Fall 2010

Identification of Gon4-like as a factor that is essential for B lymphopoiesis and capable of mediating transcriptional repression

Ping Lu
University of Iowa

IDENTIFICATION OF GON4-LIKE AS A FACTOR THAT IS ESSENTIAL FOR B LYMPHOPOIESIS AND CAPABLE OF MEDIATING TRANSCRIPTIONAL REPRESSION

by

Ping Lu

An Abstract

Of a thesis submitted in partial fulfillment
of the requirements for the Doctor of
Philosophy degree in Immunology
in the Graduate College of
The University of Iowa

December 2010

Thesis Supervisor: Professor Paul B. Rothman

ABSTRACT

The B cell population is one of the key components of the adaptive immune system, which protects the host from a tremendous variety of pathogens by producing antibodies. B cells develop from hematopoietic stem cells through a pathway known as B lymphopoiesis. This is a process accompanied by intensive gene expression reprogramming. By the end, genes appropriate for the B lineage are activated and those that are not are continuously repressed. The regulation of lineage gene expression is conferred by a network of transcriptional regulators. Although some key components have been defined, more factors, especially those orchestrating the repression of non-B lineage genes, remain to be identified.

Chemically induced mutagenesis is a potent way of identifying genes with critical biological functions. Injection of n-ethyl-n-nitrosourea, a mutagen, has generated a unique point mutation in the mouse *Gon4-like* (*Gon4l*) gene that specifically causes a loss of peripheral B cells while maintaining the T cell population. The mutation is therefore named *Justy* for Just T cells. The goal of this thesis project is to analyze the *Justy* mice and provide insights into the mechanisms underlying the regulation of B lymphopoiesis.

The work presented here demonstrates that the protein encoded by *Gon4l* is essential for early B lymphopoiesis, which is likely through the repression of non-B lineage genes. Gon4l protein contains conserved domains implicated in transcriptional repression and associates in a complex with the transcriptional repression mediators Yin Yang 1 and Sin3a/HDAC1, after these proteins are transiently expressed in cell lines. When bound to DNA, Gon4l is capable of repressing a nearby promoter and this function correlates with its ability to form a complex. Therefore, these results suggest that Gon4l may function as a transcriptional regulator by employing its associated co-factors in the identified complex. Lastly, a wide spectrum of tumors developed in *Justy* mice, indicating that Gon4l can also act as a tumor suppressor.

Abstract Approved: _____
Thesis Supervisor

_____
Title and Department

_____
Date

IDENTIFICATION OF GON4-LIKE AS A FACTOR THAT IS ESSENTIAL FOR B LYMPHOPOIESIS AND CAPABLE OF MEDIATING TRANSCRIPTIONAL REPRESSION

by

Ping Lu

A thesis submitted in partial fulfillment
of the requirements for the Doctor of
Philosophy degree in Immunology
in the Graduate College of
The University of Iowa

December 2010

Thesis Supervisor: Professor Paul B. Rothman

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

_____

PH.D. THESIS

_____

This is to certify that the Ph.D. thesis of

Ping Lu

has been approved by the Examining Committee
for the thesis requirement for the Doctor of Philosophy
degree in Immunology at the December 2010 graduation.

Thesis Committee: _____
Paul B. Rothman, Thesis Supervisor

_____
John D. Colgan

_____
Thomas J. Waldschmidt

_____
Jerrold P. Weiss

_____
Pamela K. Geyer

ACKNOWLEDGMENTS

The journey toward a Ph.D. is long and I would have never been able to reach this far if without the help and support from my mentors, teachers, friends, and family.

I would like to first express my deepest gratitude to my mentor Dr. Paul Rothman for training and guiding me, and, most import of all, providing me with the freedom to explore science. His expectation of hard work has enabled me to learn more about scientific rigor and his trust has helped me becoming more and more independent in doing research. In addition, his invaluable advice is always available at times of difficulty, which has helped me greatly to move the research forward.

I also owe my sincere appreciation to my supervisor Dr. John Colgan who has helped me with details of all the experiments as well as writing scientific articles. This thesis would not have been possible if without his kind help. I would like to express my appreciation to my thesis committee members, Drs. Thomas Waldschmidt, Jerrold Weiss and Pamela Geyer. I have always been inspired by the discussion during our committee meetings. My committee has taught me not only the experiments to do, but also and most importantly, the ways of scientific thinking.

I would like to extend my thanks to all the members of the Rothman and Colgan labs for their kind help. I am especially thankful to Dr. Bruce Hostager for his kindness in teaching me critical techniques and discussing science with me. I would like to express my deep appreciation to Isaiah Hankel, who is a colleague and a very close friend of mine. We have worked on the same project for years, during which we shared a lot of laughters and good science together. It would be very hard to imagine that the project could be kept going without his dedication. I would like to also thank Miss Janet Brenton for her diligent work in doing research with me.

Next, I would like to thank the Immunology Program at the University of Iowa for giving me an open-mind atmosphere and supporting me through the whole training process. I also want to express my thanks to my classmate and friends.

I would also like to thank Dr. Donald Auger for teaching me good science and providing me with his generous help.

Lastly I want to thank my family, especially my wife. Thank you for your support and love. I know I would not have done anything if I had not had you all.

ABSTRACT

The B cell population is one of the key components of the adaptive immune system, which protects the host from a tremendous variety of pathogens by producing antibodies. B cells develop from hematopoietic stem cells through a pathway known as B lymphopoiesis. This is a process accompanied by intensive gene expression reprogramming. By the end, genes appropriate for the B lineage are activated and those that are not are continuously repressed. The regulation of lineage gene expression is conferred by a network of transcriptional regulators. Although some key components have been defined, more factors, especially those orchestrating the repression of non-B lineage genes, remain to be identified.

Chemically induced mutagenesis is a potent way of identifying genes with critical biological functions. Injection of n-ethyl-n-nitrosourea, a mutagen, has generated a unique point mutation in the mouse *Gon4-like* (*Gon4l*) gene that specifically causes a loss of peripheral B cells while maintaining the T cell population. The mutation is therefore named *Justy* for Just T cells. The goal of this thesis project is to analyze the *Justy* mice and provide insights into the mechanisms underlying the regulation of B lymphopoiesis.

The work presented here demonstrates that the protein encoded by *Gon4l* is essential for early B lymphopoiesis, which is likely through the repression of non-B lineage genes. Gon4l protein contains conserved domains implicated in transcriptional repression and associates in a complex with the transcriptional repression mediators Yin Yang 1 and Sin3a/HDAC1, after these proteins are transiently expressed in cell lines. When bound to DNA, Gon4l is capable of repressing a nearby promoter and this function correlates with its ability to form a complex. Therefore, these results suggest that Gon4l may function as a transcriptional regulator by employing its associated co-factors in the identified complex. Lastly, a wide spectrum of tumors developed in *Justy* mice, indicating that Gon4l can also act as a tumor suppressor.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

BCR	B cell receptor
BM	Bone marrow
Cebpα	CCAAT/enhancer binding protein gene
CLP	Common lymphoid progenitor
CMP	Common myeloid progenitor
co-IP	Co-immunoprecipitation
Csf1r	Colony stimulating factor 1 receptor gene
EBF	Early B cell factor (also known as Ebf-1)
$E_{i\mu}$	IgH intronic enhancer
ELP	Early lymphoid progenitor
ENU	N-ethyl-n-nitrosourea
ETP	Early T lineage progenitor
FACS	Fluorescence activated cell sorter
Flt3	FMS-related tyrosine kinase 3
Gal4-DBD	DNA-binding domain of Gal4
GMP	Granulo-monocytic progenitor
Grg4	Groucho-related-gene 4
HAT	Histone acetyltransferase
HDAC	Histone deacetylases complex
HSC	Hematopoietic stem cell
IB	Immunoblotting
IgH	Immunoglobulin heavy-chain
IgL	Immunoglobulin light-chain
IL-7	Interleukin 7
IL-7Rα	Interleukin 7 receptor alpha chain

IP	Immunoprecipitation
Lin^-	Lineage marker negative
LMPP	Lymphoid primed multipotent progenitor
LT-HSC	Long-term hematopoietic stem cell
MEP	Megakaryo/erythroid progenitor
MPP	Multipotent progenitor
NCoR	Nuclear receptor corepressor
NLS	Nuclear localization signal
NuRD	Nucleosome remodeling and deacetylase
PAH	Paired amphipathic helices
pre-BCR	Pre-B cell receptor
Q-RT-PCR	Quantitative RT-PCR
RAG	Recombinase-activating genes
RSS	recombination signal sequences
SANT	SWI3, ADA2, N-CoR and TFIIIB
SAP30	Sin3a-associated polypeptide p30
SMRT	Silencing mediator for retinoic acid receptor and thyroid hormone receptor
ST-HSC	Short-term hematopoietic stem cell
Udu	Ugly duckling (Gon4l ortholog in zebrafish)
WB	Western blot
WCL	Whole cell lysate
WT	Wild type
YY1AP	YY1-associated protein

CHAPTER I

GENERAL INTRODUCTION

Overview of B lymphopoiesis

B lymphocytes (B cells) produce antibodies and protect the host from infection by a tremendous variety of pathogens (1). B cells are generated from self-renewing hematopoietic stem cells (HSC) through a process called B lymphopoiesis, which takes place in the fetal liver during gestation, and in the bone marrow (BM) thereafter (2-4). During B lymphopoiesis, HSCs proceed through a series of developmental stages and become immature B cells which then migrate to secondary lymphoid organs (*i.e.* lymph nodes and spleen) to undergo full maturation (5). The progression of B lymphopoiesis is characterized by distinct and intensive gene expression reprogramming at each developmental stage and the ordered recombination of the immunoglobulin heavy and light chain genes (1).

In the hematopoietic system, HSCs give rise to cells that can be categorized into either the lymphoid (cells of the lymphatic system) or the myeloid lineage (leukocytes that are not lymphocytes) (6,7). The lymphoid lineage consists of T, B, and natural killer (NK) cells, while the myeloid lineage contains megakaryocytes, granulocytes, and macrophages. Monoclonal antibodies have been identified that bind surface proteins expressed on subsets of hematopoietic cells (surface markers). When combined with analysis of lineage differentiation potential, surface marker analysis has been a powerful tool to define early hematopoiesis and B lymphopoiesis (3). Based on such analyses, the B cell development pathway can be divided into multiple stages (Figure 1). Cells of the earliest stage are the long-term HSCs (LT-HSCs) which are capable of replenishing all blood cell types (multi-potency) and undergoing unlimited self-renewal. The characteristic surface marker associated with LT-HSCs is the high expression of c-Kit (CD117), (the receptor of the stem cell factor (SCF)). LT-HSCs are devoid of markers of

any differentiated lineages (Lineage negative or Lin^-).The short-term HSC (ST-HSC) is a transition from the LT-HSCs to a stage where the cells maintain multi-potency but have limited self-renewal capacity (3,8). The multipotent progenitors (MPPs) are the immediate progeny cells of ST-HSCs. These cells have lost the ability to undergo extensive self-renewal but retain the capability of differentiating into multiple lineages. The LT-HSCs, ST-HSCs, and MPPs in adult mouse bone marrow (BM) are found in the $Lin^-Sca\text{-}1^+c\text{-}kit^{high}$ compartment (termed as the LSK subset), which constitutes about 0.1% of the whole bone marrow cells (9). The differentiation of lymphoid and myeloid lineages diverge at the MPP stage (10-12), where the MPPs can be subdivided into two populations based on the expression of FMS-related tyrosine kinase 3 (Flt3; also known as Flk2)(13). The $Flt3^-$MPPs have the potential to differentiate into megakaryocytes and erythrocytes and are designated as megakaryo/erythroid progenitors (MEP). The MPPs that have up-regulated Flt3 expression possess the potential to differentiate into lymphoid cells but have very little potential to differentiate into MEPs. The transition of MPPs from multi-potency to the acquisition of lymphoid differentiation potential is known as lymphoid lineage specification (12). Gene expression analysis reveals that transcription of granulocyte-monocyte genes and common lymphoid genes (*i.e.* IL-7Rα) is upregulated in $Flt3^+$MPPs, while that for erythroid and megakaryocyte development is downregulated (14-16). The change of gene expression profile in $Flt3^+$MPPs that facilitates differentiation along the lymphoid pathways is known as lymphoid priming and the $Flt3^+$MPPs are thus referred to as lymphoid primed MPP (LMPP) (16,17). A subpopulation of LMPP can only give rise to lymphoid cells and is thus lymphoid restricted. These cells express TdT (terminal deoxynucleotidyltransferase) and are referred to as the early lymphoid progenitors (ELPs) (3,18,19). Following the expression of TdT, the *recombinase-activating genes* (*Rag1* and *Rag2*), particularly *Rag1*, are expressed in a fraction of TdT^+ELPs, which further make these cells lymphoid biased (18). Consistent with Rag and TdT expression, gene rearrangement at the

immunoglobulin heavy-chain (*IgH*) locus (D_H-J_H rearrangement, see below) is initiated at the ELP stage, although at a low level (2,18,20). ELPs give rise to the common lymphoid progenitors (CLP), which have the potential to differentiate into B cells, natural killer (NK) cells, dendritic cells (DC), or T lymphocytes in culture (21). However, *in vivo* experiments suggest that CLPs primarily give rise to B cells and NK cells (22-24). Some ELPs, especially those that express CCR9 and CD62L, can serve as thymus-seeding progenitors and have been proposed to be the precursors of early T lineage progenitors (ETPs) (25).

For CLPs to differentiate along the B lineage pathway, genes associated with the B lineage need to be activated, a process known as B lineage specification (3,4,26,27). The earliest identifiable B lineage specified precursor is the pre-pro B cells (also known as fraction A cells in the Hardy fractionation nomenclature). These cells express a B lineage specific marker B220 (3,4). Pre-pro B cells give rise to the early pro B cells which express another B cell marker—CD19. D_H to J_H rearrangement is completed in the pro-B cell stage. At the same time, genes of other cell lineages as well as their potential to differentiate into those lineages (developmental plasticity) are repressed in pro-B cells. This process is known as B lineage commitment (2-4). In the succeeding late pro-B stage (also known as fraction C cells), V_H to DJ_H gene segment rearrangement occurs (2,4,28). A successful V_H to DJ_H rearrangement results in a functional Igμ chain expressed on the cell surface. This Ig heavy chain is paired with surrogate light chains (consisting of VpreB and λ_5) to form a pre-B cell receptor (pre-BCR). B cell precursors expressing a pre-BCR are known as the pre-B cells (2,4,28). The light chain gene is then rearranged and if it is successful, surface IgM is expressed. This cell is defined as an immature B cell. Immature B cells migrate from the bone marrow afterwards and enter the peripheral lymphoid organs where they differentiate into mature B cells (2,5).

V(D)J rearrangement during B lymphopoiesis

Along with the phenotypic changes on the cell surface, early B cell development is associated with ordered non-homologous rearrangement of the immunoglobulin heavy (*IgH*) and light (*IgL*) chain genes, which allows the generation of a BCR pool with enormous binding specificities (28). Successful rearrangement of the immunoglobulin genes is required for B cell development. Mutations affecting proteins involved in this process lead to arrest of early B cell development (5,28). The immunoglobulin heavy chain gene consists of the variable (V), diversity (D) and joining (J) segments while the light chain gene contains only V and J segments (5). All of the V, D, and J gene segments are flanked by recombination signal sequences (RSS). RSS are recognized by the recombinases, RAG1 and RAG2, which generate double stranded DNA breaks at the site (5,29). Additional nucleotides are then randomly added or cleaved by TdT or exonucleases, respectively, at the DNA ends. These ends are joined then by DNA ligase IV, irrespective of their sequence homologies.

The *IgH* locus is rearranged prior to the *IgL* locus. The heavy chain gene rearrangement can occur as early as in the ELP stage, where RAG proteins start to be expressed (5,28). The D_H to J_H gene segments are recombined first at around the CLP and pre-pro B stage. The arranged DJ_H segment is then joined by the V_H region at the later pro-B stage. Upon an in-frame VDJ rearrangement at the heavy chain gene, a pre-B cell receptor (pre-BCR) comprising the rearranged μ chain, the surrogate light chains (a heterodimer of VpreB and λ5) and the $Ig_{\alpha/\beta}$ dimer is expressed on the cell surface (28,30). The expression of the pre-BCR indicates that the cells have progressed from the pro-B stage into the pre-B-cell stage.

The pre-BCR is a major check point in B cell development (31,32). Only those cells that express a functional receptor can progress further down along the developmental pathway. Mice harboring a deletion of the exons encoding the transmembrane region of the Ig heavy chain (μMT mice) display a developmental arrest

at the pro-B to pre-B transition (33). Similarly, *λ5* deficiency leads to a block at the pro-B to pre-B transition (34). Signaling through the pre-BCR also inhibits V to DJ recombination at the other heavy chain allele (known as allelic exclusion), which ensures that only BCRs of a single specificity are expressed by a given cell. Pre-B cells that have just finished a successful heavy chain gene recombination are relatively large in size and are designated as large pre-B cells (5). Pre-BCR signaling stimulates large pre-B cells to proliferate and differentiate into cells of a smaller size which are known as the small pre-B cells. In the small pre-B cells, rearrangement of the *Ig light chain* gene occurs (35). After the light chain rearrangement is complete, the rearranged heavy and light chains, together with the $Ig_{\alpha/\beta}$ heterodimers, form a B-cell receptor, which signifies the completion of B cell development in bone marrow (1,31,32).

The regulation of V(D)J recombination is complicated and involves a variety of mechanisms. Among those, chromatin accessibility seems to play a central role (28). During V(D)J rearrangement, chromatin of the *IgH* locus is modified to adopt an "open" conformation (*i.e.* high levels of histone acetylation), and is accompanied with increased germline transcript expression (28,36). In *RAG* deficient pro-B cells, D_H to J_H recombination is blocked, due to the lack of recombinases, but the DJ region is hyperacetylated (37-39). The DJ region also associates with BRG1, one of the subunits of the chromatin remodeling complex SWI/SNF, in transformed *RAG* deficient pro-B cell lines, suggesting the region is undergoing active chromatin remodeling (39). Lastly, the *IgH* locus undergoes observable changes such as periphery-to-center nuclear repositioning and chromatin contraction mediated DNA looping, which is consistent with the profound changes of the chromatin structure (40-42). Thus, V(D)J rearrangement is a key event during B lymphopoiesis and a successful rearrangement requires regulation of the chromatin structure.

The developmental microenvironment of B lymphopoiesis

B lymphopoiesis takes place in microenvironments created by adhesive stromal cells in the bone-marrow, known as the niches. The niches are thought to provide precursor cells developmental cues by secreting several cytokines including stem cell factor, Flt3 ligand and interleukin 7. These cytokines can guide HSCs to differentiate along the B lymphopoiesis pathway (43). However, stromal cells do not have distinctive characteristics, which has made it difficult to identify the niches for different hematopoietic cells.

Stem cell factor

Stem cell factor (SCF) is a ligand for the receptor tyrosine kinase c-Kit, which is expressed at high levels in the multipotent progenitor stages and is down-regulated at the pre-B stage (Figure 1) (3,44). Activation of c-kit by SCF promotes survival and induces rapid proliferation of HSCs (45). However, the original undifferentiated state of HSCs is not maintained after such proliferation and the cells differentiate to different hematopoietic progenitors (43,46). *c-Kit* deficiency in mice (Vickid mice) leads to a severe reduction of CLPs and an almost complete loss of subsequent B lineage cells. This phenotype demonstrates that c-kit is required for CLP development (43,46).

Flt3 ligand

Flt3 ligand (Flt3L, also known as Flk2 ligand) is a ligand for the Flt3 receptor (also known as Flk2), a member of the class III growth factor receptor tyrosine kinase family (2). Flt3 bears strong homology to other family members such as the macrophage colony-stimulating factor (M-CSF) receptor and c-Kit (8). Flt3 is expressed by cells from the LMPP stage to the pre-pro-B stage in bone marrow (Figure 1). Signaling through Flt3 is crucial for the generation of CLPs and likely LMPPs (47-50). Mice deficient of *Flt3* have approximately two-fold less CLP, pre-pro-B, and pro-B cells. However, these mice have relatively normal hematopoietic populations in the periphery (43,49,50). The *Flt3l*$^{-/-}$

mice generated by targeted gene disruption, on the other hand, have a more pronounced phenotype (47). These mice have a severe reduction of B-lineage progenitors (CLP and pre-pro-B) and a great loss of leukocyte cellularity in bone marrow, peripheral blood, lymph nodes, and spleen (43,48). The different phenotypes observed between *Flt3*- and *Flt3l*-deficient mice suggest there might be an unidentified receptor for Flt3l that is also critical for B cell development. Moreover, *Flt3l*$^{-/-}$ HSCs are impaired in their ability to differentiate into myeloid and lymphoid cells (47) and *in vivo* administration of Flt3l increases the number of early progenitors with lymphoid and myeloid potential (51). Flt3l also collaborates with IL-7 in the regulation of early lymphopoiesis. The presence of Flt3l enhances expansion of pre-pro-B and pro-B cells in response to IL-7 *in vitro* (52). More strikingly, *Flt3l* and *Il7rα* (encodes IL-7 receptor α chain, see below) double knockout mice lack all stages of committed B cell progenitors in both fetal liver and bone marrow and, as a consequence, have no B cells in the periphery (53). Thus, Flt3 is not only an important marker to distinguish LMPPs (Flt3^{+}) from MPPs, but also a critical signaling molecule that determines the lymphoid and likely B lineage fate of the progenitor cells.

Interleukin-7

Interleukin-7 (IL-7) is produced by bone marrow stromal cells and is an essential cytokine for B cell development. The receptor of IL-7 (IL-7R) is composed of two chains: the ligand-specific α chain and the common γ chain (γ_c) that is shared with other cytokines (IL-2, 4, 9 and 15) (54). Signaling through IL-7R activates the JAK-STAT5a/b and PI3-kinase (PI3K) pathways which in turn activate more downstream transcription factors such as c-myc, NFAT and AP-1 (54-56). The net effect of IL7-IL7R signaling is to promote survival and VDJ recombination of the Ig heavy chain gene. In mice deficient of *IL-7* or *IL-7Rα*, the number of CLPs is reduced by about 3-fold (57,58). More importantly, *IL-7*$^{-/-}$ or *IL-7Rα*$^{-/-}$ CLPs are severely compromised in their ability to differentiate into pre-pro-B cells and to proliferate in response to cytokine stimulation

(57,58). V to DJ rearrangement at the *IgH* locus is also impaired in these mice (59,60). Mechanistic studies suggest that the defects seen in *Il7*$^{-/-}$ or *Il7rα*$^{-/-}$ CLPs is due to their inability to induce the key B lymphopoiesis transcription factor, EBF (57,58). The differentiation potential of *IL-7*$^{-/-}$ CLPs is restored when EBF is over-expressed (57,58). Therefore, IL-7 is another key cytokine that functions mainly at the CLP stage by inducing the critical transcription factor, EBF (43).

