting their genomic DNA, and subjecting it to semiquantitative PCR. For the PCR, we used previously described degenerate primers[363] that hybridise to elements belonging to some of these families and allow to semi-quantitatively measure D_H-to-J_H and V_H-to-D_HJ_H rearrangements (which take place in the pre-pro-B-to-pro-B and pro-B-to-pre-B cell transitions, respectively) (**Fig. R9b**). We did not find differences in any of the proximal rearrangements that we analysed (D_HL-J_H3 and V_HQ52), whereas distal V_H588-D_HJ_H and V_H7183-D_HJ_H rearrangements were significantly underrepresented in *IgM*$^+$ *Sirt7*$^{-/-}$ B-cells, suggesting that SIRT7 promotes distal but not proximal *Igh* rearrangements (**Fig. R9c,d**). Thus, a defect in the repair of distal V_H gene segments may underlie the pro-B cell developmental arrest in *Sirt7*$^{-/-}$ mice.

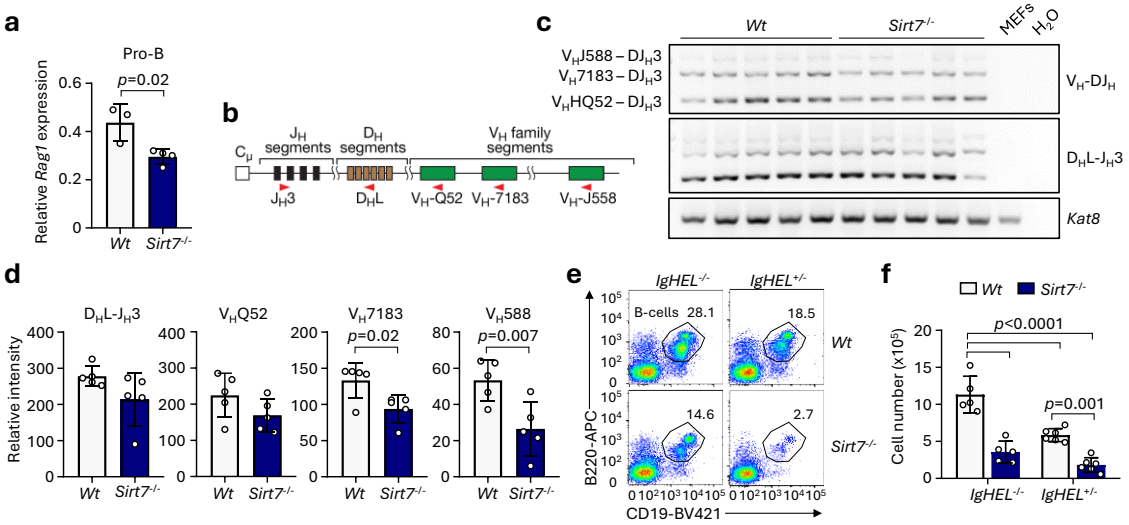


Fig. R9 | SIRT7 promotes distal V(D)J recombination but does not mediate B-cell development through this mechanism

a, RT-qPCR analysis of *Rag1* RNA expression relative to *B2m* in sorted pro-B cells. Data are shown as mean ± s.d. and were analysed by two-tailed *t*-test (*Wt*, n = 3; *Sirt7*$^{-/-}$, n = 4). **b-d,** Schematic representation (**b**), semi-quantitative genomic DNA PCR (**c**) and *Kat8*-normalised densitometric quantification (**d**) using degenerate primers targeting the regions indicated by red arrows in the mouse *IgH* locus. In panel **c**, the independent locus *Kat8* was used as a loading control. Data in panel **d** are shown as in panel **a** (n = 5). **e,f,** Representative plots (**e**) and number of B220$^+$CD19$^+$ B-cells (**f**) in the bone marrow of *IgHEL*$^{-/-}$, *Sirt7*$^{-/-}$ *IgHEL*$^{-/-}$, *IgHEL*$^{+/-}$ and *Sirt7*$^{-/-}$*IgHEL*$^{+/-}$ mice. Data are shown as mean ± s.d. and were analysed by one-way ANOVA with Tukey multiple comparison (*IgHEL*$^{-/-}$ group, n = 5; *IgHEL*$^{+/-}$ group, n = 6). MEFs, mouse embryonic fibroblasts.

Finally, to determine unambiguously whether SIRT7 promotes pro-B cell differentiation by facilitating the repair of V(D)J recombination double-strand breaks, we crossed our *Wt* or *Sirt7*$^{-/-}$ mice with mice carrying a transgenic HEL-specific immunoglobulin (IgHEL)[339]. Expression of the IgHEL transgene in B-cell progenitors prevents endogenous *Igh* and *Igk* recombination due to the constitutive expression of this functional immunoglobulin, thereby bypassing V(D)J checkpoints and rescuing B-cell developmental phenotypes attributable to V(D)J recombination defects. However, analysis of bone marrow samples from these mice revealed that the IgHEL transgene failed to rescue B-cell development in mice lacking SIRT7, as the number of B-cells was still significantly reduced in *Sirt7*$^{-/-}$ IgHEL$^{+/-}$ mice (**Fig. R9e,f**). Therefore, although SIRT7 sustains cell survival, proliferation and distal recombination in B-cell progenitors, it promotes B-cell development through a V(D)J recombination-independent mechanism.

4. SIRT7 regulates PAX5-dependent transcriptional repression during the pro-B-to-pre-B cell transition

Since impaired progression through the pre-BCR checkpoint could not explain the developmental arrest shown by pro-B cells in SIRT7-deficient mice, we next investigated the transcriptional programmes regulated by SIRT7 in B-cell progenitors. To this end, we sorted pro-B and pre-B cells from the bone marrow of *Wt* and *Sirt7*$^{-/-}$ mice and performed RNA-Seq with these samples. In line with the stronger phenotype produced by SIRT7 deficiency in pre-B cells, SIRT7 also had a stronger effect on the transcriptional programmes of pre-B cells compared to those of pro-B cells. In pro-B cells, SIRT7 significantly regulated 220 genes (FDR<0.05), including genes tightly linked with B-cell development, such as the pre-BCR component *Vpreb2* and the transcription factors *Ikzf3* (that encodes for the Ikaros homolog Aiolos) and *Irf4* (**Fig. 10a**), whereas in pre-B cells it regulated 429 genes, that also included key B-cell developmental genes, such as *Vpreb1*, the nuclear corepressor *Ncor2* and the transcription factors *Runx2*, *Mef2c* and *Erg*. Consistent with our previous observations, the top differentially regulated genes (DEGs) in *Sirt7*$^{-/-}$ pre-B cells also included key genes involved in cell survival (*Bcl2*) and pre-B cell expansion (*Myc*) and confirmed the downregulation of *Rag1*. Indeed, gene set enrichment analysis (GSEA) showed that one of the most enriched transcriptional signatures in *Sirt7*$^{-/-}$ progenitors was related to cell cycle regulation (E2F targets, FDR = 0.006) (**Fig. R10b**). Importantly, GSEA further revealed that the gene expression patterns of SIRT7-deficient

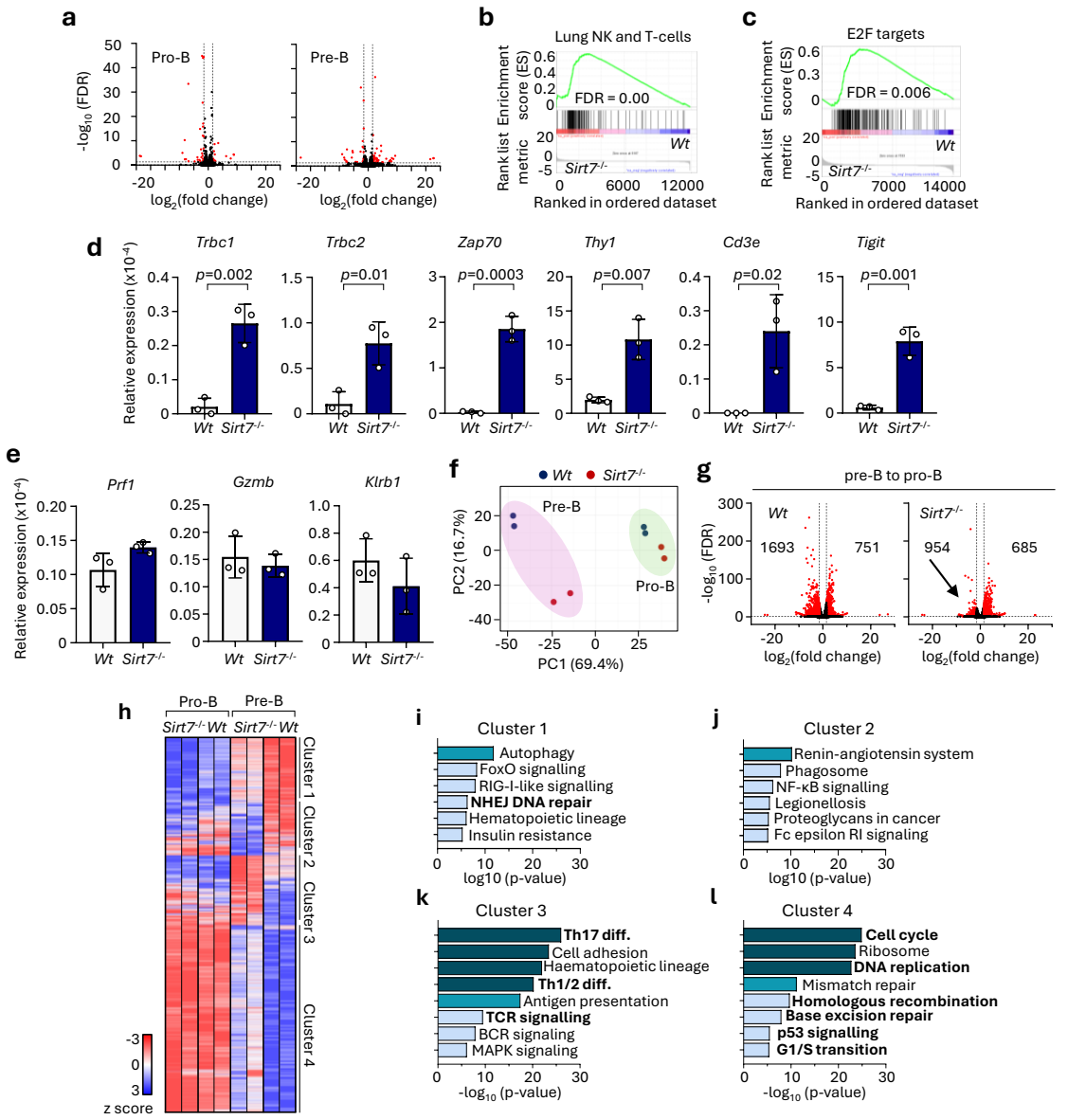


Fig. R10 | SIRT7 is involved in transcriptional repression during the pro-B-to-pre-B cell transition

a, Volcano plot of differentially expressed genes in *Sirt7*$^{-/-}$ versus *Wt* pro-B (left panel) and pre-B cells (right panel). The numbers of significantly induced and repressed genes are shown in the right and left parts of the plots, respectively. Red dots indicate genes that are significantly regulated (|$\log_2$(fold change)|>1.5, FDR<0.05). **b,c,** Gene set enrichment analysis (GSEA) of the "TRAVAGLINI_LUNG_PROLIFERATING_ NK_T_CELL" (**b**) and "HALLMARK_E2F_TARGETS" (**c**) gene sets. Pathways were enriched in *Sirt7*$^{-/-}$ pre-B cells compared to *Wt* pre-B cells. **d,e,** Reverse-transcription (RT)-qPCR of the indicated T-cell-related (**d**) or NK cell-related (**e**) genes relative to *Hprt*. Data are shown as mean ± s.d. and were analysed by two-tailed *t*-test (n = 3 replicates, representative of two independent experiments). **f,** Principal component analysis (PCA) clustering of pro-B and pre-B cells sorted from the bone marrow of *Wt* and *Sirt7*$^{-/-}$ mice, based on the top 2,000 differentially expressed genes **g,** Volcano plot of differentially expressed genes in pre-B cells versus pro-B cells from *Wt* (left panel) and *Sirt7*$^{-/-}$ (right panel) mice. Data are shown as in panel **a**. The black arrow in the right panel indicates the lack of downregulated genes. **h,** Unsupervised clustering of the top 2,000 differentially expressed genes (p<0.03). **i-l,** Gene ontology analyses of Clusters 1 (**i**), 2 (**j**), 3 (**k**) and 4 (**l**) genes. Only terms with -$\log_{10}$(p-value) > 5 are reported.

progenitors positively correlated with those of NK and T-cells, indicating inappropriate repression of alternative lineage genes and providing independent support to our observation that SIRT7 was required for early B-lineage commitment (**Fig. R10c**). To validate the derepression of NK cell and T-cell-related genes, we purified pro-B cells from *Wt* and *Sirt7*$^{-/-}$ mice, expanded them *ex vivo* and extracted their RNA to perform RT-qPCR. Intriguingly, all the T-cell-related genes that we analysed (the TCR segments *Trbc1* and *Trbc2* and genes encoding for T-cell signalling molecules *Zap70*, *Thy1*, *Cd3e*, *Tigit*) where strongly upregulated in SIRT7-deficient pro-B cells (**Fig. R10d**). However, none of the NK cell-related genes that we tested displayed major changes (**Fig. R10e**), suggesting that SIRT7 represses T-cell and NK-cell potential through different mechanisms.

In principal component analysis from these data (PCA), *Wt* and *Sirt7*$^{-/-}$ pro-B cells clustered very close, indicating that SIRT7 only had a slight effect on pro-B cell gene expression (**Fig R10f**). In contrast, *Wt* and *Sirt7*$^{-/-}$ pre-B cells clustered farther, consistent with a cumulative developmental defect. Notably, in principal component 1 (PC1), that explained 69.4% of all the variability among samples and was inferred to be the "cell type" component (because it separated pro-B and pre-B cells), *Sirt7*$^{-/-}$ pre-B cells were located closer to pro-B cells, suggesting inefficient transcriptional reprogramming during the pro-B-to-pre-B cell transition in the absence of SIRT7. To test this idea in more detail, we also compared the transcriptional changes that took place during the normal pro-B-to-pre-B cell transition and during the transition of *Sirt7*$^{-/-}$ progenitors. In *Wt* mice, 751 and 1,693 genes were significantly induced and repressed during the transition, respectively (**Fig. R10g**). Similarly, 685 genes were induced during the *Sirt7*$^{-/-}$ transition; but, in sharp contrast, only 954 genes were repressed, indicating that SIRT7 is mainly required for gene repression during the pro-B-to-pre-B cell transition and lending further support to its role in lineage commitment.

Next, we performed unsupervised clustering analysis to identify the gene clusters regulated by SIRT7 in B-cell progenitors and their associated molecular functions. We identified four different clusters: Cluster 1 encompassed genes that SIRT7 induced in both pro-B and pre-B cells, which were related to autophagy, cell signalling, haematopoiesis and NHEJ-mediated DNA repair, according to gene ontology analysis (**Fig R10h,i**). Cluster 2 genes were induced by SIRT7 especially in pre-B cells. It was enriched in genes related to autophagy, NK-κB signalling and Fc receptor signalling (**Fig. R10j**). Notably, Cluster 3 genes comprised genes that were specifically repressed by SIRT7 in pre-B cells, which included genes related to T-cell differentiation and TCR signalling, further supporting the

role of SIRT7 in the repression of T-cell potential (**Fig. R10k**). Finally, Cluster 4 accounted for nearly half of all the differentially expressed genes and included genes involved in cell cycle regulation, DNA damage response and ribosome biology. Importantly, these genes were strongly silenced during the pro-B-to-pre-B cell transition, but *Sirt7*$^{-/-}$ progenitors failed to repress them (**Fig. R10l**). Taken together, these data strongly suggest that SIRT7 participates in gene repression during early B-lymphopoiesis.

4.1. SIRT7 collaborates with PAX5 to mediate transcriptional regulation

To gain further insights into the mechanisms through which SIRT7 regulated early B-cell development, we next profiled the proteomic reprogramming triggered by SIRT7 deficiency in pre-B cells. To this end, we sorted pre-B cells from the bone marrow of *Wt* and *Sirt7*$^{-/-}$ mice and subjected these cells to low-input liquid chromatography coupled to quantitative mass spectrometry (LC-qMS/MS). First, we compared this dataset with the data obtained from our transcriptomic analysis. In total, proteomic analysis identified 3,272 unique proteins, whereas the RNA-Seq data mapped to 12,067 different protein-coding transcripts. Between both datasets, 3061 proteins and their relative transcripts were identified with both approaches (**Fig. R11a**). Comparing protein and transcript abundance by linear regression of our RNA-Seq and proteomics data showed that, in general, there was a highly significant correlation between RNA and protein expression levels, consistent with a previous report describing a similar correlation in B-ALL patient samples[346] (**Fig. R11b**).

Among all the proteins identified by proteomics, 56 were exclusively detected in *Wt* pre-B cells, and 46 in *Sirt7*$^{-/-}$ cells, in a reproducible manner. These included several interesting proteins, such as the DNA checkpoint kinase CHK2, the G_1/S-related cyclin D3 and two subunits of the RNA Pol II Mediator complex (MED21 and MED26), that were specifically detected in *Sirt7*$^{-/-}$ cells. Among the proteins detected only in *Wt* cells, we found the histone methyltransferase SETD1A, the PI3K subunit PIK3CD or the transcription factor E2F2. However, since we could not calculate a fold change for these proteins, they were excluded from downstream analyses. Notably, while the transcriptomic analysis of pre-B cells revealed differential expression in 1.6% of all the detected protein-coding genes, 5.2% of the proteome was significantly affected by SIRT7 loss, suggesting that SIRT7 not

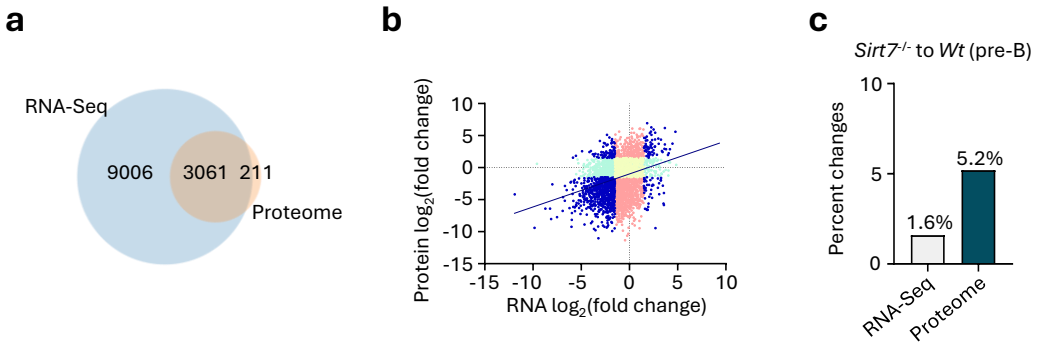


Fig. R11 | Comparative RNA-Seq / Proteome analysis
a, Venn diagram displaying the identification of proteins and their transcripts by quantitative proteomics and RNA-Seq, respectively. **b,** Scatter plot showing the linear correlation between the mean $\log_2$(fold-change) in protein levels and their corresponding transcript levels (Spearman correlation slope = 0.47, $p < 0.0001$). Protein and RNA levels with a $|\log_2(\text{fold-change})| \geq 1.5$ are shown in blue; changes ≥ 1.5 only in protein levels are shown in pale red; changes only in RNA levels are shown in pale green; and protein and RNA levels with changes ≤ 1.5 are indicated in yellow. Fold-changes were calculated relative to Nup50 expression. **c**, Percentages of significantly regulated ($|\log_2(\text{fold change})| \geq 1$, FDR < 0.1) transcripts and proteins in *Sirt7*$^{-/-}$ vs *Wt* pre-B cells, relative to all the transcripts and proteins detected by RNA-Seq and proteomics, respectively. FDR, false discovery rate; NHEJ, non-homologous end joining; diff., differentiation; BCR, B-cell receptor.

only regulates B-cell biology through transcriptional regulation but also through post-transcriptional or post-translational mechanisms (**Fig R11c**).

Among the top differentially regulated proteins, we found several transcriptional regulators, histone modifiers and chromatin remodellers, such as IRF2, the Polycomb histone methyltransferase EZH2, the SWI/SNF subunit SMARCD1 and the SET1/MLL complex member WDR48 (**Fig. R12a**). According to gene ontology analysis and similar to the ontologies regulated at the gene expression level, SIRT7 regulated proteins associated with DNA repair, cell cycle regulation and intracellular signalling (**Fig. 12b**). Importantly, one of the top downregulated proteins we found in *Sirt7*$^{-/-}$ pre-B cells was the B-cell master regulator PAX5, whose protein levels displayed a 2.8-fold reduction (**Fig. R12a**). Therefore, this finding prompted us to focus our attention on this protein. To do so, we first employed publicly available RNA-Seq data from *Pax5*$^{-/-}$ cells to compare the transcriptional programmes regulated by PAX5 and those regulated by SIRT7, using GSEA. This indicated that SIRT7 deficiency strongly correlated with the transcriptional signature caused by PAX5 deficiency (**Fig. R12c**). To further support this idea, we next performed *de novo* motif enrichment analysis with the DEGs regulated by SIRT7 to identify putative transcription factors that bind to the regulatory regions of these genes. Remarkably, we found that PAX5 was one of the top enriched factors in Cluster 3 genes, suggesting that SIRT7 and PAX5 may collaborate to regulate a common group of genes (**Fig. R12d**). We then exploited

publicly available PAX5 ChIP-Seq data[81] to get an experimental validation for this observation. We overlapped the Cluster 3 genes with the significant peaks occupied by PAX5 in pro-B cells, which confirmed that PAX5 bound to 62.1% of the Cluster 3 genes and, in many cases (43%), to the promoters of these genes (**Fig. R12e**). Finally, we aimed to directly test whether PAX5 was involved in SIRT7-mediated gene repression in B-cell progenitors. To do so, we purified *Wt* and *Sirt7*$^{-/-}$ bone marrow pro-B cells and transduced them with virus encoding either an EV or vectors expressing SIRT7 or PAX5, using the same construct as in Results section 1. After sorting the infected cells, we used these cells to analyse by RT-qPCR the expression of two of the T-cell-related genes that were derepressed in *Sirt7*$^{-/-}$ progenitors. As a result, we found that both SIRT7 and PAX5 were able to restore the repression of these genes, indicating that PAX5 bypassed SIRT7 deficiency to repress these lineage-inappropriate genes (**Fig. 12f**). Taken together, these data strongly suggest that SIRT7 and PAX5 collaborate to repress gene expression during early B-cell development.

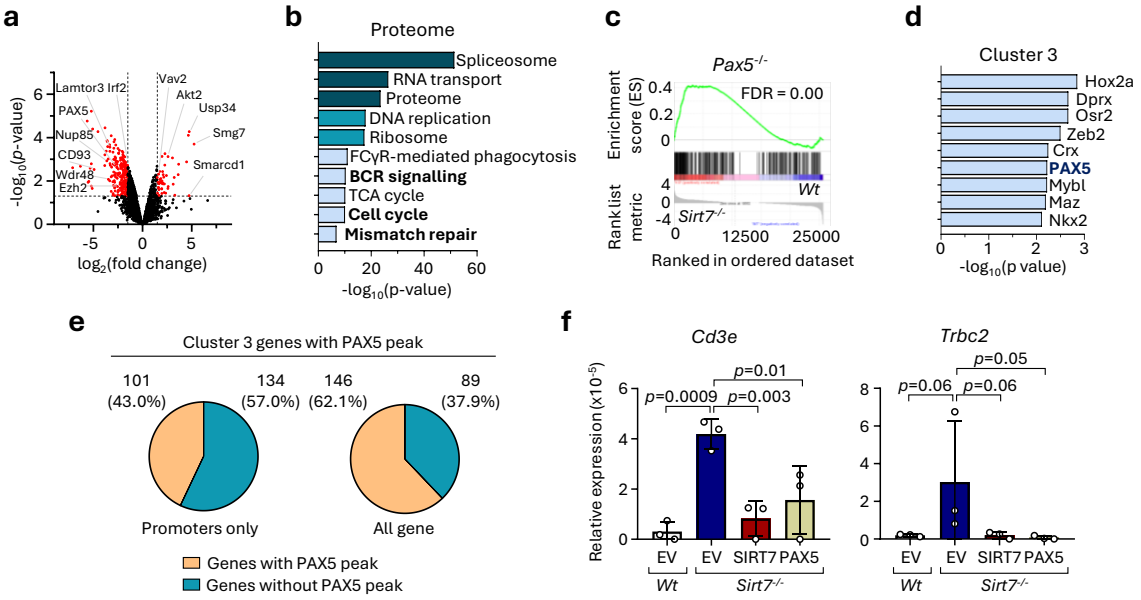


Fig. R12 | SIRT7 and PAX5 regulate a common set of genes

a, Volcano plot of differentially expressed proteins in *Sirt7*$^{-/-}$ versus *Wt* bone marrow sorted pre-B cells (red dots, proteins with changes ≥ |1.5| and *p*-value < 0.05). **b,** Gene ontology analysis of significantly regulated proteins in *Sirt7*$^{-/-}$ versus *Wt* pre-B cells. **c,** GSEA of *Sirt7*$^{-/-}$ versus *Wt* pre-B cells compared to PAX5 target genes (Ref [119]). **d,** HOMER *de novo* motif enrichment analysis showing the top transcription factors enriched in Cluster 3 genes. Motif enrichment was performed on promoter regions (TSS -2000 bp upstream). **e,** The percentage of Cluster 3 genes with significant (q < 0.05) PAX5 peaks[81] in their promoters (± 2Kb from TSS) (left panel) or gene bodies (including ± 2Kb from TSS, gene bodies and 3'UTR, right panel), is shown. **f,** RT-qPCR of T-cell *Cd3e* and *Trcb2* genes relative to *Hprt* in *ex vivo*-expanded *Wt* or *Sirt7*$^{-/-}$ pro-B cells upon retroviral transduction with empty vector (EV), SIRT7 or PAX5. Data are shown as mean ± s.d. and were analysed by one-way ANOVA with Fisher's LSD test (n = 3 replicates, representative of two separate experiments).

4.2. Epigenetic functions of SIRT7 and PAX5 in pro-B cells

As described above, SIRT7 plays an important role in epigenetic regulation through H3K18ac and H3K36ac deacetylation. Both histone marks are mainly localised in gene regulatory regions and are associated with transcriptional activation. In the case of H3K36ac, it has been reported to be predominantly enriched in the promoters of RNA Pol II-dependent actively transcribed genes in *S. cerevisiae*[278]. The genomic distribution of H3K36ac is often inversely related to that of H3K36me3, another important histone mark that is enriched in the gene bodies of actively transcribed genes and that inhibits inappropriate transcriptional initiation from cryptic start sites[364]. Furthermore, H3K36me3 sustains genome integrity, and DNA damage repair, and its hypomethylation is common in cancer[365]. Indeed, oncogenic mutations affecting this residue (mainly, K36M substitutions), as well as inactivating mutations of its methyltransferase SETD2, have been proposed to be drivers of several paediatric cancers[366-369]. Interestingly, studies in budding yeast have suggested that the presence of H3K36ac and H3K36me3 inhibits each other's deposition and that their regulation mediates a chromatin switch that is dynamically regulated along the cell cycle and that may be important for DNA repair[279]. Although PAX5 has not been related to H3K36ac, it is a well-known driver of epigenetic modifications in pro-B cells, given its ability to interact with many histone modifiers and to recruit them to its target loci to regulate gene expression[118]. Therefore, we reasoned that PAX5 may recruit SIRT7 to regulate H3K18ac and H3K36ac. Thus, we next decided to explore the impact of SIRT7 in pro-B cell epigenetic regulation and how it relates with PAX5. To this end, we first analysed H3K18ac and H3K36ac abundance in *ex vivo* expanded *Wt* and *Sirt7*$^{-/-}$ pro-B cells by immunoblotting. As we have previously observed in other cellular models, we found no differences in the overall levels of H3K18ac, suggesting that SIRT7-mediated H3K18ac deacetylation is site-specific rather than global (**Fig. R13a**).

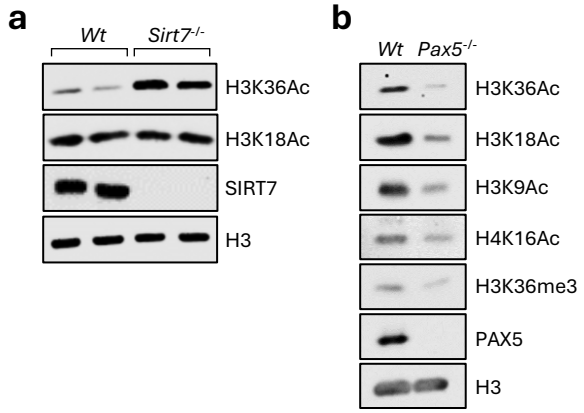


Fig. R13 | Effects of SIRT7 and PAX5 loss in active histone marks

a,b, Immunoblotting of the indicated histone marks in *Wt* and *Sirt7*$^{-/-}$ (**a**) or *Pax5*$^{-/-}$ (**b**) *ex vivo* expanded pro-B cells. Representative of 2 independent experiments.

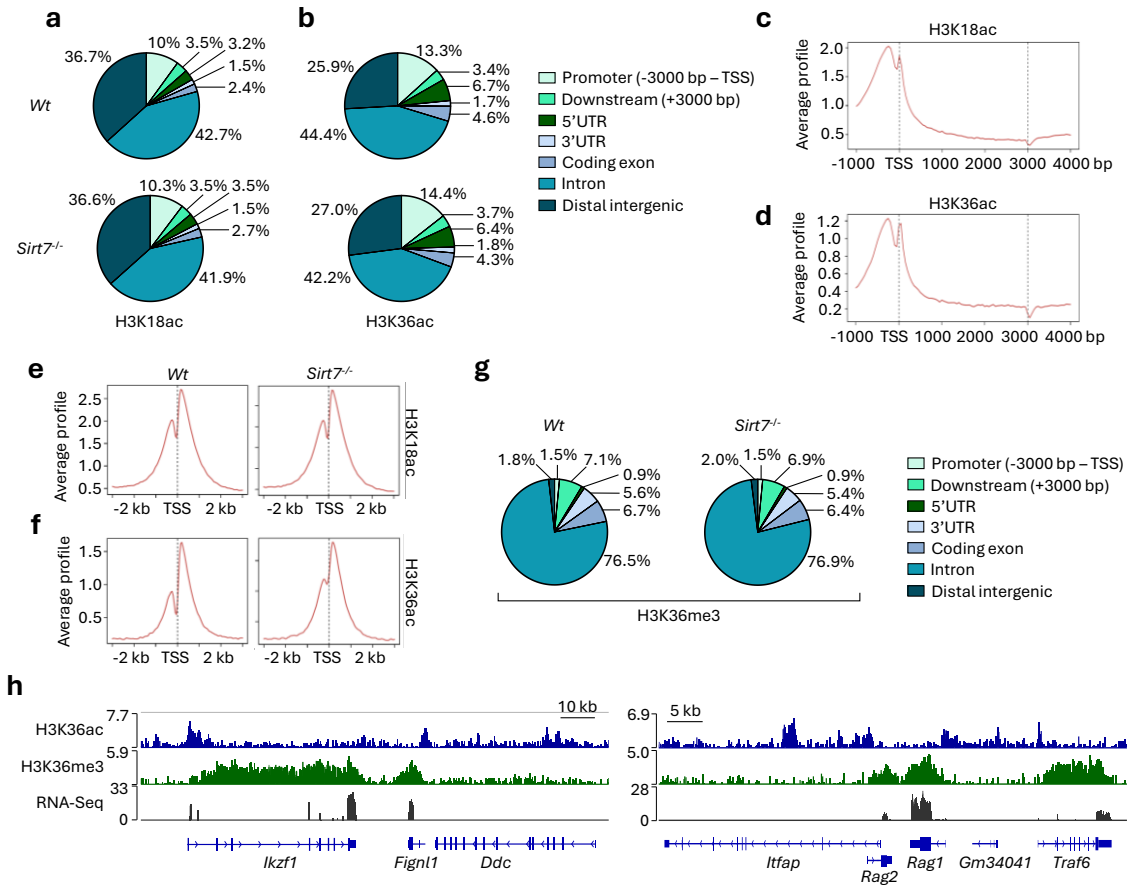


Fig. R14 | SIRT7 preferentially deacetylates H3K36ac in developing B-cells

a,b, Pie charts of ChIP-Seq data displaying the percent enrichment of H3K18ac (**a**) and H3K36ac (**b**) in the indicated genomic regions in *Wt* and *Sirt7*$^{-/-}$ proB cells. **c,d,** Average distribution at genic regions of H3K18ac (**c**) and H3K36ac (**d**) in *Wt* pro-B cells. **e,f** Average profile of H3K18ac (**e**) and H3K36ac (**f**) at the TSS in *Wt* and *Sirt7*$^{-/-}$ pro-B cells. **g,** Pie chart displaying the percent enrichment of H3K36me3 as in **a**. **h,** H3K36ac and H3K36me3 occupancy in the indicated regions. RNA-Seq tracks from *Wt* pro-B cells are shown in the lower part (n=2).