B lymphopoiesis is regulated by a network of transcription factors

The progression of B lymphopoiesis is regulated by a network of transcription factors. Microarray analysis reveals that B cell development coincides with drastic gene expression changes, including the up-regulation of lymphoid and B lineage appropriate genes (lymphoid and B lineage specification) and continuous repression of non-B lineage genes (B lineage commitment) (4,20). More than a dozen transcription factors have been demonstrated to regulate early B cell development, which include Ikaros, PU.1, E2A, EBF and Pax5 (2,4) (Figure 2).

Transcription factors that regulate lymphoid specification and commitment

Ikaros

Ikaros is a zinc finger transcription factor encoded by the *Ikzf1* gene. The Ikaros protein family also includes the transcription factors Helios and Ailolos. All the Ikaros family proteins contain multiple zinc fingers used to form homo- or hetero- dimers and to bind DNA (61).

Ikaros plays a critical role in early lymphoid specification and commitment. Deletion of the DNA-binding zinc fingers from Ikaros generates a dominant negative (DN) Ikaros protein. Mice homozygous of the *Ikaros DN* allele (*Ikaros*$^{DN/DN}$) cannot

generate any B, T, dendritic cells (DC), or NK cells (62). Targeted deletion at the C-terminus of the *Ikaros* gene removes sequences encoding the zinc fingers required for Ikaros protein dimerization and thus makes a functional null Ikaros protein (63). Mice homozygous for this *Ikaros null* allele (*Ikzf1*$^{-/-}$) do not have detectable CLP population and therefore cannot not generate any B and NK cells. The differentiation of the erythroid and myeloid lineages is not affected by the loss of Ikaros function (63-65). Lin^{-}c-kit^{low} progenitor cells from *Ikzf1*$^{-/-}$ mice, express higher levels of the granulocyte-macrophage-colony stimulating factor receptor α (*GM-Csfrα*) mRNA compared to the WT, suggesting that Ikaros promotes a bias toward lymphoid lineage by repressing myeloid genes (66). In addition, Ikaros is important for the expression of Flt3 and c-Kit in HSCs and MPPs, which are required for the generation of LMPPs and CLPs (4,66). In later stages of B cell development, Ikaros has been found to collaborate with Aiolos to regulate B cell maturation and proliferation (2,67). Mechanistically, Ikaros-mediated transcriptional regulation is likely achieved by modifying chromatin structure. Both the transcriptional co-repressors (*i.e.* Sin3/HDAC containing complexes, see below) and chromatin remodeling complexes (*i.e.* NuRD: nucleosome remodeling and deacetylase, or SWI/SNF, see below) have been shown to be recruited through protein interactions by Ikaros to bring about transcriptional activation and repression, respectively (68-71).

PU.1

PU.1 is encoded by the *Sfpi1* gene and is a member of the ETS transcription factor family. It is expressed in HSCs, CLPs and CMPs (72) and all differentiating hematopoietic cells except erythroblasts, megakaryocytes, and T cells (73-75). *Sfpi1* null mutation in mice is lethal (76). Analysis of the *Sfpi1*$^{-/-}$ embryos reveals a lack of lymphoid and myeloid progenitors (77,78). But the development of erythrocytes and megakaryocytes is not affected, suggesting that PU.1 acts within the LMPPs. Moreover, *Sfpi1*$^{-/-}$ progenitor cells do not respond to IL-7 or differentiate into pro-B cells (78), which

is probably due to the regulation of the IL-7R expression by PU.1 (76). Conditional deletion of *Sfpi1* by the IFNαβ-inducible *MxCre* transgene leads to the loss of the CLP population in bone marrow and excessive production of granulocytes (79). However, deletion of *Sfpi1* from CLPs or committed B cells has no effect on B cell development (80,81), suggesting that PU.1 is required for lymphoid specification of the early progenitors but not for those at later stages where commitment to the B lineage has been completed (2,4).

Transcription factors that regulate B lineage specification

E2A

The transcription factors E12 and E47 are splicing variants encoded by the *Tcfe2a* gene and are members of the basic helix-loop-helix (bHLH) protein family(4). E12 and E47 are collectively known as E2A. E2A proteins are expressed in a wide range of tissues where they form heterodimers with other bHLH proteins in a tissue-specific manner. In B lineage cells, however, E2A proteins need to form homodimers to function (82). Mice deficient of *Tcfe2a* have B cell development arrested at the pre-pro-B stage (83-85). In addition, the CLP population is also affected (86). A small number of $Tcfe2a^{-/-}$ B cell precursors can proceed into the pro-B stage, but they cannot undergo D to J rearrangement at the *IgH* locus (83-85). Also these cells cannot express key B lineage genes such as *CD19*, *mb-1*, *λ5*, and *Pax5* (84). These observations suggest that E2A functions at the very early stage of B cell development where B lineage specification occurs. On the molecular level, E2A binds and activates the promoter for the *Ebf-1* gene (see below), which encodes another key transcription factor of B lymphopoiesis, EBF (87-89). After being induced, EBF collaborates with E2A to activate B lineage genes such as *mb-1* and λ_5 and also to initiate VDJ recombination at the *IgH* locus. However, ectopic expression of EBF does not induce E2A, indicating that E2A functions upstream

of EBF in the pathway (88). Studies on E47 and E12 indicate that both are required for B lineage specification since disruption of either blocks pro-B cell development (83,84,90).

The function of E2A is regulated by the HLH inhibitors of the Id family (inhibitor of DNA binding). These are proteins that have the HLH motifs but do not bind DNA. The Id proteins form heterodimers with E2A, and prevent E2A from binding target genes (2,91). Expression of *Id1* from a B-cell-specific transgene results in a B cell development block at the pre-pro-B stage similar to that in the *Tcfe2a*$^{-/-}$ mice (92). Forced expression of Id2 or Id3 antagonizes the activity of E2A and promotes the differentiation of hematopoietic progenitors into DC and NK cells but not B cells (93,94). On the other hand, *Id2*$^{-/-}$ mice lack NK cells and splenic CD8α^{+} DCs and Langerhans cells (95,96). Thus E2A proteins promotes the progenitor cells to take the B lineage fate whereas the Id proteins counteract E2A and direct the progenitor cells to a DC or NK cell fate (2).

EBF

As mentioned above, one of the key B lineage transcription factors induced by E2A is EBF (early B cell factor, also known as Ebf-1). Unlike E2A, EBF is expressed only in B lineage cells including pro-B, pre-B, and mature B cells (97). Mice deficient of *EBF* (*Ebf-1*$^{-/-}$) have a B cell development arrest at the pre-pro-B stage. This phenotype is similar to that is observed in the *E2A* deficient mice (98). A small number of *EBF*$^{-/-}$ B cell precursors can progress into the pro-B stage in bone marrow. However, these cells cannot undergo VDJ rearrangement of the *IgH* locus and fail to express key B lineage genes. These B lineage genes include *Pax5* gene (see below) and those encoding components of the pre-B cell receptor such as *Igα* (also known as *Cd79a,* or *mb-1*), *Igβ* (*Cd79b*), *λ5*, and *VpreB1* (98). On the other hand, forced expression of EBF in HSCs biases these cells toward the B lineage differentiation (99). Ectopic expression of EBF also rescues B cell differentiation from multipotent progenitor cells deficient of *PU.1*,

E2A and *IL-7R* (57,58,100,101). Thus, EBF is capable of activating the B cell-lineage genes and instructing the precursor cells to take a B lineage fate.

The similar phenotypes observed in *EBF* and *E2A* deficient mice strongly suggest that the two proteins function together. In support of this, ectopic expression of EBF and E2A results in activation of endogenous *λ5* and *VpreB* genes (102). E2A and EBF both bind the promoter region of *λ5* (102,103), *VpreB* (104), and *mb-1* (*Igα*) (87,105); and synergistically activate these genes. In addition, bone marrow precursor cells from *Tcfe2a*$^{+/-}$*Ebf-1*$^{+/-}$ mice fail to express *λ5, VpreB,* and *mb-1* which overlap with those activated by ectopically expressed E2A and EBF (106). Lastly, B cell development is blocked at the early pro-B cell stage in *Tcfe2a*$^{+/-}$*Ebf-1*$^{+/-}$ mice.

E2A and EBF also regulate *Ig* rearrangement. Both E2A and EBF induce the expression of RAG recombinases (83,88,98,106). In addition, E2A binding sites are found in the immunoglobulin intronic enhancer ($E_{i\mu}$), which is essential for V(D)J rearrangement (2,107,108). Consistently, ectopic expression of E2A induces *IgH* VDJ recombination in a pre-T cell line (109). Even more strikingly, ectopic expression of E2A and EBF together with RAG recombinases can induce *Ig* recombination in a non-lymphoid cell line (110). In these settings, the chromatin accessibility of the *Ig* loci seems to be coordinately regulated by E2A and EBF (111). One possible underlying mechanism is that E2A recruits the SAGA histone acetyltransferase complex to open up the *Ig* loci (2,112). Thus, E2A and EBF seem to coordinately regulate early B lineage specification by initiating the expression of key B lineage genes and V(D)J recombination.

Transcription factors that regulate B lineage commitment

Pax5

The induction of B lineage genes by E2A and EBF is not sufficient to restrict the precursor cells to a B cell fate. The full commitment requires the transcription factor Pax5 (also known as B-cell-specific activator protein or BSAP). Pax5 contains a

conserved N-terminal paired-domain motif to bind DNA and can activate or repress transcription depending on the interaction with distinct protein partners (2,4,113). Key B lineage genes regulated by Pax5 include *mb-1* (*Igα*) and *CD19* (113).

Pax5 is exclusively expressed in B lineage cells from the pro-B stage to the mature B cell stage (2,4,114). The expression of Pax5 is coordinately regulated by E2A and EBF. *Pax5* mRNA is greatly reduced in *Tcfe2a*$^{+/-}$*Ebf-1*$^{+/-}$ pro-B cells (106). In addition, EBF binds the *Pax5* promoter and such binding is required for *Pax5* expression (106). However, ectopic expression of Pax5 cannot rescue B lymphopoiesis from *EBF* deficient cells, indicating that EBF plays more roles than just to activate *Pax5* (57,58).

Inactivation of *Pax5* results in B cell development arrest at the pro-B stage, despite the fact that these *Pax5*$^{-/-}$ cells express many B lineage genes and have completed D_H to J_H rearrangement (115,116). Although *Pax5*$^{-/-}$ pro-B cells are unable to differentiate into B cells, these cells can differentiate into other cell types including T cells, NK cells, dendritic cells, and macrophages (117,118). Consistently, gene expression analysis reveals that *Pax5*$^{-/-}$ pro-B cells express many non-B lineage genes, including the macrophage and T cell genes, *Csf1r* and *Notch1* (119-122). These observations signify the importance of alternative lineage gene repression mediated by Pax5 in the development of B cells. Restoration of *Pax5* in *Pax5*$^{-/-}$ pro-B cells rescues B cell differentiation and inhibits differentiation into other lineages (123). In addition, conditional inactivation of *Pax5* in pro-B or even in mature B cells results in upregulation of many otherwise repressed non-B lineage genes (119). Recent biochemical studies indicate that Pax5 may employ a Groucho family transcriptional co-repressor, Grg4, to promote gene repression (124,125). The Groucho proteins interact with histone deacetylases (HDACs) and play a central role in transcriptional repression (126-128). Therefore, Pax5 facilitates B lineage commitment by repressing genes associated with other lineages.

Regulation of gene expression by epigenetic modification

One important mechanism that regulates lineage gene expression is chromatin remodeling (129-134). In eukaryotic cells, DNA is wrapped around histones and packaged into higher order structures. The packaging of DNA occludes regulatory DNA elements from being bound by transcription factors and is thus inhibitory to transcription. Therefore, in order for a gene to be transcribed, its associated chromatin needs to take on an "open" conformation, where the binding sites for transcription factors are exposed. Conversely, changing the chromatin into a "compact" conformation can be utilized for transcriptional repression (61,135). Chromatin conformational change is brought about by multi-subunit protein complexes that contain enzymes that covalently modify histone tails or move nucleosomes along the DNA (129-134,136). To date, several of these enzymes have been identified. Histone acetyltransferases (HATs) covalently add acetyl groups to histone tails, which reduces the affinity of histone-DNA interaction and thus facilitates transcriptional activation (137,138). Histone deacetylases (HDACs), on the other hand, remove the acetyl groups from histones, enhancing the interaction between histones and DNA and is associated with transcriptional repression (135). The SWI/SNF complex is one of the typical chromatin remodeler complexes and has the ability to reposition nucleosomes by hydrolyzing ATP (139). In addition, DNA methylation, micro-RNA (140) and other mechanisms also act in concert with these enzymes to regulate transcription. Thus, chromatin remodeling can regulate gene expression epigenetically, without affecting the genomic DNA sequence.

The recruitment of HDACs to target gene promoters can be achieved via transcription factors that either directly associate with HDACs (*i.e.* YY1, see below) or associate with the Sin3 co-repressors that contain HDACs (141-151). During B cell development, Ikaros is likely the transcription factor responsible for restricting multipotent progenitor cells to the lymphoid lineage by repressing myeloid genes (66). Ikaros associates with HDACs, Mi-2 (an ATPase), and other components of the NURD

complex (nucleosome remodeling and histone deacetylation) to mediate nucleosome disruption as well as histone deacetylation (69). Studies show that Ikaros also associates with the Sin3 co-repressors to confer potent transcriptional repression, which is dependent on the associated HDAC activities and the hypo-deacetylation state of the target promoters (68,71). Similarly, Pax5 recruits the Grg4 co-repressor that interacts with HDACs to repress transcription (124-128,152,153). This mechanism has been implicated in its continuous repression of non-B lineage genes (113,119,154).

The Sin3 co-repressors

The mammalian Sin3 proteins, Sin3a and Sin3b, are prototypical co-repressor molecules that regulate transcription. Sin3 proteins are recruited by many different proteins to regulate gene expression. Deletion of *Sin3a* leads to defects in a wide range of biological processes including cell cycle regulation, DNA replication, DNA repair, apoptosis and chromatin modifications (155). Sin3 proteins were originally found to interact with the sequence-specific transcriptional repressors Mad1 and the related protein Mxi1, and to antagonize Myc-Max mediated transcriptional activation (141,156). Further studies revealed that Sin3 proteins can associate with HDACs as components of a series of large multimeric protein complexes, which are now known as the Sin3/HDAC co-repressor complexes (148,157,158). The Sin3 proteins contain multiple motifs that are implicated in facilitating protein-protein interactions. These include four paired amphipathic helices (PAH), a histone interaction domain and a C-terminal highly conserved region (158). Sin3 proteins are proposed to function as adaptor proteins which provide structural support to the multi-unit complexes that contain HDACs (158). The core units of the Sin3/HDAC complexes are Sin3, HDAC1, HDAC2, RbAp46, RbAp48, SAP30, and SAP18 (158).

The Sin3 proteins do not have a DNA-binding motif. Therefore, the Sin3/HDAC complex must collaborate with proteins that can bind DNA to mediate site specific

transcriptional regulation. In some cases, such interaction may require other adaptor proteins such as NCoR (nuclear receptor co-repressor) and SMRT (silencing mediator for retinoic acid receptor and thyroid hormone receptor), which provide the connection between Sin3 and the nuclear hormone receptors (143,147,159). The number of proteins associated with the Sin3 co-repressor complexes is expanding rapidly. p53, REST and E2F4 and many other critical factors are all found to associate and function together with Sin3 (160-162). In the context of transcriptional regulation in B lymphopoiesis, Sin3/HDAC complex has been demonstrated to interact with the key B lineage transcription factors Ikaros and RUNX1 (68,163). Lastly, through the interaction with other protein complexes, additional enzymatic activities could be added to the Sin3/HDAC complexes (*i.e.* histone and DNA methyltransferases) which would lead to a combinatory effect on the target chromatin (164,165).

Although conventionally believed to be a co-repressor, some evidence suggests that Sin3 also plays a role in transcriptional activation (166). Recently, Sin3a has been shown to associate with the promoters of genes that are being actively transcribed (167).

Yin Yang 1

Yin Yang 1 (YY1) is a ubiquitously expressed zinc finger transcription factor (168-170). YY1 is highly conserved and has been shown to function as a Polycomb Group (PcG) protein during development (171,172). YY1 protein contains a transcriptional activation domain and a repression domain in its N and C terminus, respectively. These domains allow YY1 to recruit HATs or HDACs, depending on the promoter context, to either activate or repress transcription (168-170,173-175). YY1 also interacts with a variety of transcriptional regulators such as TBP, TFIIB and E1A, which determine the effects that YY1 renders at a given promoter (168). The consensus YY1binding sites are found in the genome with a fairly high frequency (175-177), suggesting that YY1 exerts its effects on a large number of genes. Inactivation of *YY1*

causes embryonic lethality (178). This is consistent with its pivotal role in a variety of biological processes which include development, differentiation, replication and cell proliferation (179). Moreover, YY1 has also been shown to regulate p53 stability by blocking the interaction between p300 and p53, which is independent of the transcriptional activity of YY1 (180,181)

YY1 plays a critical role in the development of B cells. Conditional inactivation of *YY1* by *mb-1-Cre* leads to a developmental arrest at the pro-B cell stage, while the pro-B cells are still made. In addition, the frequency of V to DJ rearrangement at the *IgH* locus is significantly reduced in the $YY1^{-/-}$ pro-B cells (182). Mechanistically, YY1 binds the *IgH* intronic enhancer ($E_{i\mu}$) to regulate the contraction of the *IgH* locus (182). Whether the transcriptional activity of YY1 is required in this process is unknown, because the expression of key B lineage genes (*i.e*. *E2A*, *EBF* and *Pax5*) as well as those important for VDJ recombination (*i.e*. *Rag1*, *Rag2*, *Ku70* and *Ku80*) does not seem to be altered in $YY1^{-/-}$ pro-B cells (182). Yet, B cell development in $YY1^{-/-}$ mice is only partially rescued by expression of a prerearranged *IgH* transgene, suggesting YY1 plays more roles than just to regulate V_H to DJ_H recombination in B cell development (182). YY1 has also been shown to activate the DNase I hypersensitive site3 in the *IgH* locus, which is important for germline transcription, and hence, class switch recombination (183).

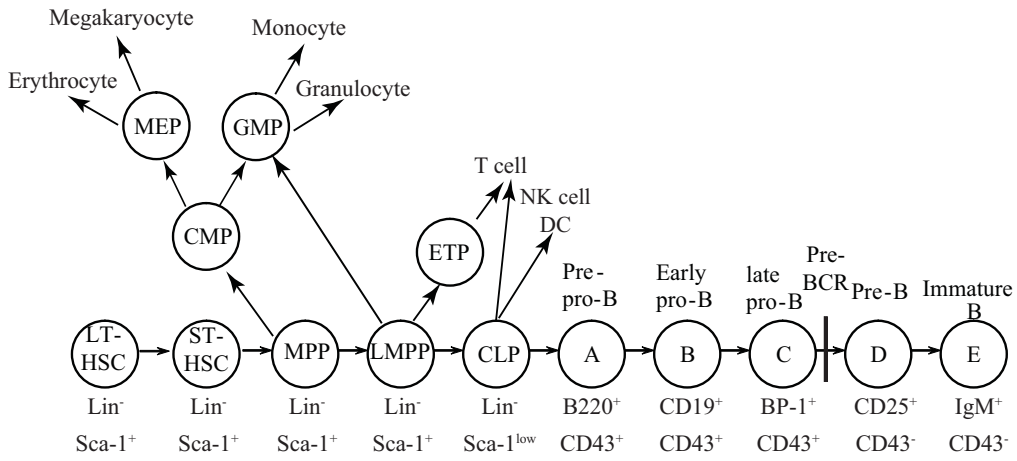
In summary, B lymphopoiesis is a tightly regulated process and is associated with dynamic gene expression changes. Transcription factors such as E2A, EBF and Pax5 as well as the co-activators and co-repressors such as HATs or HDACs play key roles in this process by initiating B lineage genes and repressing alternative lineage genes. Meanwhile, proteins like YY1 may contribute to the regulation of B lymphopoiesis independent of their transcriptional activities. It is proposed that these proteins form a network to collectively instruct the progression of B lymphopoiesis (Figure 2). However, little is known about how these proteins with distinct biological functions are orchestrated to act coordinately in the network. It is also unclear what other factors may be required.

Thus, defining new molecules especially those orchestrating the function of the known factors in the network will greatly advance the current knowledge about B lymphopoiesis. In addition, such study will shed light on the understanding of hematopoietic neoplasia which is closely related to abnormalities in this protein network.

To address these questions, a novel mutant strain of mice (*Justy* mice) with defective B cell development has been studied. The *Justy* mice were generated by chemical induced random mutagenesis and thus likely carry mutations in genes essential for B lymphopoiesis. Initial studies suggested that a T to A transversion in the *Gon4-like* gene might be the causal mutation in the *Justy* mice. The purpose of this dissertation work is to further analyze the B cell development defects in *Justy* mice, determine how the *Justy* mutation affects the expression of the protein encoded by *Gon4-like*, and explore the molecular function of this protein.

Figure 1. The current model of B lymphopoiesis in bone marrow.

Hematopoietic stem cells in the bone marrow give rise to two major lineages: the lymphoid and myeloid. The development of these two lineages diverges at the multipotent progenitor (MPP) stage. The MPPs that are $Flt3^+$ can differentiate into T, B and NK cells, as well as DCs and are known as the lymphoid primed MPPs (LMPP). The $Flt3^-$ MPPs can give rise to all types of myeloid cells and are known as the common myeloid progenitors (CMP). LMPPs that differentiate along the B lymphopoiesis pathway become IgM^+ immature B cells in the bone marrow. Both the consensus nomenclature for B cell development and the corresponding Hardy fractionation nomenclature are shown in the figure. Cell-surface marker expression and V(D)J recombination status of the *IgH* and *L* loci at each stage are shown. Detailed description of the B lymphopoiesis pathway, could be found in the text. Abbreviations: LT-HSC, long-term hematopoietic stem cells; ST-HSC, short-term hematopoietic stem cells; MPP: multipotent progenitor; LMPP, lymphoid-primed progenitor; CLP, common lymphoid progenitor; CMP, common myeloid progenitor, MEP, megakaryocyte-erythroid progenitor; GMP, granulo-monocytic progenitor; ETP, early lymphoid progenitor; Lin^-, negative for the expression of lineage-specific markers.



Adapted from:

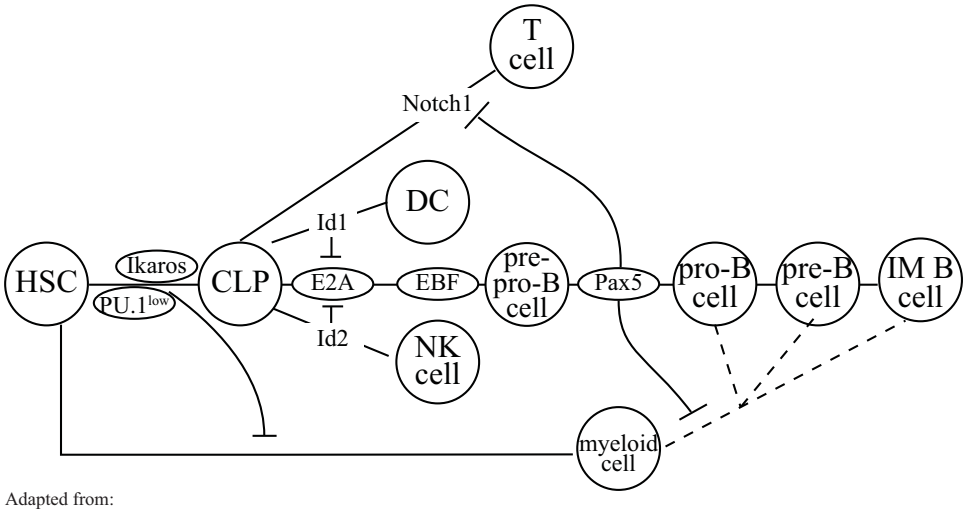
1. Busslinger, M., *Transcriptional control of early B cell development.* Annu Rev Immunol, 2004. 22: p. 55-79.

2. Hardy, R.R., P.W. Kincade, and K. Dorshkind, *The protean nature of cells in the B lymphocyte lineage*. Immunity, 2007. 26(6): p. 703-14.

3. Nagasawa, T., *Microenvironmental niches in the bone marrow required for B-cell development*. Nat Rev Immunol, 2006. 6(2): p. 107-16.

Figure 2 B cell development is regulated by a network of transcription factors.

The protein network that mediates transcriptional regulation during the progression of B lymphopoiesis. At the very early stages of B cell development, Ikaros and low levels of PU.1 promote lymphoid development by repressing the development of myeloid cells. After the cells reach the common lymphoid progenitor (CLP) stage, E2A and EBF activate B lineage appropriate genes including *Pax5*. The function of E2A can be antagonized by the Id proteins which promote the differentiation of dendritic cells (DC) and natural killer (NK) cells. Pax5 can further activate more B lineage genes after being expressed. More importantly, Pax5 mediates repression of genes required for the development of other lineages (*i.e.* myeloid and T cells), which is known as B lineage commitment. B lineage commitment mediated by Pax5 is maintained in the entire developmental pathway after the pro-B stage. The solid lines in the figure illustrate the normal progression of hematopoiesis and B lymphopoiesis. The dotted lines indicate transdifferentiation from B lineage cells to myeloid cells that is prevented by Pax5 mediated B lineage commitment. “IM B cell” denotes immature B cells.



Adapted from:

1. Busslinger, M., *Transcriptional control of early B cell development.* Annu Rev Immunol, 2004. 22: p. 55-79.

2. Nutt SL, Kee BL. *The transcriptional regulation of B cell lineage commitment*. Immunity. 2007;26(6):715-25.

CHAPTER II

THE *JUSTY* MUTATION CAUSES SEVERELY REDUCED GON4L PROTEIN EXPRESSION AND DEFECTIVE B LYMPHOPOIESIS

Abstract

Chemical induced mutagenesis is a potent way of identifying genes that are critical for B lymphopoiesis. Injection of n-ethyl-n-nitrosourea (ENU), a mutagen, generated a recessive mutation named *Justy* (for Just T cells) that specifically caused a loss of $B220^+$ B cells in mice. Analysis of B cell development in *Justy* bone marrow indicated that the generation of common lymphoid progenitors and pre-pro-B cells was normal but that of the pro-B cells was severely disrupted. Genetic studies revealed that the *Justy* mutation was a T to A transversion in the *Gon4-like* (*Gon4l*) gene. This mutation led to aberrant splicing of RNA transcribed from *Gon4l*, and greatly reduced Gon4l protein expression. Additional analysis done in collaboration with other lab members revealed that the loss of Gon4l resulted in elevated expression of myeloid genes in *Justy* pro-B cells, which, under normal conditions, are silenced for B lineage commitment at this stage. In support of this, Gon4l was found to contain homologies to classical transcriptional co-repressors. Moreover, cells taking on a myeloid lineage fate expressed none or substantially reduced Gon4l. These findings thus indicate that Gon4l plays a vital role in early B lymphopoiesis, and it is likely through transcriptional repression of the myeloid genes.

Introduction

B lymphocytes are generated from multipotent hematopoietic stem cells (HSC) through a process known as B lymphopoiesis. During this process, HSCs gradually lose developmental plasticity and eventually give rise to mature B cells that are capable of producing antibodies. B lymphopoiesis is associated with characteristic gene expression reprogramming and the ordered rearrangement of the immunoglobulin heavy and light

chain genes (1-4). By the end this process, genes appropriate for the B lineage are activated and sustained, while genes that are inappropriate for the B lineage are repressed. Meanwhile, a functional B cell receptor consisting of the rearranged immunoglobulin heavy and light chains is expressed. B lymphopoiesis is regulated by a network of transcriptional regulatory proteins (4,114) which include E2A, EBF and Pax5. These factors coordinately mediate the activation of key B lineage genes (*i.e. mb-1* (encodes Igα), *λ5* and *Vpreb* as well as V(D)J recombination (102-106,111,184).

Accumulating evidence now suggests that repression of non-B-lineage genes is equally important as the activation of B lineage genes for B cell development. In multipotent progenitor cells, chromatin associated with different lineage genes is readily accessible for transcription and those genes are expressed in a promiscuous manner (4,131,185-187). Thus, for the progenitor cells to commit to a B lineage fate, genes of other lineages need to be silenced (4). Disruption of such repression leads to abnormalities in B lymphopoiesis. This is exemplified by the studies of the B lineage transcription factor, Pax5. Deletion of the *Pax5* gene in mice results in an arrest of B cell development at the pro-B stage (115,116,120,188), and genes associated with multipotent progenitors or other lineages are not appropriately repressed (117,119). As a consequence, $Pax5^{-/-}$ pro-B cells become pluripotent and are capable of differentiating into a broad spectrum of hematopoietic cell types (117,118). More strikingly, inactivation of *Pax5* in fully differentiated B cells also leads to upregulation of alternative lineage genes such as the myeloid gene, *csf1r* (*colony-stimulating factor 1 receptor*) (119). Thus, repression of non-lineage appropriate genes by factors like Pax5 is critical in restricting differential plasticity and maintaining the B lineage identity (117,122,123).

Although some important transcription factors that guide B lymphopoiesis have been intensively studied, the mechanisms that orchestrate those factors, especially in the repression of non-B lineage genes, are unknown. Other proteins may also be required for this process. In addition, abnormalities in the hematopoietic transcription regulatory

network are closely related to development of malignancies of lymphocytes (189,190). Therefore, finding and studying other key regulators of hematopoiesis may assist the understanding of leukemia and lymphomas.

Here the characterization of a mutant strain of mice known as the *Justy* mice is described. *Justy* mice specifically lack mature B cells in the periphery. B cell development in the *Justy* mice was arrested at the B lineage commitment phase, whereas development of other lineages was not affected. Gene expression analysis done in collaboration with other lab members revealed that the repression of multiple myeloid genes was defective in *Justy* B cell precursors while the expression of the key B lineage genes was normal. Genetic studies showed that the causal mutation in *Justy* mice was a T to A transversion in the *Gon4l* gene, which led to aberrant splicing of the *Gon4l* RNA and severely reduced expression of its encoded protein. Thus, these data indicate that Gon4l is essential for early B lymphopoiesis and likely functions through repressing the myeloid genes. In support of that, Gon4l protein was found to bear homologies to known transcriptional co-repressors such as Sin3a and its expression was excluded from myeloid precursor lines.

Materials and Methods

Mice

Generation of the mutant mice was carried out by a collaborating laboratory in Ingenium Pharmaceuticals. In brief, male C3HeB/FeJ (C3H) mice were injected with n-ethyl-n-nitrosourea (ENU) to induce random mutations in the genome. The administered dosage was set to induce 1 mutation per 2.7 megabase-pairs of DNA on average (191). These mice were then mated to WT C3H females and the progenies (F1 mice) were screened for peripheral blood abnormalities. Those that were normal in the screen (but may carry recessive mutations) were mated with WT C3H females to generate the F2 progeny. The F2 female mice were then crossed with the F1 males to generate the F3

progeny. One of the F3 progeny mouse strains was found to specifically lack CD45R/B220^{+} (a B lineage marker) cells in their peripheral blood. These F3 mice were then backcrossed to WT mice to narrow down the mutated genomic region responsible for the seen phenotype and the resultant mice were named *Justy* (for Just T cells) (192).

Flow cytometric analysis

FACS analyses were performed with 4-week-old mice. Single cell suspensions of bone marrow, lymph nodes and thymus were prepared. Cells were incubated with CD16/CD32 Fc block (eBioscience) to prevent nonspecific antibody binding. $5x10^5$ cells were used per stain and stained with the indicated antibodies and analyzed on a LSR II. For staining with biotin-labeled antibodies, cells were incubated with biotinylated antibodies for 30 minutes on ice, washed and then incubated with streptavin- or advidin-conjugated fluorochromes. Collected data were analyzed using FlowJo (TreeStar). Cell yields for different fractions were calculated from post-sort analysis of collected data. Total events in the appropriate gate were divided by the total events in the forward scatter versus side scatter gate; this percentage was multiplied by the total yield of cells as determined using a hemacytometer.

The following fluorochrome-conjugated antibodies were purchased from the listed sources or generated by Dr. Thomas Waldschmidt in the University of Iowa. BioLegend: Alexafluor 700-B220 (RA3-6B2), APC-c-Kit (2B8), Pacific Blue-Sca-1 (D7); eBioscience: APC conjugates: B220 (RA3-6B2), CD3ε (145-2C11); FITC conjugates: CD11c (N418), CD49b (DX5), CD8 (H35-17.2); PE conjugates: CD135/Flt-3 (A2F10); PE-Cy7 conjugates: CD127 (IL-7Rα (A7R34), CD19 (1D3), BD Biosciences: APC-Cy7-B220 (RA3-6B2), PE-Gr-1 (RB6-8C5), Biotin-CD11b/Mac-1 (M1/70), Biotin-CD43 (S7), Biotin-CD49b (DX5); Avidin-Texas Red, Streptavidin Per-CP; Waldschmidt Lab: PE-CD4 (GK1.5), FITC-Ly6C. Invitrogen: Streptavidin Qdot 605; Sigma: FITC-Ly6d (RGRSL114.8.1).

Histology

Tissue samples were collected and fixed in 10% neutral buffered formalin as described (193). The fixed tissues were then processed, stained and analyzed by the Comparative Pathology Facility at the University of Iowa.

Isolation of lineage-negative cells

Bone marrow cells from 2-6 mice were depleted of red blood cells and resuspended in PBS containing 1% FBS (fetal bovine serum) and 2mM EDTA (PBS/FBS/EDTA). Cells were pre-incubated with anti-CD16/32 and then incubated with rat antibodies specific for mouse B220, CD5, CD19, Gr-1, IgM, Ly6C and Ter-119. Cells were washed twice with PBS/FBS/EDTA and magnetic beads coated with sheep anti-rat antibodies were added (Dynal). Cells bound to beads were removed using a magnet and unbound cells were recovered as lineage-negative (Lin^-) cells.

RT-PCR analysis

Total RNA was isolated from Lin^- cells using TRIzol (Invitrogen) and reverse-transcribed with SuperScript III First-strand cDNA Synthesis Kit (Invitrogen) according to the manufacturer's instructions. Quantitative real-time PCR was performed using POWER SYBR Green Master Mix and the ABI PRISM 7700 Detection system (Applied Biosystems). The primers used are as following: Ash1Lfw: 5'-ACTCCCAAAAGAGAC TTCTCACCTCATTA-3', Ash1Lrv: 5'-GATCGTCCTCCGTTCCTCTTGTAGTT-3'. Relative expression level was obtained using the $2^{-[\Delta\Delta Ct]}$ method with *Hprt* as the internal reference control (194). The detection of the normally spliced Gon4l RNA was performed by using a primer against the junctional sequences between exon 24 and 25, and the other primer against exon 26. The detection of the *Justy* abnormally spliced Gon4l RNA was performed by using one primer across the junctional sequences between exon 24 and the cryptic exon, and the other primer against the sequences between the cryptic exon and exon 25 (192).

Stromal cell co-cultures

Lin^- cells were isolated with the negative selection kit as described above. Pro-B cells were expanded by co-culturing Lin^- cells with confluent layers of OP9 stromal cells in the presence of SCF, Flt3L and IL-7 at 10 ng/ml each. Cultures were started in 24-well plates and after 7 days cells were transferred to 12-well plates and then 3-4 days later to 10-cm plates. This latter step was repeated twice, after which cells were harvested and used to prepare cell lysates. Flow cytometric analysis showed that > 90% of the cells obtained were $B220^+CD19^+$ (Prepared and analyzed by Dr. Colgan).

Antibody production and immunoblot analysis

The generation of anti-Gon4l antibodies against amino acids 938-1364 (N-938) or 1746-2260 (C-1746) was described in (192). Protein lysates were prepared by lysing cells in a lysis buffer containing 50mM Tris-HCl pH 8.0, 120mM NaCl, 0.5% NP-40, 0.2mM sodium orthovanadate, 100mM NaF, 50ug/ml PMSF, and protease inhibitor cocktail (Roche). Debris was removed by centrifugation at maximum speed in a desktop micro-centrifuge. Total protein was separated by electrophoresis through 7% or 3-8% Nu-PAGE gels (Invitrogen) and then transferred to PVDF membranes (Millipore). Membranes were probed with the affinity-purified anti-Gon4l antibody. Donkey anti-rabbit IgG HRP (Santa Cruz: sc-2314 and sc-2313) were used as secondary antibodies. Blots were visualized with an ECL system (SuperSignal West Pico, Thermo Scientific). The culture of the cell lines is described in chapter III.

Results

Justy mice lack peripheral B cells

Chemical induced mutagenesis was performed in a collaborating laboratory to identify genes that may be important for hematopoiesis. A novel mutant strain of mice was generated by this approach that specifically lack $B220^+$ B cells in their peripheral

blood while maintaining a seemingly normal T cell population (192). This mutant strain of mice was therefore named *Justy* (for Just T cells). The *Justy* mice are unable to produce serum immunoglobulin, but are viable and fertile.

To further characterize the phenotypes, spleens from the *Justy* mice were analyzed. *Justy* spleens were significantly smaller compared to that of the WT (Figure 3A). Hematoxylin and eosin (H&E) staining of spleen sections revealed obvious white pulp hypoplasia in *Justy* samples (192). The cellularity was correspondingly reduced in *Justy* spleens (192). Flow cytometric analysis of whole splenocytes indicated that the frequency and total number of $CD19^+$ (B cell marker) cells from *Justy* mice were severely reduced (Figure 3B). $CD19^+$ cells were not found in *Justy* lymph nodes or peritoneal lavages as well (data not shown). In contrast, the splenic $CD3^+$ T cells as well as the $CD4^+$ or $CD8^+$ T cell subpopulations seem intact in *Justy* mice (Figure.3C). Similarly, the myeloid (Mac-1^+ and/or Gr-1^+) populations in the *Justy* spleen appeared to be normal (data not shown). Taken together, these data suggest that the generation of B lymphocytes was specifically affected in *Justy* mice.

B cell development is blocked at the pre-pro-B to pro-B transition in *Justy* bone marrow

B cells originate from hematopoietic precursor cells residing in the bone marrow. Previous experiments show that adoptive transfer of *Justy* bone marrow cells into lethally irradiated wild type (WT) mice replenishes all blood cell types except B cells, while transfers of WT bone marrow into lethally irradiated *Justy* mice give rise to all cell types including B cells (experiments performed by Isaiah Hankel in the Waldschmidt lab) (192). Thus, the mutation in *Justy* mice specifically affects B cell development in the bone marrow and the defects are intrinsic to *Justy* hematopoietic progenitors.

To identify which B cell developmental stage was affected by the mutation, *Justy* bone marrow was analyzed (Figure 4). During B lymphopoiesis, B cell progenitors

proceed through the ordered pre-pro-B, pro-B, and pre-B stages to become immature B cells. The B lineage marker B220 is expressed by all B lineage cells starting from the pre-pro-B stage. Another surface marker, CD43, is expressed on pre-pro-B and pro-B cells but is downregulated when the cells enter the pre-B stage (Figure 1) (3). Fluorescence activated cell sorter (FACS) based analysis using B220 and CD43 revealed that the pre-B and immature B cell population ($B220^{+}CD43^{-}$) were almost completely lost in *Justy* bone marrow (Figure 4A upper panel). The collective total cell number of this compartment decreased by about 300 fold (192). On the other hand, the frequency and total number of the $B220^{+}CD43^{+}$ population, which contained the pre-pro-B and pro-B cells, were similar between *Justy* and WT mice (Figure 4A and data not shown). The $B220^{+}CD43^{+}$ compartment also contains precursor cells of the natural killer (NK) and plasmacytoid dendritic (pDC) cells (195,196). To refine the analysis, the non-B lineage cells were excluded by staining with CD49b (clone DX5, NK marker), CD11c (DC marker), and Ly6C (pDC markers) antibodies (Figure 4A lower panel). Cells were also stained for CD19 to resolve pro-B cells from pre-pro-B cells (see Figure 1). This analysis revealed that the frequency of pro-B cells ($CD49b^{-}CD11c^{-}Ly6C^{-}CD19^{+}$) was severely reduced in *Justy* bone marrow while that of the *Justy* pre-pro-B cells (quadruple-negative) was not affected. Similar results were obtained when total cell numbers were analyzed (these experiments were performed in collaboration with Mingyi Chiang and with supervision from Drs. John Colgan and Thomas Waldschmidt)(192). These results suggest that the generation of pro-B cells from pre-pro-B cells is disrupted in *Justy* mice.

The common lymphoid progenitors (CLPs) are the immediate precursors of the pre-pro-B cells. To confirm that the B cell development is specifically blocked at the pre-pro-B stage, CLPs in *Justy* bone marrow were analyzed (Figure 4B). CLPs and pre-pro-B cells were included in the $Lin^{-}IL7R^{+}Flt3^{+}$ population that was also $c\text{-}Kit^{+}Sca\text{-}1^{+}$ (197,198). Pre-pro-B cells expressed B220 which further distinguished them from the CLPs (Figure 4B). The CLP cells were then resolved by the expression of Ly6D into two

populations with one restricted to the B lineage ($Ly6D^+$) and the other associated with multi-lineage differentiation potential ($Ly6D^-$) (197,198). Comparison of the CLP populations between WT and *Justy* mice indicated that they were essentially the same (Figure 4B). Therefore, the generation of CLPs was not affected in the *Justy* mice.

The myeloid compartment in *Justy* bone marrow was also examined (Figure 4C). Although the ratio between lymphoid and myeloid cells was decreased, the myeloid population was present and appeared normal in *Justy* bone marrow. Because HSCs in the bone marrow give rise to early T cell progenitors (ETP) that differentiate into T cells in the thymus, T cell development was also examined (Figure 5). FACS analysis of the thymus revealed that T cell development was grossly normal in *Justy* mice, suggesting that the generation of ETPs and later T cell development were not affected by the mutation. Taken together, these data suggest that the *Justy* mutation specifically affects B cell development and causes a developmental block at the pre-pro-B to pro-B transition.

The *Justy* mutation is within a 617kb region on mouse chromosome 3

To identify the genetic locus responsible for the *Justy* phenotype, C3H *Justy* mice were crossed to wild-type B6 mice. The F1 mice were interbred and the resulting F2 mice were screened for B cell deficiency. Genomic DNA from mice with a matching phenotype was analyzed by PCR using a panel of primers to detect simple sequence length polymorphisms (SSLPs), which discriminates the C3H or B6 origin of the genomic regions. This analysis revealed that a 617 kilobase-pair (kb) region on mouse chromosome 3 between SSLPs D3Mit49 and D3Mit 175 (positions 88353984 and 88971024 in chromosome 3; UCSC assembly, July 2007) completely associated with the *Justy* phenotype (Figure 6A) (the breeding data were generated in Ingenium) (192). The identified region was then transferred to B6 background and backcrossed for 10

generations. FACS analysis of the resultant mice revealed identical block in B lymphopoiesis (192). Therefore, the causal mutation is within this identified region.

The *Justy* mutation is a T to A transversion in the *Gon4-like* gene that causes aberrant RNA splicing

Thirteen genes are located within the genetic region that correlates with the *Justy* phenotype (Figure 6A) (192). Among those genes, *Gon4-like* (*Gon4l*) and *Ash1l* encode proteins implicated in transcriptional regulation (192). Because B cell development is transcriptionally regulated, the expression of *Gon4-like* and *Ash1l* were examined (Figure 6B). Quantitative RT-PCR (Q-RT-PCR) analysis of RNA isolated from bone marrow lineage-negative (Lin^-) cells showed that the mRNA level of *Gon4l* was reduced by about two fold in *Justy* cells relative to that of the wild type, while the expression of *Ash1L* mRNA was not changed. The expression of all other genes in this region were also normal as examined later by Isaiah Hankel and Dr. John Colgan (192). Thus, the expression of *Gon4l* is specifically affected.

Further analysis of the *Justy Gon4l* locus revealed a "T" to "A" transversion in intron 24 (Figure 7 and (192)). This substitution makes the sequence flanking the point mutation more homologous to a 5' RNA splicing donor site (199). RT-PCR analysis revealed that this donor site was used in the *Justy* cells, resulting in an insertion of an 81 base-pair (bp) cryptic exon into the *Gon4l* RNA (192). More strikingly, the insertion placed two in-frame premature stop codons into the *Gon4l* RNA, which hypothetically would generate a truncated Gon4l protein. Q RT-PCR analysis of RNA collected from sort purified *Justy* pre-pro-B and pro-B cells revealed the expression of the aberrantly spliced *Gon4l* RNA species and the great reduction of the WT spliced form of Gon4l RNA (~10 fold less) (Genomic sequencing performed by Dr. Colgan and RT-PCR by Judit Knits) (192). These findings strongly correlated the B cell development defects in *Justy* mice with the expression of the aberrantly spliced *Gon4l* RNA.