In sharp contrast, H3K36ac global levels were massively increased in *Sirt7*$^{-/-}$ cells, strongly suggesting that H3K36ac is the major histone mark targeted by SIRT7 (**Fig. R13a**). Next, we also evaluated the impact of PAX5 in the global levels of these histone marks. To do so, we obtained pro-B cells from the foetal liver (FL) of *Wt* and *Pax5*$^{-/-}$ mice (in collaboration with Dr. Mikael Sigvardsson, Lund University) and expanded them *ex vivo*. Unexpectedly, the levels of both histone marks were dramatically reduced in *Pax5*$^{-/-}$ pro-B cells (**Fig. R13b**). Although this observation does not discard the possibility that PAX5 recruits SIRT7 to facilitate H3K18ac and H3K36 deacetylation, it strongly suggests that it critically contributes to the recruitment of the responsible HATs to their targets. Indeed, we also analysed the global levels of H3K36me3, H3K9ac and H4K16ac and observed the same tendency in all cases, although to a lesser extent, indicating an impressively broad function of PAX5 in the establishment of chromatin accessibility. Further studies should

address how this affects chromatin regulation at the single-locus level and whether these observations can be extrapolated to repressive marks. In this sense, PAX5 is known to participate in the removal of active histone marks but also in Polycomb-mediated deposition of the repressive H3K27me3 mark, suggesting that the establishment of facultative heterochromatin in pro-B cells may also rely on PAX5[118].

Next, we profiled the genome-wide distribution of H3K18ac, H3K36ac and H3K36me3 through ChIP-Seq of these marks in *ex vivo* expanded *Wt* and *Sirt7*$^{-/-}$ pro-B cells. As previously described, H3K18ac was mainly enriched in gene promoters (10.0%), introns (42.7%) and distal intergenic regions (36.7) (**Fig. R14a**). H3K36 was similarly distributed, although it was less frequent in distal intergenic regions and was slightly more enriched in promoters (**Fig. R14b**). No major differences were observed between *Wt* and *Sirt7*$^{-/-}$ cells regarding the genomic distribution of these marks (**Fig. R14a,b**). In genic regions, H3K18ac and H3K36ac were highly abundant in proximal promoters and near the TSS, and their abundance was progressively reduced in the gene body (**Fig. R14c**). Interestingly, H3K36ac was more skewed towards the TSS than H3K18ac, and this difference in H3K36ac was noticeably less prominent in *Sirt7*$^{-/-}$, whereas H3K18ac did not display meaningful changes in this regard (**Fig. R14d-f**). In the case of H3K36me3, it was strongly enriched in introns (76.5%), and also present in exons (6.7%) and 3' UTRs (5.6%) (**Fig. R14g**). As exemplified in **Fig. 14h,** H3K36me3 presence across the gene bodies was strongly associated with actively transcribed genes such as *Ikzf1*, *Rag1* and *Rag2*.

We then explored how SIRT7 deficiency affected the global abundance of these marks. In gene-proximal regions, *Sirt7*$^{-/-}$ progenitors displayed a modest tendency to increased levels of H3K18ac levels (**Fig. R15a**). Consistently, similar results were obtained when we analysed H3K18ac enrichment at regions overlapping with H3K4me3 peaks, which are present in the promoters of more than 80% of the genes being actively transcribed[370]. In putative enhancers marked with H3K4me1 or H3K27ac, H3K18ac was in general less abundant than in promoters, and its abundance remained unaffected upon SIRT7 loss (**Fig. R15b**). This confirmed the data obtained by immunoblotting and strongly suggested that, at least in pro-B cells, SIRT7 preferentially deacetylates H3K18ac in promoters, in a locus-specific and probably context-dependent manner. In the case of H3K36ac, our results were also consistent with those obtained by western blot. In genic regions and H3K4me3^{+} putative active promoters, SIRT7 deficiency led to a remarkable increase in H3K36ac abundance (**Fig. R15c**). Interestingly, H3K36ac occupancy was also increased in H3K4me1^{+} and, more prominently, in H3K27^{+} putative enhancers (**Fig. R15d**). Although the

effect of SIRT7 loss in putative enhancers was clearly milder compared to promoters, these data indicates that SIRT7 globally deacetylated H3K36ac across genomic regions and suggest a role of SIRT7 in the regulation of distal regulatory elements that has not been explored so far. Finally, we found a slight tendency towards increased H3K36me3 levels in the gene bodies in *Sirt7*$^{-/-}$ cells, but sample variability led us to conclude that the effect of SIRT7 in H3K36me3 levels was minor (**Fig. R15e**). Thus, these data indicate that H3K36ac is the major histone mark targeted by SIRT7 in pro-B cells.

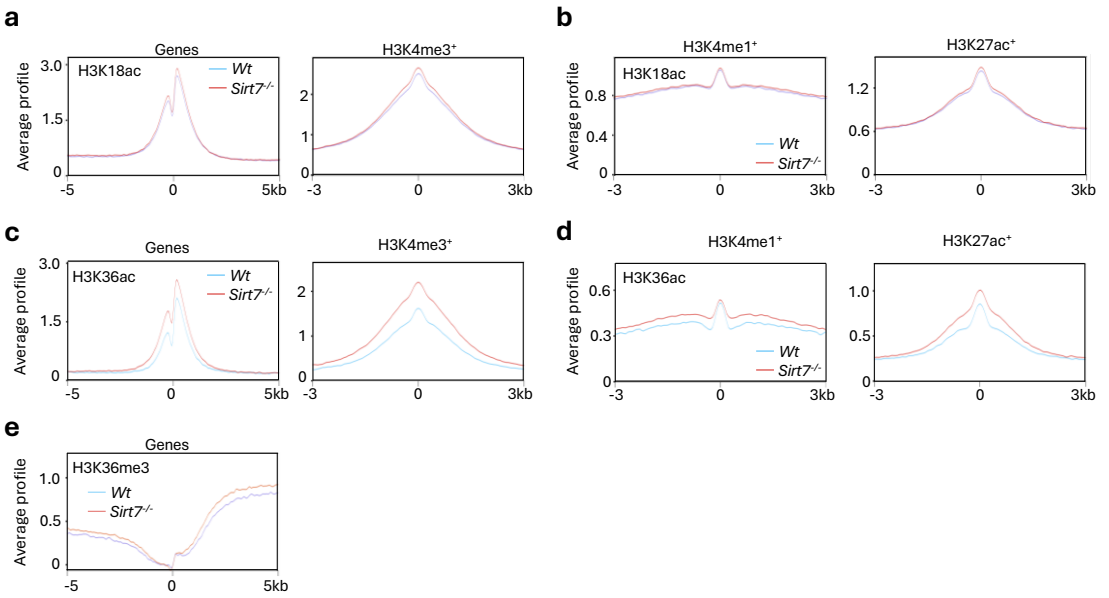


Fig. R15 | Effect of SIRT7 loss in H3K18ac and H3K36ac abundance in promoters and enhancers
a,b, Average profiles of H3K18ac at genes, H3K4me3^{+} (**a**), H3K4me1^{+} and H3K27ac^{+} regions (**b**) in *Wt* and *Sirt7*$^{-/-}$ pro-B cells. **c,d**, Average profiles of H3K36ac, as in **a,b**. **e,** Average profile of H3K36me3 at genes in *Wt* and *Sirt7*$^{-/-}$ pro-B cells (n=2).

Based on these observations, we decided to focus on H3K36ac to explore the potential interplay between SIRT7 and PAX5 in epigenetic regulation. To this end, we exploited the same previously published PAX5 ChIP-Seq data as above and interrogated the co-occupancy of PAX5 and H3K36ac. In promoters and gene bodies, PAX5 was strikingly enriched at H3K36 peaks, with more than half of PAX5 binding sites overlapping with H3K36ac peaks (**Fig. R16a,b**). Notably, this was even more evident upon SIRT7 loss, with up to 70% of PAX5 peaks in these regions coinciding with H3K36ac. Enhanced PAX5 occupancy at H3K36ac regions was also observed to some extent in intergenic regions, although their overlap was much less pronounced in these regions (**Fig. R16a**). Together with our observation that PAX5 loss led to a dramatic depletion of H3K36ac levels, these data clearly indicate that PAX5 is somehow involved in H3K36ac regulation, especially in the context of genes. Therefore, we next interrogated the impact of SIRT7 loss on H3K36ac

abundance at PAX5 binding sites (**Fig. R16c**). As expected, *Sirt7*$^{-/-}$ cells exhibited a remarkable increase of H3K36ac levels in regions occupied by PAX5, and this effect was seemingly stronger than the global effect observed in **Fig. R15c**. Similarly, SIRT7 loss also led to a more remarkable increase in H3K36ac levels in PAX5-occupied intergenic regions.

Of note, locus-specific analysis revealed that SIRT7 deficiency not only led to enhanced intensity of preexisting peaks, but also to the appearance of new peaks at PAX5 binding sites (**Fig. R16d**). Taken together, these data suggest that PAX5 recruits SIRT7 to its target loci to fine-tune H3K36ac deacetylation.

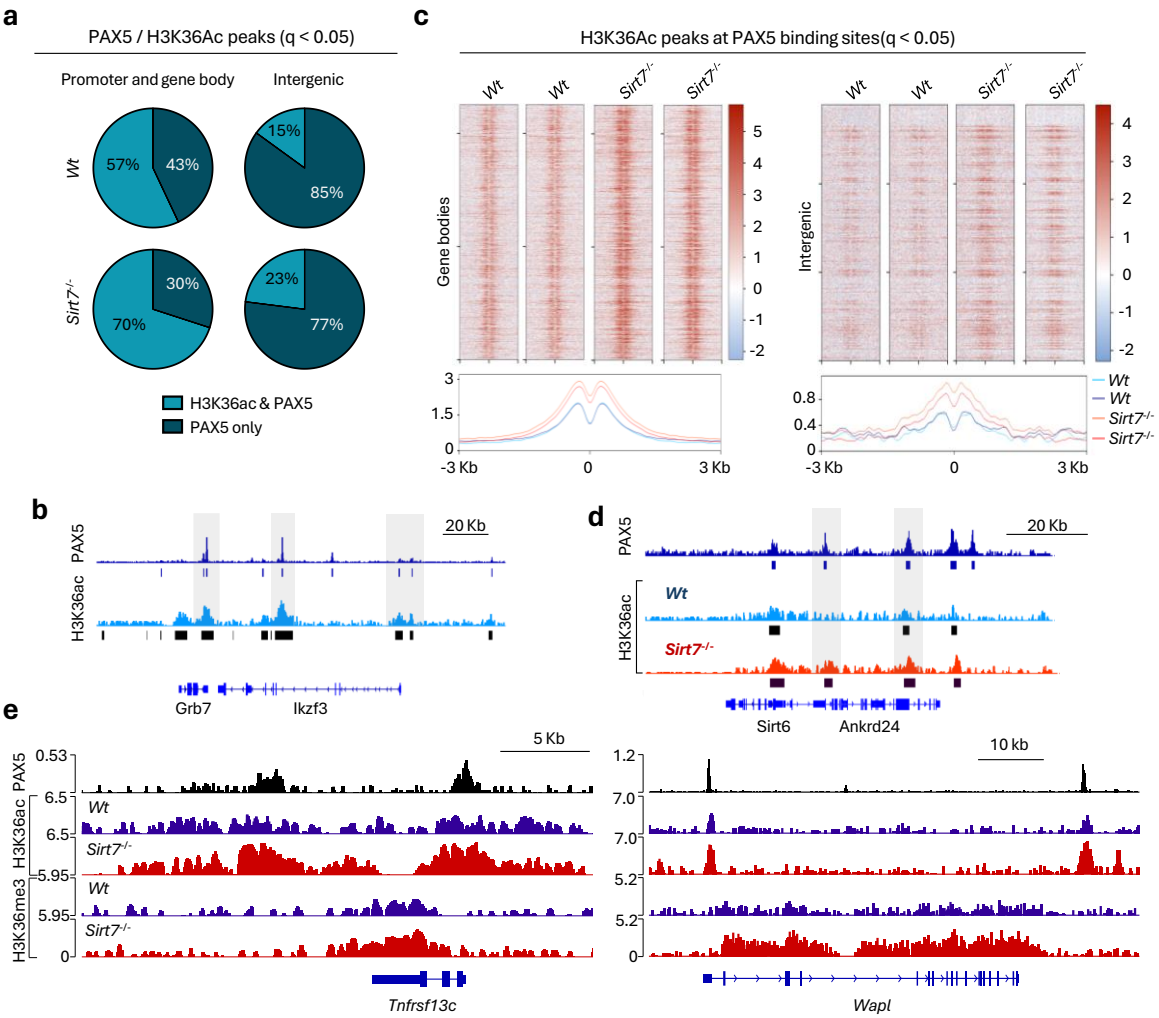


Fig. R16 | SIRT7 regulates H3K36ac at PAX5 binding sites

a, Pie charts depicting the colocalization between PAX5 and H3K36ac in *Wt* and *Sirt7*$^{-/-}$ pro-B cells. **b,** PAX5 and H3K36ac enrichment at the indicated region in *Wt* pro-B cells. **c,** Average profile of H3K36ac in *Wt* and *Sirt7*$^{-/-}$ pro-B cells at PAX5 binding sites in genes (left panels) and intergenic regions (right panels). **d,** PAX5 and H3K36ac enrichment at the indicated region *Wt* and *Sirt7*$^{-/-}$ pro-B cells. **e,** PAX5, H3K36ac and H3K36me3 enrichment at the indicated region *Wt* and *Sirt7*$^{-/-}$ pro-B cells. PAX5 ChIP-Seq data in **d,e** was retrieved from Ref [81].

Finally, we preliminarily assessed the potential effects of PAX5-mediated H3K36ac regulation on the transcriptional control of PAX5 target genes. We chose as models two PAX5 target genes that it directly represses in a stage-specific manner: *Wapl*, which PAX5 represses specifically in pro-B cells to facilitate *Igh* recombination, and *Tnfrsf13c*, which encodes for BAFFR (B-cell activating factor of the TNF family (BAFF) receptor), a receptor essential for late B-cell differentiation whose expression is induced in transitional B-cells[371]. As depicted in **Fig. R16e**, PAX5 binds to the promoters of both genes, and to a downstream region in the case of *Wapl*. In SIRT7-deficient cells, PAX5 binding correlated with increased H3K36ac at both genes. Interestingly, SIRT7 loss also led to increased levels of H3K36me3 in the bodies of both genes, probably reflecting enhanced expression of these genes, suggesting that SIRT7-mediated H3K36ac deacetylation may participate in the repression of PAX5 target genes (**Fig. R16e**). However, further research will be required to understand the connexion between PAX5, SIRT7 and H3K36 regulation and its impact on epigenetic regulation.

5. SIRT7 regulates PAX5 protein stability through deacetylation at K198

The finding that SIRT7 and PAX5 regulate a common transcriptional program, together with the 2.8-fold reduction in PAX5 protein levels displayed by *Sirt7*$^{-/-}$ pre-B cells and with our preliminary data pointing into the same direction, led us to hypothesise that SIRT7 promotes B-cell development through direct regulation of PAX5 functions. Several lines of evidence supported this hypothesis: first, we performed intracellular FACS to measure the protein levels of PAX5 in pro-B and pre-B cells from bone marrow samples. Consistent with our previous data, *Sirt7*$^{-/-}$ pro-B and pre-B cells exhibited significantly reduced PAX5 levels (**Fig. R17a,b**). This was further confirmed by immunoblotting of *ex vivo* expanded *Wt* and *Sirt7*$^{-/-}$ pro-B cells, which showed a similar reduction upon SIRT7 loss (**Fig. R17c**). Notably, this effect was fully rescued by retroviral expression of SIRT7 in *Sirt7*$^{-/-}$ cells, indicating that SIRT7 reversibly regulates PAX5 protein levels (**Fig. R17d,e**). To test whether the reduction of PAX5 protein levels was due to decreased *Pax5* gene expression, we performed RT-qPCR of *ex vivo* expanded pro-B cells, which revealed that *Sirt7*$^{-/-}$ cells expressed largely normal *Pax5* RNA levels (**Fig. 17f**). Similarly, when we analysed *Pax5* expression in our RNA-Seq data we found no differences in either pro-B or pre-B cells (**Fig. 17g)**.

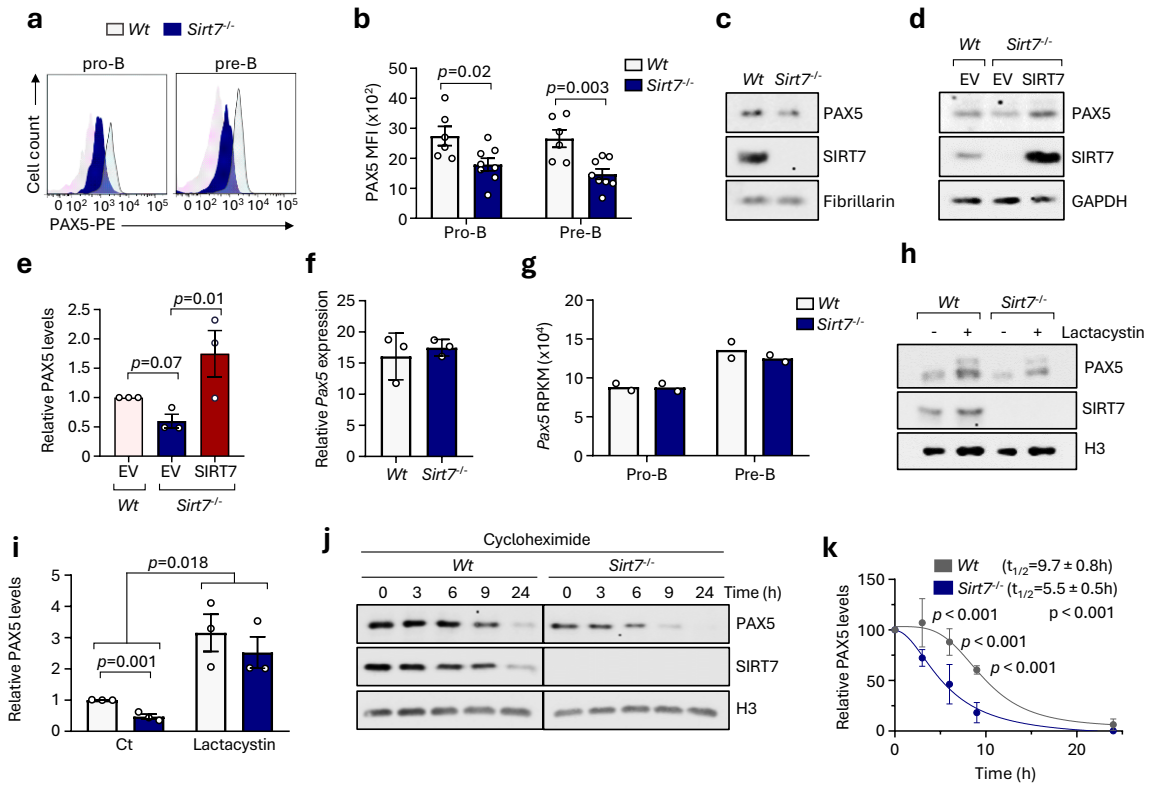


Fig. R17 | SIRT7 promotes PAX5 stability

a, Representative plots (**a**) and quantification of PAX5 MFI (**b**) measured by intracellular flow cytometry of gated bone marrow pro-B and pre-B cells. They filled grey profile in panel **a** indicates the isotype control. Data in **b** are presented as mean ± s.d. and *p*-values were calculated by multiple *t*-tests with Holm-Sidak correction (*Wt*, n = 6; *Sirt7*$^{-/-}$, n = 8). **c,** Immunoblot of PAX5 protein levels in *ex vivo*-expanded *Wt* and *Sirt7*$^{-/-}$ pro-B cells (n = 2). **d,e,** Immunoblots showing PAX5 and SIRT7 levels (**d**) and PAX5 quantification (**e**) in *ex vivo* expanded *Wt* and *Sirt7*$^{-/-}$ pro-B cells expressing the indicated retroviruses. Representative of three independent experiments. Data in **e** are shown as mean ± s.d. and were analysed by one-way ANOVA with Fisher's LSD test. **f,** RT-qPCR analysis of *Pax5* gene expression in cultured *Wt* and *Sirt7*$^{-/-}$ pro-B cells (n = 3). Data are shown as mean ± s.d. **g,** RNA-Seq analysis of *Pax5* expression in sorted pro-B and pre-B cells. Data are presented as mean (n = 2). **h,i,** Immunoblots (**h**) and densitometric quantification (**i**) of *Wt* and CRISPR-Cas9-generated *Sirt7*$^{-/-}$ HAFTL pre-B cells treated with vehicle or 2 µM lactacystin for 8 hours. In **i**, the PAX5-to-GAPDH ratio is shown as mean ± s.d., and statistical significance was calculated by one-tailed t-tests (n = 3). **j,k,** Timecourse immunoblot (**j**) and PAX5 / H3 ratio (**k**) in *Wt* and *Sirt7*$^{-/-}$ HAFTL cells treated with 100 µg/mL cycloheximide for the indicated times. PAX5-to-H3 ratios are shown as mean ± s.d, and non-linear fits with variable slope (four parameters) are depicted. The *p*-values for individual time-points were calculated by two-way ANOVA with Sidak comparison, and curve half-lives ± s.d were compared by two-tailed *t*-test (*Wt*, n = 6; *Sirt7*$^{-/-}$, n = 8).

The above data strongly suggested that SIRT7 directly regulates PAX5 turnover through a post-translational mechanism. To test this hypothesis in detail, we used the CRISPR-Cas9 system to knock-out SIRT7 in HAFTL, a Ha-Ras-transformed mouse pre-B cell line that expresses most components of the B-lymphoid transcriptional machinery and shows immunophenotypic features of pre-B cells. We transduced HAFTL cells with lentiviruses expressing the Cas9 nuclease under the control of a doxycycline-inducible promoter, selected the transduced cells with puromycin, and subsequently transduced the cells

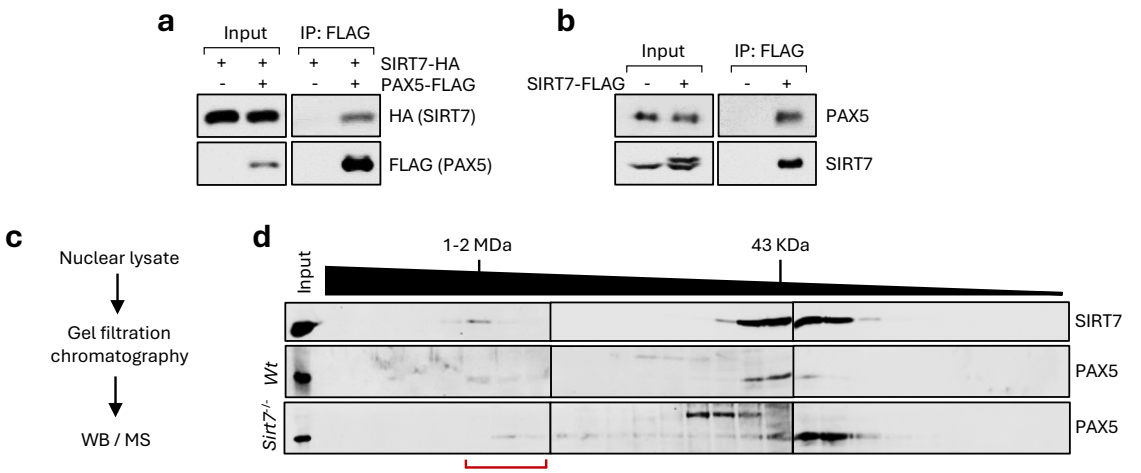


Fig. R18 | SIRT7 and PAX5 form a molecular complex

a, Anti-FLAG immunoprecipitation from HEK293F cells transiently co-expressing SIRT7-HA with EV or PAX5-Myc-FLAG. Representative of three separate experiments. **b,** Anti-FLAG co-immunoprecipitation with endogenous PAX5 in HAFTL cells transduced with empty vector (EV) or SIRT7-FLAG, followed by anti-PAX5, anti-FLAG and anti-SIRT7 immunoblotting. In the input SIRT7 panel, the lower and upper bands correspond to endogenous and FLAG-tagged SIRT7, respectively. Representative of three independent experiments. **c,d,** Schematic representation (**c**) and immunoblots (**d**) of the input and gel filtration chromatography fractions in which proteins and protein complexes from HAFTL nuclear lysates are separated as a function of their molecular weights. In the upper part, the approximate molecular weights of the indicated fractions are shown. The red brackets highlight the high-molecular-weight fractions in which PAX5 and SIRT7 co-elute.

again with lentiviral vectors containing a GFP marker and either sgRNAs or sgRNAs targeting the *Sirt7* gene. Finally, we selected the positive cells by sorting them. Although this system was designed to allow the conditional deletion of target genes, its efficiency was very low in our cells, so we decided to perform a constitutive knock-out by treating the cells with 1 µg/mL doxycycline to induce Cas9 expression, plating single clones in 96-well plates and individually testing the expression of SIRT7 in the clonally expanded cell lines by immunoblotting. Thus, we generated three HAFTL-derived clonal cell lines harbouring homozygous SIRT7 deletions (hereafter, *Sirt7*$^{-/-}$ HAFTL cells) and maintained two control clonal lines carrying doxycycline-inducible *Cas9* and unspecific sgRNAs (*Wt* HAFTL cells). We then performed biochemical studies in these cells to determine the impact of SIRT7 in PAX5 turnover. First, we evaluated the effect of SIRT7 deficiency in the proteasomal degradation of PAX5. To this end, we treated *Wt* and *Sirt7*$^{-/-}$ HAFTL cells with the irreversible proteasome inhibitor lactacystin (10 µM) for 8h and analysed the protein levels of PAX5 by immunoblotting. Importantly, SIRT7 loss led to a reduction in the levels of PAX5 also in HAFTL cells, and this effect was fully reversed by proteasome inhibition, suggesting that SIRT7 prevents PAX5 proteasomal degradation (**Fig. R17h,i**). To further support this observation, we next performed protein turnover experiments using cycloheximide, an

antifungal compound that inhibits overall protein synthesis in eukaryotes. Thus, we performed cycloheximide timecourse experiments by treating HAFTL cells with 100 µg/mL cycloheximide for 3, 6, 9 and 24h and assessed PAX5 turnover by immunoblotting. Remarkably, SIRT7 deficiency caused a strong reduction in PAX5 half-life (9.7h in *Wt* cells versus 5.5h in *Sirt7*$^{-/-}$ cells), which confirmed that SIRT7 enhances PAX5 stability by preventing its proteasomal degradation (**Fig. R17j,k**).

5.1. SIRT7 forms a molecular complex with PAX5 and other B-lymphoid transcription factors

To test whether SIRT7 directly regulates PAX5, we next examined the interactions between SIRT7 and PAX5 using co-immunoprecipitation (co-IP) and gel filtration chromatography. For the co-IP, we transiently expressed HA-tagged SIRT7 together with vectors encoding either an EV or FLAG-tagged PAX5 in HEK293F cells. Anti-FLAG immunoprecipitation followed by immunoblotting confirmed that SIRT7 and PAX5 formed a protein complex (**Fig. R18a**). Lending further support to this observation, we repeated this experiment in HAFTL cells stably expressing endogenous levels of FLAG-tagged SIRT7 and confirmed that SIRT7 interacts with endogenous PAX5, strongly suggesting that SIRT7 stabilises PAX5 through a mechanism involving direct interaction (**Fig. R18b**). Since PAX5 is known to form part of a complex transcription factor network controlling B-cell development, we hypothesised that SIRT7 may collaborate with this network, which would imply that the SIRT7/PAX5 interaction takes place in the context of a high-molecular-weight protein complex. We therefore explored this possibility by performing gel filtration chromatography of HAFTL nuclear extracts, which allows to separate proteins and protein complexes according to their molecular weight. Indeed, similar experiments have previously shown that a major fraction of Ikaros, a key component of this transcription factor network, exists in a 1-2 MDa protein complex where it interacts with, and controls the functions of, the chromatin remodelling complex NuRD[248,249]. In the same line, we found that a substantial fraction of PAX5 and SIRT7 molecules coeluted in the earliest fractions, that contained 1-2 MDa protein complexes (**Fig. R18c,d**). PAX5 fractionation was slightly delayed and showed a more dispersed pattern with an increased presence in medium-weight fractions upon SIRT7 loss, suggesting that SIRT7 may participate in the stabilisation of PAX5-containing high-molecular-weight complexes (**Fig. R18d**).

Table R1 | Putative SIRT7 interactors in high-molecular-weight protein complexes
The number of significant peptides of each protein detected by mass-spectrometry is shown. The blue squares indicate the complexes to which each of the proteins belong.

Gene names	Peptides in *Wt*	Peptides in *Sirt7*$^{-/-}$	NuRD	BAF-Type	INO80	TFIID	MLL1	SET1C
Mbd3	4	1	■					
Smarcd2	40	26	■	■				
Hdac2	5	2	■					
Hdac1	6	3	■					
Smarcc2	51	32	■	■				
Rbbp4	6	4	■					
Actl6a	21	12	■	■	■			
Chd4	24	11	■					
Mta2	10	4	■					
Gatad2b	7	2	■					
Pbrm1	44	26		■				
Smarcc2	51	32		■				
Actl6a	21	12		■				
Dpf2	21	13		■				
Brd7	16	10		■				
Smarca2	40	21		■				
Arid1b	68	38		■				
Ino80c	1	0			■		■	
Mcrs1	1	0			■		■	
Actr8	2	0			■			
Nfrkb	4	1			■			
Uchl5	1	0			■			
Taf2	3	0				■		
Taf3	1	0				■		
Taf4b	1	0				■		
Taf5	2	1				■		
Taf6	2	1				■	■	
Taf8	1	0				■		
Taf9	2	1				■	■	
Wdr5	2	0					■	■
Hcfc1	2	1						■
Dpy30	2	0						■

These findings motivated us to preliminarily explore the components of this SIRT7- and PAX5-containing 1-2 MDa complex. To do so, we repeated the chromatography with HAFTL nuclear extracts, pooled the earliest fractions containing large complexes, immunoprecipitated SIRT7 from these fractions and subjected the eluates of the IP to mass spectrometry analysis, using nuclear extracts from *Sirt7*$^{-/-}$ HAFTL cells as a negative control. Unfortunately, we could not identify any of the core B-lymphoid transcription factors using this approach. However, we found many other proteins that interacted with SIRT7 in the context of these large complexes, including several subunits of the NuRD complex (HDAC1/2, SMARCC2/D2, CHX4, MTA2 and others); proteins belonging to other chromatin remodellers (such as the INO80 complex subunits INO80C, ACTL6A and

ACTR8; the MLL complex subunit WDR5 and others); and subunits of several transcriptional regulator complexes (such as the spliceosome complex and the transcription factor TFIID), some of which are summarised in **Table R1**. Gene ontology analysis revealed significant enrichment of different complexes, but NuRD was by far the top enriched one (**Fig.19a**). Although our mass spectrometry analyses did not identify any B-cell-related factors, the observation that SIRT7 may interact with the NuRD complex, of which Ikaros is a well-known member, encouraged us to explore whether SIRT7 forms part of a broader B-lymphoid transcription factor network. For that, we transiently expressed HA-tagged SIRT7 and FLAG-tagged Ikaros or EBF1 in HEK293F cells and used the extracts to perform reciprocal anti-FLAG and anti-HA immunoprecipitations, which confirmed that SIRT7 interacted with both proteins (**Fig. R19b,c**). Furthermore, SIRT7 also co-immunoprecipitated with endogenous EBF1 in HAFTL cells, although we could not confirm its endogenous interaction with Ikaros (**Fig. R19d**). We conclude that SIRT7 interacts with PAX5 and participates in the B-lymphoid transcription factor network that controls B-cell development.