Gon4l protein is expressed in lymphoid tissues and B cell precursors

In order to determine how the aberrant Gon4l RNA splicing would affect Gon4l protein expression, two antibodies were generated. One antibody recognizes the central region of Gon4l (N-938) and the other recognizes the C-terminal regions of Gon4l (C-1746) (Figure 8A, see Figure 13 in Chapter III for detailed description of the generation of the antibodies). Consistent with its RNA expression (192), Gon4l protein is detected in thymus and spleen by both antibodies (Figure 8 B and C). But Gon4l was not detected in the bone marrow, likely due to low protein yield from harvesting bone marrow cells. To circumvent this problem, WT pro-B cells were expanded from isolated bone marrow Lin^- cells (by Dr. Colgan) (192). Immunoblotting analysis using whole cell lysate prepared from those cells revealed the expression of Gon4l protein (Figure 8D), consistent with our previous findings that Gon4l message is expressed in sort purified pro-B cells (192). Gon4l was also detected in a number of B cell and T cell lines, indicating that these could be utilized as model systems to study the molecular function of Gon4l protein (Figure 8E, and Figure 13 in Chapter III). Therefore, Gon4l protein is expressed in lymphocytes, suggesting that it plays a role in the biology of these cells. Most importantly, Gon4l is expressed in B cell precursors, implicating it in the regulation of B cell development.

Gon4l expression is severely reduced as a result of the aberrant RNA splicing

To evaluate how the *Justy* mutation affected Gon4l protein expression, *Justy* pro-B cells from Lin^- cells were cultured for immunoblot analysis. However, *Justy* Lin^- cells were unable to expand *in vitro* (Dr. John Colgan, personal communication). Thymocytes were therefore used as an alternative, since these cells also express high levels of aberrantly spliced Gon4l RNA (192). Immunoblot analysis of whole cell lysate prepared from WT and *Justy* thymocytes indicated that Gon4l protein expression was severely

reduced in *Justy* cells (Figure 9A). Q RT-PCR analysis confirmed that the reduced expression of Gon4l protein was consistent with the reduction of its WT RNA (Figure 9B). These data, thus, indicate that the aberrant splicing of Gon4l RNA, as a consequence of the *Justy* mutation, leads to severely compromised Gon4l protein expression.

Truncated Gon4l protein is not detected

The generation of truncated Gon4l proteins by the aberrant splicing of Gon4l RNA was examined. The expected truncated Gon4l protein would lack the PAH and SANT domains (Figure 10A) and have a molecular weight of ~185kDa. The N-938 Gon4l antibody would be able to recognize this form if it were expressed (Figure 10A). Immunoblot analysis of whole cell lysate prepared from WT and *Justy* thymocytes revealed a reduction of Gon4l in *Justy* cells, consistent with the results obtained by using the C-1746 Gon4l antibody (Figure 10B). However, no truncated protein species was detected. This result held true even after loading maximum amount of the *Justy* samples (Figure 10C). Thus, the mutant Gon4l RNA may not be translated. This may be because the aberrant splicing of Gon4l RNA have activated the nonsense-mediated RNA decay pathway, in which the aberrantly spliced Gon4l mRNA is degraded (200). It is also possible that the aberrantly spliced Gon4l mRNA is translated, but the mutant protein is rather unstable and cannot accumulate. Taken all together, these data support a conclusion that the severely reduced Gon4l protein as a result of the *Justy* mutation accounts for the B cell development defects seen in *Justy* mice.

Gon4l is implicated in the repression of myeloid genes

Lymphoid progenitor cells become restricted to the B lineage at the pre-pro-B and pro-B stage by actively repressing non-B lineage appropriate genes. Gene expression analysis in *Justy* pre-pro-B and pro-B cells indicated that the expression of key B lineage transcription factors *E2A*, *EBF* and *Pax5* was normal, but that of myeloid genes such as *Cebpα* (encoding CCAAT/enhancer binding protein), *Sfpi1* (encoding PU.1) and *Csf1r*

(encoding colony stimulating factor 1 receptor) was abnormally elevated (192). Myeloid genes are normally targeted for repression during B lineage commitment. Therefore these observations suggest that the loss of Gon4l protein resulted in a loss of transcriptional repression of the myeloid genes. Consistently, bioinformatic analysis indicated Gon4l bore strong homology to classical transcriptional co-repressors (Figure 12B in Chapter III). If Gon4l does mediate repression of myeloid genes, it seems reasonable that the expression of Gon4l will be repressed in hematopoietic precursor cells that have taken on a myeloid fate. Indeed, immunoblot analysis of whole cell lysates prepared from the myeloid precursor lines, 32D and Fdcp1, revealed that Gon4l was not expressed in these cells (Figure 11 B and C). In the mature macrophage line, Raw264.7, Gon4l was expressed at a low level as compared to that of the control M12 B cells (Figure 11C). Together, these data suggest that Gon4l is an essential factor that regulates B lymphopoiesis and may function by repressing myeloid genes.

Discussion

In this study *Gon4l* has been identified as a critical gene for B cell development. Chemically induced mutagenesis introduced a point mutation into an intron region of *Gon4l*, which caused aberrant splicing of *Gon4l* RNA and introduced a cryptic exon into the Gon4l RNA sequence, which led to severely reduced Gon4l protein expression. The most obvious manifestation of the mutation was a complete loss of peripheral B cells. Analysis of the *Justy* bone marrow revealed a developmental arrest at the pre-pro-B to pro-B transition, while development of other hematopoietic lineages seemed unaffected.

The *Justy* mutation caused an insertion of an exon carrying two in-frame pre-mature stop codons into the *Gon4l* RNA, which hypothetically would lead to the production of a truncated Gon4l protein. However, no such protein was detected, suggesting that either the aberrantly spliced Gon4l RNA is not translated or the translated truncated protein is very unstable. This result also suggests that the *Justy* phenotype is not

the result of the expression of a truncated and perhaps a dominant negative form of Gon4l. This is consistent with the fact that the *Justy* mutation is a recessive mutation. On the other hand, the aberrant splicing interfered with WT Gon4l mRNA expression. As a result, Gon4l protein expression was reduced. In thymocytes, the mutant Gon4l RNA was expressed exclusively in *Justy* cells, which severely disrupted the expression of WT spliced Gon4l RNA (Figure 9). Similarly, WT Gon4l RNA was reduced by about 10 fold in *Justy* pre-pro-B cells due to the occurrence of aberrant splicing (192). Thus, the loss of Gon4l protein contributes to the B cell development defects in *Justy* mice.

Although *Gon4l* RNA and protein have been detected in a variety of tissues in mammals (201-203), the development of other hematopoietic lineages as well as other tissues were not obviously affected in *Justy* mice. The reasons why the mutation selectively compromises B cell development at the pre-pro-B stage are unknown. In WT mice, *Gon4l* RNA is expressed at relatively high levels at the pre-pro-B stage as compared to other B cell developmental stages examined in bone marrow (192). This may explain why this stage is particularly sensitive to the *Justy* mutation. On the other hand, the extent of aberrant splicing may account for which tissue is affected. Indeed, the level of aberrantly spliced Gon4l RNA relative to the WT spliced form varies from tissue to tissue (192). Alternative splicing is regulated by splicing factors that are usually expressed in a dynamic and tissue specific manner (204-206). Thus, splicing factors specifically associated with the pre-pro-B stage may contribute to the tissue specific effects of the *Justy* mutation. Third, some Gon4l protein is still expressed in thymocytes and probably in other cell types, which may be enough for guiding the development of these cells. It is also possible that proteins with a redundant role as Gon4l can compensate the loss of Gon4l in non-B-lineage cells. Another possibility is that Gon4l may not play a role in the development of these non-B-lineage cells; instead Gon4l may regulate other aspects of their biology. In support of this, our current studies on T cell function have

indicated that effector T cell differentiation of *Justy* T cells is affected by the *Justy* mutation (Dr. Colgan and Isaiah Hankel unpublished data).

Progression of B lymphopoiesis requires the activation of key B lineage genes such as *E2A*, *EBF* and *Pax5*. Interestingly, the expression of these genes was all normal, or slightly elevated in *Justy* pre-pro-B and pro-B cells (192). This finding suggests that the B cell development arrest in *Justy* mice is not due to changes of these factors. More importantly, this result demonstrated that, although E2A, EBF and Pax5 are the major factors in the current B lymphopoiesis model, the expression of these factors per se is not sufficient to drive B cell development forward. More factors are involved in this process.

Successful B lymphopoiesis also requires continuous repression of myeloid genes such as *Cebpα, Sfpi1* and *Csf1r* (4,20,207,208). Overexpression of these genes promotes myeloid development and inhibits B lymphopoiesis (209-211). *Cebpα, Sfpi1* and *Csf1r* were abnormally up-regulated in *Justy* pre-pro-B and pro-B cells where Gon4l protein expression was compromised, suggesting that Gon4l negatively regulates these genes (192). In support of this, myeloid precursor cells that express high levels of myeloid genes do not express Gon4l (Figure 11). Moreover, protein sequence analysis shows that Gon4l contains homology to classical co-repressors (Figure 12 of Chapter III). Lastly, microarray analysis of *Justy* early pro-B cells showed profound gene expression changes. The majority of those changes were abnormal up-regulation of gene expression, which is indicative of a loss of gene repression (Dr. John Colgan unpublished data). Therefore, these data suggest that Gon4l is part of the pathways that regulate B lymphopoiesis by repressing the myeloid genes and probably many other genes.

Gon4l protein sequence is conserved and its orthologs are found across species (201). More importantly, the zebrafish Gon4l ortholog has been demonstrated to play a role in hematopoiesis (212,213). Therefore, Gon4l and its associated pathways may represent a general mechanism that facilitates hematopoiesis. However, some questions still remain. First, how does Gon4l repress gene expression? Does Gon4l repress gene

expression by blocking DNA binding of other transcription factors, or by epigenetic modification of the target gene associated chromatin? Second, does Gon4l function alone or does it require collaboration with other proteins? Third, what are the target genes of Gon4l? Answering these questions will help provide more insight into the molecular function of Gon4l.

Figure 3. *Justy* mice lack peripheral B cells.

(A) The spleen from *Justy* mice is smaller in size and displays obvious white pulp hypoplasia. Spleens were collected from 8 week-old WT and *Justy* mice (left panels). Sections prepared from the same sample were H&E stained (right panels). * The data were obtained by Drs. John Colgan and David Meyerholz and this panel of the figure is copied from reference (192). (B) The frequency and total number of the $CD19^+$ B cells are greatly reduced in the *Justy* spleen. Splenocytes were collected from age-matched WT and *Justy* mice. The T cell ($CD3^+$) and B cell ($CD19^+$) populations were analyzed by FACS (left panel). Collective total cell number of each population was calculated and plotted as described in Materials and Methods (right panel). (C) The $CD4^+$ and $CD8^+$ T cell sub-populations are normal in *Justy* spleen. The $CD3^+$ T cell population in (B) was further analyzed for the $CD4^+$ and $CD8^+$ T cell sub-populations by FACS (left panel). The total cell number of the two populations was similarly plotted (right panel). Data in (B) and (C) are representative of 5 independent experiments with 2 sets done by Ping Lu and 3 sets done in collaboration with Mingyi Chiang.

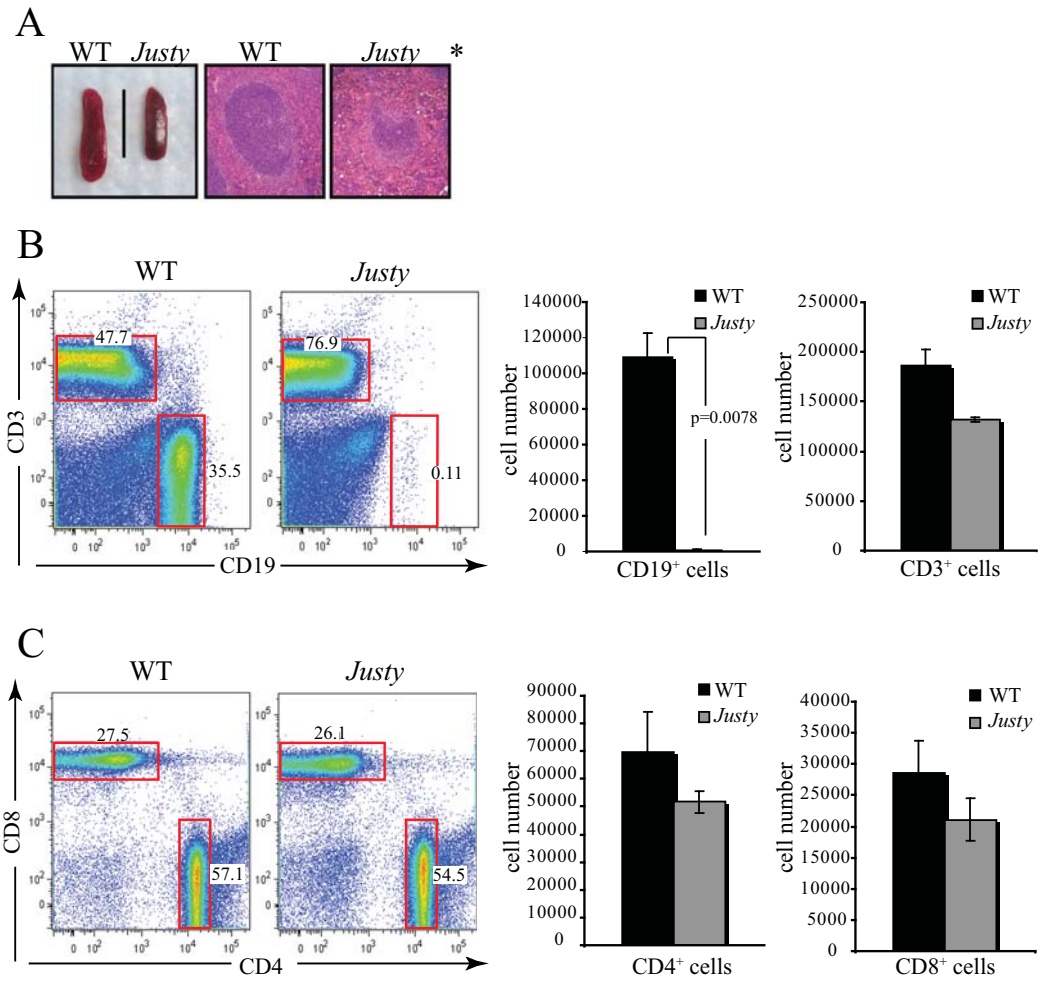

A
WT
Justy
WT
Justy
*
B
WT
Justy
CD3
CD19
47.7
35.5
76.9
0.11
cell number
WT
Justy
p=0.0078
CD19+ cells
CD3+ cells
C
WT
Justy
CD8
CD4
27.5
57.1
26.1
54.5
CD4+ cells
CD8+ cells

Figure 4. B lymphopoiesis is blocked in *Justy* bone marrow.

(A) A B cell developmental block at the pre-pro-B to pro-B transition in *Justy* bone marrow. Bone marrow cells from WT and *Justy* mice were stained with the indicated antibodies and analyzed by FACS. The right panels denote the outlines of the gating strategies. The arrow heads indicate the direction of B cell development. The $B220^{+}CD43^{-}$ population which includes the pre-B cells and immature B cells is greatly reduced in *Justy* mice (upper left panel). The $B220^{+}CD43^{+}$ population which contained the pre-pro-B and pro-B cells were further stained with the indicated antibodies and analyzed (lower left panel). NK cell precursors and plasmacytoid dendritic cells were gated out by a combination of Ly6C, CD11C and CD49b staining. The $CD19^{+}Ly6C^{-}CD11c^{-}CD49b^{-}$ pro-B cell population is greatly reduced in *Justy* mice while the quadruple-negative pre-pro-B cell population remains normal. (B) The common lymphoid progenitor (CLP) population which gives rise to the pre-pro-B cells is not affected by the *Justy* mutation. *Justy* and WT bone marrow cells were stained with the indicated antibodies and analyzed by FACS. The $Lin^{-}IL\text{-}7R^{+}Flt3^{+}$ population contains the CLP and pre-pro-B cells and are also Sca-1^{+} and c-Kit^{+}. The $B220^{+}Ly6D^{+}$ cells are the pre-pro-B population. The $B220^{-}$ fraction is the CLP population which is further divided into the $Ly6D^{+}$ CLP (B lineage committed) and $Ly6D^{-}$ CLP (B lineage uncommitted) population. (C) The myeloid population is present in *Justy* bone marrow but its ratio over the lymphoid cells is altered. Whole bone marrow cells from WT and *Justy* mice were stained with the myeloid markers Gr-1 and Mac-1. The $Mac\text{-}1^{-}Gr\text{-}1^{-}$ cells are the lymphoid cells. Data are representative of 5 independent experiments for (A) with 2 sets done by Ping Lu and 3 sets done in collaboration with Mingyi Chiang. Data are representative of 2 independent experiments for (B) and (C).

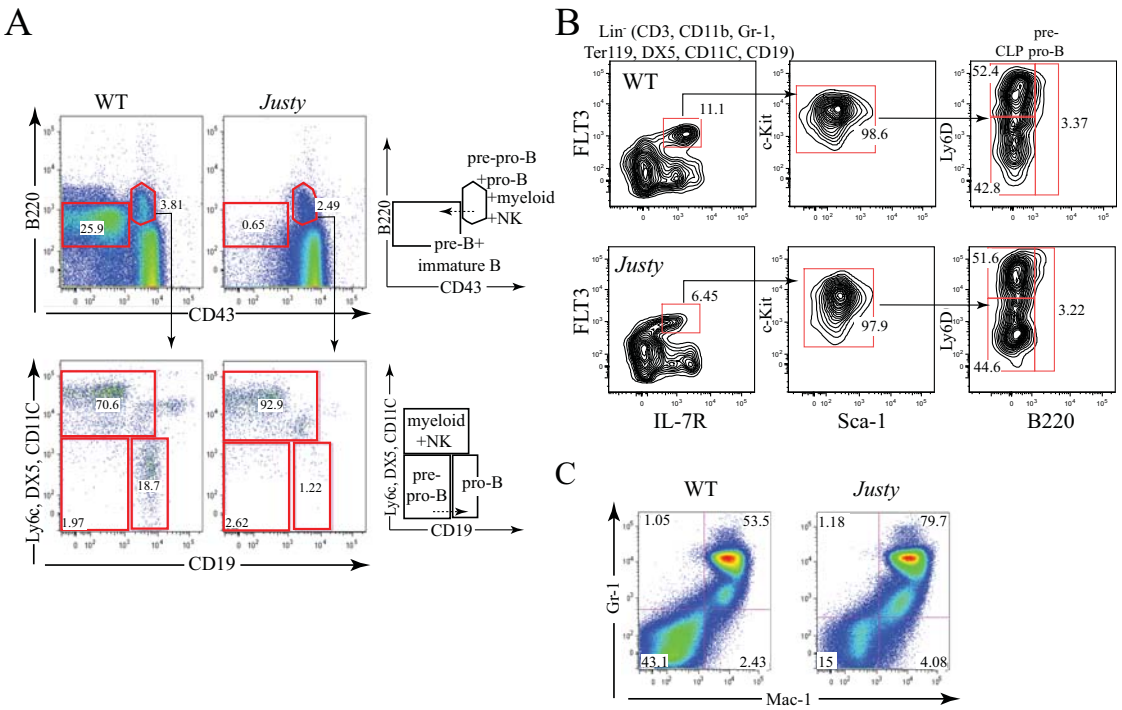

A
WT
Justy
B220
CD43
3.81
25.9
2.49
0.65
pre-pro-B
+pro-B
+myeloid
+NK
pre-B+
immature B
CD43
Ly6c, DX5, CD11C
CD19
70.6
18.7
1.97
92.9
1.22
2.62
myeloid
+NK
pre-
pro-B
pro-B
CD19
B
Lin- (CD3, CD11b, Gr-1,
Ter119, DX5, CD11C, CD19)
pre-
CLP pro-B
WT
FLT3
11.1
c-Kit
98.6
Ly6D
52.4
3.37
42.8
Justy
6.45
97.9
51.6
3.22
44.6
IL-7R
Sca-1
B220
C
WT
Justy
Gr-1
Mac-1
1.05
53.5
43.1
2.43
1.18
79.7
15
4.08

Figure 5. T cell development is grossly unaffected in *Justy* mice.

(A) FACS analysis of whole thymocytes. Whole thymocytes were collected from age-matched WT and *Justy* mice. The cells were then stained with CD3, CD4 and CD8 antibodies. The plot shows the analysis of $CD4^+$ and $CD8^+$ populations within the $CD3^+$ cells. (B) Plot of calculated total numbers of $CD4^+CD8^+$ double positive (DP), $CD4^-CD8^-$ double negative (DN), CD8 single positive (SP) and CD4SP cells obtained from individual thymocyte preparations. Cell numbers were calculated as described in Materials and Methods. Data are representative of 2 independent experiments.