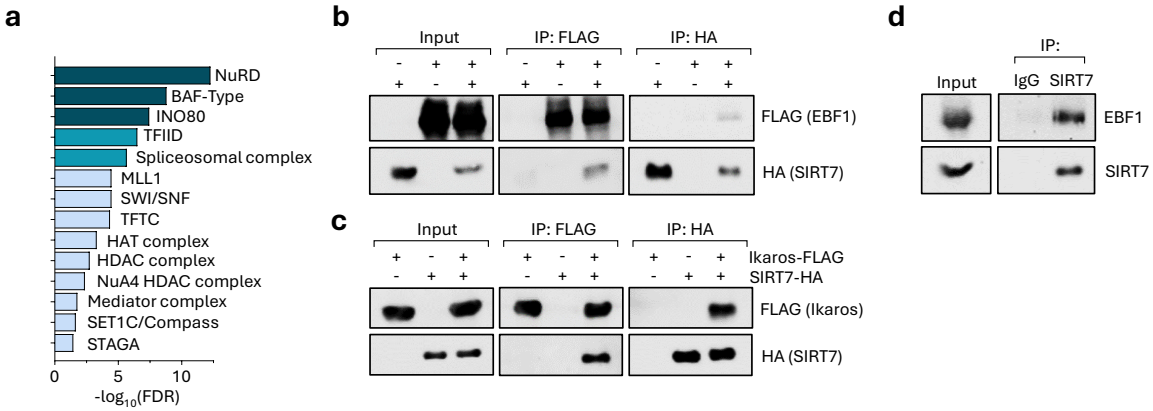


Fig. R19 | SIRT7 interacts with several high-molecular-weight protein complexes and forms part of a B-lymphoid transcription factor network

a, Gene ontology analysis depicting the most enriched protein complexes among the significant SIRT7 interactors identified by mass spectrometry. **b,c,** Anti-FLAG immunoprecipitation from HEK293F cells transiently co-expressing SIRT7-HA with EV and EBF1-FLAG (**b**) or Ikaros-FLAG (**c**). Representative of three separate experiments. **d,** Anti-SIRT7 co-immunoprecipitation in HAFTL cell lysates followed by anti-EBF1 and anti-SIRT7 immunoblotting. Representative of 3 independent experiments.

5.2. SIRT7 controls PAX5 stability by deacetylating PAX5^{K198}

Acetylation of several PAX5 residues by p300 has been previously reported and linked to its transcriptional activity[372]. However, whether PAX5 acetylation regulates its stability is not known. Furthermore, we previously found in Section 1 of Results that the deacetylase activity of SIRT7 is indispensable for B-cell development, suggesting that SIRT7 may control B-cell development through PAX5 deacetylation. Consistent with this idea, treatment with the pan-sirtuin inhibitor NAM induced PAX5 degradation in both HAFTL cells and KOPN8 human B-ALL cells (**Fig. R20a,b**). Conversely, NAM did not affect *Pax5* gene expression (**Fig. R20c**). To directly test whether SIRT7 deacetylates PAX5, we next expressed FLAG-tagged PAX5 and SIRT7 in separated HEK293F plates and purified both proteins through anti-FLAG immunoprecipitation followed by elution in native conditions. We then used both proteins to perform *in vitro* deacetylation assays and monitored PAX5 acetylation using a pan-anti-acetyl-lysine (AcK) antibody. Importantly, PAX5 global acetylation was partially reduced when we incubated PAX5 and SIRT7 in the presence of

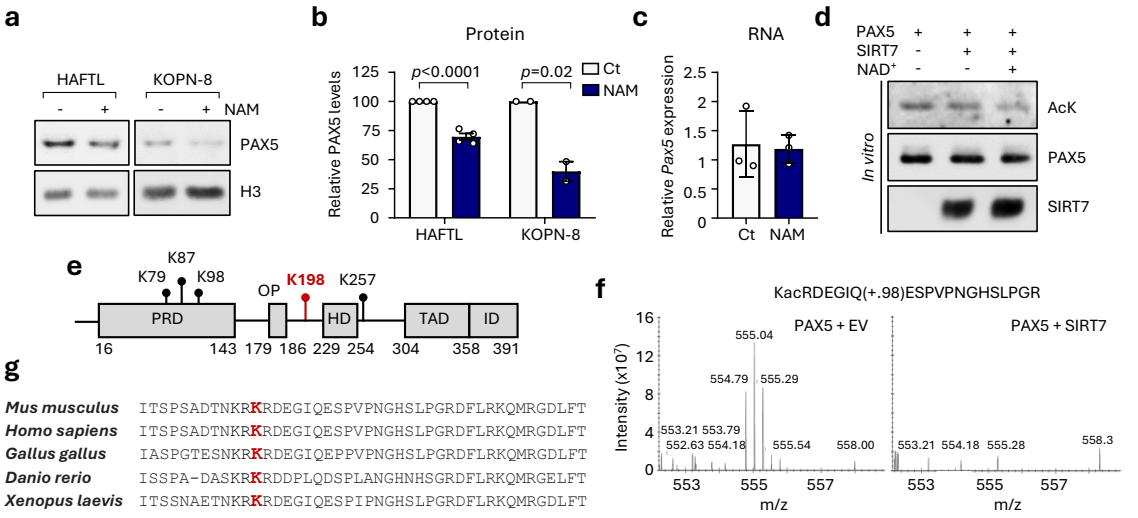


Fig. R20 | SIRT7 deacetylates PAX5 at K198

a, Immunoblots (**a**) and quantification (**b**) of PAX5 protein levels in HAFTL mouse pre-B cells and KOPN-8 human B-ALL cells treated with vehicle or 5 mM NAM for 48 hours (HAFTL, n = 4; KOPN-8, n = 2). Data in **b** are presented as mean ± s.d. and were analysed by multiple *t*-tests (HAFTL, n = 4; KOPN-8, n = 2). **c,** RT-qPCR analysis of *Pax5* expression relative to *Hprt* in HAFTL cells (n = 3 replicates). Data are presented as mean ± s.d. **d,** *In vitro* PAX5 deacetylation assay. PAX5 acetylation was measured by anti-pan-acetyl lysine (AcK) immunoblotting. **e,** Diagram of mouse PAX5 functional domains consisting of the paired-box (PRD), octapeptide, partial homeodomain (HD), transactivation (TAD) and inhibitory (ID) domains. The numbers in the lower part indicate the start and end of each domain, and the detected acetyl-lysine (K) residues are indicated in the upper part. SIRT7-targeted K198 is shown in red. **f,** Fragmentation MS/MS spectra of the indicated peptide, as determined by proteomic analysis of PAX5 immunoprecipitated from *SIRT7*$^{-/-}$ HEK293F cells transiently expressing PAX5 and either an EV or SIRT7. Representative of two independent experiments. Kac indicates the acetylated K198 residue. **g,** Conservation of the PAX5 186-229 peptide in the indicated species. K198 is shown in red. NAM, nicotinamide.

NAD^+, suggesting that SIRT7 specifically deacetylates some PAX5 residues (**Fig. R20d**). To identify such residues, we transfected *SIRT7*$^{-/-}$ HEK293F cells previously generated in our lab with plasmids encoding FLAG-tagged PAX5, together with an EV or a vector encoding SIRT7. We then purified PAX5 from these samples and subjected the eluates to mass spectrometry. We detected five different acetylated lysines in PAX5, all of which had been previously identified by He *et al.*[372] (**Fig. R20e**). Of note, SIRT7 expression abrogated the acetylation of only one of these lysines, K198, as evidenced by the loss of the cluster of peaks spanning from 554.79 to 555.54 m/z, corresponding to the acetylated peptide. This indicates that SIRT7 specifically deacetylates PAX5 at this residue (**Fig. R20f**). Interestingly, K198 is located within a putative intrinsically disordered region of unknown function between the conserved octapeptide (OP) and the partial homeodomain (HD), and it is conserved across PAX5 orthologs in chordates (**Fig. R20e and R20g**).

We next aimed to determine the functional relevance of PAX5^{K198} acetylation. To do it, we mutated this lysine to Gln (PAX5^{K198Q}) or Arg (PAX5^{K198R}) to mimic its acetylated and deacetylated forms, respectively. Ectopic expression of these PAX5 mutants in HEK293F cells largely recapitulated the presence or absence of SIRT7, since PAX5^{K198Q} showed markedly reduced levels, whereas PAX5^{K198R} displayed enhanced expression (**Fig. R21a,b**). Next, we tested whether these differences in expression were due to differential proteasomal degradation. To this end, we transiently expressed the PAX5 mutants in HEK293F cells together with HA-tagged ubiquitin, whose covalent addition to target proteins marks them for proteasomal degradation, immunoprecipitated PAX5 and analysed its ubiquitination levels by western blot. As a result, PAX5^{K198Q} underwent increased ubiquitination, whereas reduced ubiquitination was observed in PAX5^{K198R} (**Fig. R21c,d**). Consistently, cycloheximide timecourse experiments confirmed a two-fold reduction in PAX5^{K198Q} half-life, in contrast to the enhanced stability displayed by PAX5^{K198R} (**Fig. R21e**). To exclude the possibility that the increased stability shown by PAX5^{K198R} was caused by potential arginine modifications, we also mutated K198 to alanine (PAX5^{K198A}), which also mimics unmodified lysine. This showed that both the protein levels and the stability of this mutant closely resembled those of PAX5^{K198R}, indicating that deacetylation of this lysine rather than its mutation causes increased PAX5 stability (**Fig. R21f-h**). Finally, we also explored the effect of K198 acetylation on PAX5 subcellular distribution, by performing subcellular fractionations in HEK293F cells transiently expressing the PAX5 mutants. As expected, none of the mutants localised in the cytoplasm. Furthermore, the levels of both mutants were indistinctly regulated in nucleoplasm and chromatin

(**Fig. R21i**). Collectively, our data demonstrate that SIRT7-mediated K198 deacetylation prevents PAX5 proteasomal degradation and enhances its stability.

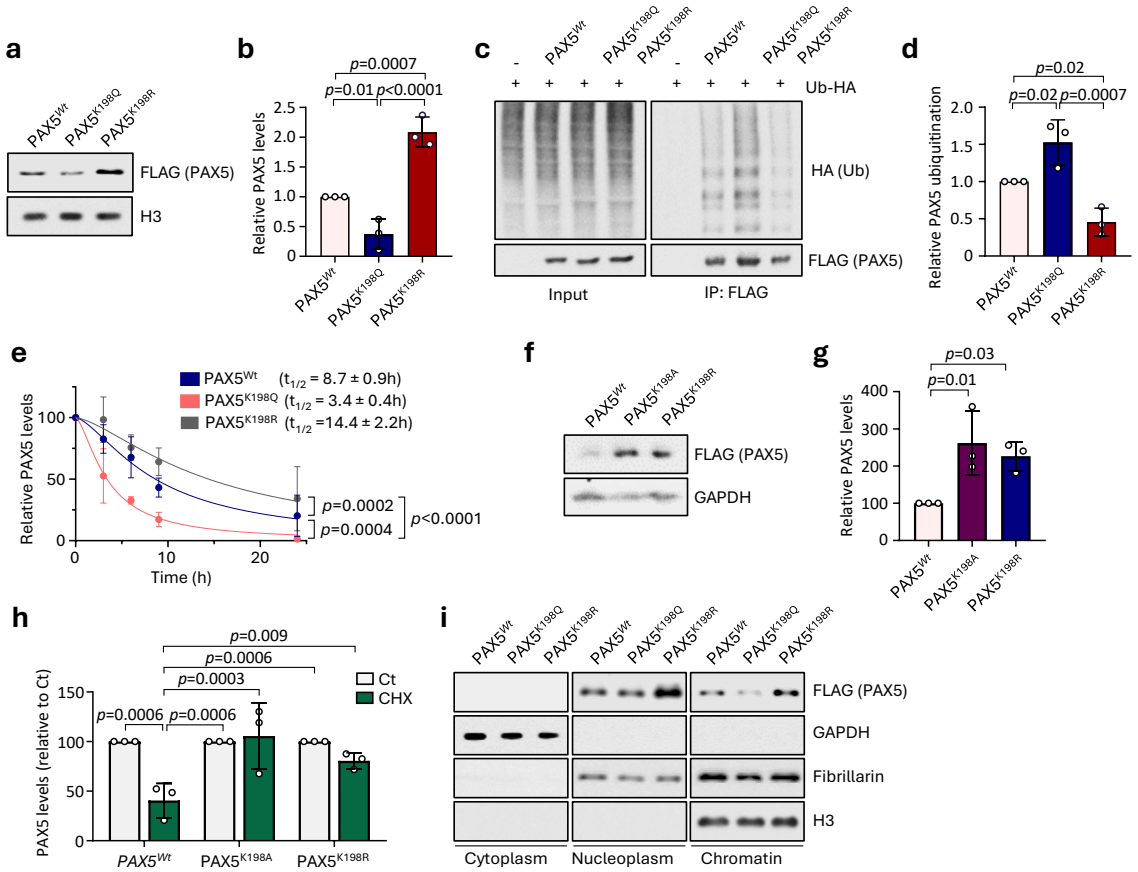


Fig. R21 | SIRT7-mediated PAX5^{K198} deacetylation enhances PAX5 stability

a,b, Immunoblots (**a**) and quantification (**b**) of the levels of FLAG-tagged PAX5-K198 mutants transiently expressed in HEK293F cells. Data in **b** are presented as mean ± s.d. Statistical significance was assessed by two-way ANOVA with Fisher's LSD test (n = 3). **c,** Immunoblots (**c**) and quantification (**d**) of the ubiquitination of FLAG-purified PAX5 mutants in HEK293F cells transiently expressing HA-Ubiquitin (HA-Ub). Data in **d** are presented as in **b** (n = 3). **e,** Cycloheximide timecourse (100 µg/mL) of HEK293F cells transiently expressing the PAX5 K198 mutants. PAX5-to-H3 ratios are shown as mean ± s.d, and non-linear fits with variable slope (four parameters) are depicted. Half-lives ± s.d were compared by two-tailed *t*-test (n = 4). **f,g,** Immunoblots (**f**) and quantification (**g**) of the indicated PAX5 forms. Data in **g** are shown as in panel **b** (n = 3). **h,** Relative levels of the indicated PAX5 forms in HEK293F cells treated with vehicle or 100 µg/mL cycloheximide for 8h. Data are presented mean ± s.d. Statistical significance was assessed by two-way ANOVA with Fisher's LSD test (n = 3). **i,** Subcellular fractionation of HEK293F cells expressing the indicated PAX5 forms. GAPDH, Fibrillarin and H3 are shown as cytoplasmic, nuclear and chromatin controls, respectively. Representative of two separate experiments.

5.3. PCAF is the major PAX5^{K198} acetyltransferase

Next, we sought to identify the acetyltransferase/s that catalyse PAX5 acetylation at K198. p300 was evidently a good candidate, since it has been previously shown to acetylate several PAX5 lysines, including K198[372]. However, we decided to take an unbiased approach based on a two-step screening of potential candidates: first, we obtained a list of all mammalian proteins with HAT activity and searched for these proteins among all reported PAX5 interactors (which comprised the results of a previously published PAX5 interactome and the experimentally validated interactors compiled in the BioGRID repository). This led to the identification of four PAX5-interacting proteins that harbour HAT activity: p300, PCAF, NCOA3 and GTF3C4 (**Fig. R22a**). In the second step, we employed publicly available quantitative proteomics data from 27 human B-ALL cases to correlate the protein levels of PAX5 with those of the candidate HATs, based on the assumption that the HAT responsible for K198 acetylation should reduce PAX5 stability and therefore its protein levels should negatively correlate with those of PAX5. According to linear regression analysis, only PCAF protein levels were inversely correlated with those of PAX5, which strongly suggested that PCAF may be the main K198 HAT (**Fig. R22b**).

To validate these observations, we obtained expression vectors encoding p300, PCAF, NCOA3 and GTF3C4 and co-transfected each of these vectors with PAX5 in HEK293F cells to investigate their effects on PAX5 protein levels. Consistent with the proteomics data, PAX5 was dramatically downregulated upon PCAF expression, whereas the other HATs had no effect on PAX5 (**Fig. R22c**). Most importantly, PAX5^{K198R} was completely unaffected by PCAF expression, suggesting that PCAF reduces PAX5 stability by acetylating K198 (**Fig. R22d,e**). To test whether it was the case, we finally co-expressed PAX5^{Wt} or PAX5^{K198R} together the candidate HATs, immunoprecipitated PAX5 and measured its global acetylation levels using an anti-AcK antibody. Notably, PCAF was the only HAT that significantly increased PAX5^{Wt} acetylation, but this effect was abrogated by PAX5^{K198R} mutation (**Fig. R22f,g**). Altogether, these observations suggest that PCAF opposes SIRT7-mediated PAX5 stabilization by specifically acetylating PAX5^{K198}.

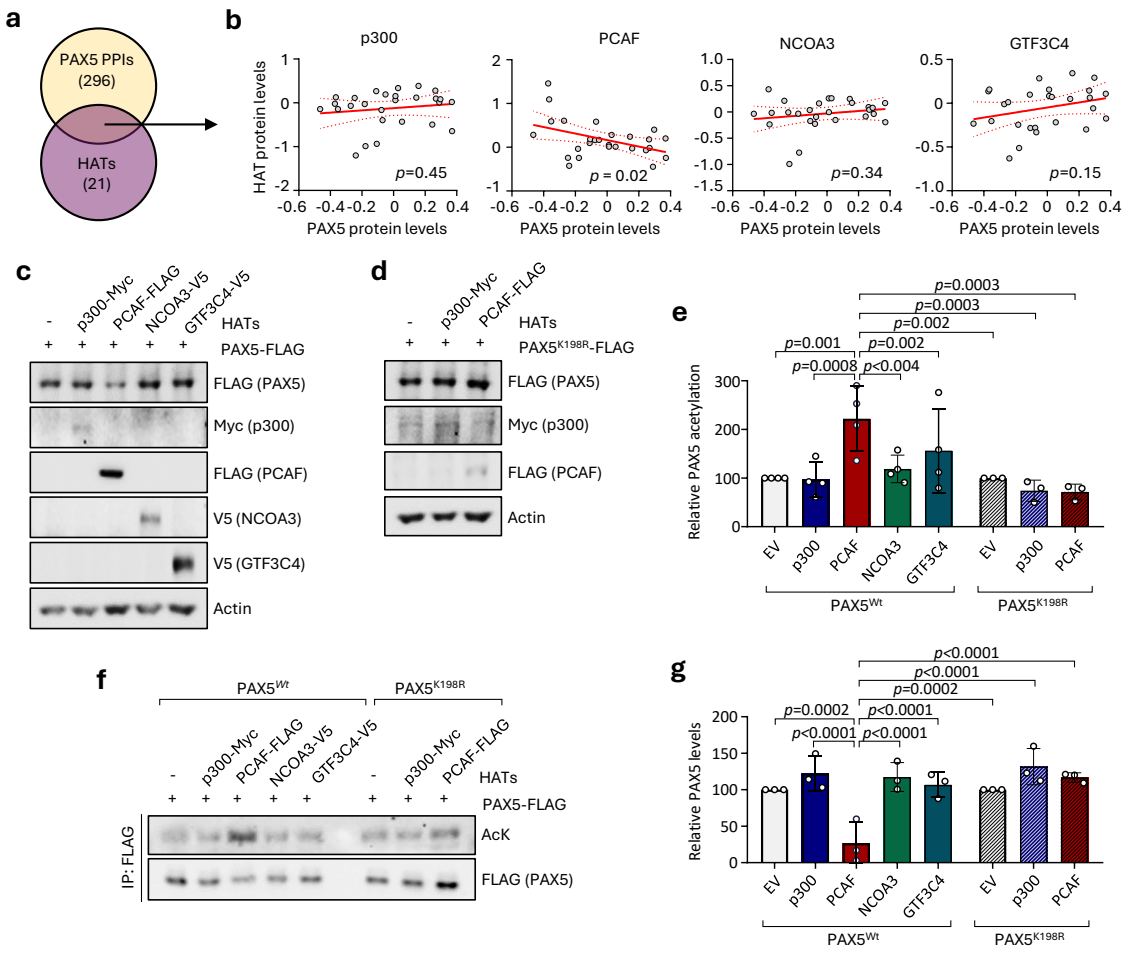
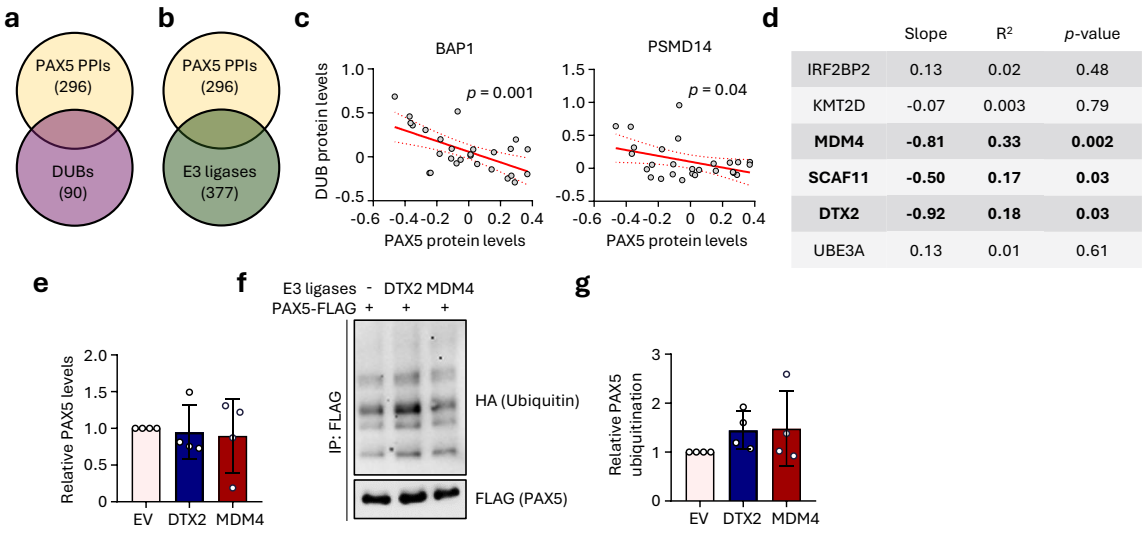


Fig. R22 | PCAF is the major PAX5^{K198} acetyltransferase

a, Venn diagram depicting the overlap between PAX5 PPIs and mammalian HATs. PAX5 protein-protein interactions (PPIs) integrate the interactors described in Ref [345] and those compiled in the BioGRID repository. The list of mammalian HATs was retrieved from Ref [174]. **b,** Scatter plots showing the correlation between the protein levels of PAX5 and the indicated HATs, as determined by proteomics in human B-ALL patients samples[346]. Each point corresponds to one sample. Linear regression and 95% confidence intervals (dashed lines) are shown (n = 27). **c-e,** Representative immunoblots (**c,d**) and quantification (**e**) of PAX5 protein levels in HEK293F cells expressing the indicated constructs. Data in **e** are shown as in **b** and were analysed by one-way ANOVA with Fisher's LSD test (PAX5^{Wt}, n = 4; PAX5^{K198R}, n= 3). **f,g,** Immunoblots (**f**) and quantification (**g**) of the global of PAX5 immunoprecipitated from HEK293F cells expressing the indicated constructs. Data in **g** are presented as in **e**. AcK, pan-acetyl lysine.

5.4. E3 ligases and deubiquitinases regulating PAX5 turnover

To identify the E3 ubiquitin ligases and deubiquitinases (DUBs) regulating PAX5 turnover, we followed a similar approach. We cross-referenced the same list of PAX5 interactors as above with lists of all mammalian E3 ligases and DUBs and subsequently compared the protein levels of the candidates with those of PAX5 in B-ALL patient samples. Among 90 DUBs and 377 E3 ligases, we found two PAX5 interactors with DUB activity (BAP1 and PSMD14) and six PAX5-interacting E3 ligases (IRF2BP2, KMT2D, MDM4, SCAF11, DTX2 and UBE3A) (**Fig. R23a,b**). In the case of the DUBs, both BAP1 and PSMD14 negatively correlated with PAX5, indicating that although they may deubiquitinate PAX5, they are unlikely to regulate its proteasomal degradation, since it is associated with increased ubiquitination (**Fig. R23c**). This prevented us from performing downstream analyses and identifying any DUBs involved in controlling PAX5 turnover. In contrast, three ubiquitin ligases (DTX2, SCAF11 and MDM4) displayed a significant inverse correlation with PAX5, suggesting a role in regulating its turnover (**Fig. R23d**). We therefore tried to obtain expression vectors for these proteins and achieved vectors encoding DTX and MDM4, but we were unable to clone the *Scaf11* cDNA, probably due to its very large size and the existence of multiple isoforms. Thus, we focused on DTX2 and MDM4 and co-expressed them with PAX5 and HA-Ubiquitin to evaluate their ability to promote PAX5 ubiquitination and degradation. Overall, DTX2, and to a lesser extent MDM4, slightly reduced PAX5 levels and increased its ubiquitination, suggesting that both targeted PAX5 (**Fig. R23e-g**). However, the effects of both proteins on PAX5 were very weak and variable, so we could not conclude that either of them are *bona fide* regulators of PAX5 ubiquitination. Although our previous experiments in HEK293F cells strongly suggested that PAX5 degradation is regulated by ubiquitous E3 ligases and DUBs, these new data suggested a more complex regulation. Nevertheless, further research will be required to better understand the mechanisms that regulate PAX5 degradation and to determine the potential involvement of SCAF11 in this process.



	Slope	R^2	*p*-value
IRF2BP2	0.13	0.02	0.48
KMT2D	-0.07	0.003	0.79
MDM4	**-0.81**	**0.33**	**0.002**
SCAF11	**-0.50**	**0.17**	**0.03**
DTX2	**-0.92**	**0.18**	**0.03**
UBE3A	0.13	0.01	0.61

Fig. R23 | Identification of putative PAX5 E3 ligases and DUBs

a,b, Venn diagram depicting the overlap between PAX5 PPIs and mammalian deubiquitinases (DUBs) (**a**) or E3 ligases (**b**). PAX5 PPIs integrate the interactors described by Okuyama *et al*. (2019, PLOS Genetics)[345] and those compiled in the BioGRID repository. The list of mammalian DUBs was obtained from https://esbl.nhlbi.nih.gov/Databases/KSBP2/Targets/Lists/DUBs/, and the list of mammalian E3 ligases was from https://esbl.nhlbi.nih.gov/Databases/KSBP2/Targets/Lists/E3-ligases/ **c,** Scatter plots showing the correlation between the protein levels of PAX5, and the indicated DUBs determined by proteomics in human B-ALL patients samples[346]. Each point corresponds to one sample. Linear regression and 95% confidence intervals (dashed lines) and Spearman's rank correlation coefficient (R^2) and significance are shown (n = 27). **d,** Table summarizing the correlation between the protein levels of PAX5 and the indicated E3 ligases. **e,** Densitometric quantification of the relative protein levels of PAX5 in HEK293F cells expressing the indicated E3 ligases. Data are shown as mean ± s.d. **f,g,** Representative immunoblots (**f**) and quantification (**g**) of the ubiquitination levels of FLAG-tagged PAX5 purified from HEK293F cells expressing the indicated constructs. In **g**, data are presented as in mean ± s.d.

6. PAX5^{K198} dynamic deacetylation is required for B-cell lymphopoiesis

6.1. K198 deacetylation regulates PAX5 binding to chromatin and PAX5-mediated transcriptional regulation

The above findings raised the compelling hypothesis that SIRT7 may regulate early B-cell development and commitment through deacetylation of PAX5 at K198. However, PAX5 haploinsufficiency is known to have a minimal impact in B-cell progenitors[114], so PAX5 downregulation in *Sirt7*$^{-/-}$ B-cell progenitors alone is unlikely to explain the role of SIRT7 in this process. In contrast, the finding that SIRT7 deacetylates PAX5 may provide a rationale for understanding how SIRT7 promotes B-cell development. Therefore, we next investigated the impact of PAX5^{K198} acetylation on its transcriptional functions and on B-cell lymphopoiesis. To this end, we used pro-B cells from the foetal livers of *Pax5*$^{-/-}$ mice, obtained through a collaboration with Dr. Mikael Sigvardsson (Lund University, Sweden).

We reintroduced either PAX5^{Wt} or PAX5^{K198} mutants into these cells and subjected them to PAX5 ChIP-Seq, RNA-Seq and *in vivo* transplantation experiments (**Fig. R24a**). Profiling global binding intensities of ChIP-Seq experiments revealed that acetylated PAX5^{K198Q} poorly bound to its target loci, whereas deacetylated PAX5^{K198R} exhibited enhanced binding (**Fig. R24b**). Notably, when we analysed the regions occupied by each of the PAX5 forms we found substantial differences among them (**Fig. R24c**). These differences were due to three reasons, which provided important insights into the genomic functions of the mutants: first, acetylated PAX5^{K198Q} only bound to 20.8% of all the peaks detected, indicating a dramatic impairment of its functions in chromatin (**Fig. R24c**). Second, up to 3,093 (28.2%) of the detected peaks were newly bound by deacetylated PAX5^{K198R} but not by PAX5^{Wt}, suggesting that K198 deacetylation led to a remarkable PAX5 genomic redistribution. Third, PAX5^{K198R} not only bound to new regions but also lost 1,659 (15.1%) canonical PAX5 binding sites, which also supported that K198 deacetylation leads to PAX5 redistribution. In the light of these findings, we decided to carefully analyse this possibility. We used publicly available PAX5 ChIP-Seq data which had identified roughly 30,000 peaks, in contrast to the ≈10,000 peaks detected in our analysis, to determine whether the unique PAX5^{K198R} binding sites were previously described to be bound by PAX5 or were indeed newly bound by PAX5^{K198R}. As a result, more than 90% of the peaks occupied by PAX5^{K198R} overlapped with previously reported PAX5 peaks (**Fig. R24d**). Therefore, we inferred that PAX5 acetylation regulates its ability to bind to specific loci, rather than causing its redistribution. Consistent with this idea, locus-specific analyses of regions containing unique PAX5^{K198R} or PAX5^{Wt} peaks confirmed that these regions were weakly bound by the PAX5 forms that lacked significant peaks, as exemplified by the *Actr6*, *Egfl6*, *Funcd1* and *Nus1* loci (**Fig. R24e**). Thus, the differential peaks identified by the peak calling were due to quantitative rather than qualitative differences. Consistently, *de novo* motif enrichment analysis indicated that the PAX5 DNA binding motif was minimally affected by K198 acetylation (**Fig. R24f**). We conclude that K198 acetylation affects PAX5 binding strength in a locus-specific manner, rather than causing its redistribution.

We next interrogated the potential outcomes of the differential binding by deacetylated PAX5^{K198R} by performing gene ontology analysis on its “unique” binding sites. This revealed that the peaks preferentially bound by PAX5^{K198R} were associated to lineage-inappropriate genes (helper and cytotoxic T-cell genes and monocyte genes), as well as genes related to B-cell development, cell cycle and metabolism (**Fig. R24g**). Therefore, the acetylation dynamics of PAX5 impact its functions in genome regulation and may affect its ability to

control specific aspects of developing B-cells.

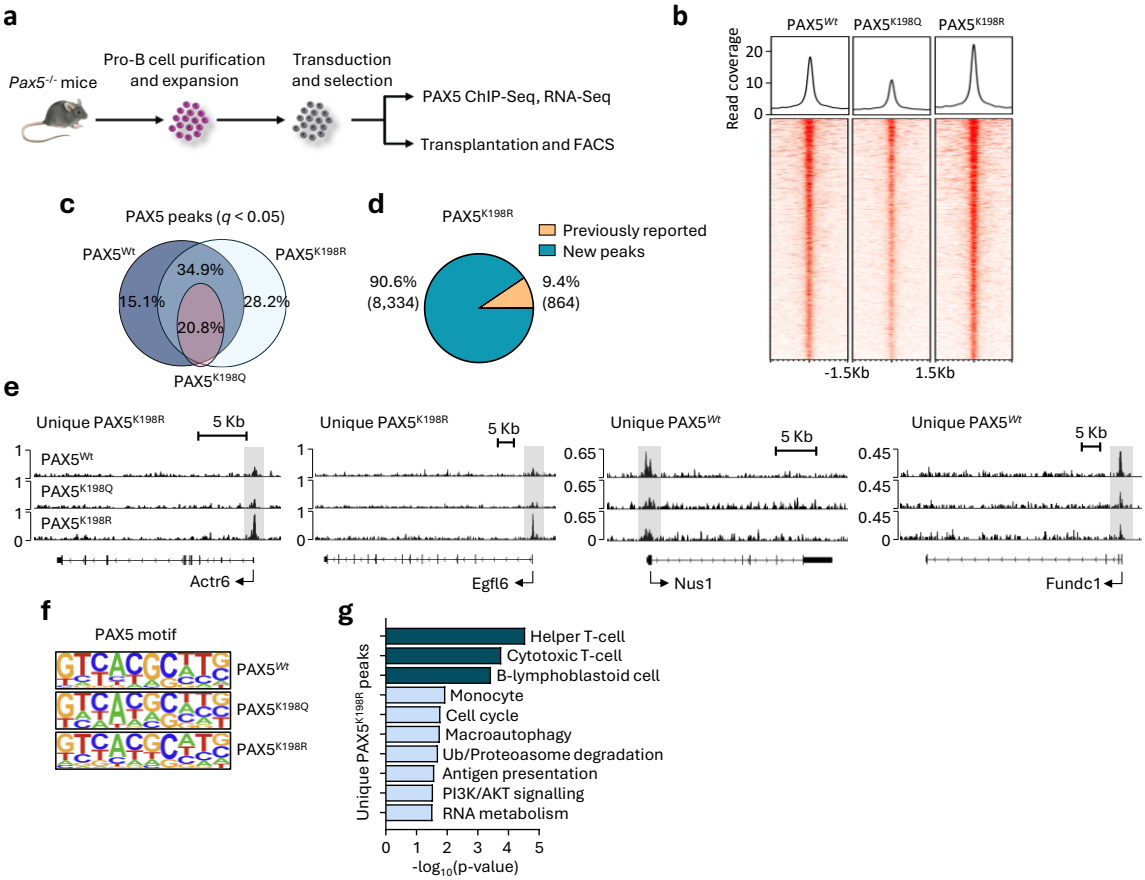


Fig. R24 | Impact of PAX5^{K198} acetylation on PAX5 genomic occupancy

a, Schematic representation of the purification, retroviral transduction and selection of foetal liver *Pax5*$^{-/-}$ pro-B cells followed by ChIP-Seq, RNA-Seq and congenic transplantation experiments with the generated cells expressing an empty vector (EV) or the PAX5^{K198} mutant forms. **b,** PAX5 ChIP-Seq analysis showing the genomic occupancies of PAX5^{Wt}, PAX5^{K198Q} and PAX5^{K198R} in *ex vivo*-expanded Pax5$^{-/-}$ pro-B cells. 1000 random significant peaks are displayed. Top panels, read coverage profiles. Representative of two replicates. **c,** Venn diagram showing the overlap between the significant peaks (q < 0.05) detected by ChIP-Seq of the indicated PAX5 forms. **d,** Comparison of the significant peaks bound by PAX5^{K198R} and those reported in Ref [81]. **e,** Binding of PAX5 mutants at four representative regions (**d,** Actr6; **e,** Nus1). The y-axis represents read coverage. **f,** Motif enrichment analysis of the significant peaks (q<0.05) detected by ChIP-Seq of PAX5 mutants. **g,** Gene ontology terms of the unique PAX5^{K198R} peaks. Representative of two replicates.

RNA-Seq analysis of pro-B cells reconstituted with PAX5 mutants confirmed that K198 acetylation dynamics strongly affected the transcriptional programmes induced by PAX5. PCA clustering revealed that pro-B cells expressing PAX5^{K198Q} clustered closer to *Pax5*$^{-/-}$ cells (EV) than those expressing PAX5^{Wt} did, indicating that acetylated PAX5 insufficiently regulated its target genes (**Fig. R25a**). Surprisingly, although deacetylated PAX5^{K198R} clustered farther away in PC1, it failed to regulate those features belonging to PC2. Since

PAX5^{K198Q} was much more efficient in regulating these features, the PCA suggested an unexpected role for acetylated PAX5 in transcriptional regulation. Unsupervised clustering analysis revealed that all three PAX5 forms regulated the same sets of genes, because these genes clustered into only two categories: cluster 1, containing the genes that were downregulated by PAX5 expression; and cluster 2, containing the genes that PAX5 induced (**Fig. R25b**). As previously described, PAX5 repressed alternative lineage genes, as well as genes related to cytokine signalling and upstream B-cell differentiation stages, while activating genes important for B-cell development, such as those related to V(D)J recombination or signalling by pre-BCR and IL7R. Notably, while PAX5 mutants regulated the same set of genes, they did it with very different efficiency, with PAX5^{K198R} massively repressing and inducing its targets, and PAX5^{K198Q} regulating them poorly. When we looked specifically at the genes that each PAX5 mutant induced and repressed significantly, the activation of PAX5 targets was relatively unaffected by K198 acetylation. In sharp contrast, PAX5 acetylation almost completely disrupted its ability to repress its target genes, whereas deacetylated PAX5^{K198R} stringently repressed target gene expression (**Fig. R25c**). Comparing transcript abundance with the binding of PAX5 mutants showed that, in general, the differential binding of the mutants to the promoter or to putative regulatory regions of its target genes strongly correlated with their regulation (**Fig. R25d**). However, in addition to these general differences, PAX5^{K198R} aberrantly regulated specific genes related to proliferation (e.g. *Il7r*, *Dusp4* and *Mycn*) and V(D)J recombination (*Blnk*, *Irf4* and *Rag1*), whereas PAX5^{K198Q}-mediated regulation of these genes resembled more to that of PAX5^{Wt}, again supporting a role for PAX5^{K198Q} in differentiation (**Fig. R25e**). Therefore, SIRT7-mediated K198 deacetylation strongly enhances PAX5's ability to control its target genes but also triggers the abnormal regulation of a subset of PAX5 targets.

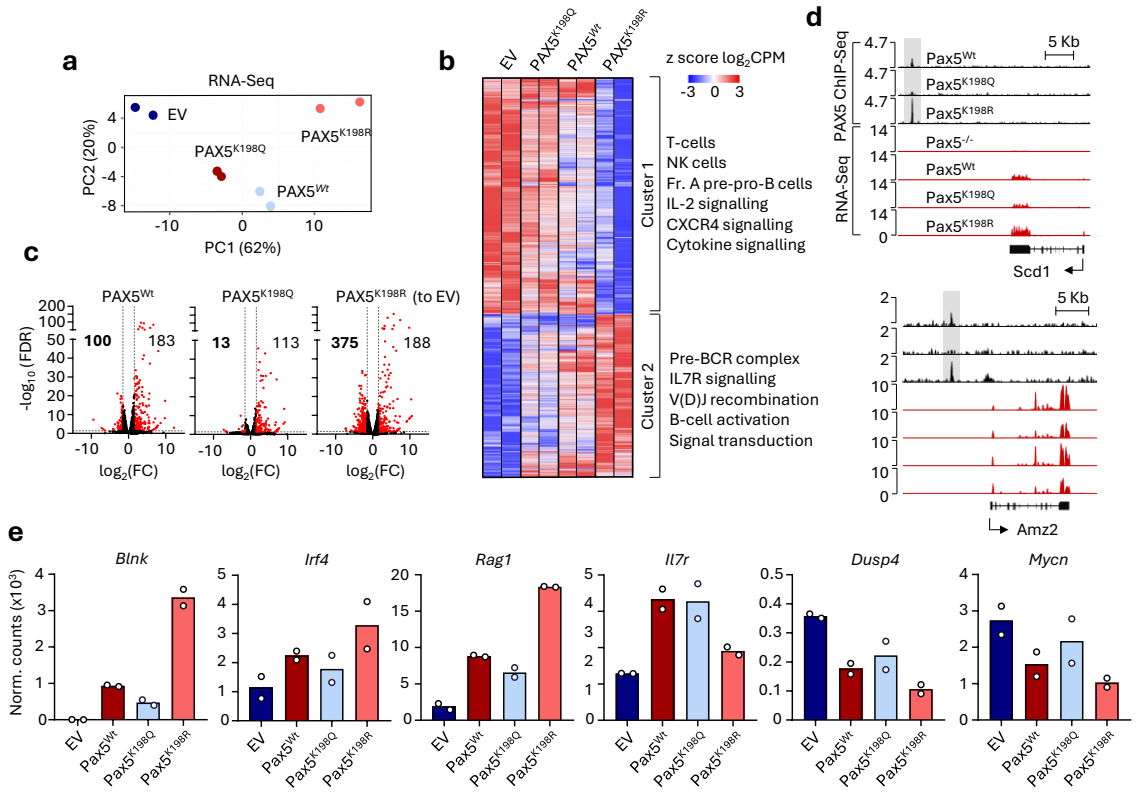


Fig. R25 | Impact of PAX5^{K198} acetylation on PAX5-mediated transcriptional regulation

a, PCA clustering of RNA-Seq data from *Pax5*$^{-/-}$ pro-B cells expressing the indicated retroviruses. **b,** Unsupervised clustering of differentially expressed genes (FDR < 0.05). The significant gene ontology terms for Clusters 1 and 2 are shown. **c,** Volcano plots of differentially expressed genes of PAX5^{K198} mutants versus EV-expressing *Pax5*$^{-/-}$ pro-B cells. The black arrow in the middle panel indicates the absence of downregulated genes, and significantly regulated genes ($|\log_2(\text{fold-change})| \geq 1.5$, FDR < 0.05) are shown in red. The numbers of significantly genes induced and repressed by each PAX5 form are shown in the right and left parts of the plots, respectively. **d,** Binding of PAX5 mutants (upper tracks, black) at two representative genes and RNA-Seq expression (lower tracks, red) of these genes. The y-axis represents read coverage. Representative of two replicates. **e,** RNA-Seq analysis of the expression of the indicated genes in pro-B cells expressing an empty vector (EV) or the PAX5^{K198} mutants. Data are presented as mean (n = 2).

6.2. Impact of PAX5^{K198} deacetylation on B-cell development and commitment

The above findings strongly supported the hypothesis that PAX5 deacetylation by SIRT7 is required for B-cell development. Furthermore, PAX5 is the major mediator of lineage commitment in pro-B cells, and we found that SIRT7 was also required for this process and that PAX5^{K198Q} failed to repress lineage-inappropriate genes. Therefore, our data also suggested that PAX5 deacetylation may be important for lineage commitment. As anticipated, we tested these hypotheses by performing *in vivo* transplantation experiments with the CD45.1 system. Since *Pax5*$^{-/-}$ cells were purified from C57BL/6 CD45.2 mice, we used congenic C57BL/6 CD45.1 mice for these experiments. We

injected *Pax5*$^{-/-}$ pro-B cells expressing either an EV or the different PAX5 forms and analysed their ability to repopulate the B-cell lineage and their T-cell lineage potential. Unexpectedly, while PAX5^{Wt} efficiently rescued B-cell development, neither PAX5^{K198Q} nor PAX5^{K198R} were able to do so (**Fig. R26a**). In sharp contrast, PAX5^{K198R} completely abrogated lineage potential, as deduced by the fact that no pro-B cell derived CD45.2^{+}CD4^{+} or CD45.2^{+}CD8^{+} T-cells were detected in the thymus of mice injected with pro-B cells expressing PAX5^{K198R} (**Fig. R26b**). Conversely, PAX5^{K198Q} was completely unable to induce lineage commitment, since pro-B cells expressing this mutant generated as many T-cells as *Pax5*$^{-/-}$ cells did. Therefore, PAX5 deacetylation by SIRT7 promotes B-cell lineage commitment but is clearly not sufficient to drive B-cell development.

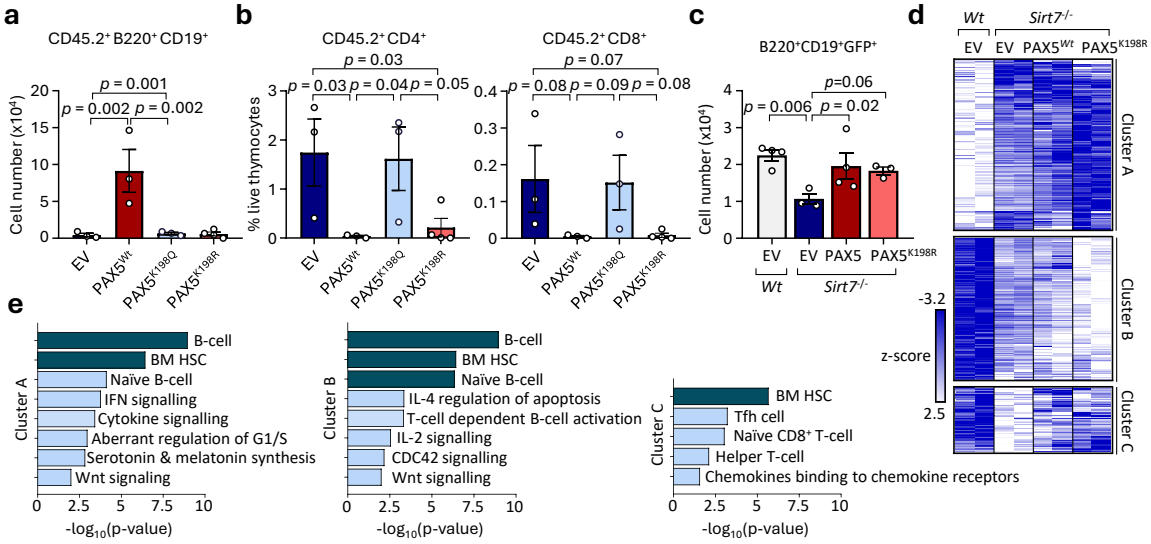


Fig. R26 | Impact of PAX5K198 deacetylation on B-cell

a,b, Number of CD45.2^{+}CD19^{+}B220^{+} cells in the bone marrow (**a**) and percentages of CD45.2^{+}CD4^{+} and CD45.2^{+}CD8^{+} (**b**) cells in the thymuses of CD45.1 recipient mice 4 weeks after transplantation of *Pax5*$^{-/-}$ pro-B cells expressing the indicated retroviruses. Data are presented as mean ± SEM and were analysed by one-way ANOVA with Fisher's LSD test (EV, PAX5^{Wt} and PAX5^{K198Q}, n = 3; PAX5^{K198R}, n = 4). **c,** Number of B220^{+}CD19^{+}GFP^{+} cells in the spleens of recipient mice injected with *Wt* or *Sirt7*$^{-/-}$ pro-B cells expressing the indicated retroviruses. Data are presented as mean ± s.d. and were analysed by one-way ANOVA with Fisher's LSD test. **d,** Unsupervised clustering of the differentially expressed genes ($p < 0.05$) between *Wt* and *Sirt7*$^{-/-}$ pro-B cells. **e, ,** RNA-Seq analysis of the expression of the indicated genes in pro-B cells expressing an empty vector (EV) or the PAX5^{K198R} mutants. Data are presented as mean ± s.d. and were analysed by one-way ANOVA with Fisher's LSD test (n = 2).

The finding that both PAX5^{K198Q} and PAX5^{K198R} failed to revert the B-cell developmental arrest suggested that dynamic acetylation and deacetylation of K198, rather than its deacetylation alone, may be required for optimal B-cell development. This was further supported by the observation that PAX5^{K198Q} was important for the physiological regulation of several key B-cell developmental genes that PAX5^{K198R} failed to properly regulate.

Further supporting this idea, it was recently reported that the conditional deletion of PCAF and its close homolog GCN5 in B-cell progenitors impairs B-lymphopoiesis[373], and here we found that PCAF is the major K198 acetyltransferase. Therefore, acetylated and deacetylated PAX5 may both be indispensable and perform non-redundant functions to orchestrate B-cell development. To interrogate this possibility, we leveraged the fact that *Sirt7*$^{-/-}$ pro-B cells express reduced levels of hyperacetylated PAX5, and reintroduced PAX5^{Wt} or deacetylated PAX5^{K198R} into these cells to test their ability to bypass the developmental defect induced by SIRT7 deficiency. Thus, we transduced *Wt* and *Sirt7*$^{-/-}$ pro-B cells with retroviruses encoding either an EV, PAX5^{Wt} or PAX5^{K198R} followed by a cassette containing an IRES and the coding sequence of GFP as a selection marker. After sorting the transduced GFP^{+} cells, we injected them into CD45.1CD45.2 and determined their ability to reconstitute the B-cell compartment. Importantly, both PAX5^{Wt} or PAX5^{K198R} rescued B-cell development in *Sirt7*$^{-/-}$ cells to levels comparable to those of *Wt* cells, which confirmed that SIRT7 promotes this process through deacetylation of PAX5^{K198}, and that both acetylated and deacetylated PAX5 are required for normal B-cell development (**Fig. R26c)**. Finally, we aimed to understand how PAX5 rescued the transcriptional programmes of SIRT7-deficient pro-B cells, so we subjected these cells to RNA-Seq. Notably, unsupervised clustering analysis of the genes regulated by SIRT7 indicated that PAX5^{K198R} and, to a lesser extent PAX5^{Wt}, strongly repressed 93 (31.3%) of the genes that were upregulated after SIRT7 loss. In sharp contrast, neither of the PAX5 forms affected the genes induced by SIRT7 (**Fig. R26d**). Consistent with its role in lineage commitment, PAX5 mainly repressed lineage-inappropriate genes, as well as genes related to chemokine signalling (**Fig. R26e**). Importantly, some of these genes were more efficiently rescued by PAX5^{Wt}, and PAX5^{K198R} abnormally regulated several Cluster B genes, which further supported the relevance of PAX5 dynamic rather than constitutive deacetylation. Taken together, these results indicate that PAX5 deacetylation by SIRT7 is required for lineage commitment and that B-cell development relies on the dynamic deacetylation of PAX5.

7. The PAX5/SIRT7 interplay is conserved in human B-ALL

PAX5 is a strong tumour suppressor in human B-ALL. In this disease, inactivating mutations of one allele of the *PAX5* gene are frequent, but biallelic *PAX5* mutations are rarely observed[145]. Given the ability of SIRT7 to stabilise PAX5 and increase its protein levels, it may also behave as a tumour suppressor in B-ALL patients harbouring *PAX5* mutations. This prompted us to explore whether the PAX5/SIRT7 interplay is functionally relevant in human B-ALL. We established a panel of human B-ALL cell lines and analysed the protein and RNA levels of PAX5 and SIRT7 in these cells by immunoblotting and RT-qPCR, respectively. Notably, despite the diverse genetic backgrounds of these cell lines, SIRT7 and PAX5 protein but not RNA levels were significantly correlated, strongly suggesting that SIRT7 also promotes PAX5 stability in human B-ALL (**Fig. R27a-c**). To further explore the ability of SIRT7 to stabilise PAX5 in B-ALL, we next transduced two of the cell lines expressing the lowest PAX5 levels (NALM-20 and TANOUE) with retroviruses to overexpress SIRT7 in these cells. Importantly, SIRT7 overexpression significantly increased PAX5 protein levels (**Fig. R27d,e**). This raised the intriguing possibility that

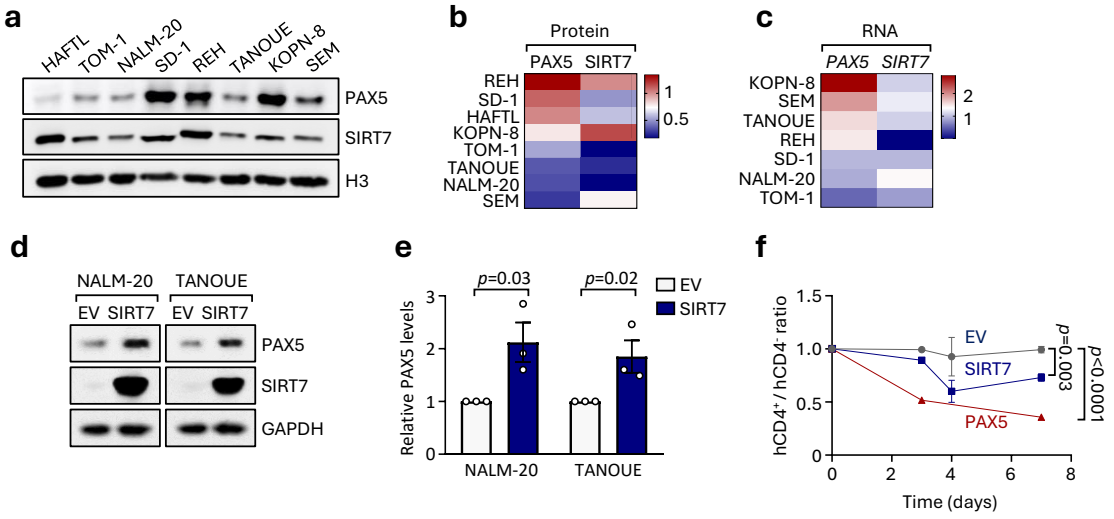


Fig. R27 | The PAX5/SIRT7 interplay is conserved in human B-ALL

a-c, Immunoblot (**a**) and heatmaps showing the correlation between PAX5 and SIRT7 protein (**b**) and RNA (**c**) levels in B-ALL cell lines. In **b** and **c**, the mean z-scores for the housekeeper-normalised PAX5 and SIRT7 levels are shown (Protein, PAX5 to H3; RNA, *Pax5* to *Hprt*). The *p*-values were determined by Spearman's rank correlation (**a,b**, three separate experiments; **c**, two separate experiments). **d,** Immunoblots showing the levels of PAX5 and SIRT7 in NALM-20 and TANOUE B-ALL cells expressing the indicated retroviruses. Representative of three separate experiments. **e,** PAX5-to-GAPDH ratio from the data in panel **d**. Data are shown as mean ± s.d. and were analysed by one-tailed t-tests (n = 3). **f,** *In vitro* competitive growth assay of TANOUE cells stably transduced with the indicated constructs. The ratio between infected versus non-infected cells ($hCD4^+/hCD4^-$ ratio) was determined by FACS in the pool of transduced cells. Data are shown as mean ± s.d. and were analysed by two-way ANOVA (n = 3 replicates, representative of 2 independent experiments).

SIRT7-mediated PAX5 stabilisation enhances its tumour suppressor functions and leads to B-ALL regression. We preliminarily tested this possibility by stably expressing an EV, PAX5 or SIRT7 in TANOUE cells through retroviral transduction of constructs containing the $hCD4^+$ selection marker. Instead of selecting the infected cells, we used the pooled cells to follow by FACS the relative abundance of the cells expressing the plasmid ($hCD4^+$), as a measure of the anti-leukaemic activity of PAX5 and SIRT7. Remarkably, the transduced cells expressing SIRT7 or PAX5 were significantly depleted after 3 or 7 days, whereas no significant differences were observed in cells expressing the EV (**Fig. R27f**). These competitive growth assays confirmed that both SIRT7 and PAX5 behaved as tumour suppressors in TANOUE B-ALL cells.

Next, we sought to expand these observations to human patients. To do so, we analysed the protein and RNA expression of PAX5 and SIRT7 using publicly available proteogenomic data[346] from a cohort of 27 ETV6::RUNX1 and High Hyperploidy (HeH) B-ALL patient samples, two of the most common B-ALL with good prognosis. These data included proteomics and RNA-Seq data. In these patients, PAX5 and SIRT7 protein levels again displayed a striking positive relation (**Fig. R28a**). In this case, *PAX5* and *SIRT7* RNA levels were also slightly correlated in HeH B-ALL cases, probably due to chromosomal amplifications involving their genes (**Fig. R28b**). However, this correlation was stronger at the protein level (**Fig. R28c**). In addition, we also correlated the levels of PAX5 with those of the different nuclear sirtuins in the same samples. Consistent with a specific role of SIRT7 in B-cell development and malignancy, neither of the other nuclear sirtuins was associated with PAX5 protein levels (**Fig. R28d,e**). Furthermore, we found an equivalent connection between SIRT7 and PAX5 when we similarly analysed a panel of B-CLL cell lines[347], which strongly suggested that SIRT7 regulates PAX5 in B-cell leukaemia regardless of the genetic background and differentiation stage (**Fig. R28f**). Thus, PAX5 protein stabilization by SIRT7 is not only relevant during normal B-lymphopoiesis but also in human B-ALL, which may provide an unprecedented rationale to target this disease.

The above findings strongly suggested that SIRT7 might be a therapeutically relevant tumour suppressor in B-ALL. Therefore, we finally explored whether *SIRT7* expression was a predictor of clinical outcome in B-ALL patients. For that purpose, we analysed publicly available gene expression, overall survival and genomic alteration data from the COG9906 trial, which enrolled paediatric patients with high-risk B-cell precursor B-ALL. Stratification of these patients according to their median level of SIRT7 expression revealed that

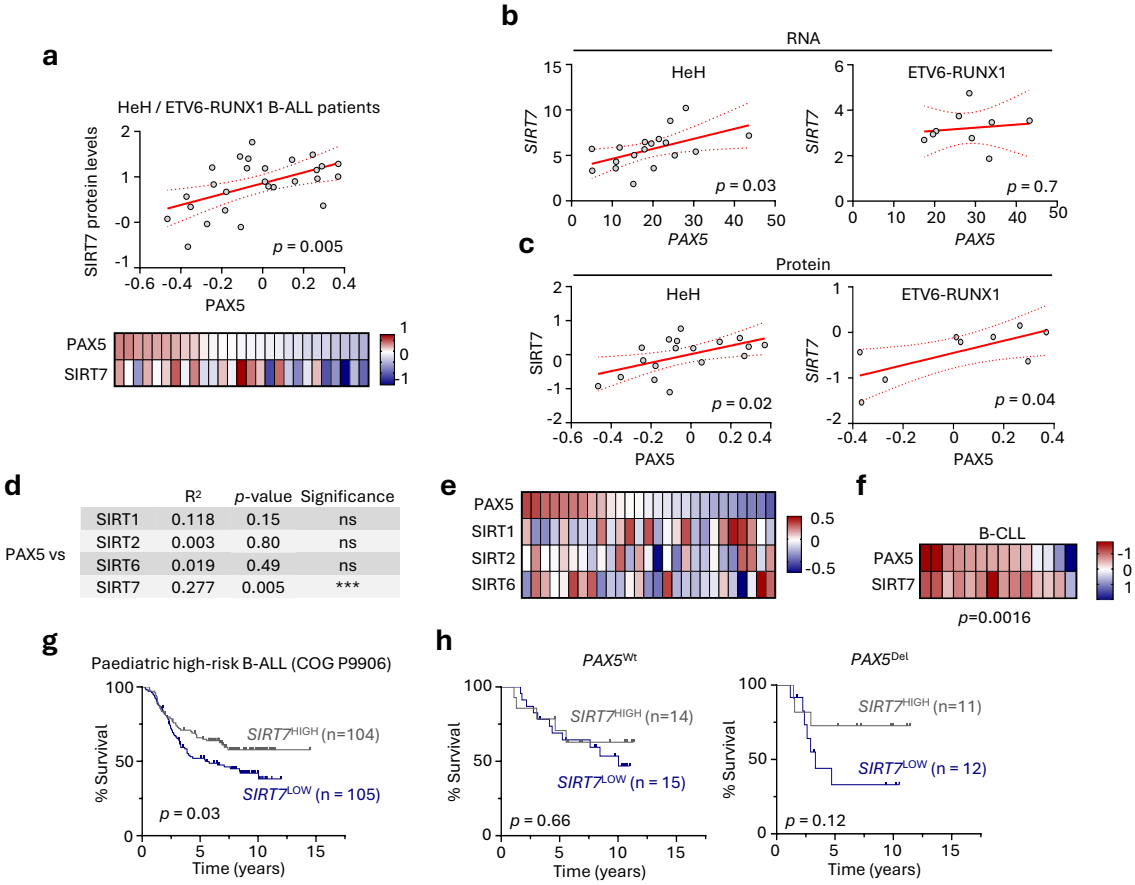


Fig. R28 | SIRT7 promotes PAX5 expression and is a tumour suppressor in human B-ALL

a, Scatter plot (upper panel) and heatmap (lower panel, z-score) showing PAX5 and SIRT7 protein levels determined by proteomics in human B-ALL patient samples[346]. Each point corresponds to one sample. Linear regression, 95% confidence intervals (dashed lines) and Spearman's rank correlation coefficient (R^2) are shown (n = 27). **b,c,** Scatter plots of PAX5 and SIRT7 RNA (**b**) and protein (**c**) levels derived from proteomics and RNA-Seq, respectively, of human B-ALL HeH and ETV6-RUNX1 patient samples[346]. Data are shown as in **a** (HeH, n = 18; ETV6-RUNX1, n = 9). **d,e,** Table (**d**) and heatmap (**e**) summarizing the correlation between the protein levels of PAX5 and nuclear Sirtuins as in **a**. **f,** Heatmap showing PAX5 and SIRT7 $\log_2$(fold-changes) in B-CLL cell lines versus healthy donor B-cells from a proteomics experiment[347]. The *p*-values were determined using Spearman's rank correlation. **g,** Kaplan-Meier survival curves for children with high-risk B-ALL (COG-P9906 study, n = 209) stratified into two groups based on higher- or lower-than-median *SIRT7* RNA expression. Statistical significance was determined by log-rank test. **h,** Kaplan-Meier survival curves of children with high-risk B-ALL (COG-P9906 study) stratified into four groups based on higher- or lower-than-median *SIRT7* RNA expression levels and the occurrence of *PAX5* deletions. Genomic data of PAX5 deletions were retrieved from Ref [348] Data are shown as in **g** (*PAX5*^{Wt}*SIRT7*^{HIGH}, n = 14; *PAX5*^{Wt}*SIRT7*^{LOW}, n = 15; *PAX5*^{Del}*SIRT7*^{HIGH}, n = 11; *PAX5*^{Del}*SIRT7*^{LOW}, n = 12). HeH, high hyperploidy. Patient survival and gene expression data in panels **g,f** were retrieved from the Genomic Data Commons database (GSE11877).

patients who expressed higher-than-median *SIRT7* levels (*SIRT7*^{HIGH}) were associated with good clinical outcomes, indicating that SIRT7 is an independent good prognosis factor in human B-ALL (**Fig. R28g**). According to our data, SIRT7 probably performs its tumour suppressor functions in B-ALL by enhancing PAX5 functions, suggesting that it may be particularly important in B-ALL cases showing monoallelic *PAX5* inactivating lesions. We

tested this hypothesis by further stratifying patients according to the presence or absence of *PAX5* deletions. Although the sample size was relatively low due to the limited availability of genomic data, our analysis clearly showed that *SIRT7*^{HIGH} expression was a predictor of good outcome only in patients harbouring *PAX5* deletions (**Fig. R28h**). Overall, our data strongly supports the notion that SIRT7 exerts its tumour suppressive functions in B-ALL by increasing the stability and functions of PAX5 in patients with hypomorphic *PAX5* mutations.