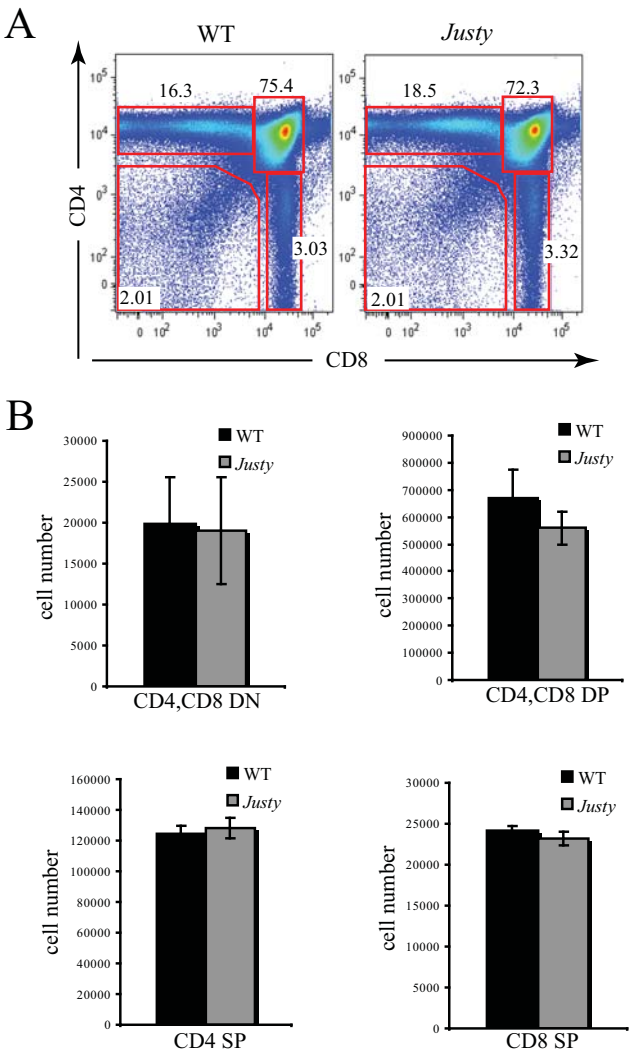
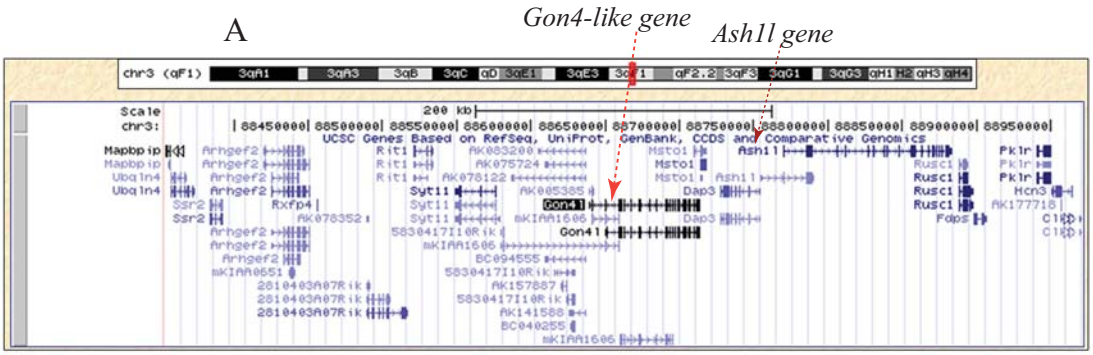
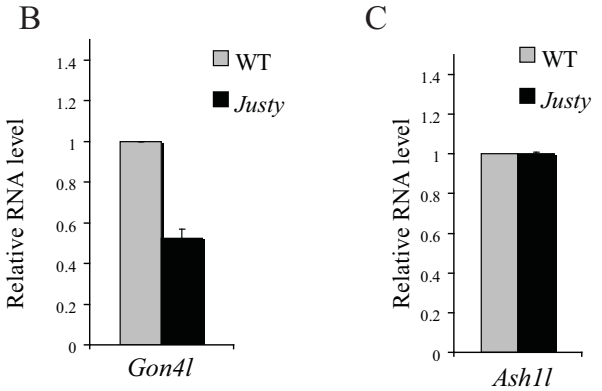

A
WT
Justy
16.3
75.4
3.03
2.01
18.5
72.3
3.32
2.01
CD4
CD8
B
WT
Justy
cell number
CD4,CD8 DN
CD4,CD8 DP
CD4 SP
CD8 SP

Figure 6. The *Justy* mutation affects *Gon4-like* gene expression.

(A) Schematic illustration of the 617kb region in chromosome 3 that contains the *Justy* mutation. The plot was copied from the UCSC gene browser. *Gon4l-like* and *Ash1l* are the two genes that encode putative transcriptional regulators and their relative positions are indicated by arrow heads. (B) The expression of *Gon4-like* is substantially reduced in *Justy* bone marrow Lin^- cells (lineage marker negative). Lin^- cells were isolated from WT and *Justy* bone marrow as described in Materials and Methods. RNA was then isolated and analyzed by quantitative RT-PCR. The expression values were first normalized to that of the internal control *Hprt*. The resultant values were then normalized again to the obtained WT/Hprt values. The graph shows the average of the ratios obtained and the standard deviation between them. This set of data is kindly provided by Isaiah Hankel. Data are representative of 3 independent experiments. (C) The expression of *Ash1l* is not affected by the *Justy* mutation. Bone marrow Lin^- cells were isolated and analyzed as in (B). Data are representative of 2 independent experiments.





Data from Isaiah Hankel

Figure 7. The *Justy* mutation causes aberrant splicing of RNA transcribed from *Gon4l*.

Schematic showing the location of the *Justy* mutation in the *Gon4l* gene. The plot is copied from UCSC gene browser (upper panel). The *Justy* mutation is within intron 24 of *Gon4l* as indicated by the dashed lines (middle panel). This mutation strengthens the homology to the consensus donor splice site sequence for RNA splicing (lower panel). The red vertical line among the nucleotides denotes a match to the consensus donor splice site sequence.

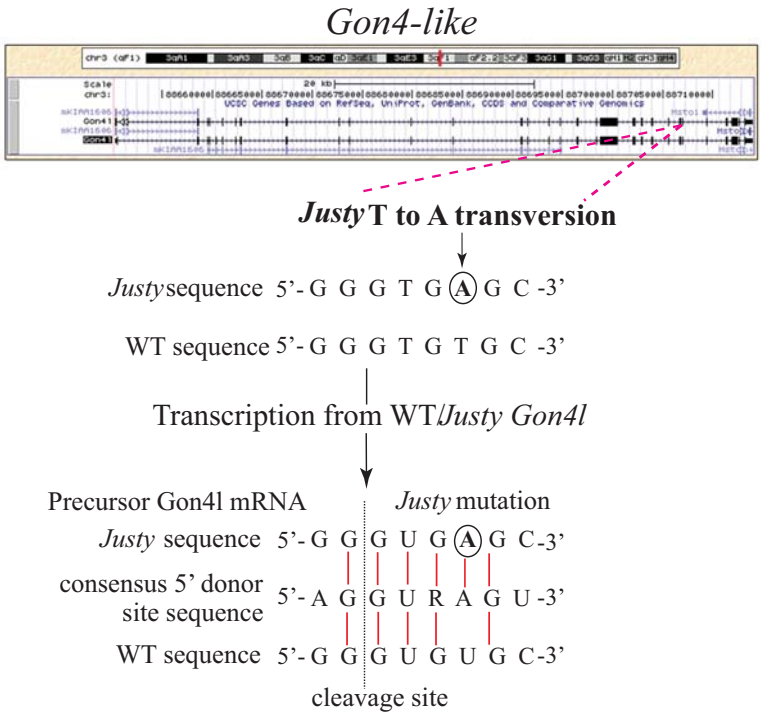

Gon4-like
***Justy* T to A transversion**
Justy sequence 5'- G G G T G A G C -3'
WT sequence 5'- G G G T G T G C -3'
Transcription from WT/*Justy Gon4l*
Precursor Gon4l mRNA
Justy mutation
Justy sequence 5'- G G G U G A G C -3'
consensus 5' donor site sequence 5'- A G G U R A G U -3'
WT sequence 5'- G G G U G U G C -3'
cleavage site

Figure 8. Gon4l protein is expressed in multiple tissues.

(A) Schematic showing the regions recognized by the two Gon4l antibodies (see Figure 13 in Chapter III for description of the antibodies). (B) Screening of Gon4l expression in tissues using the anti-Gon4l C terminus antibody (C-1746). Whole cell lysate (WCL) were prepared from brain, heart, thymus, spleen and bone marrow as indicated in Material and Methods. WCL from 293T cells transiently transfected to express a FLAG epitope tagged Gon4l (Gon4l-293T) was included as a positive control for Gon4l signal. Western blot (WB) of GAPDH was included as a loading control. The arrow head indicates Gon4l. (C) Screening of Gon4l expression in tissues using the anti-Gon4l central region antibody (N-938). WCL preparation and WB were carried out as in (B). (D) The expression of Gon4l in cultured WT pro-B cells. Lin^- bone marrow cells were isolated, differentiated and expanded in culture with addition of appropriate cytokines. Cells were then lysed and immunoblotted as indicated. WCL from M12 B cells was included as a positive control for Gon4l. Immunoblotting of GAPDH was included as an indication of equal protein loading. (E) Gon4l is expressed in the Jurkat T cell line. WCL from M12 cells was included as a positive control for Gon4l. WB of actin was included as a control of equal loading. The actin signal from the blank lane is due to leakage of the sample well in the protein gel. Data are representative of 3-5 independent experiments.

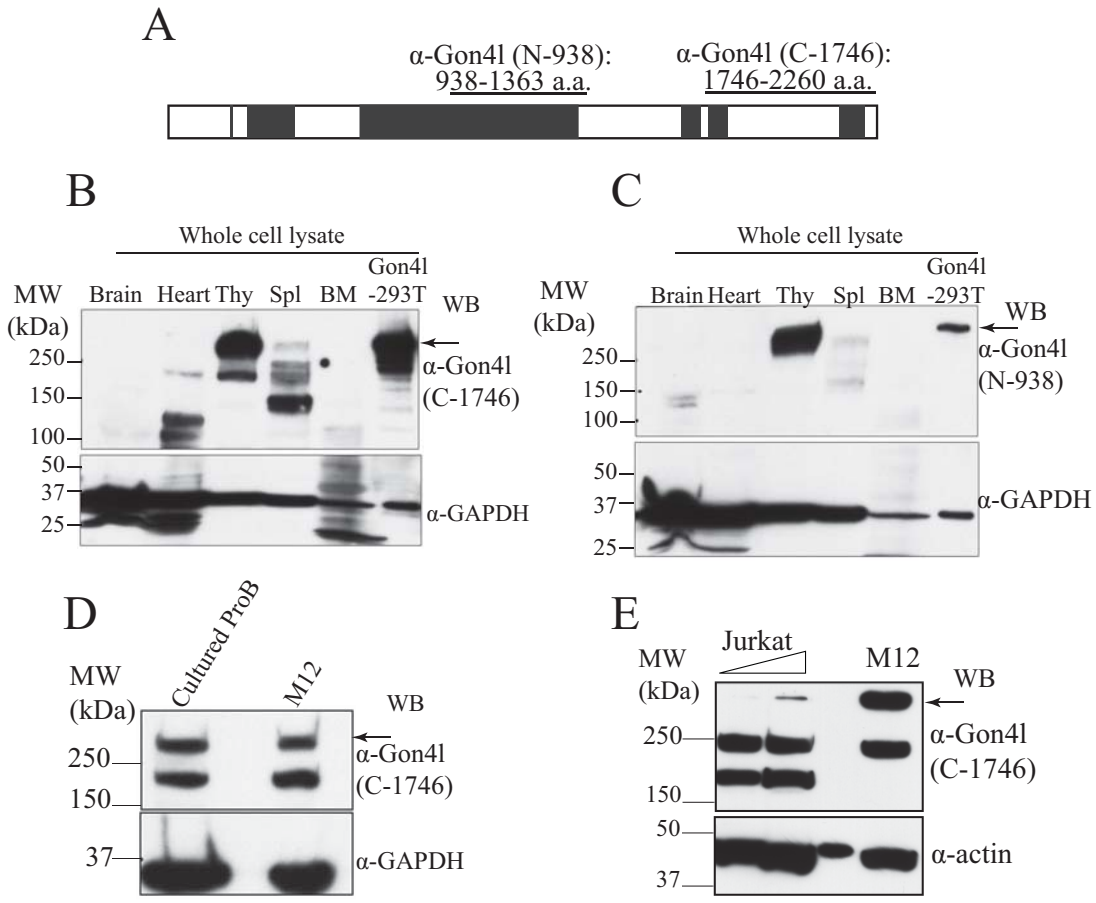
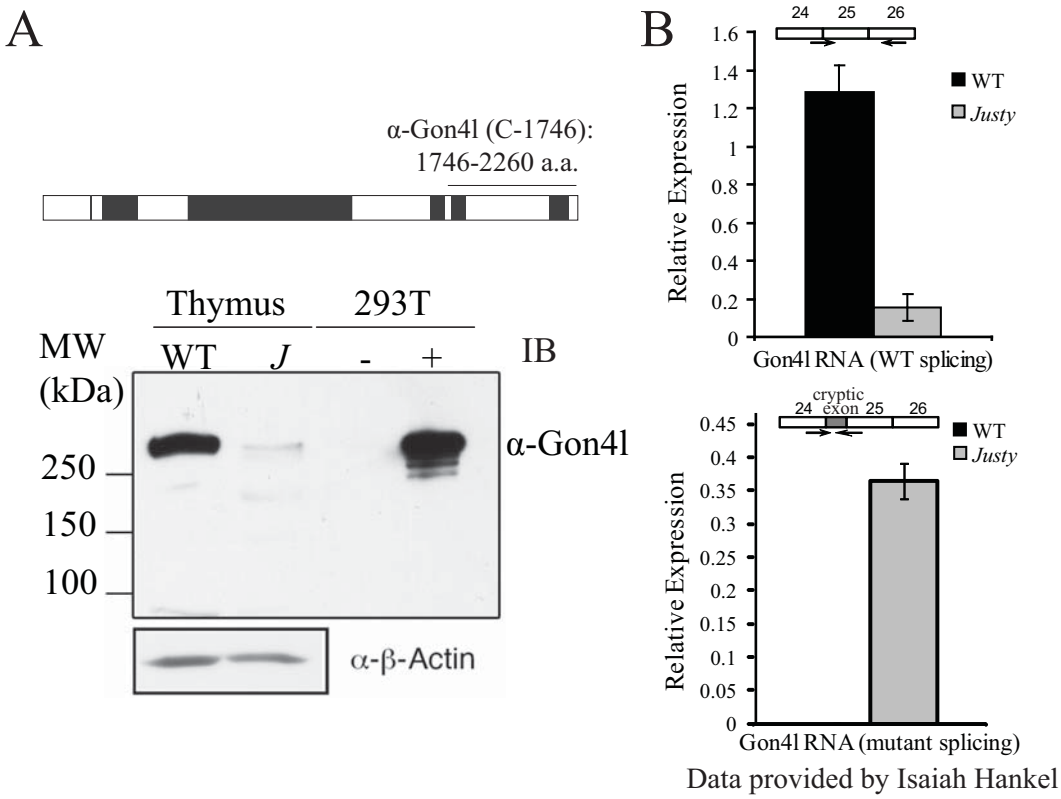

A
α-Gon4l (N-938):
938-1363 a.a.
α-Gon4l (C-1746):
1746-2260 a.a.
B
Whole cell lysate
Gon4l
MW
(kDa)
Brain
Heart
Thy
Spl
BM
-293T
WB
250
150
100
50
37
25
α-Gon4l
(C-1746)
α-GAPDH
C
Whole cell lysate
Gon4l
MW
(kDa)
Brain
Heart
Thy
Spl
BM
-293T
WB
250
150
100
50
37
25
α-Gon4l
(N-938)
α-GAPDH
D
Cultured ProB
M12
MW
(kDa)
WB
250
150
37
α-Gon4l
(C-1746)
α-GAPDH
E
Jurkat
M12
MW
(kDa)
WB
250
150
50
37
α-Gon4l
(C-1746)
α-actin

Figure 9. The aberrant splicing of *Gon4l* RNA as a result of the *Justy* mutation severely disrupts Gon4l protein expression.

(A) Gon4l protein expression is severely reduced in *Justy* thymocytes. Whole thymocytes were collected from WT and *Justy* mice, lysed and subjected to immunoblotting using the Gon4l C-1746 antibody. The schematic in the upper panel shows the region that is recognized by the Gon4l antibody. Whole cell lysate prepared from 293T cells or FLAG-Gon4l expressing 293T cells were included as controls for the Gon4l signal. Immunoblotting of β-actin was included as an indication of equal protein loading. Data are representative of 5 independent experiments. (B) The level of WT spliced *Gon4l* RNA is similarly reduced in *Justy* thymocytes. Q RT-PCR analysis of the levels of the normally-spliced form (upper panel) and *Justy* aberrantly spliced form (lower panel) of *Gon4l* RNA in thymocytes. Schematic showing the exon sequences detected by the primers is shown. Primers are indicated as arrow heads. The expression values were normalized to that for *Hprt* and plotted. Data shown are the average and standard deviation of the values collected. This set of data is kindly provided by Isaiah Hankel. Data are representative of 3 independent experiments.



Data provided by Isaiah Hankel

Figure 10. The hypothetical truncated Gon4l protein is not detected.

(A) Schematic of the hypothetical truncated Gon4l protein translated from the *Justy* aberrantly spliced *Gon4l* mRNA. The regions that are recognized by the two Gon4l antibodies are indicated. (B) The truncated Gon4l protein (~185kDa) is not detected by the N938 antibody. WCL was prepared from WT and *Justy* thymocytes. WB was carried out using the N-938 Gon4l antibody as in Figure 9. WCL from 293T and FLAG-Gon4l expressing 293T cells were included as controls for the Gon4l signal. Immunoblotting of β-actin was included as loading controls. A total of 15μg protein is loaded. (C) WB of WCL prepared from WT and *Justy* thymocytes with maximum loading. 50μl (maximum volume held by each well of the gel) of WT and *Justy* WCL (67μg and 138μg total protein respectively) was run on a SDS-PAGE gel followed by immunoblotting analysis. Data are representative of 3 independent experiments for (B) and 2 for (C).

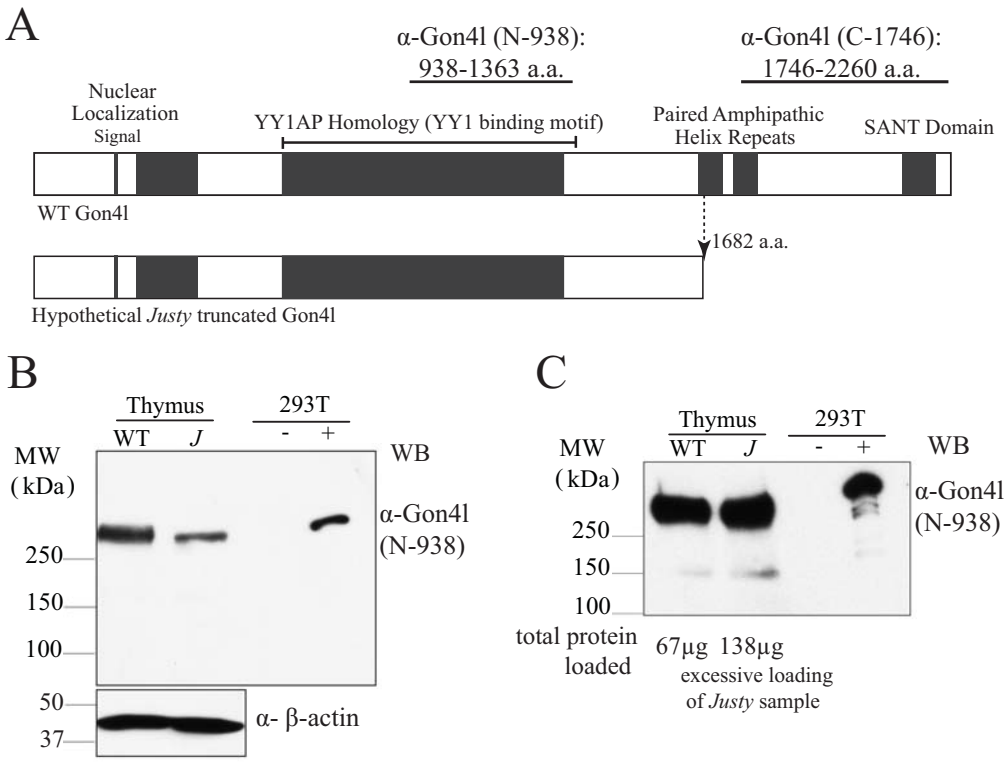
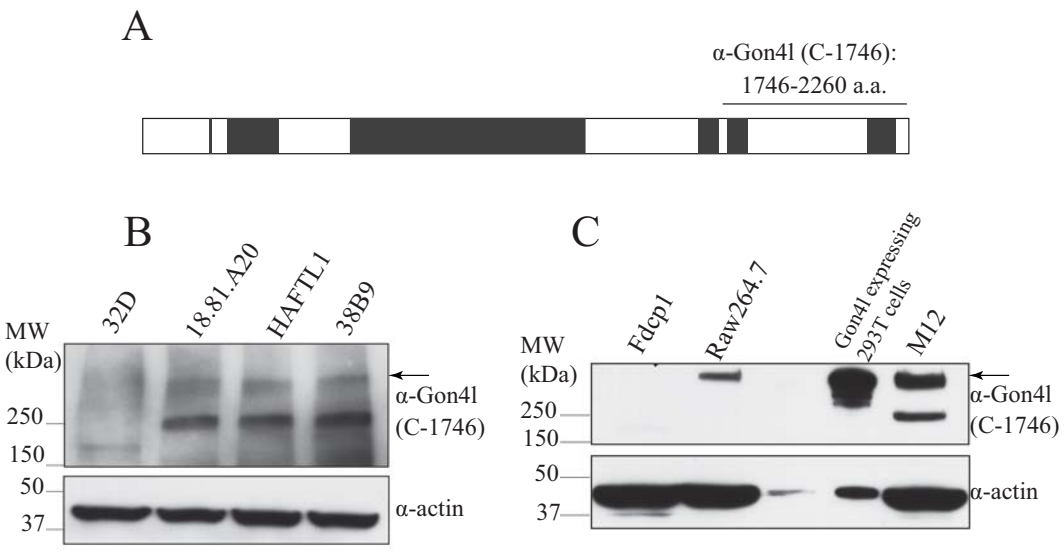

A
α-Gon4l (N-938):
938-1363 a.a.
α-Gon4l (C-1746):
1746-2260 a.a.
Nuclear
Localization
Signal
YY1AP Homology (YY1 binding motif)
Paired Amphipathic
Helix Repeats
SANT Domain
WT Gon4l
1682 a.a.
Hypothetical Justy truncated Gon4l
B
Thymus
293T
WT
J
-
+
MW
(kDa)
WB
α-Gon4l
(N-938)
250
150
100
50
37
α- β-actin
C
Thymus
293T
WT
J
-
+
MW
(kDa)
WB
α-Gon4l
(N-938)
250
150
100
total protein
loaded
67µg 138µg
excessive loading
of Justy sample

Figure 11. Gon4l is not expressed or expressed at very low levels in myeloid cell lines.