VII. DISCUSSION

In this study, we originally aimed to explore new functions of nuclear sirtuins during haematopoietic differentiation and how these functions could relate to the malignant transformation of haematopoietic progenitors. By combining experimental validation with unbiased approaches based on the profiling of sirtuin expression in immune and progenitor cells, we showed that SIRT7 is specifically upregulated in B-cell progenitors and sustains their normal development. Mechanistically, we identified a functional interplay between SIRT7 and the B-cell master regulator PAX5, that coordinated the diverse functions of PAX5 in genome regulation and strongly impacted the biology of B-cell progenitors and B-ALL cells. We propose a "dynamic PAX5 acetylation switch" model, whereby, during the pro-B-to-pre-B cell transition, SIRT7-mediated deacetylation of $PAX5^{K198}$ enhances its protein stability and genome occupancy to drive efficient chromatin and transcriptional regulation, but also leads to abnormal regulation of several PAX5 target genes. To accurately balance between complex processes such as lineage restriction, *Igh* recombination and cell proliferation, PCAF-mediated acetylation of $PAX5^{K198}$ globally inhibits PAX5 function but facilitates the proper expression of specific PAX5 targets. Ultimately, deacetylated PAX5 efficiently prompts B-cell lineage commitment alone, whereas B-cell differentiation is orchestrated by the coordinated action of both proteoforms (**Fig. D1**).

1. Sirtuin heterogeneity in haematopoiesis

Although here we focused on the role of SIRT7 in B-cells, our analysis of sirtuin expression in haematopoietic and immune cells provided some clues for the identification of potential novel functions of sirtuins in haematopoiesis and leukaemia. We found an unexpected heterogeneous expression pattern of nuclear sirtuins across haematopoietic progenitors, suggesting that they may operate in a highly context-specific manner. For instance, SIRT2 expression considerably paralleled that of SIRT7, since it not only displayed the highest overall abundance after SIRT7 in mouse haematopoietic progenitors but also underwent upregulation as B-cell development progressed. Although this observation clearly rises the possibility that SIRT2 and SIRT7 may play partially redundant functions in B-cell lymphopoiesis, the fact that SIRT2 displayed no correlation with PAX5 in B-ALL seems to indicate that SIRT2 involvement in this process, if any, is independent of PAX5 regulation. Another interesting observation concerned SIRT6, that was nearly undetectable in most of the cell types we profiled. Although this was shocking given its evolutionary proximity with SIRT7, a previous report already suggested that SIRT6 may be

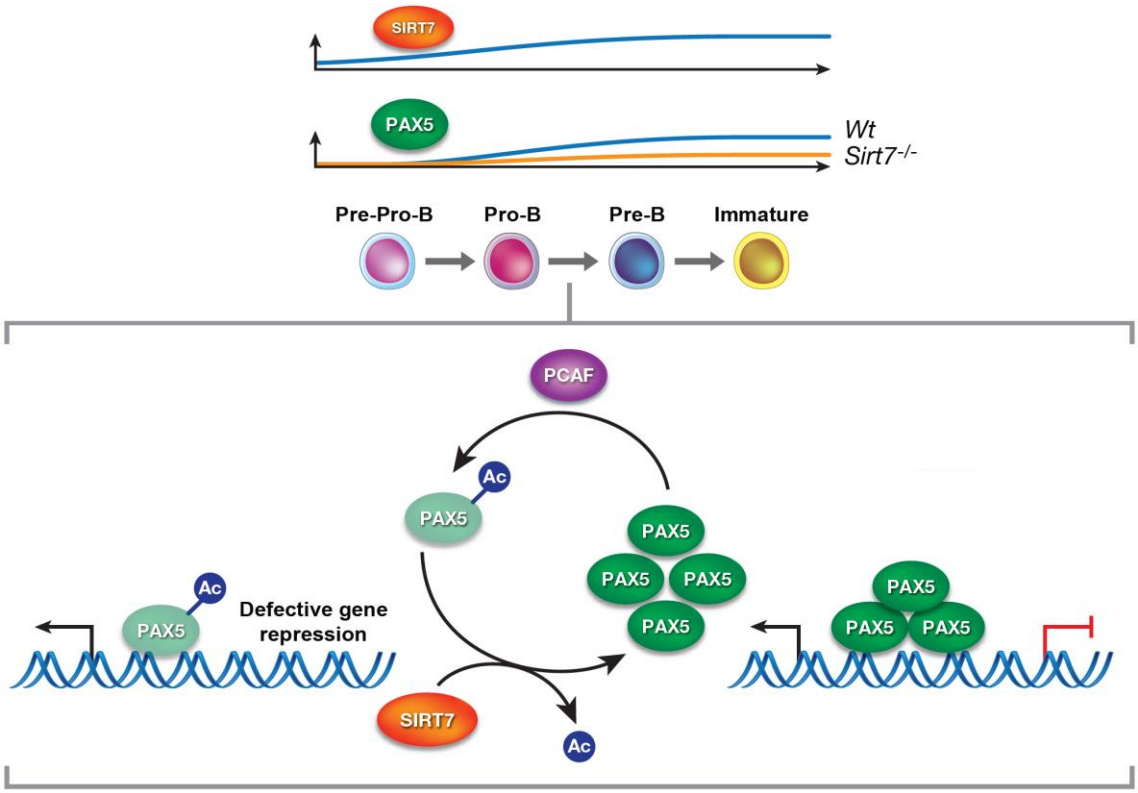


Fig. D1 | Schematic model for the SIRT7-mediated PAX5 acetylation switch
SIRT7 is upregulated during B lymphopoiesis to enhance the functional dose of PAX5. In B-cell progenitors, SIRT7 establishes a PAX5^{K198} acetylation switch that controls PAX5 stability and functions to ensure optimal B-cell development and commitment. K198 deacetylation by SIRT7 enhances PAX5 protein stability, thereby increasing its genomic occupancy and its ability to repress gene expression. Contrarily, K198 acetylation by PCAF reduces PAX5 levels and binding to chromatin, which prevents PAX5-mediated gene repression.

dispensable at least for lymphopoiesis, since *Sirt6*$^{-/-}$ bone marrow cells normally developed T-cells and B-cells in competitive transplantation experiments[286]. Finally, human SIRT7 seems to be a lymphoid-biased sirtuin within the immune compartment, since its expression was particularly high not only in B-cells but also in T-cells and NK cells. It will be interesting to explore whether the functions of SIRT7 in haematopoiesis and immunity extend beyond its interplay with PAX5.

Since the original discovery of sirtuins, a wide range of sirtuin studies have focused mainly on their relevance in the cellular stress response and in safeguarding genome integrity. Sirtuins certainly play fundamental roles in sensing and responding to various stresses, which typically increase their activity and alter their subcellular distribution. However, their evolutionary diversification and the strong phenotypes produced by some sirtuin deficiencies in mice suggested several years ago that they also have important functions in physiology in the absence of overt insults[285-288]. Nevertheless, their roles in organismal processes under normal conditions have been less attended. In particular, a large body of

literature has implicated sirtuins in immune responses through many different mechanisms, especially by regulating inflammation in myeloid cells and T-cells, but their roles in haematopoiesis remain largely unexplored[304]. In this sense, our findings here directly implicate a sirtuin family member, SIRT7, in the regulation of several physiological aspects of B-cell lymphopoiesis, including the establishment of early B-cell commitment.

To the best of our knowledge, sirtuins have not previously been shown to repress lineage plasticity in haematopoietic progenitors, but some of them are actively involved in cell fate decisions of non-haematopoietic cell types. One interesting example comes from the observation that in reducing conditions SIRT1 mediates neuronal versus astroglial differentiation in neural progenitor cells, whereas it promotes astroglial differentiation in response to oxidative stress, a common feature of brain pathologies[374]. Our findings suggest that SIRT7 may similarly link physiology with stress in B-cell progenitors. Indeed, the bone marrow is a naturally hypoxic environment, highlighting the need for haematopoietic progenitors to orchestrate an efficient stress response in the steady state, in which sirtuins are likely to participate[375]. Several sirtuins, including SIRT7, have been shown to negatively regulate HIF1, a fundamental sensor of hypoxia whose deficiency has been associated with strong defects in haematopoiesis in general and B-cell development in particular[327,376-378]. Thus, the interplay of sirtuins with stress factors such as HIF1 might underlie an intimate relation between cellular stress and haematopoietic fate decisions.

2. The various functions of SIRT7 in B-cell biology

Our results suggest that SIRT7 combines its ubiquitous functions with B-cell-specific roles to drive B-cell differentiation. SIRT7 represses lineage plasticity, in line with its role as a transcriptional repressor[379], while contributing to the dynamic coordination of proliferation and V(D)J recombination transcriptional programs. Furthermore, we show that SIRT7 is indispensable for cell survival and proliferation after the pre-BCR checkpoint, and that it may partly assist these processes by regulating STAT5, a central mediator of cell proliferation in pre-B cells. Although we did not investigate this hypothetical SIRT7/STAT5 axis in detail, SIRT7 directly interacted with STAT5 and appeared to be required for its complete activation, since STAT5 phosphorylation at Y694 was significantly reduced in *Sirt7*$^{-/-}$ progenitors *in vivo*. Interestingly, SIRT1 negatively regulates STAT5 activation through deacetylation, although in a completely different context[380]. From the perspective of genome integrity, the ability of SIRT7 to facilitate NHEJ by recruiting the DNA repair

factor 53BP1[288] acquired especial relevance in the repair of distal *Igh* recombination events, which may be important to sustain B-cell clonal diversity by ensuring a balanced usage of proximal and distal elements.

Another interesting aspect raised by our work comes from the observation that a significant fraction of SIRT7 and PAX5 forms part of large protein complexes. SIRT7 was previously reported to interact with several high-molecular-weight protein complexes, including the SWI/SNF and B-WICH chromatin remodeller complexes and RNA Pol I- and RNA Pol II-associated complexes[321,323]. However, these proteomic studies were performed in HEK293 cells overexpressing SIRT7, which can force non-physiological interactions. Here, we weight fractionated nuclear lysates from immortalised HAFTL pre-B cells and specifically immunoprecipitated the SIRT7 molecules associated to high-molecular-weight fractions. Our analysis supported the interaction of SIRT7 with many of the proteins identified in these reports, especially members of the SWI/SNF complex, and extended the list of protein complexes potentially containing SIRT7 to NuRD, INO80, MLL1, SET1C/COMPASS, STAGA, TFIID and others, pointing to SIRT7 as a common member of transcription factor, chromatin remodeller and histone modifier complexes, at least in B-lymphoid cells. Interestingly, PAX5 was recently reported to interact with many of these complexes, but we could not detect it in our proteomic analysis[381]. Although we did not investigate the functions of SIRT7 in the context of these complexes, we speculate that it may play both regulatory and catalytic functions through different mechanisms, including the deacetylation or mono-ADP-Ribosylation of their members; the recruitment to specific loci to regulate H3K18ac and H3K36ac; or even a direct contribution to the genomic distribution of these complexes. In any case, it will be interesting to explore in the future the functions of SIRT7 in these complexes and the extent to which these functions are B-cell-specific.

Our evidence not only identified SIRT7 as a member of several high-molecular-weight protein complexes but also seemed to implicate SIRT7 in the intricate transcription factor network governing B-cell development. Beyond its molecular interplay with PAX5, SIRT7 interacted with at least two other key members of this network, EBF1 and Ikaros, suggesting that it may play a broader role in the transcriptional control of B-cell development by regulating other B-lymphoid factors. Interestingly, this network has been associated with malignant transformation, since several of its members are frequently mutated in B-ALL[345]. Furthermore, B-ALL cases with PAX5 mutations often display complex karyotypes, suggesting that SIRT7 may collaborate with this network to protect

genome integrity[145,382]. Thus, determining the genome-wide occupancy of SIRT7 could provide valuable insights into its participation in the developmental and tumour suppressive functions of this network.

From an epigenetic perspective, SIRT7 is considered a transcriptional repressor though its ability to deacetylate H3K18ac and H3K36ac[276,277], although how H3K36ac abundance affects chromatin biology remains largely unexplored. We showed here that in pro-B cells SIRT7 stringently regulates H3K36ac while modestly affecting H3K18ac, which is consistent with other unpublished data from our laboratory pointing into the same direction in other cellular paradigms. Importantly, our findings also revealed a molecular interplay between PAX5 and H3K36ac, that is likely to be mediated in part by SIRT7. In *Pax5*$^{-/-}$ cells, H3K36ac dramatically collapsed, as other histone marks did, which strongly suggest that PAX5 may recruit GCN5 (the only known H3K36ac HAT[278]) and other histone modifiers to deposit these marks. Indeed, PAX5 regulates many of its target genes by recruiting chromatin modifying complexes to the regulatory elements of these genes[118]. Interestingly, our finding that PCAF, a close GCN5 homolog, acetylates PAX5 at K198, may underlie a more complex circuitry governing the cross-regulation between PAX5, PCAF, SIRT7 and H3K36ac. Given the considerable homology that PCAF and GCN5 display, GCN5 possibly forms part of this network as well, and it is reasonable to speculate that it might also target PAX5^{K198}. Furthermore, PCAF has been shown to bind H3K36ac itself, suggesting that it may also catalyse its deposition[383]. Although we did not formally demonstrate that PAX5 recruits SIRT7 to regulate H3K36ac levels, our data strongly suggest that this may be the case, since PAX5 and H3K36ac displayed an impressive genomic colocalization (57% shared peaks in genes), that was even more prominent upon SIRT7 loss (70% shared peaks in genes). However, the involvement of H3K36ac in the functions of PAX5 and SIRT7 in B-cell development and commitment remains unresolved.

Despite these observations, many questions remain open regarding the relevance of SIRT7-mediated H3K36ac deacetylation in genome regulation. For instance, although we did not show these data here, our preliminary analyses did not find any clear connexion between H3K36ac deacetylation and the genes that SIRT7 repressed, suggesting that other mechanisms may account for the repression of these genes and that H3K36ac may be involved in other chromatin regulatory processes. Furthermore, we did not explore whether H3K36ac also colocalised with other B-lymphoid transcription factors such as Ikaros or EBF1. Therefore, a more comprehensive analysis of our data, in combination with

more genomic data, should shed some light on the effects of H3K36ac in genome regulation.

The B-cell developmental defect caused by the loss of SIRT7 eventually contracted the peripheral B-cell compartment and directly implicated SIRT7 in B-cell immunity. Although we did not directly test the proficiency of *Sirt7*$^{-/-}$ mice to confront a pathogenic infection, the humoral response was largely impaired in these mice, that displayed very low levels of antigen-specific IgM, IgG1 and IgG3 immunoglobulins upon immunization, indicating that SIRT7-deficient mice failed to mount a sufficient B-cell response. Intriguingly, the effect of SIRT7 loss in B-cell subsets was quite heterogeneous. This was especially clear in marginal zone B-cells, that tolerated better the loss of SIRT7 than other subsets did. Notably, similar differences have been observed in other models. For example, a recent study exploring the specific functions of the different domains of PAX5 found that the deletion of each PAX5 domain produced largely different effects in late B-cell development, which in some cases resulted even in increased absolute numbers of marginal zone B-cells, together with a reduction of follicular B-cells[381]. Given the shared developmental origin of marginal zone and follicular B-cells, these differences are most probably produced at the level of transitional B-cells, their closest upstream precursors. One of the main determinants of the marginal-zone-versus-follicular B-cell fate choice is the strength of BCR signal, that promotes the marginal zone B-cell fate or the follicular B-cell fate when its signal is weak or relatively strong, respectively[384]. Therefore, SIRT7 and PAX5 may play, together or independently, a role in regulating the cellular signalling downstream of BCR activation.

Thus, although our focus here was on the earliest stages of B-cell development, it is likely that SIRT7 also participates in specific aspects of later B-cell subsets. The most surprising phenotype that we observed was that of germinal center B-cells, whose relative numbers were increased by almost two-fold in *Sirt7*$^{-/-}$ mice. This might reflect enhanced susceptibility to infection, increased clonal proliferation after T-cell-mediated activation or, perhaps more probably, a second developmental arrest in terminal B-cell maturation. In this sense, it should be mentioned that, whereas we normally observed significant reductions in spleen cellularity and volume in *Sirt7*$^{-/-}$ mice, approximately 25% of these mice paradoxically experienced splenomegaly (data not shown). Although we did not investigate this in detail, we did not find any evidence of malignant transformation in these spleens, since we never observed any expansion of cells with B-lymphoid traits. Instead, these enlarged spleens contained abundant dead cells, which may reflect a chronic

inflammatory state and could be related to the expansion of germinal center cells that we observed. However, these observations clearly require further examination to understand the underlying phenotypes. In any case, SIRT7 seems to play a complex role in B-cells and their progenitors, and more research should clarify in more detail the stage-specific functions that it plays.

Interestingly, our characterisation of *Sirt7*^{N189Q} mutant mice (in which SIRT7 preserves its deacetylase activity but lacks mADPRT activity) revealed that the mADPRT activity of SIRT7 was not required for normal B-cell development, since these mice had normal numbers of B-cells and their progenitors. This was further supported by the observation that the deacetylase-dead SIRT7^{H187Y} mutant completely failed to rescue B-cell development upon introduction to *Sirt7*$^{-/-}$ cells, which attributed the main functions of SIRT7 in this context to its deacetylase activity. However, there are a few reasons to believe that this observation does not completely exclude the possibility that SIRT7-mediated mono-ADP-Ribosylation operates in specific aspects of haematopoiesis or B-cell immunity. First, the vast majority of our mouse work was performed in naïve, unchallenged mice, whereas SIRT7 mADPRT activity is strongly boosted under stress conditions[263]. It would be worth testing whether SIRT7-mediated mono-ADP-Ribosylation is of particular relevance under specific organismal insults such as nutrient starvation, infections or in the context of emergency haematopoiesis. Second, given the crucial importance of sirtuins during ageing, it is also possible that the mADPRT activity of SIRT7 specifically contributes to B-cell development or immunity in older mice. Indeed, unpublished RNA-Seq data generated by our lab has shown a potential protective effect of the N189Q point mutation in the liver. In particular, bulk RNA-Seq analysis of the livers of *Sirt7*^{N189Q} mice revealed a strong downregulation of genes associated with the adaptive immune response in these mice, suggesting reduced age-associated inflammation. Third, a comprehensive characterization of peripheral B-cell subsets would be required to properly explore the functions of SIRT7 mADPRT activity in B-cell immunity, which could provide information to explain some of the phenotypes that we observed in these subsets, such as the germinal center B-cell expansion. Finally, our studies only focused on B-cell development, whereas the potential functions of SIRT7 in other haematopoietic lineages remained unexplored. Thus, it will be of great interest to investigate in a broader manner the specific HDAC- and mADPRT-dependent functions that SIRT7 might play in haematopoietic progenitors.

3. A dynamic switch for a dynamic pathway

Our findings concerning the SIRT7/PAX5 interplay were of particular interest given the enormous functional relevance of PAX5 and the considerable lack of knowledge on the molecular mechanisms that regulate its protein dynamics. In addition to several lines of evidence implicating SIRT7 in different molecular aspects of B-cell development, we have shown that its main functions in B-cell development were mediated by its molecular interplay with PAX5. This was clearly shown by the experiment in **Fig. R26c**, in which the reintroduction of either PAX5^{Wt} or deacetylated PAX5^{K198R} completely rescued the B-cell developmental defect exhibited by *Sirt7*$^{-/-}$ pro-B cells, attributing the major roles of SIRT7 in these cells to its participation in the dynamic acetylation switch governing PAX5 function.

At a first glance, deacetylated PAX5 appeared to be the active form of PAX5, since it occupied most of the canonical PAX5 binding sites and regulated most of its target genes more efficiently than the *Wt* form did. Since SIRT7 catalyses K198 deacetylation and its deficiency impaired B-cell development and negatively affected transcriptional regulation, we initially hypothesised that PAX5 deacetylation would be a constitutive mechanism to enhance its activity and that it would be sufficient to drive B-cell development. In the same line, the observation that K198 deacetylation by SIRT7 enhanced PAX5 protein stability originally led us to think that the function of SIRT7 was to simply promote a sufficient PAX5 dosage. However, we were surprised to find that deacetylated PAX5^{K198R} completely failed to rescue B-cell development, which rejected our hypothesis and suggested a more complex mechanism. This conundrum was solved by the proposal of a dynamic deacetylation model that we experimentally validated and that indicated that not only deacetylated PAX5 but also acetylated PAX5 were biologically active.

While acetylated PAX5 was clearly required for B-cell development, the specific mechanisms through which it promotes B-cell development remain a mystery. We observed that PAX5^{K198Q} regulated to near-physiological levels several genes that deacetylated PAX5^{K198R} failed to regulate, especially genes related to V(D)J recombination and cell proliferation. This may partly explain the relevance of K198 acetylation, but it should be stressed that, even if acetylated PAX5 was proficient in regulating these genes, its genomic occupancy was profoundly reduced genome-wide. Therefore, acetylated PAX5 may indirectly regulate its target genes through mechanisms independent of DNA binding. Furthermore, acetylated PAX5 displayed a very short half-life, suggesting that it is physiologically underrepresented compared to the deacetylated form, which raises the

question of how both forms coordinate B-cell biology. One possibility is that PAX5 deacetylation is spatially regulated. Thus, pools of acetylated and deacetylated PAX5 may work at the same time in different nuclear regions to orchestrate different biological processes independently (**Fig. D2a**). The other possibly is that PAX5 deacetylation is regulated in time. In this case, different cellular states may be associated with different PAX5 acetylation status. For instance, our data suggested that deacetylated PAX5^{K198R} strongly promoted the expression of genes associated to V(D)J recombination, whereas acetylated PAX5^{K198Q} was more associated with proliferation programmes, which may reflect that PAX5 undergoes waves of acetylation and deacetylation at specific moments of development (**Fig. D2b**). An interesting way to address this issue would be to develop specific antibodies against the acetylated and the deacetylated forms, which should determine the relative abundance of both forms in different nuclear regions and progenitor stages.

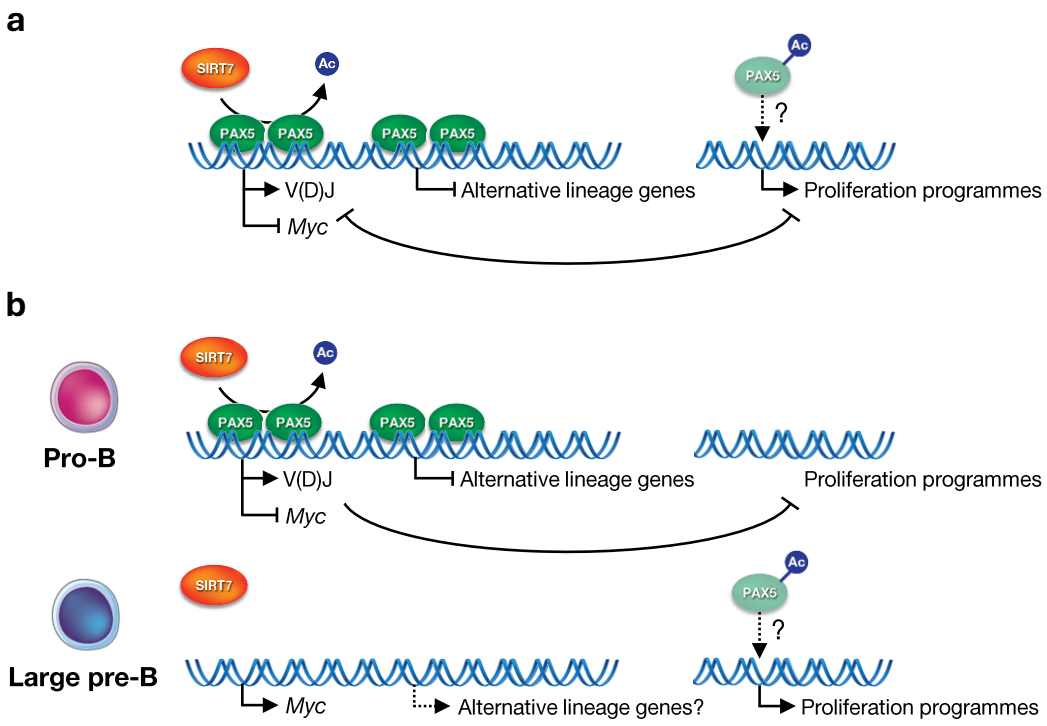


Fig. D2 | Tentative models for the spatial or temporal regulation of PAX5 acetylation
a, Pools of deacetylated and acetylated PAX5 simultaneously control distinct gene expression programmes in B-cell progenitors. In this model, deacetylated PAX5 represses lineage-inappropriate genes and the *Myc* locus while promoting V(D)J recombination. In contrast, acetylated PAX5 participates in the activation of proliferation programmes, presumably through an indirect mechanism. In parallel, V(D)J and proliferation programmes are mutually repressed by the concerted action of surface receptors. **b,** In pro-B cells, SIRT7 deacetylates PAX5 to facilitate *Igh* recombination, *Myc* repression and lineage commitment. Upon differentiation into large pre-B cells, PAX5 acetylation prevents PAX5 binding to the *Myc* locus, among others, to reactivate its expression, while contributing indirectly to the activation of proliferation programmes. PAX5 acetylation may weaken the repression of alternative lineage genes in large pre-B cells.

The identification of PCAF as the major acetyltransferase responsible for K198 acetylation provided independent support for the relevance of acetylated PAX5 in developing B-cells. PCAF has been shown to promote early B-cell development, since the combined knock-out of *Kat2b* (the gene encoding for PCAF) and its close homologue *Kat2a* (encoding for GCN5) synergistically impaired this process at the pre-B cell stage[373]. Indeed, the B-cell developmental phenotype exhibited by *Kat2a*$^{-/-}$*Kat2b*$^{-/-}$ mice notably resembled that of our *Sirt7*$^{-/-}$ mice. While the authors did not explore the mechanism, our finding that acetylated PAX5 was also required for normal B-cell differentiation may provide an explanation for this observation. In this sense, it is interesting to highlight that an interplay between SIRT7 and PCAF has been described in HCT116 colon cancer cells. Upon glucose starvation, SIRT7 deacetylates PCAF to facilitate its binding to the E3 ubiquitin ligase MDM2, leading to MDM2 degradation, upregulation of its target p53 and cell cycle arrest[302]. Here, we failed to identify any E3 ligase or DUB that robustly regulated PAX5 ubiquitination and proteasomal degradation. Although MDM2 was not one of the candidates revealed by our screening, its interplay with SIRT7 and PCAF suggests that it may be the E3 ligase catalysing PAX5 ubiquitination. Notably, one of our candidates, MDM4, is known to possess very weak E3 ligase activity and to perform its functions by binding to MDM2 and enhancing its activity[385]. Thus, one possibility is that MDM4 indirectly regulates PAX5 turnover by promoting MDM2-mediated PAX5 ubiquitination, which may also be subjected to further regulation by SIRT7 and PCAF. The mechanisms controlling PAX5 degradation may therefore be more complex than we originally thought.

4. PAX5^{K198} acetylation beyond gene expression regulation

Another matter of considerable interest would be to explore the effect of K198 acetylation in the three-dimensional structure of PAX5. Unfortunately, the whole crystal structure of PAX5 has not been resolved to this date, which prevented us from performing structural studies. However, the machine learning method AlphaFold provides protein structure prediction with high accuracy[386]. Indeed, its prediction of PAX5 structure fits well with the only PAX5 domain whose structure has been resolved, the paired domain, suggesting that it properly predicts the whole PAX5 structure[387]. According to AlphaFold, PAX5 is a highly disordered protein with very few folded regions, which are mostly found in the paired domain (**Fig. D3**). K198 is located within a putative unfolded region between the homeodomain (HD) and the conserved octapeptide (OP), so it is not likely that K198 acetylation status affects PAX5 conformation. It might however affect the ability of PAX5 to

interact with some of its transcriptional regulator partners. In particular, the HD domain of PAX5 interacts with the TFIID transcription factor complex, whereas the OP interacts with Groucho family corepressors[388,389]. Therefore, K198 acetylation may affect these interactions and thereby PAX5-mediated transcriptional regulation.

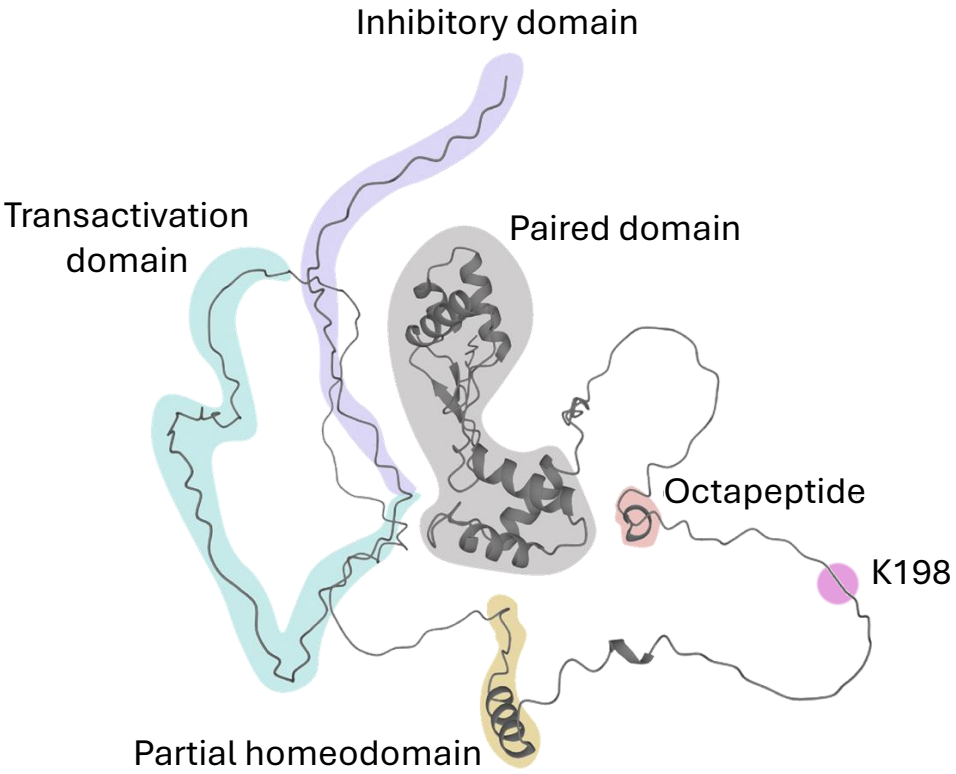


Fig. D3 | AlphaFold prediction of mouse PAX5 three-dimensional structure.
The functional domains of PAX5, as well as K198, are shaded and indicated. Adapted from Gruenbacher *et al*. (2023)[381].

Since PAX5 is a putative disordered protein and works in the context of a large network involving many other transcription factors and chromatin regulators, it is likely that PAX5 and its partners work in the context of phase separation condensates. Indeed, EBF1 is known to undergo phase separation together with BRG1 to mediate chromatin remodelling[110]. Phase separation or LLPS has emerged as a crucial mechanism of transcriptional control in which specific proteins undergo liquid-liquid partitioning to form membrane-less *organelles* or *transcriptional condensates* that assemble and disassemble within a timescale of seconds to minutes[390]. In these condensates, many transcriptional regulators are gathered to control several key steps of transcription, from gene activation or repression to transcriptional elongation and mRNA splicing. LLPS droplets are typically formed by weak multivalent interactions between intrinsically disordered regions (IDR)[391]. In the case of PAX5, K198 seems to lie in one of these regions,

suggesting that it may be involved in the ability of PAX5 to undergo LLPS. Although the possibility that PAX5 forms part of these kind of condensates has not been explored to this date, it would be worth testing whether this is the case, and whether K198 affects this phenomenon. In this sense, SIRT1-mediated deacetylation of IRF3 and IRF7 is essential for their ability to undergo LLPS during viral infection, which drives the consequent transcriptional response[312]. Further research should determine whether SIRT7 similarly regulates PAX5 LLPS under specific conditions or in a broader manner.

As introduced above, a recent report by the Busslinger lab elegantly analysed the *in vivo* functions of the different domains of PAX5 in early B-lymphopoiesis and in gene expression regulation[381]. In their work, the authors individually deleted several PAX5 functional domains by directly targeting the endogenous *Pax5* gene. They found that the C-terminal domain (that comprised the transactivation (TAD) and the inhibitory (ID) canonical PAX5 domains) were fundamental for B-cell development and for the regulation of PAX5 target genes, whereas the OP and HD domains independently regulated gene expression and also contributed to development. Since the region containing K198 does not belong to any of the functional domains of PAX5, it was not investigated in this report. Given its proximity to the OP and HD domains, it would be conceivable that K198 regulation somehow collaborates with these regions. However, in contrast with the global effect of K198 acetylation in gene expression, the deletion of the OP and HD domains affected a relatively low number of PAX5 target genes and led to a phenotype that was very different to the one that we observed. For instance, OP and HD independent deletion led to an accumulation and a reduction of pro-B cells, respectively, whereas SIRT7 deletion did not affect the number of pro-B cells. Furthermore, the deletion of these regions also increased the numbers of marginal zone B-cells, which was not observed in our mice. Our phenotype was also different to the one that the authors observed upon deletion of the C-terminal domain, that dramatically reduced the numbers of pre-B cells and all downstream stages. This was expectable, since the C-terminal domain of PAX5 seems to be very distant to K198, which discards any proximity-related interaction (see **Fig. D3**). However, these remarkable differences are puzzling at this moment, and further investigation will be needed to clarify how K198 acetylation relates to the independent functions of PAX5 domains.

5. A single residue for an alternative fate?

One of the critical functions that PAX5 plays in developing and mature B-cells is to abrogate lineage plasticity, which was originally thought to be an irreversible process. B-cell conversion into alternative fates has been previously observed in several systems, based on the deletion of PAX5[3,117] and EBF1[38,114], reflecting their key relevance in this process, or on the ectopic expression of alternative lineage transcription factors, such as C/EBPα[6] and, more recently, HOXB5[392]. These seminal reports challenged the classical dogma that haematopoiesis is a unidirectional process, and strongly suggested the intriguing possibility that lineage conversion may take place under physiological conditions. However, both the silencing of PAX5 or EBF1 and the activation of C/EBPα and HOXB5 expression are unlikely to happen under physiological conditions, so no mechanisms that could disrupt lineage commitment in normal progenitors had been identified so far. Importantly, our finding that the post-translational modification of a single PAX5 residue is sufficient to disrupt lineage commitment may represent a mechanism for lineage commitment reversal during normal differentiation.

According to our data, while deacetylated PAX5^{K198R} was sufficient to block T-cell lineage potential, acetylated PAX5^{K198Q} completely failed to repress lineage inappropriate genes and to commit pro-B cells to the B-cell lineage. This suggests a provocative scenario in which PAX5 acetylation may transiently facilitate lineage conversion in normal B-cell progenitors. In this sense, an elegant lineage tracing system recently showed for the first time that pro-B cells can develop into macrophages during normal differentiation, in a process that was enhanced by inflammation[4]. Furthermore, another study identified that, in the context of cancer immunosurveillance, tumour-secreted M-CSF (macrophage colony-stimulating factor) stimulated pre-B cell transdifferentiation into macrophages[5]. Importantly, pre-B cells that responded to M-CSF expressed low levels of PAX5, suggesting that this cytokine promotes transdifferentiation by downregulating PAX5, although the mechanisms were not detailed. Given the key role of sirtuins in inflammation and in the coordination of the cellular stress response, we speculate that extracellular clues might target the stabilization of PAX5 by SIRT7 to facilitate lineage conversion.

6. A therapeutic opportunity?

From a more translational perspective, the developmental defect imposed by SIRT7 deficiency may predispose for malignant transformation. Although we did not formally show that SIRT7 prevents leukaemic transformation in B-cell progenitors, SIRT7 behaved as a tumour suppressor in cell lines, and its high expression was associated with better prognosis in patients. Given the fundamental role that SIRT7 plays in genome integrity and DNA repair, it is conceivable that its activity contributes to prevent malignancy. More importantly, our data strongly suggests that the interplay between SIRT7 and PAX5 is largely conserved across B-lymphoblastoid cells, including in B-ALL and CLL. While no SIRT7 mutations have been reported in B-ALL, the *PAX5* gene is frequently targeted by monoallelic inactivating mutations[141,145,146]. PAX5 plays a crucial tumour suppressor function by limiting glucose metabolism to levels that are insufficient for malignant transformation[125]. In B-ALL, PAX5 downregulation relieves this metabolic constraint, which critically contributes to sustain the energetic demands of leukaemic cells. Accordingly, reconstitution of PAX5 physiological levels imposes an energetic crisis that leads to leukaemic cell death[125]. These observations open an attractive window to pharmacologically enhance the tumour suppressor functions of PAX5 in B-ALL. However, the lack of knowledge on the upstream regulators of its protein dynamics have impeded the possibility of exploring such strategy. Here, we have described for the first time a molecular mechanism controlling PAX5 stability in B-cell progenitors, which may provide a novel rationale to indirectly target PAX5 in B-ALL. Since sirtuins have the unique property of being pharmacologically actionable[393], developing specific SIRT7 agonists to enhance SIRT7 and PAX5 functions may provide a novel strategy to tackle human B-ALL.

VIII. CONCLUSIONS

1. SIRT7 is upregulated during early B-cell differentiation in mice.
2. SIRT7 promotes B-cell development, commitment and immunity through a pre-B cell-intrinsic mechanism that depends on its deacetylase activity.
3. SIRT7 is required for pre-B cell survival and proliferation.
4. Distal *Igh* recombination is impaired in mice lacking SIRT7 but does not explain the B-cell developmental defect exhibited by these mice.
5. Transcriptional repression during the pro-B-to-pre-B cell transition depends on SIRT7 and is partly mediated by its collaboration with PAX5.
6. SIRT7 regulates pro-B cell epigenetics by preferentially deacetylating H3K36ac at PAX5 binding sites.
7. SIRT7 sustains PAX5 protein levels and stability through deacetylation of $PAX5^{K198}$.
8. SIRT7 interacts with several components of the transcription factor network that drives B-cell development and with various chromatin remodeller, histone modifiers and transcription factor complexes.
9. Deacetylation of $PAX5^{K198}$ globally enhances its genomic occupancy and the control of its target genes, but also impairs its binding to specific loci and the regulation of some of its targets. $PAX5^{K198}$ is acetylated by PCAF, and this event reduces its protein stability, genomic occupancy and ability to regulate gene expression, while facilitating the control of a particular subset of its target genes.
10. B-cell lineage commitment requires $PAX5^{K198}$ deacetylation, whereas B-cell differentiation depends on the dynamic acetylation and deacetylation of $PAX5^{K198}$.
11. $PAX5^{K198}$ deacetylation by SIRT7 mediates the main functions of SIRT7 in early B-cell development.
12. The PAX5/SIRT7 interplay is conserved in human B-ALL cell lines and patients and CLL cell lines. In B-ALL, SIRT7 is a tumour suppressor, and its high RNA expression is a good prognosis marker, especially in patients with *PAX5* deletions.

IX. PUBLISHED ARTICLES

1. **Gamez-Garcia, A.**, Espinosa-Alcantud, M., Bueno-Costa, A., Alari-Pahissa, E., Marazuela-Duque, A., Thackray, J.K., Ray, C., Berenguer, C., Kumari, P., Bech, J.J., Braun, T., Ianni, A., Tischfield, J.A., Serrano, L., Esteller, M., Sardina, J.L., de la Torre, C., Sigvardsson, M., Vazquez, B.N. & Vaquero, A. A SIRT7-dependent acetylation switch regulates early B-cell differentiation and lineage commitment through Pax5**.** ***Nature Immunology (in press)*****.**

2. Kumari, P., Tarighi, S., Fuchshuber, E., Li, L., Fernandez-Duran, I., Wang, M., Ayoson, J., Castello-Garcia, J.M., **Gamez-Garcia, A.**, Espinosa Alcantud, M., Sreenivasan, K., Guenther, S., Olivella, M., Savai, R., Yue, S., Vaquero, A., Braun, T., & Ianni, A. SIRT7 promotes lung cancer progression by destabilizing the tumor suppressor ARF. ***Proc Natl Acad Sci USA*** **121,** e2409269121 (2024). https://doi.org/10.1073/pnas.2409269121

3. Rasti, G., Becker, M., Vazquez, B.N., Espinosa-Alcantud, M., Fernández-Duran, I., **Gamez-Garcia, A.**, Ianni, A., Gonzalez, J., Bosch-Presegue, L., Marazuela-Duque, A., Guitart-Solanes, A., Segura-Bayona, S., Bech-Serra J.J., Scher, M., Serrano, L., Shankavaram, U., Erdjument-Bromage, H., Tempst, P., Reinberg, D., Olivella, M., Stracker, T.H., de la Torre, C. & Vaquero, A. SIRT1 regulates DNA damage signaling through the PP4 phosphatase complex. ***Nucleic Acids Res*** **51**, 6754-6769 (2023). https://doi.org/10.1093/nar/gkad504

4. Gamez-Garcia, A. & Vazquez, B.N. Nuclear Sirtuins and the Aging of the Immune System. ***Genes (Basel)*** **12** (2021). https://doi.org/10.3390/genes12121856

X. REFERENCES

1. Pinho, S. & Frenette, P.S. Haematopoietic stem cell activity and interactions with the niche. *Nat Rev Mol Cell Biol* **20**, 303-320 (2019). https://doi.org/10.1038/s41580-019-0103-9
2. Passegue, E. *et al.* Normal and leukemic hematopoiesis: are leukemias a stem cell disorder or a reacquisition of stem cell characteristics? *Proc Natl Acad Sci U S A* **100 Suppl 1**, 11842-11849 (2003). https://doi.org/10.1073/pnas.2034201100
3. Cobaleda, C. *et al.* Conversion of mature B cells into T cells by dedifferentiation to uncommitted progenitors. *Nature* **449**, 473-477 (2007). https://doi.org/10.1038/nature06159
4. Audzevich, T. *et al.* Pre/pro-B cells generate macrophage populations during homeostasis and inflammation. *Proc Natl Acad Sci U S A* **114**, E3954-E3963 (2017). https://doi.org/10.1073/pnas.1616417114
5. Chen, C. *et al.* Cancer co-opts differentiation of B-cell precursors into macrophage-like cells. *Nat Commun* **13**, 5376 (2022). https://doi.org/10.1038/s41467-022-33117-y
6. Xie, H. *et al.* Stepwise reprogramming of B cells into macrophages. *Cell* **117**, 663-676 (2004). https://doi.org/10.1016/s0092-8674(04)00419-2
7. Stass, S. *et al.* Lineage switch in acute leukemia. *Blood* **64**, 701-706 (1984).
8. Dorantes-Acosta, E. & Pelayo, R. Lineage switching in acute leukemias: a consequence of stem cell plasticity? *Bone Marrow Res* **2012**, 406796 (2012). https://doi.org/10.1155/2012/406796
9. Perna, F. & Sadelain, M. Myeloid leukemia switch as immune escape from CD19 chimeric antigen receptor (CAR) therapy. *Transl Cancer Res* **5**, S221-S225 (2016). https://doi.org/10.21037/tcr.2016.08.15
10. Velten, L. *et al.* Human haematopoietic stem cell lineage commitment is a continuous process. *Nat Cell Biol* **19**, 271-281 (2017). https://doi.org/10.1038/ncb3493
11. Naik, S.H. *et al.* Diverse and heritable lineage imprinting of early haematopoietic progenitors. *Nature* **496**, 229-232 (2013). https://doi.org/10.1038/nature12013
12. Wilson, N.K. *et al.* Combined Single-Cell Functional and Gene Expression Analysis Resolves Heterogeneity within Stem Cell Populations. *Cell Stem Cell* **16**, 712-724 (2015). https://doi.org/10.1016/j.stem.2015.04.004
13. Paul, F. *et al.* Transcriptional Heterogeneity and Lineage Commitment in Myeloid Progenitors. *Cell* **163**, 1663-1677 (2015). https://doi.org/10.1016/j.cell.2015.11.013
14. Cheng, H. *et al.* New paradigms on hematopoietic stem cell differentiation. *Protein Cell* **11**, 34-44 (2020). https://doi.org/10.1007/s13238-019-0633-0
15. Moreno, M. & Wiegand, T.J. Blood. In: Wexler, P. (ed). *Encyclopedia of Toxicology (Third Edition)*. Academic Press: Oxford, 2014, pp 526-532. https://doi.org/https://doi.org/10.1016/B978-0-12-386454-3.00683-7
16. Swann, J.W. *et al.* Made to order: emergency myelopoiesis and demand-adapted innate immune cell production. *Nat Rev Immunol* **24**, 596-613 (2024). https://doi.org/10.1038/s41577-024-00998-7
17. Li, D. & Wu, M. Pattern recognition receptors in health and diseases. *Signal Transduct Target Ther* **6**, 291 (2021). https://doi.org/10.1038/s41392-021-00687-0
18. Nutt, S.L. *et al.* Dynamic regulation of PU.1 expression in multipotent hematopoietic progenitors. *J Exp Med* **201**, 221-231 (2005). https://doi.org/10.1084/jem.20041535
19. Back, J. *et al.* Visualizing PU.1 activity during hematopoiesis. *Exp Hematol* **33**, 395-402 (2005). https://doi.org/10.1016/j.exphem.2004.12.010

20. DeKoter, R.P. & Singh, H. Regulation of B lymphocyte and macrophage development by graded expression of PU.1. *Science* **288**, 1439-1441 (2000) . https://doi.org/10.1126/science.288.5470.1439
21. Scott, E.W. *et al*. Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages. *Science* **265**, 1573-1577 (1994). https://doi.org/10.1126/science.8079170
22. Iwasaki, H. *et al*. Distinctive and indispensable roles of PU.1 in maintenance of hematopoietic stem cells and their differentiation. *Blood* **106**, 1590-1600 (2005). https://doi.org/10.1182/blood-2005-03-0860
23. Dakic, A. *et al*. PU.1 regulates the commitment of adult hematopoietic progenitors and restricts granulopoiesis. *J Exp Med* **201**, 1487-1502 (2005). https://doi.org/10.1084/jem.20050075
24. Zhang, P. *et al*. Enhancement of hematopoietic stem cell repopulating capacity and self-renewal in the absence of the transcription factor C/EBP alpha. *Immunity* **21**, 853-863 (2004). https://doi.org/10.1016/j.immuni.2004.11.006
25. Zhang, D.E. *et al*. Absence of granulocyte colony-stimulating factor signaling and neutrophil development in CCAAT enhancer binding protein alpha-deficient mice. *Proc Natl Acad Sci U S A* **94**, 569-574 (1997). https://doi.org/10.1073/pnas.94.2.569
26. Dahl, R. *et al*. Regulation of macrophage and neutrophil cell fates by the PU.1:C/EBPalpha ratio and granulocyte colony-stimulating factor. *Nat Immunol* **4**, 1029-1036 (2003). https://doi.org/10.1038/ni973
27. Tamura, T. *et al*. ICSBP directs bipotential myeloid progenitor cells to differentiate into mature macrophages. *Immunity* **13**, 155-165 (2000) . https://doi.org/10.1016/s1074-7613(00)00016-9
28. 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). https://doi.org/10.1016/j.molcel.2010.05.004
29. Stavast, C.J. *et al*. The interplay between critical transcription factors and microRNAs in the control of normal and malignant myelopoiesis. *Cancer Lett* **427**, 28-37 (2018). https://doi.org/10.1016/j.canlet.2018.04.010
30. Nagasawa, M. *et al*. Innate Lymphoid Cells (ILCs): Cytokine Hubs Regulating Immunity and Tissue Homeostasis. *Cold Spring Harb Perspect Biol* **10** (2018). https://doi.org/10.1101/cshperspect.a030304
31. Diefenbach, A. *et al*. Development, differentiation, and diversity of innate lymphoid cells. *Immunity* **41**, 354-365 (2014). https://doi.org/10.1016/j.immuni.2014.09.005
32. Samson, S.I. *et al*. GATA-3 promotes maturation, IFN-gamma production, and liver-specific homing of NK cells. *Immunity* **19**, 701-711 (2003). https://doi.org/10.1016/s1074-7613(03)00294-2
33. Yagi, R. *et al*. The transcription factor GATA3 is critical for the development of all IL-7Ralpha-expressing innate lymphoid cells. *Immunity* **40**, 378-388 (2014). https://doi.org/10.1016/j.immuni.2014.01.012
34. Klose, C.S.N. *et al*. Differentiation of type 1 ILCs from a common progenitor to all helper-like innate lymphoid cell lineages. *Cell* **157**, 340-356 (2014). https://doi.org/10.1016/j.cell.2014.03.030
35. Bain, G. *et al*. E2A deficiency leads to abnormalities in alphabeta T-cell development and to rapid development of T-cell lymphomas. *Mol Cell Biol* **17**, 4782-4791 (1997). https://doi.org/10.1128/MCB.17.8.4782
36. Kee, B.L. E and ID proteins branch out. *Nat Rev Immunol* **9**, 175-184 (2009). https://doi.org/10.1038/nri2507

37. Roessler, S. *et al.* Distinct promoters mediate the regulation of Ebf1 gene expression by interleukin-7 and Pax5. *Mol Cell Biol* **27**, 579-594 (2007). https://doi.org/10.1128/MCB.01192-06
38. Nechanitzky, R. *et al.* Transcription factor EBF1 is essential for the maintenance of B cell identity and prevention of alternative fates in committed cells. *Nat Immunol* **14**, 867-875 (2013). https://doi.org/10.1038/ni.2641
39. Yang, C.Y. *et al.* The transcriptional regulators Id2 and Id3 control the formation of distinct memory CD8+ T cell subsets. *Nat Immunol* **12**, 1221-1229 (2011). https://doi.org/10.1038/ni.2158
40. Adolfsson, J. *et al.* Identification of Flt3+ lympho-myeloid stem cells lacking erythro-megakaryocytic potential a revised road map for adult blood lineage commitment. *Cell* **121**, 295-306 (2005). https://doi.org/10.1016/j.cell.2005.02.013
41. Inlay, M.A. *et al.* Ly6d marks the earliest stage of B-cell specification and identifies the branchpoint between B-cell and T-cell development. *Genes Dev* **23**, 2376-2381 (2009). https://doi.org/10.1101/gad.1836009
42. Pui, J.C. *et al.* Notch1 expression in early lymphopoiesis influences B versus T lineage determination. *Immunity* **11**, 299-308 (1999). https://doi.org/10.1016/s1074-7613(00)80105-3
43. Radtke, F. *et al.* Deficient T cell fate specification in mice with an induced inactivation of Notch1. *Immunity* **10**, 547-558 (1999). https://doi.org/10.1016/s1074-7613(00)80054-0
44. Wilson, A. *et al.* Notch 1-deficient common lymphoid precursors adopt a B cell fate in the thymus. *J Exp Med* **194**, 1003-1012 (2001). https://doi.org/10.1084/jem.194.7.1003
45. Rothenberg, E.V. *et al.* Launching the T-cell-lineage developmental programme. *Nat Rev Immunol* **8**, 9-21 (2008). https://doi.org/10.1038/nri2232
46. Hosokawa, H. & Rothenberg, E.V. How transcription factors drive choice of the T cell fate. *Nat Rev Immunol* **21**, 162-176 (2021). https://doi.org/10.1038/s41577-020-00426-6
47. Weber, B.N. *et al.* A critical role for TCF-1 in T-lineage specification and differentiation. *Nature* **476**, 63-68 (2011). https://doi.org/10.1038/nature10279
48. Germar, K. *et al.* T-cell factor 1 is a gatekeeper for T-cell specification in response to Notch signaling. *Proc Natl Acad Sci U S A* **108**, 20060-20065 (2011). https://doi.org/10.1073/pnas.1110230108
49. Hosoya, T. *et al.* GATA-3 is required for early T lineage progenitor development. *J Exp Med* **206**, 2987-3000 (2009). https://doi.org/10.1084/jem.20090934
50. Johnson, J.L. *et al.* Lineage-Determining Transcription Factor TCF-1 Initiates the Epigenetic Identity of T Cells. *Immunity* **48**, 243-257 e210 (2018). https://doi.org/10.1016/j.immuni.2018.01.012
51. Garcia-Ojeda, M.E. *et al.* GATA-3 promotes T-cell specification by repressing B-cell potential in pro-T cells in mice. *Blood* **121**, 1749-1759 (2013). https://doi.org/10.1182/blood-2012-06-440065
52. Garcia-Perez, L. *et al.* Functional definition of a transcription factor hierarchy regulating T cell lineage commitment. *Sci Adv* **6**, eaaw7313 (2020). https://doi.org/10.1126/sciadv.aaw7313
53. Wei, G. *et al.* Genome-wide analyses of transcription factor GATA3-mediated gene regulation in distinct T cell types. *Immunity* **35**, 299-311 (2011). https://doi.org/10.1016/j.immuni.2011.08.007
54. Li, P. *et al.* Reprogramming of T cells to natural killer-like cells upon Bcl11b deletion. *Science* **329**, 85-89 (2010). https://doi.org/10.1126/science.1188063
55. Ikawa, T. *et al.* An essential developmental checkpoint for production of the T cell lineage. *Science* **329**, 93-96 (2010). https://doi.org/10.1126/science.1188995

56. Li, L. *et al*. An early T cell lineage commitment checkpoint dependent on the transcription factor Bcl11b. *Science* **329**, 89-93 (2010). https://doi.org/10.1126/science.1188989
57. Hosokawa, H. *et al*. Bcl11b sets pro-T cell fate by site-specific cofactor recruitment and by repressing Id2 and Zbtb16. *Nat Immunol* **19**, 1427-1440 (2018). https://doi.org/10.1038/s41590-018-0238-4
58. Longabaugh, W.J.R. *et al*. Bcl11b and combinatorial resolution of cell fate in the T-cell gene regulatory network. *Proc Natl Acad Sci U S A* **114**, 5800-5807 (2017). https://doi.org/10.1073/pnas.1610617114
59. Zhang, J.A. *et al*. Dynamic transformations of genome-wide epigenetic marking and transcriptional control establish T cell identity. *Cell* **149**, 467-482 (2012). https://doi.org/10.1016/j.cell.2012.01.056
60. 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). https://doi.org/10.1016/s1074-7613(00)80269-1
61. Nichogiannopoulou, A. *et al*. Defects in hemopoietic stem cell activity in Ikaros mutant mice. *J Exp Med* **190**, 1201-1214 (1999). https://doi.org/10.1084/jem.190.9.1201
62. Yoshida, T. *et al*. Early hematopoietic lineage restrictions directed by Ikaros. *Nat Immunol* **7**, 382-391 (2006). https://doi.org/10.1038/ni1314
63. Ng, S.Y. *et al*. Genome-wide lineage-specific transcriptional networks underscore Ikaros-dependent lymphoid priming in hematopoietic stem cells. *Immunity* **30**, 493-507 (2009). https://doi.org/10.1016/j.immuni.2009.01.014
64. Geimer Le Lay, A.S. *et al*. The tumor suppressor Ikaros shapes the repertoire of notch target genes in T cells. *Sci Signal* **7**, ra28 (2014). https://doi.org/10.1126/scisignal.2004545
65. Oravecz, A. *et al*. Ikaros mediates gene silencing in T cells through Polycomb repressive complex 2. *Nat Commun* **6**, 8823 (2015). https://doi.org/10.1038/ncomms9823
66. Winandy, S. *et al*. A dominant mutation in the Ikaros gene leads to rapid development of leukemia and lymphoma. *Cell* **83**, 289-299 (1995). https://doi.org/10.1016/0092-8674(95)90170-1
67. Urban, J.A. & Winandy, S. Ikaros null mice display defects in T cell selection and CD4 versus CD8 lineage decisions. *J Immunol* **173**, 4470-4478 (2004). https://doi.org/10.4049/jimmunol.173.7.4470
68. Wong, L.Y. *et al*. Ikaros sets the potential for Th17 lineage gene expression through effects on chromatin state in early T cell development. *J Biol Chem* **288**, 35170-35179 (2013). https://doi.org/10.1074/jbc.M113.481440
69. Avitahl, N. *et al*. Ikaros sets thresholds for T cell activation and regulates chromosome propagation. *Immunity* **10**, 333-343 (1999) . https://doi.org/10.1016/s1074-7613(00)80033-3
70. Anderson, M.K. *et al*. Constitutive expression of PU.1 in fetal hematopoietic progenitors blocks T cell development at the pro-T cell stage. *Immunity* **16**, 285-296 (2002). https://doi.org/10.1016/s1074-7613(02)00277-7
71. Ungerback, J. *et al*. Pioneering, chromatin remodeling, and epigenetic constraint in early T-cell gene regulation by SPI1 (PU.1). *Genome Res* **28**, 1508-1519 (2018). https://doi.org/10.1101/gr.231423.117
72. Champhekar, A. *et al*. Regulation of early T-lineage gene expression and developmental progression by the progenitor cell transcription factor PU.1. *Genes Dev* **29**, 832-848 (2015). https://doi.org/10.1101/gad.259879.115

73. Seki, M. *et al*. Recurrent SPI1 (PU.1) fusions in high-risk pediatric T cell acute lymphoblastic leukemia. *Nat Genet* **49**, 1274-1281 (2017). https://doi.org/10.1038/ng.3900
74. Rosenbauer, F. *et al*. Lymphoid cell growth and transformation are suppressed by a key regulatory element of the gene encoding PU.1. *Nat Genet* **38**, 27-37 (2006). https://doi.org/10.1038/ng1679
75. Clark, M.R. *et al*. Orchestrating B cell lymphopoiesis through interplay of IL-7 receptor and pre-B cell receptor signalling. *Nat Rev Immunol* **14**, 69-80 (2014). https://doi.org/10.1038/nri3570
76. Puel, A. *et al*. Defective IL7R expression in T(-)B(+)NK(+) severe combined immunodeficiency. *Nat Genet* **20**, 394-397 (1998). https://doi.org/10.1038/3877
77. Jensen, C.T. *et al*. FLT3 ligand and not TSLP is the key regulator of IL-7-independent B-1 and B-2 B lymphopoiesis. *Blood* **112**, 2297-2304 (2008). https://doi.org/10.1182/blood-2008-04-150508
78. Yao, Z. *et al*. Stat5a/b are essential for normal lymphoid development and differentiation. *Proc Natl Acad Sci U S A* **103**, 1000-1005 (2006). https://doi.org/10.1073/pnas.0507350103
79. Mandal, M. *et al*. CXCR4 signaling directs Igk recombination and the molecular mechanisms of late B lymphopoiesis. *Nat Immunol* **20**, 1393-1403 (2019). https://doi.org/10.1038/s41590-019-0468-0
80. Herzog, S. *et al*. Regulation of B-cell proliferation and differentiation by pre-B-cell receptor signalling. *Nat Rev Immunol* **9**, 195-205 (2009) . https://doi.org/10.1038/nri2491
81. Hill, L. *et al*. Wapl repression by Pax5 promotes V gene recombination by Igh loop extrusion. *Nature* **584**, 142-147 (2020). https://doi.org/10.1038/s41586-020-2454-y
82. Fuxa, M. *et al*. Pax5 induces V-to-DJ rearrangements and locus contraction of the immunoglobulin heavy-chain gene. *Genes Dev* **18**, 411-422 (2004). https://doi.org/10.1101/gad.291504
83. Zhang, Y. *et al*. Molecular basis for differential Igk versus Igh V(D)J joining mechanisms. *Nature* **630**, 189-197 (2024). https://doi.org/10.1038/s41586-024-07477-y
84. Chi, X. *et al*. V(D)J recombination, somatic hypermutation and class switch recombination of immunoglobulins: mechanism and regulation. *Immunology* **160**, 233-247 (2020). https://doi.org/10.1111/imm.13176
85. Stavnezer, J. *et al*. Mechanism and regulation of class switch recombination. *Annu Rev Immunol* **26**, 261-292 (2008). https://doi.org/10.1146/annurev.immunol.26.021607.090248
86. Teng, G. & Papavasiliou, F.N. Immunoglobulin somatic hypermutation. *Annu Rev Genet* **41**, 107-120 (2007). https://doi.org/10.1146/annurev.genet.41.110306.130340
87. Schatz, D.G. & Ji, Y. Recombination centres and the orchestration of V(D)J recombination. *Nat Rev Immunol* **11**, 251-263 (2011). https://doi.org/10.1038/nri2941
88. McLean, K.C. & Mandal, M. It Takes Three Receptors to Raise a B Cell. *Trends Immunol* **41**, 629-642 (2020). https://doi.org/10.1016/j.it.2020.05.003
89. Georgopoulos, K. Haematopoietic cell-fate decisions, chromatin regulation and ikaros. *Nat Rev Immunol* **2**, 162-174 (2002). https://doi.org/10.1038/nri747
90. Ng, S.Y. *et al*. Ikaros and chromatin regulation in early hematopoiesis. *Curr Opin Immunol* **19**, 116-122 (2007). https://doi.org/10.1016/j.coi.2007.02.014
91. Dahl, R. *et al*. The transcriptional repressor GFI-1 antagonizes PU.1 activity through protein-protein interaction. *J Biol Chem* **282**, 6473-6483 (2007). https://doi.org/10.1074/jbc.M607613200

92. Spooner, C.J. *et al.* A recurrent network involving the transcription factors PU.1 and Gfi1 orchestrates innate and adaptive immune cell fates. *Immunity* **31**, 576-586 (2009). https://doi.org/10.1016/j.immuni.2009.07.011
93. Dias, S. *et al.* E2A proteins promote development of lymphoid-primed multipotent progenitors. *Immunity* **29**, 217-227 (2008). https://doi.org/10.1016/j.immuni.2008.05.015
94. Welinder, E. *et al.* 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). https://doi.org/10.1073/pnas.1111766108
95. Amin, R.H. & Schlissel, M.S. Foxo1 directly regulates the transcription of recombination-activating genes during B cell development. *Nat Immunol* **9**, 613-622 (2008). https://doi.org/10.1038/ni.1612
96. Mansson, R. *et al.* Positive intergenic feedback circuitry, involving EBF1 and FOXO1, orchestrates B-cell fate. *Proc Natl Acad Sci U S A* **109**, 21028-21033 (2012). https://doi.org/10.1073/pnas.1211427109
97. Zandi, S. *et al.* EBF1 is essential for B-lineage priming and establishment of a transcription factor network in common lymphoid progenitors. *J Immunol* **181**, 3364-3372 (2008). https://doi.org/10.4049/jimmunol.181.5.3364
98. Lin, H. & Grosschedl, R. Failure of B-cell differentiation in mice lacking the transcription factor EBF. *Nature* **376**, 263-267 (1995). https://doi.org/10.1038/376263a0
99. Zhang, Z. *et al.* Enforced expression of EBF in hematopoietic stem cells restricts lymphopoiesis to the B cell lineage. *EMBO J* **22**, 4759-4769 (2003). https://doi.org/10.1093/emboj/cdg464
100. Kikuchi, K. *et al.* IL-7 receptor signaling is necessary for stage transition in adult B cell development through up-regulation of EBF. *J Exp Med* **201**, 1197-1203 (2005). https://doi.org/10.1084/jem.20050158
101. Reynaud, D. *et al.* Regulation of B cell fate commitment and immunoglobulin heavy-chain gene rearrangements by Ikaros. *Nat Immunol* **9**, 927-936 (2008). https://doi.org/10.1038/ni.1626
102. Pongubala, J.M. *et al.* Transcription factor EBF restricts alternative lineage options and promotes B cell fate commitment independently of Pax5. *Nat Immunol* **9**, 203-215 (2008). https://doi.org/10.1038/ni1555
103. Medina, K.L. *et al.* Assembling a gene regulatory network for specification of the B cell fate. *Dev Cell* **7**, 607-617 (2004). https://doi.org/10.1016/j.devcel.2004.08.006
104. Sigvardsson, M. *et al.* EBF and E47 collaborate to induce expression of the endogenous immunoglobulin surrogate light chain genes. *Immunity* **7**, 25-36 (1997). https://doi.org/10.1016/s1074-7613(00)80507-5
105. Sigvardsson, M. *et al.* Early B-cell factor, E2A, and Pax-5 cooperate to activate the early B cell-specific mb-1 promoter. *Mol Cell Biol* **22**, 8539-8551 (2002). https://doi.org/10.1128/MCB.22.24.8539-8551.2002
106. Treiber, T. *et al.* Early B cell factor 1 regulates B cell gene networks by activation, repression, and transcription- independent poising of chromatin. *Immunity* **32**, 714-725 (2010). https://doi.org/10.1016/j.immuni.2010.04.013
107. Lin, Y.C. *et al.* A global network of transcription factors, involving E2A, EBF1 and Foxo1, that orchestrates B cell fate. *Nat Immunol* **11**, 635-643 (2010). https://doi.org/10.1038/ni.1891
108. Boller, S. *et al.* Pioneering Activity of the C-Terminal Domain of EBF1 Shapes the Chromatin Landscape for B Cell Programming. *Immunity* **44**, 527-541 (2016). https://doi.org/10.1016/j.immuni.2016.02.021