(A) Schematic of the region recognized by the C-1746 Gon4l antibody. (B) Gon4l is not expressed in the 32D cell line (myeloid precursor cells). WCL was prepared and WB was carried out using the C-1746 Gon4l antibody as in Figure 9. The B lineage lines 18.81.A20, HAFTL1 and 38B9 were included as positive controls for Gon4l. Lysates were also probed with antibodies specific for β-actin to confirm equal protein loading. (C) Gon4l is not expressed in Fdcp1 cells (myeloid precursor cells) and is expressed in greatly reduced levels in Raw264.7 cells (mature macrophage cells). Lysates were prepared and analyzed as in B. WCL from M12 and FLAG-Gon4l expressing 293T cells were included as positive controls for Gon4l. Arrow heads denote the Gon4l protein. All data are representative of 2 independent experiments.


A
α-Gon4l (C-1746):
1746-2260 a.a.
B
32D
18.81.A20
HAFTL1
38B9
MW
(kDa)
250
150
50
37
α-Gon4l
(C-1746)
α-actin
C
Fdcp1
Raw264.7
Gon4l expressing
293T cells
M12
MW
(kDa)
250
150
50
37
α-Gon4l
(C-1746)
α-actin

CHAPTER III

GON4L AND THE TRANSCRIPTIONAL REGULATORS YIN YANG 1, SIN3A AND HDAC1 ASSOCIATE IN A COMPLEX

Abstract

The *Justy* mutation is a recessive point mutation in mice that causes aberrant splicing of the *Gon4l* RNA and disrupts its protein expression. As a consequence, B cell development in *Justy* mice is blocked at a very early stage, indicating that Gon4l is a critical factor for B cell development. In *Justy* B cell precursors, the absence of Gon4l results in ineffective repression of myeloid genes as well as many other genes, connecting Gon4l with the regulation of transcriptional repression. In this study, Gon4l protein was shown to bear homologies to classical transcriptional regulators such as the co-repressors which often form multimeric-protein-complexes to function. In the search for Gon4l co-factors, Gon4l was found to interact with the DNA-binding transcription factor Yin Yang 1 (YY1) and co-localized with it in the nucleus. Density gradient centrifugation of protein lysates from a mouse B cell line showed that Gon4l and YY1 both co-sedimented with the transcriptional co-repressor Sin3a and its protein partner HDAC1. Further analysis demonstrated that YY1 and Sin3a/HDAC1 all associated with Gon4l as components of a complex. Both YY1 and Sin3a/HDAC1 are well-documented mediators of transcriptional repression. Thus, these findings suggest a model where Gon4l can mediate transcriptional repression by employing its co-factors in the identified complex.

Introduction

B lymphocytes are generated from bone marrow hematopoietic stem cells through a process called B lymphopoiesis (3). In this process, gene expression is intensively reprogrammed in progenitor cells to favor differentiation toward the B lineage pathway. As a result, genes associated with the B lineage are activated (known as B lineage specification) and at the same time, genes associated with other lineages are repressed to

prevent the differentiation into other lineages (known as lineage commitment) (26,214-216). The repression of alternative lineage genes is active and continuous, and is the key for maintaining a B lineage identity. This repression is required even after the cells have become fully differentiated (113).

In eukaryotic cells, gene repression can be achieved by epigenetic modification of the chromatin (132,133,139,217). Acetylation of histone tails neutralizes the positive charge carried by the histones and reduces the interaction affinity between histones and the negatively charged DNA. Therefore histone acetylation is associated with active transcription. On the other hand, removing the acetyl groups from histone tails by histone deacetylase complexes (HDACs) is associated with formation of a closed chromatin structure and transcriptional repression (132,133,135,139,217). HDACs are often recruited by adaptor proteins known as transcriptional co-repressors, which often form large multimeric complexes that contain DNA binding transcription factors and other subunits. Transcription factors are thought to provide the co-repressor complexes with gene targeting specificity. Higher eukaryotes express several other proteins that are classified as co-repressors. These include Sin3a/Sin3b, Mi-2β, N-CoR and SMRT (69,218-221).

In B lymphopoiesis, transcriptional co-repressors containing HDACs are recruited to confer gene repression. For example, Ikaros, one of the earliest transcription factors required for lymphoid specification, associates with components of the NURD complex (nucleosome remodeling and histone deacetylation) and HDACs (69). Ikaros also associates with the Sin3 co-repressors and acts as a potent transcriptional repressor (68,71). Later in B cell development, the transcription factor Pax5 is key for repressing non-B lineage genes. Pax5 recruits Grg4 (Groucho-related-gene), a member of the Groucho co-repressor family that interacts with HDACs, to mediate transcriptional repression (124,127,152,153). Although the importance of the co-repressors in regulating

B lymphopoiesis has been more and more appreciated, the underlying mechanisms still remain to be elucidated.

Previous analysis of the *Justy* mice revealed that the *Gon4l* gene and its encoded protein are critical for B lymphopoiesis (192). The *Justy* mutation causes aberrant splicing of *Gon4l* RNA and disrupts Gon4l protein expression. B cell development in *Justy* mice is blocked at the B lineage commitment phase (pre-pro-B to pro-B transition), which correlated with severe loss of transcriptional repression of key myeloid genes and many other genes (192). Those genes are normally targeted for repression in WT cells at the same stages. This finding suggests that Gon4l is involved in pathways that regulate transcriptional repression important for successful B lymphopoiesis. However, the underlying molecular mechanisms are unknown.

In this study, the Gon4l protein sequence was analyzed and was shown to bear multiple conserved motifs that are also found in transcriptional co-repressors. Those motifs have been shown to mediate protein-protein interactions, especially those implicated in transcriptional regulation. The most conserved region in Gon4l suggested that Gon4l interact with the DNA-binding transcription factor Yin Yang 1 (YY1). Indeed, Gon4l physically associated with YY1 and co-localized with it in the nucleus. In the search for more Gon4l co-factors, the co-repressors Sin3a/HDAC1 were found to co-sediment with Gon4l and YY1 in gradient centrifugation assays. Further, Sin3a/HDAC1 associated with Gon4l and YY1 as components of a complex when expressed in transfected cells. These findings suggest that Gon4l could function by associating and employing its co-factors to mediate transcriptional repression.

Materials and Methods

Analyses of the protein homologies

Putative protein structural domains were identified using the SMART databases and NCBI BLAST tool. Protein sequence conservation was analyzed by using the

Vector-NTI software (Invitrogen). NCBI accession number for the protein sequences: Human YY1AP (NP_620829.1), Human GON4L (NP_001032622), mouse Gon4l (Q9DB00), zebra fish ugly duckling (ABP65284), C.elegans Gon4 (NP_001122786.1), and Drosophila CG34415 (NP_001097352, previously known as Cdp1, CG33456) (201,222).

Cell culture and transient transfection

M12, HAFTL1, 38B9, 18.81A20 and Jurkat cells were cultured in RPMI 1640 medium containing 10% FBS, 100-U/ml penicillin, 100-µg/ml streptomycin, and 2 mM L-glutamine. HEK (Human embryonic kidney) -293T cells were maintained in Dulbecco's modified Eagle's medium containing 10% fetal bovine serum, 100-U/ml penicillin, 100-µg/ml streptomycin, 1mM sodium pyruvate and 2 mM L-glutamine.

Transfection and plasmids

Transient transfections were carried out using Lipofectamine 2000 (Invitrogen) following the instructions provided by the manufacturer. pcDNA3.1 vector was used as the backbone for all the Gon4l constructs. The FLAG tag and *Gon4l* cDNA or its derivatives were cloned in frame into pcDNA3.1 vector by standard methods. HDAC1 cDNA was PCR amplified from pCMV-SPORT6-HDAC1 (Open-biosystems: MMM1013-98478662) and then subcloned into pCMV-HA (Clontech: PT3283-5) to generate pCMV-HA-HDAC1. pSPORT-YY1 was purchased from Open-biosystems (MMM1013-9200816). pCS2+MT-mSin3A and its truncated mutants are kind gifts from Dr. Donald E. Ayer at the University of Utah.

Immunoprecipitation assay

Cells were lysed in lysis buffer containing 50 mM Tris pH 8.0, 120 mM NaCl, 0.5% (v/v) NP-40, 0.2 mM sodium orthovanadate, 100 mM NaF, 50 µg/ml PMSF, and protease inhibitor cocktail (Roche: 11697498001). Cell lysates were incubated for 15min

at 4ºC with gentle rotation and centrifuged at 16.1×g for 10 min at 4°C. Supernatants were collected and incubated with anti-FLAG (Sigma: M2), or anti-Sin3a (Santa Cruz: K-20), or anti-HDAC1 (Santa Cruz: H51) antibodies bound with Dynal protein G beads (Invitrogen: 100.03D). The binding of antibodies to the beads was performed following the manufacturer's instructions. In the co-immunoprecipitation of endogenous YY1 and Gon4l, whole cell lysate was incubated with anti-YY1 antibody (Santa Cruz: H414) and protein G agarose beads (Santa Cruz: sc-2002) as described (223). Immunoprecipitates were all eluted with 1X Nu-PAGE LDS-sample buffer (Invitrogen: NP0007), unless otherwise specified. In all sequential IPs, immunoprecipitates were first eluted with 100ug/ml FLAG-peptide (Sigma: F3290) after the anti-FLAG IP and then subject to secondary IPs using the indicated antibodies. In all experiments, IP with equivalent amount of corresponding normal IgG (Santa Cruz: mouse/rabbit sc-2025; sc-2027) were used as a negative control.

Immunoblot analysis

Samples were electrophoresed on 7% or 3-8% Nu-PAGE gels (Invitrogen) and transferred to PVDF membranes (Millipore). Membranes were probed with the following antibodies: anti-Gon4l (C-1746 and N-938), anti-FLAG (Sigma: M2), anti-YY1, anti-Myc, anti-HDAC1, anti-Sin3a, anti-STAT5, anti-Ebf-1 (Santa Cruz: H-414, 9E10, H-51, K-20, RK5C1, C-17, H-300) and anti-HA (Roche: 3F10). Protein bands were visualized by using SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific).

Sucrose Gradient centrifugation

M12 cells (5×10^6-10^7 cells) were lysed in PBS containing 0.5%Triton X-100 and 1mM EDTA. Cell lysates were separated by centrifugation at 41,000 rpm in a swinging bucket rotor (Beckman: SW41Ti) for 16 hours (Beckman: L8-70M ultracentrifuge). Sedimentation standards (Bio-Rad: precision-plus protein standards, and Sigma: Thyroglobulin) were analyzed in parallel gradients. Fractions of 500 µl were collected,

and 33 µl of each fraction was analyzed by SDS-PAGE (224). Thymocytes from WT and *Justy* mice were analyzed similarly as the M12 cells.

Immunofluorescence and confocal microscopy

HEK 293T cells were transfected with pcDNA-FLAG-Gon4l and pSPORT-YY1 for 24 hours, re-plated onto poly-L-lysine mounted glass coverslips (BD: 354085) followed by another overnight incubation. Cells were fixed with ice-cold methanol at -20ºC for 10 minutes and permeabilized with 0.2% Triton X-100 in PBS. After being blocked with 5% normal goat serum (Sigma: G9023), the cells were incubated overnight simultaneously with rabbit anti-YY1 (Santa Cruz: C-20) and mouse anti-FLAG (Sigma: M2) antibodies diluted in the blocking solution. Cells were then extensively washed, and incubated with Goat anti-Rabbit conjugated to Alexa Fluor 568 and Goat anti-Mouse conjugated to Alexa Fluor 488 (Invitrogen: A21069 and A11017). Nuclei were counterstained with TO-PRO-3 iodide (Invitrogen: T3605). Images were captured using a Zeiss 510 confocal microscope.

Results

Gon4l protein contains homologies to classical transcriptional regulators

To help unravel the molecular function of Gon4l, its amino acid sequence was analyzed. Consistent with its critical role in B lymphopoiesis, Gon4l was phylogenetically conserved (Figure 12A). Using the SMART database, multiple structural motifs in mouse Gon4l protein were revealed, some of which were also reported for other Gon4l orthologs (Figure 12B) (202,213). The N-terminal region of Gon4l contained a putative nuclear localization sequence (NLS), suggesting Gon4l is a nuclear protein. In the same region, a weak homology to nucleoplasmin (256-410a.a.) was found. Nucleoplasmin is a protein important for chromatin assembly and gene

expression regulation (225-227). The central portion of Gon4l (611-1364a.a.) was found 78% similar to a human protein known as YY1-associated protein (YY1AP). YY1AP interacts with the transcription factor YY1 (Yin Yang 1) and facilitate YY1-mediated transcriptional regulation (228). Overlapping with the C-terminus of the YY1AP domain is an α-amylase inhibitor (AAI) homology (1011-1231a.a.), another motif known to bind various macromolecules (229). The C-terminus of Gon4l contains 2 regions (amino acid: 1670-1718, 1752-1799) homologous to the paired amphipathic helix (PAH) repeat sequences which are also found in the transcriptional co-repressors, Sin3a and Sin3b. This motif is thought to mediate protein-protein interactions, including those with HDACs (230). Lastly, nuclear magnetic resonance (NMR) analysis (protein data bank accession code 1UG2) showed that the most C-terminal residues of Gon4l form a helix-loop-helix structure called a SANT domain (SWI3, ADA2, N-CoR and TFIIIB) (231). The SANT motif is found in many transcription factors like c-Myb which utilizes the SANT domain to bind histone H3 and facilitate histone H3 acetylation (232). Thus, these findings suggest that Gon4l is a nuclear protein which contains multiple motifs implicated in interacting with factors involved in transcriptional regulation.

The YY1AP homology is the most conserved domain.

To refine our analysis of the Gon4l protein, the amino acid sequences of 6 Gon4l orthologs from invertebrate and vertebrate species were subjected to pair-wise ClustalW comparisons to identify the most conserved residues. Sequences spanning the putative domains described above had higher levels of conservation relative to other regions (Figure 12 B-D). The most conserved regions (amino acids 723-754 and 825-856; > 90% similarity) were found near the N-terminus of the YY1AP homology overlapping with residues required for the binding of YY1AP to YY1 *in vitro* (228). The *YY1AP* gene is unique to primates and was apparently generated by duplication of *Gon4l* (201). Therefore, these findings suggest that Gon4l would likely interact with YY1 as does

YY1AP. Moreover, the high conservation of the residues also suggests that such an interaction may represent the most important aspect of Gon4l's function.

Gon4l is expressed in B lineage cell lines

In order to analyze the function of Gon4l, two anti-Gon4l polyclonal antibodies were generated with one specific for the C-terminal region of Gon4l (C-1746) and the other for a portion of the YY1AP homology (N-938) (Figure 13A). Screening of cell lines representing different stages of B cell development (pro-B to mature B stage) revealed abundant Gon4l expression in all the lines tested, suggesting Gon4l may play a role in both B cell precursors as well as in mature B cells (Figure 13B). Notably, a slightly smaller protein with a molecular weight of around 200 kDa was recognized by the C-1746 antibody. Both signals were also detected in lysates prepared from pro-B cells generated in culture (Figure 8D). However, the N-938 antibody detected only the full-length Gon4l, not the smaller protein (Figure 13A, right panel). Thus, the smaller protein is either a Gon4l isoform that lacks a portion of the YY1AP homology, or a protein encoded by another gene which cross-reacts with the C-1746 antibody.

Gon4l associates with YY1

The high conservation of the YY1AP homology domain in Gon4l suggests Gon4l interact with YY1. Thus, the existence of such interactions was examined. YY1 antibody or control IgG were used to immunoprecipitate (IP) proteins out of lysates prepared from mouse M12 B cells which express both Gon4l and YY1 from the endogenous genes (Figure 14A). Immunoblot analysis of the recovered material using the C-1746 Gon4l antibodies revealed that Gon4l was specifically co-immunoprecipitated (co-IP) with YY1. The same Gon4l signal was detected when the N-938 Gon4l antibody was used for immunoblotting (data not shown). Further, IPs done with lysates from Jurkat T cells revealed the same association between YY1 and Gon4l, suggesting this interaction may be a general mechanism by which Gon4l renders its function (Figure 14C).

The association between YY1 and Gon4l expressed from plasmids was assessed (Figure 14B). Plasmids encoding mouse YY1 or FLAG epitope-tagged Gon4l were transiently co-transfected into 293T cells. Immunoblot analysis of material recovered by IP with anti-FLAG antibodies demonstrated that YY1 and FLAG-Gon4l could be co-immunoprecipitated. This result suggests that the system could be used to define the amino acids in Gon4l required for interacting with YY1 (see Figure 21 in Chapter IV). Whether YY1 and FLAG-Gon4l co-localized in cells was then determined. The YY1 and Gon4l expression plasmids described above were again co-transfected into 293T cells which were then stained with the appropriate antibodies and analyzed by immunofluorescent confocal microscopy (Figure 14D). FLAG-Gon4l was detected in the nucleus while YY1 was seen in both the cytoplasm and the nucleus as previously reported (233). Among 60 cells staining positive for both proteins, 57 (~95%) showed overlapping nuclear localization of YY1 and Gon4l. These data demonstrate that Gon4l and YY1 can associate with each other and co-localize in the nucleus.

Gon4l co-sedimented with YY1, Sin3a and HDAC1

In order to identify more factors that might associate with Gon4l and YY1, density gradient co-sedimentation assays were employed. Whole cell lysates prepared from M12 B cells were separated in a 20-50% linear sucrose gradient by centrifugation. Protein standards were separated in parallel in identical gradients. The gradient was collected in fractions after the centrifugation and analyzed by immunoblot analysis using anti-Gon4l or anti-YY1 antibodies (Figure 15A). Some of the full-length Gon4l was detected in fractions corresponding to a molecular weight of ~250kDa, which is equivalent to the mass of monomeric Gon4l. Gon4l was also found in fractions containing molecules of a molecular weight more than 670kDa. YY1 was detected in the low molecular weight fractions but was also found in the same high molecular weight

fractions where Gon4l was detected, consistent with the finding that Gon4l and YY1 are in association with each other in M12 cells.

Gon4l contains two PAH repeats that may be utilized to interact with HDACs. Thus HDACs and their co-factors were screened. Immunoblot analysis revealed that HDAC1 and its associated co-repressor—Sin3a co-sedimented at the same high molecular weight regions as YY1 and Gon4l (Figure 15A). Screening of B lineage transcription factors, EBF (also known as Ebf-1), STAT5, Ikaros and Pax5, indicated that none of these proteins co-sedimented with Gon4l in the high molecular weight fractions (Figure 15A, B and data not shown). Therefore, these results suggest that, in addition to YY1, Sin3a and HDAC1 may be co-factors of Gon4l.

Gon4l interacts with the transcriptional co-repressor Sin3a

To further explore the relationship between Gon4l and Sin3a, the association between the two proteins was examined. Expression plasmids encoding FLAG-Gon4l and myc-tagged-Sin3a were transiently co-transfected into 293T cells. WCL were prepared from the transfected cells and subject to IPs using anti-FLAG antibodies (Figure 16). Immunoblot analysis demonstrated that Sin3a specifically co-immunoprecipitated with Gon4l. This association required regions in Sin3a spanning from the PAH2 domain to the C-terminal end (Fig. 16B and C), consistent with other reports that this region mediates most of the protein-interaction known for Sin3a (158,230). Furthermore, immunofluorescent confocal microscopy analysis indicated that Gon4l and Sin3a co-localized in the nucleus when both were expressed in 293T cells (Dr. Colgan unpublished data). Collectively, these data indicate that Gon4l associates with Sin3a.

Gon4l, YY1 and Sin3a/HDAC1 associate within the same protein complex

Whether YY1 and Sin3a could interact with Gon4l at the same time was tested. Expression plasmids encoding FLAG-Gon4l, myc-Sin3a and YY1 were co-transfected

into 293T cells and IPs were performed using the anti-FLAG antibodies. Subsequent immunoblot analysis detected Gon4l, Sin3a and YY1 in the recovered material (Figure 17A). Sequential IPs were then performed. Anti-FLAG antibodies were used to first isolate complexes containing FLAG-Gon4l; and the resultant immunoprecipitates were subjected to a secondary IP against the myc tag on the Sin3a protein to isolate Sin3a-containing complexes within the FLAG-Gon4l complex pool. Immunoblot analysis showed that Gon4l, Sin3a and YY1 were all recovered by this procedure (Figure 17B), demonstrating that these three proteins could associate as subunits of a complex.

Because HDAC1 is a co-factor of Sin3a and co-sedimented with Gon4l, sequential IP analysis was performed to determine if HDAC1 could associate with the complex containing Gon4l, YY1, and Sin3a. Expression plasmids encoding the four proteins were co-transfected into 293T cells and sequential IPs were carried out. Anti-FLAG antibodies were used first and either anti-HDAC1 or Sin3a antibodies were used for the secondary IPs (Figure 17C and D). In both experiments, all four proteins were recovered by both the first and second IPs. Thus, combining with the co-sedimentation data, these results indicate that the four proteins are components of the same complex/s.

Screening of Gon4l complex in WT and *Justy* thymocytes

Because Gon4l co-immunoprecipitated with YY1 in Jurkat T cells, whether Gon4l forms a similar complex with YY1 and Sin3a/HDAC1 in thymocytes was examined. Whole cell lysates were prepared from WT thymocytes and separated in a 20-50% linear sucrose gradient by centrifugation as mentioned above. Gradient fractions were collected and analyzed by immunoblot analysis (Figure 18A). As expected, a portion of full-length Gon4l was detected in fractions containing the monomeric Gon4l which has a molecular weight of 250kDa. Gon4l was also recovered in fractions with a molecular weight slightly less than 670 kDa but much higher than 250kDa, indicating that Gon4l forms certain undefined complexes in thymocytes.

YY1, Sin3a and HDAC1 were all recovered from the high molecular weight fractions that contained Gon4l (Figure 18A). However, when the overall patterns were considered, the co-sedimentation of the four proteins was only partial and there was not a region where the signals from the four proteins peaked together as seen in Figure 15A. Thus, it was difficult to determine whether YY1, Sin3a/HDAC1 were components of the Gon4l-complexes or not. Gon4l protein is severely reduced in *Justy* thymocytes. Whole cell lysate prepared from *Justy* thymocytes was analyzed in parallel to see if the lack of Gon4l protein would disrupt the co-sedimentation of the other three proteins (Figure 18B). However, no obvious sedimentation pattern changes were observed for YY1, Sin3a or HDAC1. Thus, these data suggest that although Gon4l may form complexes in the thymocytes, their composition is likely different from those in the M12 B cells.

Discussion

The *Justy* mutation causes aberrant splicing of *Gon4l* RNA and reduces the expression of its encoded protein (192). As a result, B cell development in *Justy* bone marrow is blocked at the lineage commitment phase, which correlates with defects in the repression of key myeloid genes and many other genes, implicating Gon4l in the regulation of transcriptional repression. However the molecular mechanisms underlying these effects are unclear. In this study, evidence is provided to support a model where Gon4l could mediate transcriptional repression by forming a protein complex with the transcription factor YY1 and the classical co-repressor Sin3a/HDAC1.