109. Li, R. *et al*. Dynamic EBF1 occupancy directs sequential epigenetic and transcriptional events in B-cell programming. *Genes Dev* **32**, 96-111 (2018). https://doi.org/10.1101/gad.309583.117
110. Wang, Y. *et al*. A Prion-like Domain in Transcription Factor EBF1 Promotes Phase Separation and Enables B Cell Programming of Progenitor Chromatin. *Immunity* **53**, 1151-1167 e1156 (2020). https://doi.org/10.1016/j.immuni.2020.10.009
111. Ma, F. *et al*. Three-dimensional chromatin reorganization regulates B cell development during ageing. *Nat Cell Biol* **26**, 991-1002 (2024). https://doi.org/10.1038/s41556-024-01424-9
112. Decker, T. *et al*. Stepwise activation of enhancer and promoter regions of the B cell commitment gene Pax5 in early lymphopoiesis. *Immunity* **30**, 508-520 (2009). https://doi.org/10.1016/j.immuni.2009.01.012
113. Medvedovic, J. *et al*. Pax5: a master regulator of B cell development and leukemogenesis. *Adv Immunol* **111**, 179-206 (2011). https://doi.org/10.1016/B978-0-12-385991-4.00005-2
114. Ungerback, J. *et al*. Combined heterozygous loss of Ebf1 and Pax5 allows for T-lineage conversion of B cell progenitors. *J Exp Med* **212**, 1109-1123 (2015). https://doi.org/10.1084/jem.20132100
115. Nutt, S.L. *et al*. Essential functions of Pax5 (BSAP) in pro-B cell development: difference between fetal and adult B lymphopoiesis and reduced V-to-DJ recombination at the IgH locus. *Genes Dev* **11**, 476-491 (1997). https://doi.org/10.1101/gad.11.4.476
116. Urbanek, P. *et al*. Complete block of early B cell differentiation and altered patterning of the posterior midbrain in mice lacking Pax5/BSAP. *Cell* **79**, 901-912 (1994). https://doi.org/10.1016/0092-8674(94)90079-5
117. Nutt, S.L. *et al*. Commitment to the B-lymphoid lineage depends on the transcription factor Pax5. *Nature* **401**, 556-562 (1999). https://doi.org/10.1038/44076
118. McManus, S. *et al*. The transcription factor Pax5 regulates its target genes by recruiting chromatin-modifying proteins in committed B cells. *EMBO J* **30**, 2388-2404 (2011). https://doi.org/10.1038/emboj.2011.140
119. Revilla, I.D.R. *et al*. The B-cell identity factor Pax5 regulates distinct transcriptional programmes in early and late B lymphopoiesis. *EMBO J* **31**, 3130-3146 (2012). https://doi.org/10.1038/emboj.2012.155
120. Johanson, T.M. *et al*. Transcription-factor-mediated supervision of global genome architecture maintains B cell identity. *Nat Immunol* **19**, 1257-1264 (2018). https://doi.org/10.1038/s41590-018-0234-8
121. Somasundaram, R. *et al*. EBF1 and PAX5 control pro-B cell expansion via opposing regulation of the Myc gene. *Blood* **137**, 3037-3049 (2021). https://doi.org/10.1182/blood.2020009564
122. Medvedovic, J. *et al*. Flexible long-range loops in the VH gene region of the Igh locus facilitate the generation of a diverse antibody repertoire. *Immunity* **39**, 229-244 (2013). https://doi.org/10.1016/j.immuni.2013.08.011
123. Schebesta, M. *et al*. Control of pre-BCR signaling by Pax5-dependent activation of the BLNK gene. *Immunity* **17**, 473-485 (2002). https://doi.org/10.1016/s1074-7613(02)00418-1
124. Ramamoorthy, S. *et al*. EBF1 and Pax5 safeguard leukemic transformation by limiting IL-7 signaling, Myc expression, and folate metabolism. *Genes Dev* **34**, 1503-1519 (2020). https://doi.org/10.1101/gad.340216.120
125. Chan, L.N. *et al*. Metabolic gatekeeper function of B-lymphoid transcription factors. *Nature* **542**, 479-483 (2017). https://doi.org/10.1038/nature21076

126. Tang, L. *et al*. Immunotherapy in hematologic malignancies: achievements, challenges and future prospects. *Signal Transduct Target Ther* **8**, 306 (2023). https://doi.org/10.1038/s41392-023-01521-5
127. Alexander, T.B. & Mullighan, C.G. Molecular Biology of Childhood Leukemia. *Annual Review of Cancer Biology* **5**, 95-117 (2021). https://doi.org/https://doi.org/10.1146/annurev-cancerbio-043020-110055
128. Dohner, H. *et al*. Acute Myeloid Leukemia. *N Engl J Med* **373**, 1136-1152 (2015). https://doi.org/10.1056/NEJMra1406184
129. Schafer, D. *et al*. Five percent of healthy newborns have an ETV6-RUNX1 fusion as revealed by DNA-based GIPFEL screening. *Blood* **131**, 821-826 (2018). https://doi.org/10.1182/blood-2017-09-808402
130. Martin-Lorenzo, A. *et al*. Infection Exposure is a Causal Factor in B-cell Precursor Acute Lymphoblastic Leukemia as a Result of Pax5-Inherited Susceptibility. *Cancer Discov* **5**, 1328-1343 (2015). https://doi.org/10.1158/2159-8290.CD-15-0892
131. Rodriguez-Hernandez, G. *et al*. Infection Exposure Promotes ETV6-RUNX1 Precursor B-cell Leukemia via Impaired H3K4 Demethylases. *Cancer Res* **77**, 4365-4377 (2017). https://doi.org/10.1158/0008-5472.CAN-17-0701
132. Braun, T.P. *et al*. Response and Resistance to BCR-ABL1-Targeted Therapies. *Cancer Cell* **37**, 530-542 (2020). https://doi.org/10.1016/j.ccell.2020.03.006
133. Maude, S. & Barrett, D.M. Current status of chimeric antigen receptor therapy for haematological malignancies. *Br J Haematol* **172**, 11-22 (2016) . https://doi.org/10.1111/bjh.13792
134. Miller, B.C. & Maus, M.V. CD19-Targeted CAR T Cells: A New Tool in the Fight against B Cell Malignancies. *Oncol Res Treat* **38**, 683-690 (2015). https://doi.org/10.1159/000442170
135. Terwilliger, T. & Abdul-Hay, M. Acute lymphoblastic leukemia: a comprehensive review and 2017 update. *Blood Cancer J* **7**, e577 (2017). https://doi.org/10.1038/bcj.2017.53
136. Kimura, S. & Mullighan, C.G. Molecular markers in ALL: Clinical implications. *Best Pract Res Clin Haematol* **33**, 101193 (2020) . https://doi.org/10.1016/j.beha.2020.101193
137. Rowe, J.M. Prognostic factors in adult acute lymphoblastic leukaemia. *Br J Haematol* **150**, 389-405 (2010). https://doi.org/10.1111/j.1365-2141.2010.08246.x
138. Davis, K. *et al*. Emerging molecular subtypes and therapies in acute lymphoblastic leukemia. *Semin Diagn Pathol* **40**, 202-215 (2023) . https://doi.org/10.1053/j.semdp.2023.04.003
139. Arber, D.A. *et al*. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood* **127**, 2391-2405 (2016). https://doi.org/10.1182/blood-2016-03-643544
140. Migita, N.A. *et al*. Classification and genetics of pediatric B-other acute lymphoblastic leukemia by targeted RNA sequencing. *Blood Adv* **7**, 2957-2971 (2023). https://doi.org/10.1182/bloodadvances.2022009179
141. Gu, Z. *et al*. PAX5-driven subtypes of B-progenitor acute lymphoblastic leukemia. *Nat Genet* **51**, 296-307 (2019). https://doi.org/10.1038/s41588-018-0315-5
142. Alaggio, R. *et al*. The 5th edition of the World Health Organization Classification of Haematolymphoid Tumours: Lymphoid Neoplasms. *Leukemia* **36**, 1720-1748 (2022). https://doi.org/10.1038/s41375-022-01620-2
143. Tasian, S.K. *et al*. Philadelphia chromosome-like acute lymphoblastic leukemia. *Blood* **130**, 2064-2072 (2017). https://doi.org/10.1182/blood-2017-06-743252

144. Yasuda, T. *et al*. Oncogenic lesions and molecular subtypes in adults with B-cell acute lymphoblastic leukemia. *Cancer Sci* **114**, 8-15 (2023). https://doi.org/10.1111/cas.15583
145. Mullighan, C.G. *et al*. Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature* **446**, 758-764 (2007). https://doi.org/10.1038/nature05690
146. Mullighan, C.G. *et al*. BCR-ABL1 lymphoblastic leukaemia is characterized by the deletion of Ikaros. *Nature* **453**, 110-114 (2008). https://doi.org/10.1038/nature06866
147. Prasad, M.A. *et al*. Ebf1 heterozygosity results in increased DNA damage in pro-B cells and their synergistic transformation by Pax5 haploinsufficiency. *Blood* **125**, 4052-4059 (2015). https://doi.org/10.1182/blood-2014-12-617282
148. Heltemes-Harris, L.M. *et al*. Ebf1 or Pax5 haploinsufficiency synergizes with STAT5 activation to initiate acute lymphoblastic leukemia. *J Exp Med* **208**, 1135-1149 (2011). https://doi.org/10.1084/jem.20101947
149. Martin-Lorenzo, A. *et al*. Loss of Pax5 Exploits Sca1-BCR-ABL(p190) Susceptibility to Confer the Metabolic Shift Essential for pB-ALL. *Cancer Res* **78**, 2669-2679 (2018). https://doi.org/10.1158/0008-5472.CAN-17-3262
150. Liu, G.J. *et al*. Pax5 loss imposes a reversible differentiation block in B-progenitor acute lymphoblastic leukemia. *Genes Dev* **28**, 1337-1350 (2014) . https://doi.org/10.1101/gad.240416.114
151. Jia, Z. & Gu, Z. PAX5 alterations in B-cell acute lymphoblastic leukemia. *Front Oncol* **12**, 1023606 (2022). https://doi.org/10.3389/fonc.2022.1023606
152. Waddington, C.H. The epigenotype. *Endeavour* **1**, 18-20 (1942).
153. Farooqi, A.A. *et al*. Epigenetic deregulation in cancer: Enzyme players and non-coding RNAs. *Semin Cancer Biol* **83**, 197-207 (2022). https://doi.org/10.1016/j.semcancer.2020.07.013
154. Luger, K. *et al*. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389**, 251-260 (1997). https://doi.org/10.1038/38444
155. McGinty, R.K. & Tan, S. Nucleosome structure and function. *Chem Rev* **115**, 2255-2273 (2015). https://doi.org/10.1021/cr500373h
156. Kornberg, R.D. Chromatin structure: a repeating unit of histones and DNA. *Science* **184**, 868-871 (1974). https://doi.org/10.1126/science.184.4139.868
157. Robinson, P.J. *et al*. EM measurements define the dimensions of the "30-nm" chromatin fiber: evidence for a compact, interdigitated structure. *Proc Natl Acad Sci U S A* **103**, 6506-6511 (2006). https://doi.org/10.1073/pnas.0601212103
158. Robinson, P.J. & Rhodes, D. Structure of the '30 nm' chromatin fibre: a key role for the linker histone. *Curr Opin Struct Biol* **16**, 336-343 (2006). https://doi.org/10.1016/j.sbi.2006.05.007
159. Klein, D.C. & Hainer, S.J. Chromatin regulation and dynamics in stem cells. *Curr Top Dev Biol* **138**, 1-71 (2020). https://doi.org/10.1016/bs.ctdb.2019.11.002
160. Trojer, P. & Reinberg, D. Facultative heterochromatin: is there a distinctive molecular signature? *Mol Cell* **28**, 1-13 (2007). https://doi.org/10.1016/j.molcel.2007.09.011
161. Saksouk, N. *et al*. Constitutive heterochromatin formation and transcription in mammals. *Epigenetics Chromatin* **8**, 3 (2015). https://doi.org/10.1186/1756-8935-8-3
162. Moore, L.D. *et al*. DNA methylation and its basic function. *Neuropsychopharmacology* **38**, 23-38 (2013) . https://doi.org/10.1038/npp.2012.112
163. Jin, B. *et al*. DNA methylation: superior or subordinate in the epigenetic hierarchy? *Genes Cancer* **2**, 607-617 (2011). https://doi.org/10.1177/1947601910393957

164. Feinberg, A.P. *et al*. The epigenetic progenitor origin of human cancer. *Nat Rev Genet* **7**, 21-33 (2006). https://doi.org/10.1038/nrg1748
165. Jeong, M. & Goodell, M.A. New answers to old questions from genome-wide maps of DNA methylation in hematopoietic cells. *Exp Hematol* **42**, 609-617 (2014). https://doi.org/10.1016/j.exphem.2014.04.008
166. Millan-Zambrano, G. *et al*. Histone post-translational modifications - cause and consequence of genome function. *Nat Rev Genet* **23**, 563-580 (2022). https://doi.org/10.1038/s41576-022-00468-7
167. Bannister, A.J. & Kouzarides, T. Regulation of chromatin by histone modifications. *Cell Res* **21**, 381-395 (2011). https://doi.org/10.1038/cr.2011.22
168. Lim, P.S. *et al*. Epigenetic control of inducible gene expression in the immune system. *Epigenomics* **2**, 775-795 (2010). https://doi.org/10.2217/epi.10.55
169. Ali, I. *et al*. Lysine Acetylation Goes Global: From Epigenetics to Metabolism and Therapeutics. *Chem Rev* **118**, 1216-1252 (2018). https://doi.org/10.1021/acs.chemrev.7b00181
170. Carrozza, M.J. *et al*. The diverse functions of histone acetyltransferase complexes. *Trends Genet* **19**, 321-329 (2003). https://doi.org/10.1016/S0168-9525(03)00115-X
171. Creyghton, M.P. *et al*. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci U S A* **107**, 21931-21936 (2010). https://doi.org/10.1073/pnas.1016071107
172. Wang, Z. *et al*. Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat Genet* **40**, 897-903 (2008). https://doi.org/10.1038/ng.154
173. Kouzarides, T. Chromatin modifications and their function. *Cell* **128**, 693-705 (2007). https://doi.org/10.1016/j.cell.2007.02.005
174. Hyndman, K.A. & Knepper, M.A. Dynamic regulation of lysine acetylation: the balance between acetyltransferase and deacetylase activities. *Am J Physiol Renal Physiol* **313**, F842-F846 (2017). https://doi.org/10.1152/ajprenal.00313.2017
175. Seto, E. & Yoshida, M. Erasers of histone acetylation: the histone deacetylase enzymes. *Cold Spring Harb Perspect Biol* **6**, a018713 (2014). https://doi.org/10.1101/cshperspect.a018713
176. Li, L. *et al*. SIRT7 is a histone desuccinylase that functionally links to chromatin compaction and genome stability. *Nat Commun* **7**, 12235 (2016). https://doi.org/10.1038/ncomms12235
177. Tan, M. *et al*. Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell* **146**, 1016-1028 (2011). https://doi.org/10.1016/j.cell.2011.08.008
178. Chen, Y. *et al*. Lysine propionylation and butyrylation are novel post-translational modifications in histones. *Mol Cell Proteomics* **6**, 812-819 (2007). https://doi.org/10.1074/mcp.M700021-MCP200
179. Lee, K.K. & Workman, J.L. Histone acetyltransferase complexes: one size doesn't fit all. *Nat Rev Mol Cell Biol* **8**, 284-295 (2007). https://doi.org/10.1038/nrm2145
180. Yang, X.J. & Seto, E. HATs and HDACs: from structure, function and regulation to novel strategies for therapy and prevention. *Oncogene* **26**, 5310-5318 (2007). https://doi.org/10.1038/sj.onc.1210599
181. Park, S.Y. & Kim, J.S. A short guide to histone deacetylases including recent progress on class II enzymes. *Exp Mol Med* **52**, 204-212 (2020). https://doi.org/10.1038/s12276-020-0382-4
182. Xiong, B. *et al*. Hos1 is a lysine deacetylase for the Smc3 subunit of cohesin. *Curr Biol* **20**, 1660-1665 (2010). https://doi.org/10.1016/j.cub.2010.08.019
183. Parra, M. Class IIa HDACs - new insights into their functions in physiology and pathology. *FEBS J* **282**, 1736-1744 (2015). https://doi.org/10.1111/febs.13061

184. Liu, S.S. *et al.* HDAC11: a rising star in epigenetics. *Biomed Pharmacother* **131**, 110607 (2020). https://doi.org/10.1016/j.biopha.2020.110607
185. Sheikh, B.N. *et al.* MOZ (KAT6A) is essential for the maintenance of classically defined adult hematopoietic stem cells. *Blood* **128**, 2307-2318 (2016). https://doi.org/10.1182/blood-2015-10-676072
186. Chan, W.I. *et al.* The transcriptional coactivator Cbp regulates self-renewal and differentiation in adult hematopoietic stem cells. *Mol Cell Biol* **31**, 5046-5060 (2011). https://doi.org/10.1128/MCB.05830-11
187. Valerio, D.G. *et al.* Histone acetyltransferase activity of MOF is required for adult but not early fetal hematopoiesis in mice. *Blood* **129**, 48-59 (2017) . https://doi.org/10.1182/blood-2016-05-714568
188. Azagra, A. *et al.* In vivo conditional deletion of HDAC7 reveals its requirement to establish proper B lymphocyte identity and development. *J Exp Med* **213**, 2591-2601 (2016). https://doi.org/10.1084/jem.20150821
189. de Barrios, O. *et al.* HDAC7 is a major contributor in the pathogenesis of infant t(4;11) proB acute lymphoblastic leukemia. *Leukemia* **35**, 2086-2091 (2021). https://doi.org/10.1038/s41375-020-01097-x
190. Hyun, K. *et al.* Writing, erasing and reading histone lysine methylations. *Exp Mol Med* **49**, e324 (2017). https://doi.org/10.1038/emm.2017.11
191. Black, J.C. *et al.* Histone lysine methylation dynamics: establishment, regulation, and biological impact. *Mol Cell* **48**, 491-507 (2012). https://doi.org/10.1016/j.molcel.2012.11.006
192. Heintzman, N.D. *et al.* Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet* **39**, 311-318 (2007). https://doi.org/10.1038/ng1966
193. Krivtsov, A.V. & Armstrong, S.A. MLL translocations, histone modifications and leukaemia stem-cell development. *Nat Rev Cancer* **7**, 823-833 (2007). https://doi.org/10.1038/nrc2253
194. Cao, R. *et al.* Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* **298**, 1039-1043 (2002). https://doi.org/10.1126/science.1076997
195. Kuzmichev, A. *et al.* Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. *Genes Dev* **16**, 2893-2905 (2002). https://doi.org/10.1101/gad.1035902
196. Mochizuki-Kashio, M. *et al.* Ezh2 loss in hematopoietic stem cells predisposes mice to develop heterogeneous malignancies in an Ezh1-dependent manner. *Blood* **126**, 1172-1183 (2015). https://doi.org/10.1182/blood-2015-03-634428
197. Simon, C. *et al.* A key role for EZH2 and associated genes in mouse and human adult T-cell acute leukemia. *Genes Dev* **26**, 651-656 (2012). https://doi.org/10.1101/gad.186411.111
198. Musselman, C.A. *et al.* Perceiving the epigenetic landscape through histone readers. *Nat Struct Mol Biol* **19**, 1218-1227 (2012). https://doi.org/10.1038/nsmb.2436
199. Chang, B. *et al.* JMJD6 is a histone arginine demethylase. *Science* **318**, 444-447 (2007). https://doi.org/10.1126/science.1145801
200. Agger, K. *et al.* UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. *Nature* **449**, 731-734 (2007). https://doi.org/10.1038/nature06145
201. Lan, F. *et al.* A histone H3 lysine 27 demethylase regulates animal posterior development. *Nature* **449**, 689-694 (2007). https://doi.org/10.1038/nature06192

202. Hsia, D.A. *et al.* KDM8, a H3K36me2 histone demethylase that acts in the cyclin A1 coding region to regulate cancer cell proliferation. *Proc Natl Acad Sci U S A* **107**, 9671-9676 (2010). https://doi.org/10.1073/pnas.1000401107
203. Agger, K. *et al.* The KDM4/JMJD2 histone demethylases are required for hematopoietic stem cell maintenance. *Blood* **134**, 1154-1158 (2019). https://doi.org/10.1182/blood.2019000855
204. Li, J. *et al.* Programmable human histone phosphorylation and gene activation using a CRISPR/Cas9-based chromatin kinase. *Nat Commun* **12**, 896 (2021). https://doi.org/10.1038/s41467-021-21188-2
205. Foster, E.R. & Downs, J.A. Histone H2A phosphorylation in DNA double-strand break repair. *FEBS J* **272**, 3231-3240 (2005). https://doi.org/10.1111/j.1742-4658.2005.04741.x
206. Thiriet, C. & Hayes, J.J. Linker histone phosphorylation regulates global timing of replication origin firing. *J Biol Chem* **284**, 2823-2829 (2009). https://doi.org/10.1074/jbc.M805617200
207. Rogakou, E.P. *et al.* DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *J Biol Chem* **273**, 5858-5868 (1998) . https://doi.org/10.1074/jbc.273.10.5858
208. Chowdhury, D. *et al.* gamma-H2AX dephosphorylation by protein phosphatase 2A facilitates DNA double-strand break repair. *Mol Cell* **20**, 801-809 (2005). https://doi.org/10.1016/j.molcel.2005.10.003
209. Chowdhury, D. *et al.* A PP4-phosphatase complex dephosphorylates gamma-H2AX generated during DNA replication. *Mol Cell* **31**, 33-46 (2008). https://doi.org/10.1016/j.molcel.2008.05.016
210. Banerjee, T. & Chakravarti, D. A peek into the complex realm of histone phosphorylation. *Mol Cell Biol* **31**, 4858-4873 (2011). https://doi.org/10.1128/MCB.05631-11
211. Crosio, C. *et al.* Mitotic phosphorylation of histone H3: spatio-temporal regulation by mammalian Aurora kinases. *Mol Cell Biol* **22**, 874-885 (2002). https://doi.org/10.1128/MCB.22.3.874-885.2002
212. Komander, D. & Rape, M. The ubiquitin code. *Annu Rev Biochem* **81**, 203-229 (2012). https://doi.org/10.1146/annurev-biochem-060310-170328
213. Damgaard, R.B. The ubiquitin system: from cell signalling to disease biology and new therapeutic opportunities. *Cell Death Differ* **28**, 423-426 (2021). https://doi.org/10.1038/s41418-020-00703-w
214. Khaminets, A. *et al.* Ubiquitin-Dependent And Independent Signals In Selective Autophagy. *Trends Cell Biol* **26**, 6-16 (2016). https://doi.org/10.1016/j.tcb.2015.08.010
215. Zhang, Y. Transcriptional regulation by histone ubiquitination and deubiquitination. *Genes Dev* **17**, 2733-2740 (2003). https://doi.org/10.1101/gad.1156403
216. Barbour, H. *et al.* Polycomb group-mediated histone H2A monoubiquitination in epigenome regulation and nuclear processes. *Nat Commun* **11**, 5947 (2020). https://doi.org/10.1038/s41467-020-19722-9
217. Lee, H.G. *et al.* Genome-wide activities of Polycomb complexes control pervasive transcription. *Genome Res* **25**, 1170-1181 (2015). https://doi.org/10.1101/gr.188920.114
218. Oguro, H. *et al.* Poised lineage specification in multipotential hematopoietic stem and progenitor cells by the polycomb protein Bmi1. *Cell Stem Cell* **6**, 279-286 (2010). https://doi.org/10.1016/j.stem.2010.01.005
219. Henry, K.W. *et al.* Transcriptional activation via sequential histone H2B ubiquitylation and deubiquitylation, mediated by SAGA-associated Ubp8. *Genes Dev* **17**, 2648-2663 (2003). https://doi.org/10.1101/gad.1144003

220. Pavri, R. *et al*. Histone H2B monoubiquitination functions cooperatively with FACT to regulate elongation by RNA polymerase II. *Cell* **125**, 703-717 (2006). https://doi.org/10.1016/j.cell.2006.04.029
221. Nakanishi, S. *et al*. Histone H2BK123 monoubiquitination is the critical determinant for H3K4 and H3K79 trimethylation by COMPASS and Dot1. *J Cell Biol* **186**, 371-377 (2009). https://doi.org/10.1083/jcb.200906005
222. Fradet-Turcotte, A. *et al*. 53BP1 is a reader of the DNA-damage-induced H2A Lys 15 ubiquitin mark. *Nature* **499**, 50-54 (2013). https://doi.org/10.1038/nature12318
223. van der Veen, A.G. & Ploegh, H.L. Ubiquitin-like proteins. *Annu Rev Biochem* **81**, 323-357 (2012). https://doi.org/10.1146/annurev-biochem-093010-153308
224. Shiio, Y. & Eisenman, R.N. Histone sumoylation is associated with transcriptional repression. *Proc Natl Acad Sci U S A* **100**, 13225-13230 (2003). https://doi.org/10.1073/pnas.1735528100
225. Ryu, H.Y. & Hochstrasser, M. Histone sumoylation and chromatin dynamics. *Nucleic Acids Res* **49**, 6043-6052 (2021). https://doi.org/10.1093/nar/gkab280
226. Pearson, C.K. Chapter 13 ADP-ribosylation reactions. In: Bittar, E.E. & Bittar, N. (eds). *Principles of Medical Biology*, vol. 4. Elsevier, 1995, pp 305-322. https://doi.org/https://doi.org/10.1016/S1569-2582(06)80015-2
227. Messner, S. & Hottiger, M.O. Histone ADP-ribosylation in DNA repair, replication and transcription. *Trends Cell Biol* **21**, 534-542 (2011). https://doi.org/10.1016/j.tcb.2011.06.001
228. Boulikas, T. DNA strand breaks alter histone ADP-ribosylation. *Proc Natl Acad Sci U S A* **86**, 3499-3503 (1989). https://doi.org/10.1073/pnas.86.10.3499
229. Hottiger, M.O. Nuclear ADP-Ribosylation and Its Role in Chromatin Plasticity, Cell Differentiation, and Epigenetics. *Annu Rev Biochem* **84**, 227-263 (2015). https://doi.org/10.1146/annurev-biochem-060614-034506
230. Hottiger, M.O. *et al*. Toward a unified nomenclature for mammalian ADP-ribosyltransferases. *Trends Biochem Sci* **35**, 208-219 (2010). https://doi.org/10.1016/j.tibs.2009.12.003
231. Krishnakumar, R. & Kraus, W.L. The PARP side of the nucleus: molecular actions, physiological outcomes, and clinical targets. *Mol Cell* **39**, 8-24 (2010). https://doi.org/10.1016/j.molcel.2010.06.017
232. Althaus, F.R. *et al*. Poly ADP-ribosylation: a DNA break signal mechanism. *Mol Cell Biochem* **193**, 5-11 (1999).
233. Ray Chaudhuri, A. & Nussenzweig, A. The multifaceted roles of PARP1 in DNA repair and chromatin remodelling. *Nat Rev Mol Cell Biol* **18**, 610-621 (2017). https://doi.org/10.1038/nrm.2017.53
234. Bryant, H.E. *et al*. Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. *Nature* **434**, 913-917 (2005). https://doi.org/10.1038/nature03443
235. Farmer, H. *et al*. Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature* **434**, 917-921 (2005). https://doi.org/10.1038/nature03445
236. Bruin, M.A.C. *et al*. Pharmacokinetics and Pharmacodynamics of PARP Inhibitors in Oncology. *Clin Pharmacokinet* **61**, 1649-1675 (2022). https://doi.org/10.1007/s40262-022-01167-6
237. Tyagi, M. *et al*. Chromatin remodelers: We are the drivers!! *Nucleus* **7**, 388-404 (2016). https://doi.org/10.1080/19491034.2016.1211217
238. Clapier, C.R. *et al*. Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nat Rev Mol Cell Biol* **18**, 407-422 (2017). https://doi.org/10.1038/nrm.2017.26

239. Kadoch, C. *et al*. Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. *Nat Genet* **45**, 592-601 (2013). https://doi.org/10.1038/ng.2628
240. Murawska, M. & Brehm, A. CHD chromatin remodelers and the transcription cycle. *Transcription* **2**, 244-253 (2011). https://doi.org/10.4161/trns.2.6.17840
241. Boeger, H. *et al*. Removal of promoter nucleosomes by disassembly rather than sliding in vivo. *Mol Cell* **14**, 667-673 (2004). https://doi.org/10.1016/j.molcel.2004.05.013
242. Mizuguchi, G. *et al*. ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. *Science* **303**, 343-348 (2004). https://doi.org/10.1126/science.1090701
243. Goldberg, A.D. *et al*. Distinct factors control histone variant H3.3 localization at specific genomic regions. *Cell* **140**, 678-691 (2010). https://doi.org/10.1016/j.cell.2010.01.003
244. Centore, R.C. *et al*. Mammalian SWI/SNF Chromatin Remodeling Complexes: Emerging Mechanisms and Therapeutic Strategies. *Trends Genet* **36**, 936-950 (2020). https://doi.org/10.1016/j.tig.2020.07.011
245. Church, M.C. *et al*. The Swi-Snf chromatin remodeling complex mediates gene repression through metabolic control. *Nucleic Acids Res* **51**, 10278-10291 (2023). https://doi.org/10.1093/nar/gkad711
246. Morse, K. *et al*. Swi/Snf chromatin remodeling regulates transcriptional interference and gene repression. *Mol Cell* **84**, 3080-3097 e3089 (2024). https://doi.org/10.1016/j.molcel.2024.06.029
247. Allen, H.F. *et al*. The NuRD architecture. *Cell Mol Life Sci* **70**, 3513-3524 (2013). https://doi.org/10.1007/s00018-012-1256-2
248. Kim, J. *et al*. Ikaros DNA-binding proteins direct formation of chromatin remodeling complexes in lymphocytes. *Immunity* **10**, 345-355 (1999). https://doi.org/10.1016/s1074-7613(00)80034-5
249. Zhang, J. *et al*. Harnessing of the nucleosome-remodeling-deacetylase complex controls lymphocyte development and prevents leukemogenesis. *Nat Immunol* **13**, 86-94 (2011). https://doi.org/10.1038/ni.2150
250. Bosch-Presegue, L. & Vaquero, A. Sirtuin-dependent epigenetic regulation in the maintenance of genome integrity. *FEBS J* **282**, 1745-1767 (2015). https://doi.org/10.1111/febs.13053
251. Hirschey, M.D. Old enzymes, new tricks: sirtuins are NAD(+)-dependent de-acylases. *Cell Metab* **14**, 718-719 (2011). https://doi.org/10.1016/j.cmet.2011.10.006
252. Bosch-Presegue, L. & Vaquero, A. Sirtuins in stress response: guardians of the genome. *Oncogene* **33**, 3764-3775 (2014). https://doi.org/10.1038/onc.2013.344
253. Frye, R.A. Phylogenetic classification of prokaryotic and eukaryotic Sir2-like proteins. *Biochem Biophys Res Commun* **273**, 793-798 (2000). https://doi.org/10.1006/bbrc.2000.3000
254. Michishita, E. *et al*. Evolutionarily conserved and nonconserved cellular localizations and functions of human SIRT proteins. *Mol Biol Cell* **16**, 4623-4635 (2005). https://doi.org/10.1091/mbc.e05-01-0033
255. Tanno, M. *et al*. Nucleocytoplasmic shuttling of the NAD+-dependent histone deacetylase SIRT1. *J Biol Chem* **282**, 6823-6832 (2007). https://doi.org/10.1074/jbc.M609554200
256. Buler, M. *et al*. Who watches the watchmen? Regulation of the expression and activity of sirtuins. *FASEB J* **30**, 3942-3960 (2016). https://doi.org/10.1096/fj.201600410RR