The protein sequence of Gon4l was first analyzed for some clues of its molecular function (Figure 12). Multiple conserved domains were identified in Gon4l. Most of those domains are also found in classical transcriptional co-repressors in which the domains are utilized to mediate interactions with proteins critical for transcriptional regulation (*i.e.* histones, HDACs and DNA-binding transcription factors). Furthermore, Gon4l was found to have a functional nuclear localization signal (NLS) which facilitates

Gon4l to localize to the nucleus where transcriptional regulation takes place (Dr. Colgan unpublished data). Therefore, the domains found in Gon4l strongly suggest that this protein may play a role in transcriptional repression by interacting with other transcriptional regulators, just like the co-repressors.

Gon4l protein is expressed in cell lines representing different developmental stages, indicating that this protein may play a role in all these stages (Figure 13). The C-1746 Gon4l antibody recognizes a smaller protein species in all the B cell lines tested as well as the pro-B cells expanded *in vitro* (Figure 8D, chapter II). This smaller band, however, is not detected by the N-938 Gon4l antibody, which recognizes the YY1AP homology domain in Gon4l. This result suggests that the smaller protein species may be a Gon4l isoform that lacks the YY1AP homology domain. If this is true, the two isoforms likely share the same domains on both the N- and C-terminus, which can be utilized to mediate similar protein-protein interactions. By competing for the same protein partners, the two Gon4l isoforms may be able to regulate the function of each other. In *Justy* cells, the loss of the larger form Gon4l could lead to the imbalance between these two isoforms, which may contribute to the B cell developmental defects in *Justy* mice.

Alternative splicing of GON4L RNA occurs in human, and that leads to the generation of two different protein species (234). Yet, the alternative splicing of murine Gon4l RNA has not been reported. Further, the possibility that the smaller band detected is due to cross-reactivity of the antibody still could not be ruled out. The final determination of the identity of this band may rely on mass spectrometry analysis.

The interaction between Gon4l and other proteins was then examined. The most conserved region in Gon4l suggested that it may interact with the transcription factor YY1. Indeed, Gon4l was found associated with YY1 in the nucleus. The fact that YY1 is a zinc finger protein that binds DNA explains how Gon4l could mediate transcriptional repression while it does not contain a DNA-binding domain to allow target gene binding. More importantly, multiple reports have shown that YY1 is capable of recruiting co-

factors to regulate transcriptional repression (168,170,235). Thus, Gon4l could directly participate in the repression of a gene via its interaction with YY1.

In the further search for more Gon4l co-factors, the co-repressor complex Sin3a/HDAC1 was found to interact with Gon4l and YY1 simultaneously. In fact, the four proteins associated as components of a complex (Figure 17). Sin3a/HDAC1 is the core-complex of the many Sin3a co-repressor complexes, which are critical for a variety of biological processes. A common mechanism for Sin3a complexes to mediate site specific transcriptional repression is through the interaction with transcription factors (158,221,230). Known transcription factors that interact with Sin3a complexes include Mad family proteins, Sp1-like transcription factors, and the tumor suppressor P53 (158). Our finding that YY1 and Sin3a associate suggests that Sin3a may play a role in B cell development, since YY1 is required for this process (182). In support of this, conditional inactivation of *Sin3a* by *Mx-cre*, where Cre expression is under the control of the Type I interferon (IFN)-inducible Mx1 promoter (236), leads to significant B cell development arrest (Ping Lu and Dr. Colgan personal communication). However, whether and how the Gon4l-complex regulates B cell development remains to be examined.

The interaction between Gon4l and Sin3a requires a region in Sin3a that has been repeatedly reported for mediating interactions between Sin3a and other proteins (230). Thus, Gon4l may represent a yet unknown factor of the Sin3a complexes. Gon4l contains a SANT domain in its C-terminus end. Interestingly, a SANT-containing protein has long been proposed to be in the Sin3a complexes (237).

So far, there has not been any report suggesting YY1 and Sin3a function together, although SAP30, a component of the Sin3a/HDAC complexes, has been shown to associate with YY1 and facilitate YY1 mediated transcriptional repression (235). Our study, thus, for the first time, connected Sin3a/HDAC1 and YY1 into a single complex. The studies on Sin3 proteins started decades ago and many complexes containing Sin3a/HDACs have been isolated (158). Why the Gon4l-YY1-Sin3a/HDAC1 complex

has been missed is unclear. One possible reason is that this complex is B lineage specific. Our analysis showed difference in Gon4l-complexes formed in thymocytes compared to those formed in M12 B cells. It is also unclear whether Gon4l is required for YY1 to associate with Sin3a/HDAC1. If not, Gon4l may be required to connect YY1 and Sin3a/HDAC1 to other yet unidentified factor/s. But if Gon4l does serve as a link between YY1 and Sin3a/HDAC1, the severe reduction of Gon4l protein in *Justy* B progenitor cells would lead to defects in the recruitment of Sin3a/HDAC1 to YY1 and disrupts the repression of YY1 regulated genes.

The key myeloid genes *Sfpi1* (encodes PU.1) and *C/EBPα* are dysregulated and expressed at highly elevated levels in *Justy* pro-B cells (192). Bioinformatic analysis of the promoter regions of these two genes revealed multiple YY1 consensus binding sites (Ping Lu searches using MmPD and the CONSITE databases), suggesting that YY1 could directly repress these genes. This is consistent with previous reports indicating that YY1 is a negative regulator of G-CSF (granulocyte -colony-stimulating factor) induced myeloid differentiation in both primary bone marrow cells and myeloid precursor lines (238). However, the current model that Gon4l facilitates the recruitment of Sin3a/HDAC1 to YY1 to mediate transcriptional repression needs to be tested in more detail, especially in a system where its impact on B lymphopoiesis could be examined.

In addition to the regulation of gene expression, successful B cell development also requires ordered rearrangement of the *Ig H* and *L* gene (V(D)J recombination) (28,239). Analysis of the *Justy* B progenitor cells revealed defective D_H to J_H rearrangement (Mingyi Chiang and Dr. Colgan unpublished data), suggesting a role of Gon4l in this process. Whether Gon4l participates in V(D)J rearrangement in the form of the Gon4l-YY1-Sin3a/HDAC1 complex has to be tested. However, some evidence exists to support this notion. First, deletion of YY1 leads to defective *IgH* gene rearrangement, suggesting that YY1 is required for this process. On the molecular level, YY1 directly binds the intronic enhancer ($E_{i\mu}$) and mediates DNA looping in the *IgH* locus (182).

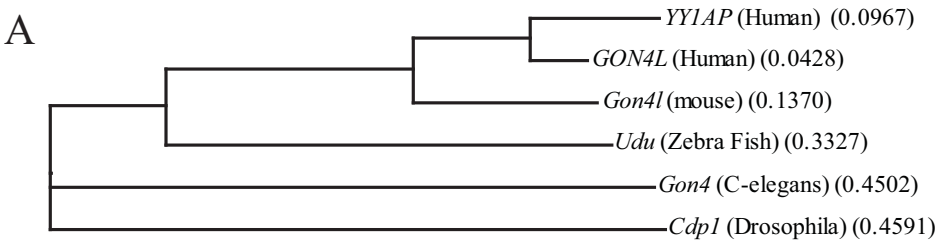
Second, the central region of the D_H gene segment is maintained in a heterochromatic configuration by continuous histone deacetylation during D_H to J_H rearrangement (36). Therefore, both YY1 and HDAC activities are required for the *IgH* gene rearrangement. Third, the D_H and $E_{i\mu}$ gene loci are in close vicinity which makes possible for the two regions to be regulated at the same time by a complex that contains both YY1 and HDAC1.

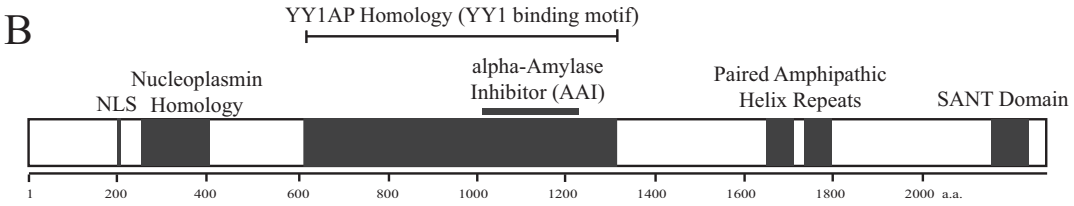
How the Gon4l complex specifically regulates B cell development is unclear. All four proteins in the Gon4l complex are not expressed in a tissue specific manner. Thus, some B lineage specific factors may be involved. In support of this, the total molecular weight of the Gon4l complex in M12 B cells is more than 670 kDa by sucrose gradient while the sum of the molecular weights of the four proteins in the complex is only ~550 kDa. It is also likely that other components of the Sin3a/HDAC complex such as SAP30 or Sds3 are in this complex (235,237).

In summary, a complex formed by Gon4l, YY1 and Sin3a/HDAC1 have been identified. While Gon4l is clearly required for B lymphopoiesis, the identification of the Gon4l complex provided a potential model as how Gon4l functions.

Figure 12. Gon4l is a conserved protein and contains putative structure domains.

(A) Gon4l protein orthologs are found across species. Phylogenetic tree of mouse Gon4l and its orthologs were generated by the software VECTOR NTI. The length of the branches represent the number of substitutions (or mutations) required to get from one node to the next. The calculated distance among the sequences are indicated in the parentheses. (B) Schematic showing the locations of structural domains predicted to be formed by mouse Gon4l based on bioinformatic analysis of the primary amino acid sequence. Descriptions of the domains are provided in the text. NLS: nuclear localization signal. (C) Similarity and absolute complexity plots indicating the extent of homology between human YY1AP and Gon4l proteins from human, mouse, zebrafish, *Drosophila melanogaster* and *Caenorhabditis elegans*. The similarity plot was generated by giving each of the residues in the alignment a value depending on how similar it is to the consensus sequence. The absolute complexity is the average of the pairwise alignment score derived from the substitution matrix. Red bars denote regions with the highest similarity among the Gon4l proteins compared. (D) Sequence alignment of the YY1AP homology regions in Gon4l proteins from different species**.** Human YY1AP and the YY1AP homology regions encoded by Gon4l proteins from the indicated species were aligned using VECTOR-NTI. Residues that are identical or similar are shaded. Conserved identical amino acids are shown in dark black letters. The bottom rows show the identified consensus sequence. The most conserved regions are labeled with the box.


A
YY1AP (Human) (0.0967)
GON4L (Human) (0.0428)
Gon4l (mouse) (0.1370)
Udu (Zebra Fish) (0.3327)
Gon4 (C-elegans) (0.4502)
Cdp1 (Drosophila) (0.4591)


B
YY1AP Homology (YY1 binding motif)
Nucleoplasmin
NLS Homology
alpha-Amylase
Inhibitor (AAI)
Paired Amphipathic
Helix Repeats
SANT Domain
1
200
400
600
800
1000
1200
1400
1600
1800
2000 a.a.

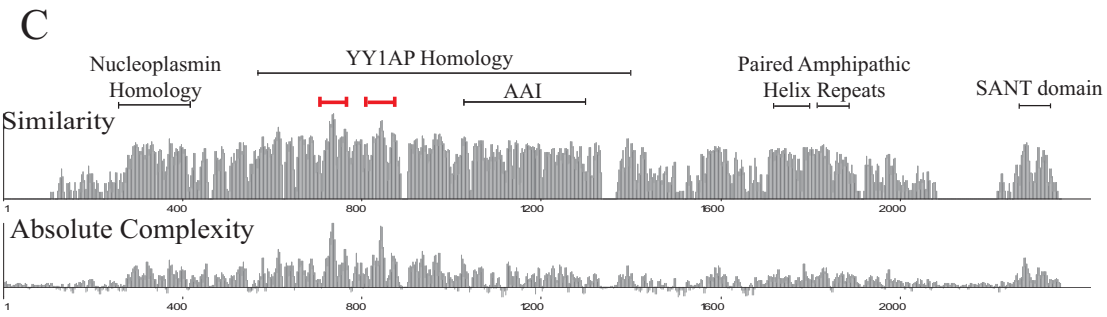

C
Nucleoplasmin
Homology
YY1AP Homology
AAI
Paired Amphipathic
Helix Repeats
SANT domain
Similarity
Absolute Complexity
1
400
800
1200
1600
2000

Figure 12 continued.

D

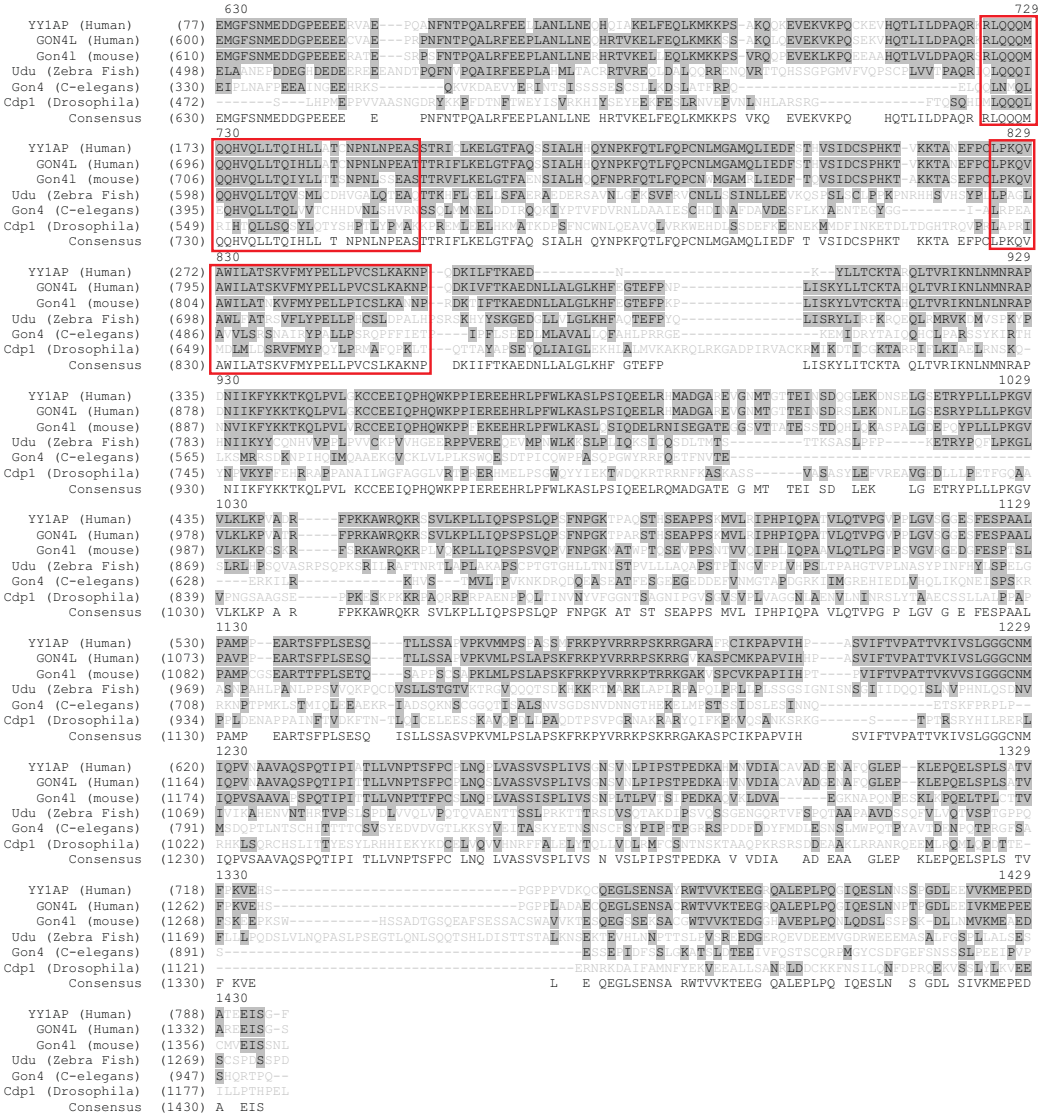


Figure 13. Gon4l is expressed in B cell lines of different developmental stages.

(A) Schematic showing the regions of mouse Gon4l used as immunogens to generate the rabbit polyclonal antibodies. (B) Immunoblot analysis of Gon4l expression in B cell lines of different developmental stages (M12: mature B; HAFTL1: pro-B; 38B9: pre-B and 18.81A20: pre-B). Whole cell lysates from these lines were probed with antibodies specific for the C-terminus of Gon4l (C-1746). Increasing amounts (indicated by the wedge) of whole cell lysate (WCL) from 293T cells expressing FLAG-Gon4l were used as a positive control. Lysates were also probed with antibodies specific for the B cell transcription factor Pax5 and GAPDH to confirm B-lineage identity and equal protein loading, respectively. The expression of the transcription factors important for B lymphopoiesis such as YY1, PU.1 and Ebf-1 was also analyzed. The arrow head indicates the Gon4l protein (C) Immunoblot analysis of Gon4l expression in B cell lines using antibodies specific for a portion of the YY1AP homology region (N-938). WCL from 293T cells expressing FLAG-Gon4l were used as a positive control. Lysates were probed with antibodies specific for GAPDH to confirm equal loading (bottom panel). All data are representative of at least 3 independent experiments.

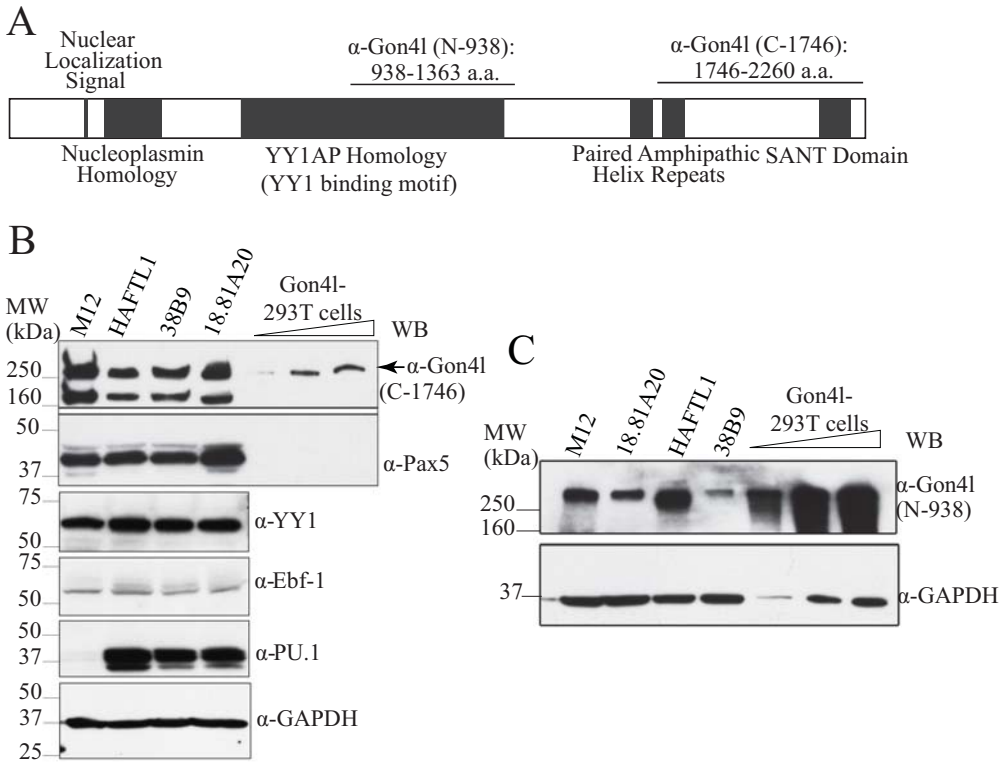

A
Nuclear Localization Signal
α-Gon4l (N-938): 938-1363 a.a.
α-Gon4l (C-1746): 1746-2260 a.a.
Nucleoplasmin Homology
YY1AP Homology (YY1 binding motif)
Paired Amphipathic Helix Repeats
SANT Domain
B
MW (kDa)
M12
HAFTL1
38B9
18.81A20
Gon4l-293T cells
WB
250
160
α-Gon4l (C-1746)
50
37
α-Pax5
75
50
α-YY1
75
50
α-Ebf-1
50
37
α-PU.1
50
37
25
α-GAPDH
C
MW (kDa)
M12
18.81A20
HAFTL1
38B9
Gon4l-293T cells
WB
250
160
α-Gon4l (N-938)
37
α-GAPDH

Figure 14. Mouse Gon4l associates with the zinc finger transcription factor YY1.

(A) Co-IP of endogenous YY1 and Gon4l from M12 cells. Whole cell lysates (WCL) were prepared from M12 cells. YY1 was immunoprecipitated by using the anti-YY1 antibodies. Non-specific IgG was used for control IPs. Recovered material was immunoblotted with the C-1746 antibody specific for the C-terminus of Gon4l (top panel) or YY1 (bottom panel). The arrow head indicates Gon4l. (B) Co-IP of YY1 and FLAG-Gon4l co-expressed in transfected 293T cells. Expression plasmids transfected into 293T cells are shown at the bottom of the panel. Anti-FLAG and nonspecific IgG were used to IP proteins from lysates prepared transfected 293T cells. Recovered material was immunoblotted with antibodies specific for the FLAG epitope to detect Gon4l (top panel) or antibodies specific for YY1 (bottom panel). (C) Co-IP of YY1 and Gon4l from WCL of Jurkat T cells. IP and immunoblotting were done as in (A). The arrow head indicates Gon4l. (D) Gon4l and YY1 co-localized in the nucleus. Expression plasmids encoding both FLAG-Gon4l and YY1 were co-transfected into 293T cells. Cells were then fixed and stained with anti-FLAG and anti-YY1 antibodies. Nuclei were counter-stained with TO-PRO-3. 293T cells transfected with expression plasmid lacking a cDNA insert were used as a negative control. The stained cells were then examined under a Zeiss 510 confocal microscope. The yellow color in the merged image indicates the co-localization of FLAG-Gon4l and YY1. Data are representative of 3 independent experiments for (A) and (C); 4 for (B) and (D).