257. Landry, J. *et al*. The silencing protein SIR2 and its homologs are NAD-dependent protein deacetylases. *Proc Natl Acad Sci U S A* **97**, 5807-5811 (2000). https://doi.org/10.1073/pnas.110148297
258. Sauve, A.A. *et al*. The biochemistry of sirtuins. *Annu Rev Biochem* **75**, 435-465 (2006). https://doi.org/10.1146/annurev.biochem.74.082803.133500
259. Bheda, P. *et al*. The Substrate Specificity of Sirtuins. *Annu Rev Biochem* **85**, 405-429 (2016). https://doi.org/10.1146/annurev-biochem-060815-014537
260. Zhu, A.Y. *et al*. Plasmodium falciparum Sir2A preferentially hydrolyzes medium and long chain fatty acyl lysine. *ACS Chem Biol* **7**, 155-159 (2012). https://doi.org/10.1021/cb200230x
261. Van Meter, M. *et al*. SIRT6 represses LINE1 retrotransposons by ribosylating KAP1 but this repression fails with stress and age. *Nat Commun* **5**, 5011 (2014). https://doi.org/10.1038/ncomms6011
262. Mao, Z. *et al*. SIRT6 promotes DNA repair under stress by activating PARP1. *Science* **332**, 1443-1446 (2011). https://doi.org/10.1126/science.1202723
263. Simonet, N.G. *et al*. SirT7 auto-ADP-ribosylation regulates glucose starvation response through mH2A1. *Sci Adv* **6**, eaaz2590 (2020). https://doi.org/10.1126/sciadv.aaz2590
264. Ianni, A. *et al*. SIRT7: a novel molecular target for personalized cancer treatment? *Oncogene* **43**, 993-1006 (2024). https://doi.org/10.1038/s41388-024-02976-8
265. Fraga, M.F. *et al*. Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. *Nat Genet* **37**, 391-400 (2005). https://doi.org/10.1038/ng1531
266. Dang, W. *et al*. Histone H4 lysine 16 acetylation regulates cellular lifespan. *Nature* **459**, 802-807 (2009). https://doi.org/10.1038/nature08085
267. Imai, S. *et al*. Transcriptional silencing and longevity protein Sir2 is an NAD-dependent histone deacetylase. *Nature* **403**, 795-800 (2000). https://doi.org/10.1038/35001622
268. Vaquero, A. *et al*. SirT2 is a histone deacetylase with preference for histone H4 Lys 16 during mitosis. *Genes Dev* **20**, 1256-1261 (2006). https://doi.org/10.1101/gad.1412706
269. Krauss, V. Glimpses of evolution: heterochromatic histone H3K9 methyltransferases left its marks behind. *Genetica* **133**, 93-106 (2008). https://doi.org/10.1007/s10709-007-9184-z
270. Michishita, E. *et al*. SIRT6 is a histone H3 lysine 9 deacetylase that modulates telomeric chromatin. *Nature* **452**, 492-496 (2008). https://doi.org/10.1038/nature06736
271. McCord, R.A. *et al*. SIRT6 stabilizes DNA-dependent protein kinase at chromatin for DNA double-strand break repair. *Aging (Albany NY)* **1**, 109-121 (2009). https://doi.org/10.18632/aging.100011
272. Kawahara, T.L. *et al*. SIRT6 links histone H3 lysine 9 deacetylation to NF-kappaB-dependent gene expression and organismal life span. *Cell* **136**, 62-74 (2009). https://doi.org/10.1016/j.cell.2008.10.052
273. Michishita, E. *et al*. Cell cycle-dependent deacetylation of telomeric histone H3 lysine K56 by human SIRT6. *Cell Cycle* **8**, 2664-2666 (2009). https://doi.org/10.4161/cc.8.16.9367
274. Yuan, J. *et al*. Histone H3-K56 acetylation is important for genomic stability in mammals. *Cell Cycle* **8**, 1747-1753 (2009). https://doi.org/10.4161/cc.8.11.8620
275. Topal, S. *et al*. Distinct transcriptional roles for Histone H3-K56 acetylation during the cell cycle in Yeast. *Nat Commun* **10**, 4372 (2019). https://doi.org/10.1038/s41467-019-12400-5

276. Wang, W.W. *et al.* A Click Chemistry Approach Reveals the Chromatin-Dependent Histone H3K36 Deacylase Nature of SIRT7. *J Am Chem Soc* **141**, 2462-2473 (2019). https://doi.org/10.1021/jacs.8b12083
277. Barber, M.F. *et al.* SIRT7 links H3K18 deacetylation to maintenance of oncogenic transformation. *Nature* **487**, 114-118 (2012). https://doi.org/10.1038/nature11043
278. Morris, S.A. *et al.* Identification of histone H3 lysine 36 acetylation as a highly conserved histone modification. *J Biol Chem* **282**, 7632-7640 (2007). https://doi.org/10.1074/jbc.M607909200
279. Pai, C.C. *et al.* A histone H3K36 chromatin switch coordinates DNA double-strand break repair pathway choice. *Nat Commun* **5**, 4091 (2014). https://doi.org/10.1038/ncomms5091
280. Kim, H. *et al.* Chd1p recognizes H3K36Ac to maintain nucleosome positioning near the transcription start site. *Biochem Biophys Res Commun* **503**, 1200-1206 (2018). https://doi.org/10.1016/j.bbrc.2018.07.025
281. Vaquero, A. *et al.* SIRT1 regulates the histone methyl-transferase SUV39H1 during heterochromatin formation. *Nature* **450**, 440-444 (2007). https://doi.org/10.1038/nature06268
282. Bouras, T. *et al.* SIRT1 deacetylation and repression of p300 involves lysine residues 1020/1024 within the cell cycle regulatory domain 1. *J Biol Chem* **280**, 10264-10276 (2005). https://doi.org/10.1074/jbc.M408748200
283. Black, J.C. *et al.* The SIRT2 deacetylase regulates autoacetylation of p300. *Mol Cell* **32**, 449-455 (2008). https://doi.org/10.1016/j.molcel.2008.09.018
284. Lu, L. *et al.* Modulations of hMOF autoacetylation by SIRT1 regulate hMOF recruitment and activities on the chromatin. *Cell Res* **21**, 1182-1195 (2011). https://doi.org/10.1038/cr.2011.71
285. Serrano, L. *et al.* The tumor suppressor SirT2 regulates cell cycle progression and genome stability by modulating the mitotic deposition of H4K20 methylation. *Genes Dev* **27**, 639-653 (2013). https://doi.org/10.1101/gad.211342.112
286. Mostoslavsky, R. *et al.* Genomic instability and aging-like phenotype in the absence of mammalian SIRT6. *Cell* **124**, 315-329 (2006). https://doi.org/10.1016/j.cell.2005.11.044
287. Cheng, H.L. *et al.* Developmental defects and p53 hyperacetylation in Sir2 homolog (SIRT1)-deficient mice. *Proc Natl Acad Sci U S A* **100**, 10794-10799 (2003). https://doi.org/10.1073/pnas.1934713100
288. Vazquez, B.N. *et al.* SIRT7 promotes genome integrity and modulates non-homologous end joining DNA repair. *EMBO J* **35**, 1488-1503 (2016). https://doi.org/10.15252/embj.201593499
289. Yuan, Z. *et al.* SIRT1 regulates the function of the Nijmegen breakage syndrome protein. *Mol Cell* **27**, 149-162 (2007). https://doi.org/10.1016/j.molcel.2007.05.029
290. Li, K. *et al.* Regulation of WRN protein cellular localization and enzymatic activities by SIRT1-mediated deacetylation. *J Biol Chem* **283**, 7590-7598 (2008). https://doi.org/10.1074/jbc.M709707200
291. Jeong, J. *et al.* SIRT1 promotes DNA repair activity and deacetylation of Ku70. *Exp Mol Med* **39**, 8-13 (2007). https://doi.org/10.1038/emm.2007.2
292. Rasti, G. *et al.* SIRT1 regulates DNA damage signaling through the PP4 phosphatase complex. *Nucleic Acids Res* **51**, 6754-6769 (2023). https://doi.org/10.1093/nar/gkad504
293. Tang, Y. *et al.* Acetylation is indispensable for p53 activation. *Cell* **133**, 612-626 (2008). https://doi.org/10.1016/j.cell.2008.03.025
294. Vaziri, H. *et al.* hSIR2(SIRT1) functions as an NAD-dependent p53 deacetylase. *Cell* **107**, 149-159 (2001). https://doi.org/10.1016/s0092-8674(01)00527-x

295. Langley, E. *et al.* Human SIR2 deacetylates p53 and antagonizes PML/p53-induced cellular senescence. *EMBO J* **21**, 2383-2396 (2002). https://doi.org/10.1093/emboj/21.10.2383
296. Wood, M. *et al.* Trichostatin A inhibits deacetylation of histone H3 and p53 by SIRT6. *Arch Biochem Biophys* **638**, 8-17 (2018). https://doi.org/10.1016/j.abb.2017.12.009
297. Ghosh, S. *et al.* Haploinsufficiency of Trp53 dramatically extends the lifespan of Sirt6-deficient mice. *Elife* **7** (2018). https://doi.org/10.7554/eLife.32127
298. Wang, R.H. *et al.* Impaired DNA damage response, genome instability, and tumorigenesis in SIRT1 mutant mice. *Cancer Cell* **14**, 312-323 (2008). https://doi.org/10.1016/j.ccr.2008.09.001
299. Peck, B. *et al.* SIRT inhibitors induce cell death and p53 acetylation through targeting both SIRT1 and SIRT2. *Mol Cancer Ther* **9**, 844-855 (2010). https://doi.org/10.1158/1535-7163.MCT-09-0971
300. Li, S. *et al.* p53-induced growth arrest is regulated by the mitochondrial SirT3 deacetylase. *PLoS One* **5**, e10486 (2010). https://doi.org/10.1371/journal.pone.0010486
301. Zhao, J. *et al.* SIRT7 regulates hepatocellular carcinoma response to therapy by altering the p53-dependent cell death pathway. *J Exp Clin Cancer Res* **38**, 252 (2019). https://doi.org/10.1186/s13046-019-1246-4
302. Lu, Y.F. *et al.* SIRT7 activates p53 by enhancing PCAF-mediated MDM2 degradation to arrest the cell cycle. *Oncogene* **39**, 4650-4665 (2020). https://doi.org/10.1038/s41388-020-1305-5
303. Vazquez, B.N. *et al.* SIRT7 and p53 interaction in embryonic development and tumorigenesis. *Front Cell Dev Biol* **11**, 1281730 (2023). https://doi.org/10.3389/fcell.2023.1281730
304. Gamez-Garcia, A. & Vazquez, B.N. Nuclear Sirtuins and the Aging of the Immune System. *Genes (Basel)* **12** (2021). https://doi.org/10.3390/genes12121856
305. Wang, F. *et al.* SIRT2 deacetylates FOXO3a in response to oxidative stress and caloric restriction. *Aging Cell* **6**, 505-514 (2007). https://doi.org/10.1111/j.1474-9726.2007.00304.x
306. van der Horst, A. *et al.* FOXO4 is acetylated upon peroxide stress and deacetylated by the longevity protein hSir2(SIRT1). *J Biol Chem* **279**, 28873-28879 (2004). https://doi.org/10.1074/jbc.M401138200
307. Motta, M.C. *et al.* Mammalian SIRT1 represses forkhead transcription factors. *Cell* **116**, 551-563 (2004). https://doi.org/10.1016/s0092-8674(04)00126-6
308. Jeng, M.Y. *et al.* Metabolic reprogramming of human CD8(+) memory T cells through loss of SIRT1. *J Exp Med* **215**, 51-62 (2018). https://doi.org/10.1084/jem.20161066
309. Lim, J.H. *et al.* Sirtuin 1 modulates cellular responses to hypoxia by deacetylating hypoxia-inducible factor 1alpha. *Mol Cell* **38**, 864-878 (2010). https://doi.org/10.1016/j.molcel.2010.05.023
310. Yeung, F. *et al.* Modulation of NF-kappaB-dependent transcription and cell survival by the SIRT1 deacetylase. *EMBO J* **23**, 2369-2380 (2004). https://doi.org/10.1038/sj.emboj.7600244
311. Rothgiesser, K.M. *et al.* SIRT2 regulates NF-kappaB dependent gene expression through deacetylation of p65 Lys310. *J Cell Sci* **123**, 4251-4258 (2010). https://doi.org/10.1242/jcs.073783
312. Qin, Z. *et al.* Deactylation by SIRT1 enables liquid-liquid phase separation of IRF3/IRF7 in innate antiviral immunity. *Nat Immunol* **23**, 1193-1207 (2022). https://doi.org/10.1038/s41590-022-01269-0

313. Vakhrusheva, O. *et al.* Sirt7 increases stress resistance of cardiomyocytes and prevents apoptosis and inflammatory cardiomyopathy in mice. *Circ Res* **102**, 703-710 (2008). https://doi.org/10.1161/CIRCRESAHA.107.164558
314. Shin, J. *et al.* SIRT7 represses Myc activity to suppress ER stress and prevent fatty liver disease. *Cell Rep* **5**, 654-665 (2013). https://doi.org/10.1016/j.celrep.2013.10.007
315. Mohrin, M. *et al.* Stem cell aging. A mitochondrial UPR-mediated metabolic checkpoint regulates hematopoietic stem cell aging. *Science* **347**, 1374-1377 (2015). https://doi.org/10.1126/science.aaa2361
316. Wyman, A.E. *et al.* SIRT7 deficiency suppresses inflammation, induces EndoMT, and increases vascular permeability in primary pulmonary endothelial cells. *Sci Rep* **10**, 12497 (2020). https://doi.org/10.1038/s41598-020-69236-z
317. Liu, X. *et al.* SIRT7 Facilitates CENP-A Nucleosome Assembly and Suppresses Intestinal Tumorigenesis. *iScience* **23**, 101461 (2020). https://doi.org/10.1016/j.isci.2020.101461
318. Kaiser, A. *et al.* SIRT7: an influence factor in healthy aging and the development of age-dependent myeloid stem-cell disorders. *Leukemia* **34**, 2206-2216 (2020). https://doi.org/10.1038/s41375-020-0803-3
319. Ford, E. *et al.* Mammalian Sir2 homolog SIRT7 is an activator of RNA polymerase I transcription. *Genes Dev* **20**, 1075-1080 (2006). https://doi.org/10.1101/gad.1399706
320. Chen, S. *et al.* SIRT7-dependent deacetylation of the U3-55k protein controls pre-rRNA processing. *Nat Commun* **7**, 10734 (2016). https://doi.org/10.1038/ncomms10734
321. Zhang, C. *et al.* Quantitative proteome-based systematic identification of SIRT7 substrates. *Proteomics* **17** (2017). https://doi.org/10.1002/pmic.201600395
322. Tsai, Y.C. *et al.* Functional proteomics establishes the interaction of SIRT7 with chromatin remodeling complexes and expands its role in regulation of RNA polymerase I transcription. *Mol Cell Proteomics* **11**, 60-76 (2012). https://doi.org/10.1074/mcp.A111.015156
323. Blank, M.F. *et al.* SIRT7-dependent deacetylation of CDK9 activates RNA polymerase II transcription. *Nucleic Acids Res* **45**, 2675-2686 (2017). https://doi.org/10.1093/nar/gkx053
324. Chen, S. *et al.* Repression of RNA polymerase I upon stress is caused by inhibition of RNA-dependent deacetylation of PAF53 by SIRT7. *Mol Cell* **52**, 303-313 (2013). https://doi.org/10.1016/j.molcel.2013.10.010
325. Iyer-Bierhoff, A. *et al.* SIRT7-Dependent Deacetylation of Fibrillarin Controls Histone H2A Methylation and rRNA Synthesis during the Cell Cycle. *Cell Rep* **25**, 2946-2954 e2945 (2018). https://doi.org/10.1016/j.celrep.2018.11.051
326. Tsai, Y.C. *et al.* Sirtuin 7 plays a role in ribosome biogenesis and protein synthesis. *Mol Cell Proteomics* **13**, 73-83 (2014). https://doi.org/10.1074/mcp.M113.031377
327. Hubbi, M.E. *et al.* Sirtuin-7 inhibits the activity of hypoxia-inducible factors. *J Biol Chem* **288**, 20768-20775 (2013). https://doi.org/10.1074/jbc.M113.476903
328. Kiran, S. *et al.* Sirtuin 7 promotes cellular survival following genomic stress by attenuation of DNA damage, SAPK activation and p53 response. *Exp Cell Res* **331**, 123-141 (2015). https://doi.org/10.1016/j.yexcr.2014.11.001
329. Houtkooper, R.H. *et al.* Sirtuins as regulators of metabolism and healthspan. *Nat Rev Mol Cell Biol* **13**, 225-238 (2012). https://doi.org/10.1038/nrm3293
330. Picard, F. *et al.* Sirt1 promotes fat mobilization in white adipocytes by repressing PPAR-gamma. *Nature* **429**, 771-776 (2004). https://doi.org/10.1038/nature02583

331. Fang, J. *et al.* Sirt7 promotes adipogenesis in the mouse by inhibiting autocatalytic activation of Sirt1. *Proc Natl Acad Sci U S A* **114**, E8352-E8361 (2017). https://doi.org/10.1073/pnas.1706945114
332. Yan, W.W. *et al.* Arginine methylation of SIRT7 couples glucose sensing with mitochondria biogenesis. *EMBO Rep* **19** (2018). https://doi.org/10.15252/embr.201846377
333. Sun, L. *et al.* Regulation of energy homeostasis by the ubiquitin-independent REGgamma proteasome. *Nat Commun* **7**, 12497 (2016). https://doi.org/10.1038/ncomms12497
334. Ryu, D. *et al.* A SIRT7-dependent acetylation switch of GABPbeta1 controls mitochondrial function. *Cell Metab* **20**, 856-869 (2014). https://doi.org/10.1016/j.cmet.2014.08.001
335. Jiang, L. *et al.* Ubiquitin-specific peptidase 7 (USP7)-mediated deubiquitination of the histone deacetylase SIRT7 regulates gluconeogenesis. *J Biol Chem* **292**, 13296-13311 (2017). https://doi.org/10.1074/jbc.M117.780130
336. Carafa, V. *et al.* Dual Tumor Suppressor and Tumor Promoter Action of Sirtuins in Determining Malignant Phenotype. *Front Pharmacol* **10**, 38 (2019). https://doi.org/10.3389/fphar.2019.00038
337. Tang, X. *et al.* SIRT7 antagonizes TGF-beta signaling and inhibits breast cancer metastasis. *Nat Commun* **8**, 318 (2017). https://doi.org/10.1038/s41467-017-00396-9
338. Li, W. *et al.* SIRT7 suppresses the epithelial-to-mesenchymal transition in oral squamous cell carcinoma metastasis by promoting SMAD4 deacetylation. *J Exp Clin Cancer Res* **37**, 148 (2018). https://doi.org/10.1186/s13046-018-0819-y
339. Mason, D.Y. *et al.* Development and follicular localization of tolerant B lymphocytes in lysozyme/anti-lysozyme IgM/IgD transgenic mice. *Int Immunol* **4**, 163-175 (1992). https://doi.org/10.1093/intimm/4.2.163
340. Ng, A.P. *et al.* An Erg-driven transcriptional program controls B cell lymphopoiesis. *Nat Commun* **11**, 3013 (2020). https://doi.org/10.1038/s41467-020-16828-y
341. Wolf, F.A. *et al.* SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol* **19**, 15 (2018). https://doi.org/10.1186/s13059-017-1382-0
342. Badia, I.M.P. *et al.* decoupleR: ensemble of computational methods to infer biological activities from omics data. *Bioinform Adv* **2**, vbac016 (2022). https://doi.org/10.1093/bioadv/vbac016
343. Franzen, O. *et al.* PanglaoDB: a web server for exploration of mouse and human single-cell RNA sequencing data. *Database (Oxford)* **2019** (2019). https://doi.org/10.1093/database/baz046
344. Heng, T.S. *et al.* The Immunological Genome Project: networks of gene expression in immune cells. *Nat Immunol* **9**, 1091-1094 (2008). https://doi.org/10.1038/ni1008-1091
345. Okuyama, K. *et al.* PAX5 is part of a functional transcription factor network targeted in lymphoid leukemia. *PLoS Genet* **15**, e1008280 (2019). https://doi.org/10.1371/journal.pgen.1008280
346. Yang, M. *et al.* Proteogenomics and Hi-C reveal transcriptional dysregulation in high hyperdiploid childhood acute lymphoblastic leukemia. *Nat Commun* **10**, 1519 (2019). https://doi.org/10.1038/s41467-019-09469-3
347. Johnston, H.E. *et al.* Proteomics Profiling of CLL Versus Healthy B-cells Identifies Putative Therapeutic Targets and a Subtype-independent Signature of Spliceosome Dysregulation. *Mol Cell Proteomics* **17**, 776-791 (2018). https://doi.org/10.1074/mcp.RA117.000539

348. Roberts, K.G. *et al*. Targetable kinase-activating lesions in Ph-like acute lymphoblastic leukemia. *N Engl J Med* **371**, 1005-1015 (2014). https://doi.org/10.1056/NEJMoa1403088
349. Patro, R. *et al*. Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods* **14**, 417-419 (2017). https://doi.org/10.1038/nmeth.4197
350. Love, M.I. *et al*. Tximeta: Reference sequence checksums for provenance identification in RNA-seq. *PLoS Comput Biol* **16**, e1007664 (2020). https://doi.org/10.1371/journal.pcbi.1007664
351. Love, M.I. *et al*. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014). https://doi.org/10.1186/s13059-014-0550-8
352. Langmead, B. & Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357-359 (2012). https://doi.org/10.1038/nmeth.1923
353. Li, H. *et al*. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009). https://doi.org/10.1093/bioinformatics/btp352
354. Tarasov, A. *et al*. Sambamba: fast processing of NGS alignment formats. *Bioinformatics* **31**, 2032-2034 (2015). https://doi.org/10.1093/bioinformatics/btv098
355. Ramirez, F. *et al*. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res* **42**, W187-191 (2014). https://doi.org/10.1093/nar/gku365
356. Thorvaldsdottir, H. *et al*. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform* **14**, 178-192 (2013). https://doi.org/10.1093/bib/bbs017
357. Subramanian, A. *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). https://doi.org/10.1073/pnas.0506580102
358. Kuleshov, M.V. *et al*. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res* **44**, W90-97 (2016). https://doi.org/10.1093/nar/gkw377
359. Zhang, Y. *et al*. Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137 (2008). https://doi.org/10.1186/gb-2008-9-9-r137
360. Nutt, S.L. & Tarlinton, D.M. Germinal center B and follicular helper T cells: siblings, cousins or just good friends? *Nat Immunol* **12**, 472-477 (2011). https://doi.org/10.1038/ni.2019
361. Zebedee, S.L. *et al*. Comparison of mouse Ly5a and Ly5b leucocyte common antigen alleles. *Dev Immunol* **1**, 243-254 (1991). https://doi.org/10.1155/1991/52686
362. Difilippantonio, S. *et al*. 53BP1 facilitates long-range DNA end-joining during V(D)J recombination. *Nature* **456**, 529-533 (2008). https://doi.org/10.1038/nature07476
363. Schlissel, M.S. *et al*. Virus-transformed pre-B cells show ordered activation but not inactivation of immunoglobulin gene rearrangement and transcription. *J Exp Med* **173**, 711-720 (1991). https://doi.org/10.1084/jem.173.3.711
364. Carrozza, M.J. *et al*. Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* **123**, 581-592 (2005). https://doi.org/10.1016/j.cell.2005.10.023
365. Klein, B.J. *et al*. Recognition of cancer mutations in histone H3K36 by epigenetic writers and readers. *Epigenetics* **13**, 683-692 (2018). https://doi.org/10.1080/15592294.2018.1503491
366. Schwartzentruber, J. *et al*. Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. *Nature* **482**, 226-231 (2012). https://doi.org/10.1038/nature10833

367. Wu, G. *et al*. Somatic histone H3 alterations in pediatric diffuse intrinsic pontine gliomas and non-brainstem glioblastomas. *Nat Genet* **44**, 251-253 (2012). https://doi.org/10.1038/ng.1102
368. Behjati, S. *et al*. Distinct H3F3A and H3F3B driver mutations define chondroblastoma and giant cell tumor of bone. *Nat Genet* **45**, 1479-1482 (2013). https://doi.org/10.1038/ng.2814
369. Strahl, B.D. *et al*. Set2 is a nucleosomal histone H3-selective methyltransferase that mediates transcriptional repression. *Mol Cell Biol* **22**, 1298-1306 (2002). https://doi.org/10.1128/MCB.22.5.1298-1306.2002
370. Zhao, X.D. *et al*. Whole-Genome Mapping of Histone H3 Lys4 and 27 Trimethylations Reveals Distinct Genomic Compartments in Human Embryonic Stem Cells. *Cell Stem Cell* **1**, 286-298 (2007). https://doi.org/10.1016/j.stem.2007.08.004
371. Hsu, B.L. *et al*. Cutting Edge: BLyS Enables Survival of Transitional and Mature B Cells Through Distinct Mediators1. *The Journal of Immunology* **168**, 5993-5996 (2002). https://doi.org/10.4049/jimmunol.168.12.5993
372. He, T. *et al*. Histone acetyltransferase p300 acetylates Pax5 and strongly enhances Pax5-mediated transcriptional activity. *J Biol Chem* **286**, 14137-14145 (2011). https://doi.org/10.1074/jbc.M110.176289
373. Oksenych, V. *et al*. Acetyltransferases GCN5 and PCAF Are Required for B Lymphocyte Maturation in Mice. *Biomolecules* **12** (2021). https://doi.org/10.3390/biom12010061
374. Hisahara, S. *et al*. Histone deacetylase SIRT1 modulates neuronal differentiation by its nuclear translocation. *Proc Natl Acad Sci U S A* **105**, 15599-15604 (2008). https://doi.org/10.1073/pnas.0800612105
375. Morikawa, T. & Takubo, K. Hypoxia regulates the hematopoietic stem cell niche. *Pflugers Arch* **468**, 13-22 (2016). https://doi.org/10.1007/s00424-015-1743-z
376. Prozorovski, T. *et al*. Sirt1 contributes critically to the redox-dependent fate of neural progenitors. *Nat Cell Biol* **10**, 385-394 (2008). https://doi.org/10.1038/ncb1700
377. Takubo, K. *et al*. Regulation of the HIF-1alpha level is essential for hematopoietic stem cells. *Cell Stem Cell* **7**, 391-402 (2010). https://doi.org/10.1016/j.stem.2010.06.020
378. Kojima, H. *et al*. Abnormal B lymphocyte development and autoimmunity in hypoxia-inducible factor 1alpha -deficient chimeric mice. *Proc Natl Acad Sci U S A* **99**, 2170-2174 (2002). https://doi.org/10.1073/pnas.052706699
379. Blank, M.F. & Grummt, I. The seven faces of SIRT7. *Transcription* **8**, 67-74 (2017). https://doi.org/10.1080/21541264.2016.1276658
380. Yamamoto, M. *et al*. SIRT1 regulates adaptive response of the growth hormone--insulin-like growth factor-I axis under fasting conditions in liver. *Proc Natl Acad Sci U S A* **110**, 14948-14953 (2013). https://doi.org/10.1073/pnas.1220606110
381. Gruenbacher, S. *et al*. Essential role of the Pax5 C-terminal domain in controlling B cell commitment and development. *J Exp Med* **220** (2023). https://doi.org/10.1084/jem.20230260
382. Familiades, J. *et al*. PAX5 mutations occur frequently in adult B-cell progenitor acute lymphoblastic leukemia and PAX5 haploinsufficiency is associated with BCR-ABL1 and TCF3-PBX1 fusion genes: a GRAALL study. *Leukemia* **23**, 1989-1998 (2009). https://doi.org/10.1038/leu.2009.135
383. Zeng, L. *et al*. Structural basis of site-specific histone recognition by the bromodomains of human coactivators PCAF and CBP/p300. *Structure* **16**, 643-652 (2008). https://doi.org/10.1016/j.str.2008.01.010

384. Pillai, S. & Cariappa, A. The follicular versus marginal zone B lymphocyte cell fate decision. *Nat Rev Immunol* **9**, 767-777 (2009). https://doi.org/10.1038/nri2656
385. Yang, J. *et al*. MDMX Recruits UbcH5c to Facilitate MDM2 E3 Ligase Activity and Subsequent p53 Degradation In Vivo. *Cancer Res* **81**, 898-909 (2021). https://doi.org/10.1158/0008-5472.CAN-20-0790
386. Jumper, J. *et al*. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589 (2021). https://doi.org/10.1038/s41586-021-03819-2
387. Garvie, C.W. *et al*. Structural studies of Ets-1/Pax5 complex formation on DNA. *Mol Cell* **8**, 1267-1276 (2001). https://doi.org/10.1016/s1097-2765(01)00410-5
388. Eberhard, D. *et al*. Transcriptional repression by Pax5 (BSAP) through interaction with corepressors of the Groucho family. *EMBO J* **19**, 2292-2303 (2000). https://doi.org/10.1093/emboj/19.10.2292
389. Eberhard, D. & Busslinger, M. The partial homeodomain of the transcription factor Pax-5 (BSAP) is an interaction motif for the retinoblastoma and TATA-binding proteins. *Cancer Res* **59**, 1716s-1724s; discussion 1724s-1725s (1999).
390. Stortz, M. *et al*. Transcriptional condensates: a blessing or a curse for gene regulation? *Commun Biol* **7**, 187 (2024). https://doi.org/10.1038/s42003-024-05892-5
391. Wagh, K. *et al*. Phase separation in transcription factor dynamics and chromatin organization. *Curr Opin Struct Biol* **71**, 148-155 (2021). https://doi.org/10.1016/j.sbi.2021.06.009
392. Zhang, M. *et al*. Transcription factor Hoxb5 reprograms B cells into functional T lymphocytes. *Nat Immunol* **19**, 279-290 (2018). https://doi.org/10.1038/s41590-018-0046-x
393. Dai, H. *et al*. Sirtuin activators and inhibitors: Promises, achievements, and challenges. *Pharmacol Ther* **188**, 140-154 (2018). https://doi.org/10.1016/j.pharmthera.2018.03.004