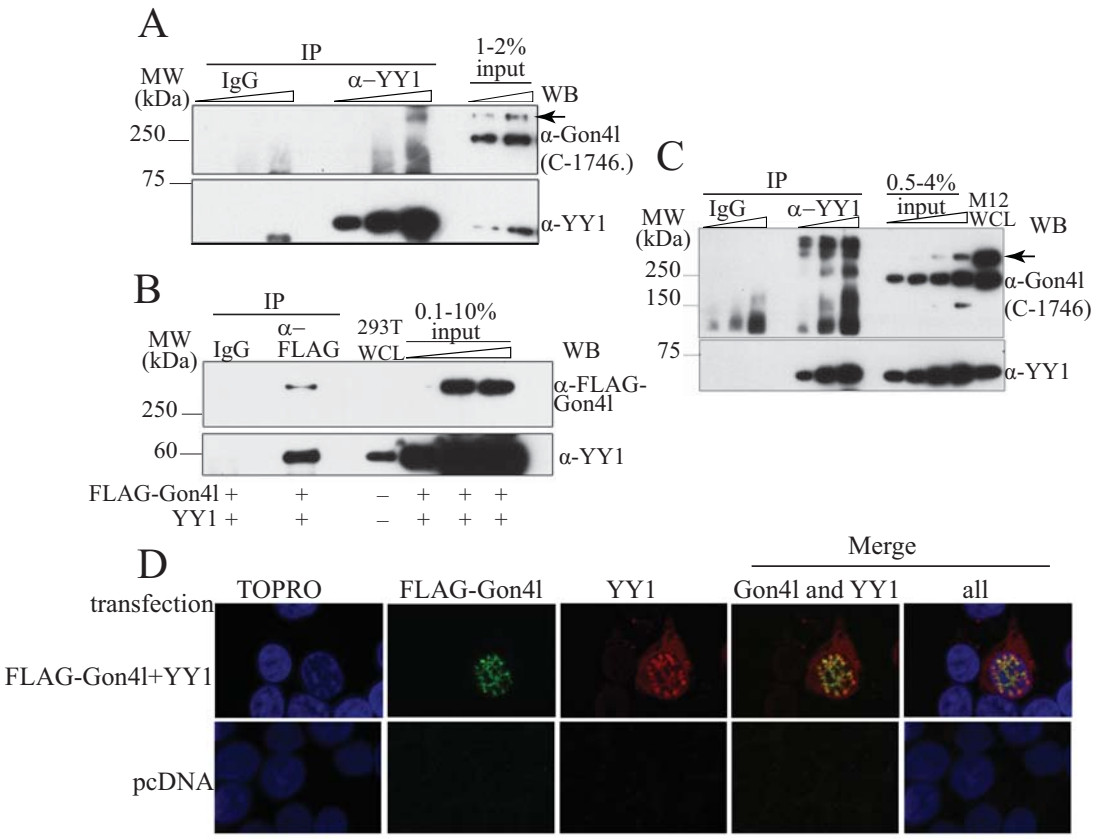

A
IP
1-2% input
MW (kDa)
IgG
α–YY1
WB
250
α-Gon4l (C-1746.)
75
α-YY1
B
IP
α–FLAG
293T WCL
0.1-10% input
MW (kDa)
IgG
WB
250
α-FLAG-Gon4l
60
α-YY1
FLAG-Gon4l + + – + + +
YY1 + + – + + +
C
IP
0.5-4% input
M12 WCL
MW (kDa)
IgG
α–YY1
WB
250
150
α-Gon4l (C-1746)
75
α-YY1
D
Merge
TOPRO
FLAG-Gon4l
YY1
Gon4l and YY1
all
transfection
FLAG-Gon4l+YY1
pcDNA

Figure 15. Gon4l co-sedimented with Sin3a and HDAC1 in sucrose gradient centrifugation.

(A) Co-sedimentation of Gon4l, YY1, Sin3a and HDAC1 in sucrose density gradients. Lysates prepared from the M12 B cells were loaded onto a 20-50% linear sucrose gradient. After centrifugation, the indicated fractions were collected and analyzed by immunoblot using the antibodies listed on the right side of the panel. “Input” denotes whole cell lysate (WCL; left side of the panel) of M12 B cells used as a positive control for immunoblots. Protein standards were separated in parallel gradients and fractions containing the protein standards are indicated. Fractions containing Gon4l in complexes are outlined with the dashed line box. (B) Ikaros does not co-sediment with Gon4l. Lysates from M12 cells were separated and analyzed as in (A). Fractions containing Gon4l in complexes are outlined as in (A). All data are representative of at least 3 independent experiments.

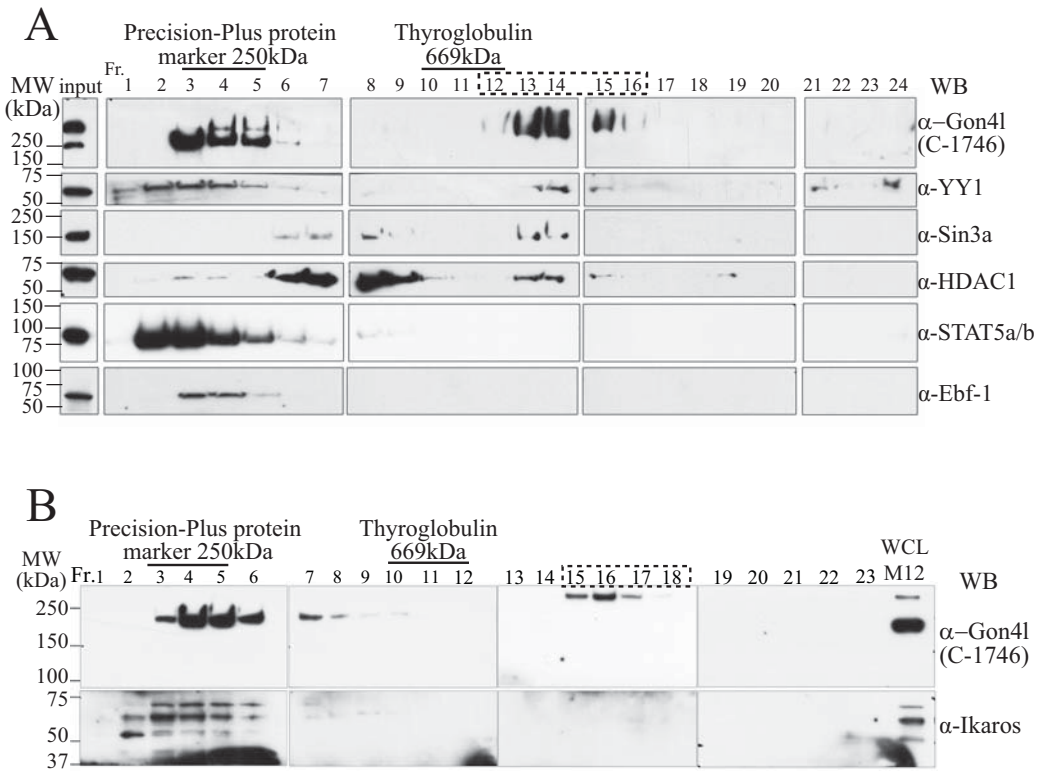
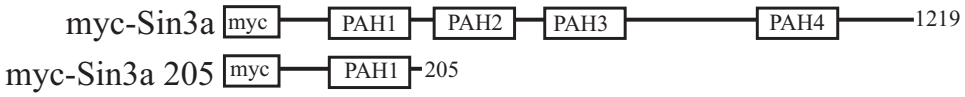

A
Precision-Plus protein marker 250kDa
Thyroglobulin 669kDa
Fr.
MW (kDa)
input
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
WB
250
150
75
50
250
150
75
50
150
100
75
100
75
50
α–Gon4l (C-1746)
α-YY1
α-Sin3a
α-HDAC1
α-STAT5a/b
α-Ebf-1
B
Precision-Plus protein marker 250kDa
Thyroglobulin 669kDa
MW (kDa)
Fr.
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
WCL M12
WB
250
150
100
75
50
37
α–Gon4l (C-1746)
α-Ikaros

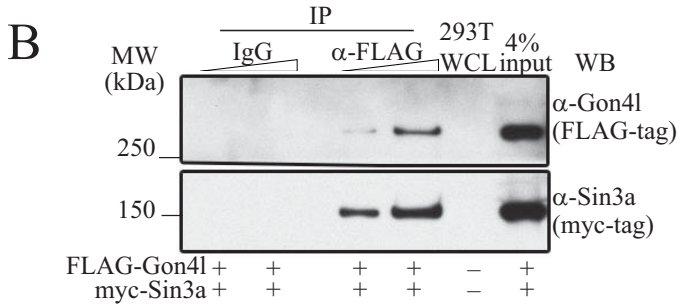
Figure 16. Gon4l associates with Sin3a.

(A) Schematics of the myc-tagged Sin3a and its truncated mutant, myc-Sin3a 205. (B) co-immunoprecipitation of Gon4l and Sin3a. 293T cells were co-transfected with the expression plasmids indicated at the bottom of the panel. Lysates from the transfected cells were then subjected to anti-FLAG-Gon4l IP. Non-specific IgG was used to perform control IPs in parallel. Recovered material was analyzed by immunoblotting using anti-FLAG and anti-myc antibodies as indicated. (C) Sin3a mutant that lacks PAH2, 3 and 4 did not co-IP with Gon4l. Experiments were performed as in B. Data are representative of 2 independent experiments.

A


myc-Sin3a
myc
PAH1
PAH2
PAH3
PAH4
1219
myc-Sin3a 205
myc
PAH1
205

B


IP
MW
(kDa)
IgG
α-FLAG
293T
WCL
4%
input
WB
250
150
α-Gon4l
(FLAG-tag)
α-Sin3a
(myc-tag)
FLAG-Gon4l + + + + – +
myc-Sin3a + + + + – +

C

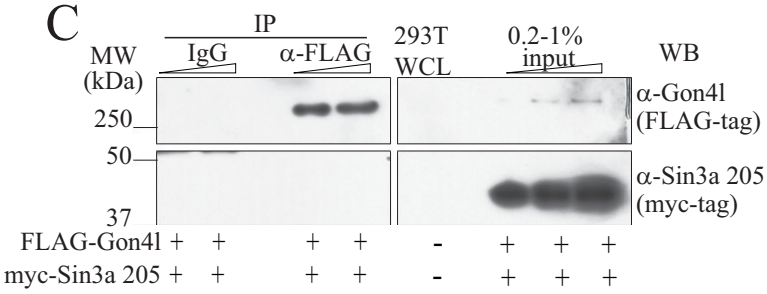

IP
MW
(kDa)
IgG
α-FLAG
293T
WCL
0.2-1%
input
WB
250
50
37
α-Gon4l
(FLAG-tag)
α-Sin3a 205
(myc-tag)
FLAG-Gon4l + + + + - + + +
myc-Sin3a 205 + + + + - + + +

Figure 17. Gon4l associates with YY1, Sin3a and HDAC1 simultaneously.

(A-B) Gon4l, YY1 and Sin3a are in the same complex. 293T cells were transfected with the expression plasmids as indicated at the bottom of the panels. WCL was prepared and subject to an anti-FLAG IP to isolate complexes containing Gon4l (A). The isolated Gon4l-containing complexes in the immunoprecipitates from the first IP were eluted with FLAG peptides and then subjected to a secondary anti-Sin3a IP (B). Recovered materials from each IP were immunoblotted with antibodies detecting FLAG-Gon4l, myc-Sin3a, and YY1. Non-specific IgG was used to perform control IPs. (C-D) Gon4l, YY1, Sin3a and HDAC1 can associate as parts of a single complex. 293T cells were transfected with expression plasmids encoding FLAG-Gon4l, YY1, myc-Sin3a and HA-HDAC1. Lysates from the transfected cells were prepared and then subjected to sequential IPs as in (A) and (B). For both (C) and (D), FLAG antibodies were used for the first IP. Anti- HDAC1 antibodies (C) or Sin3a antibodies (D) were used for the second IP. Recovered samples from each IP were immunoblotted with antibodies detecting FLAG (Gon4l), myc (Sin3a), YY1 or HA (HDAC1). For both (C) and (D), upper panel: immunoblot analysis of proteins recovered by the first IP. Lower panel: immunoblot analysis of proteins recovered by the second IP. Immunoblotting for HA-HDAC1 was performed using membranes already bound by YY1 antibodies, which generated the signals marked with asterisks. Data are representative of at least 3 independent experiments.

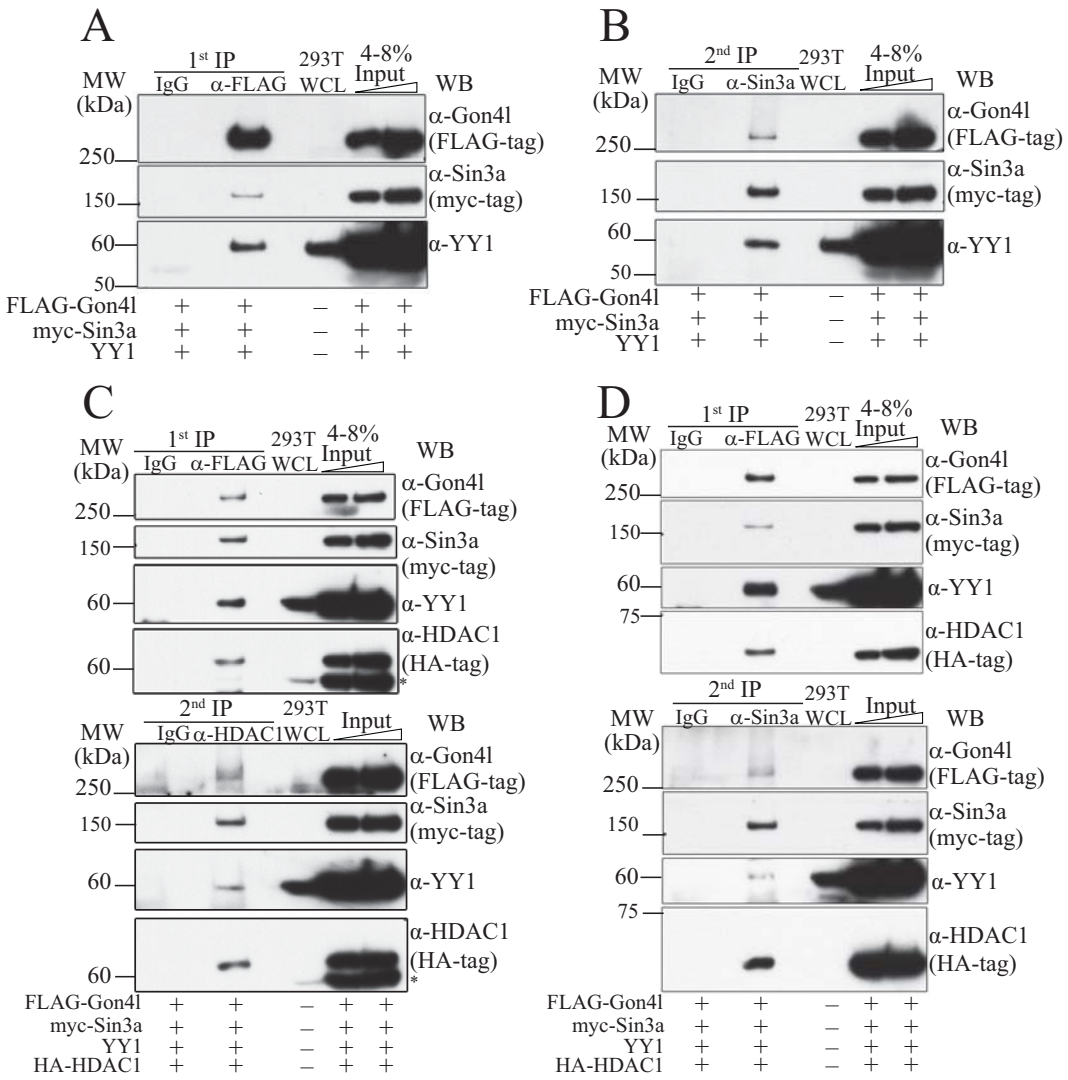
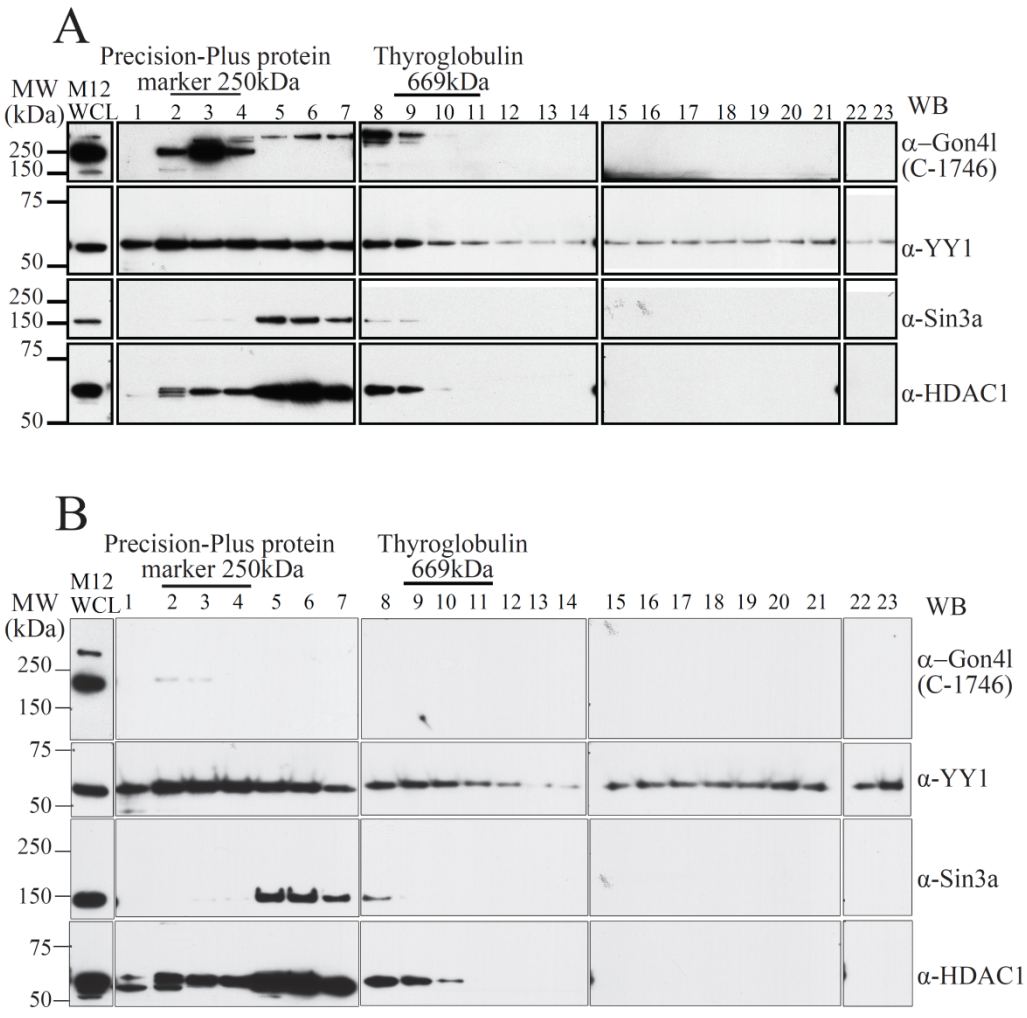

A
1st IP
IgG
α-FLAG
293T WCL
4-8% Input
MW (kDa)
WB
α-Gon4l (FLAG-tag)
α-Sin3a (myc-tag)
α-YY1
250
150
60
50
FLAG-Gon4l + + – + +
myc-Sin3a + + – + +
YY1 + + – + +
B
2nd IP
IgG
α-Sin3a
293T WCL
4-8% Input
MW (kDa)
WB
α-Gon4l (FLAG-tag)
α-Sin3a (myc-tag)
α-YY1
250
150
60
50
FLAG-Gon4l + + – + +
myc-Sin3a + + – + +
YY1 + + – + +
C
1st IP
IgG
α-FLAG
293T WCL
4-8% Input
MW (kDa)
WB
α-Gon4l (FLAG-tag)
α-Sin3a (myc-tag)
α-YY1
α-HDAC1 (HA-tag)
250
150
60
60
*
2nd IP
IgG
α-HDAC1
293T WCL
Input
MW (kDa)
WB
α-Gon4l (FLAG-tag)
α-Sin3a (myc-tag)
α-YY1
α-HDAC1 (HA-tag)
250
150
60
60
*
FLAG-Gon4l + + – + +
myc-Sin3a + + – + +
YY1 + + – + +
HA-HDAC1 + + – + +
D
1st IP
IgG
α-FLAG
293T WCL
4-8% Input
MW (kDa)
WB
α-Gon4l (FLAG-tag)
α-Sin3a (myc-tag)
α-YY1
α-HDAC1 (HA-tag)
250
150
60
75
2nd IP
IgG
α-Sin3a
293T WCL
Input
MW (kDa)
WB
α-Gon4l (FLAG-tag)
α-Sin3a (myc-tag)
α-YY1
α-HDAC1 (HA-tag)
250
150
60
75
FLAG-Gon4l + + – + +
myc-Sin3a + + – + +
YY1 + + – + +
HA-HDAC1 + + – + +

Figure 18. Screening of the Gon4l-complex in thymocytes.

Whole thymocytes from WT and *Justy* mice were lysed and loaded onto a 20-50% linear sucrose gradient as in Figure 15. After centrifugation, the fractions were collected and analyzed by immunoblot using the antibodies indicated. Whole cell lysate from M12 cells (WCL; left side of the panel) was used as a positive control for immunoblots. Protein standards were separated in parallel gradients and fractions containing the protein standards are indicated. (A) Sucrose gradient centrifugation analysis of WT whole thymocyte lysate. (B) Sucrose gradient centrifugation of *Justy* whole thymocyte lysate. Note that Gon4l protein is greatly reduced in *Justy* thymocytes. Data are representative of 2 independent experiments.


A
Precision-Plus protein marker 250kDa
Thyroglobulin 669kDa
MW (kDa)
M12 WCL
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
WB
250
150
75
50
α–Gon4l (C-1746)
α-YY1
α-Sin3a
α-HDAC1
B
Precision-Plus protein marker 250kDa
Thyroglobulin 669kDa
MW (kDa)
M12 WCL
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
WB
250
150
75
50
α–Gon4l (C-1746)
α-YY1
α-Sin3a
α-HDAC1

CHAPTER IV

GON4L CAN MEDIATE TRANSCRIPTIONAL REPRESSION

Abstract

The *Gon4l* gene is essential for B lineage commitment. Mice harboring the *Justy* mutation in *Gon4l* have severely reduced number of pro-B cells and a complete loss of all downstream progeny cells due to the loss of Gon4l protein. In the few remaining *Justy* pro-B cells, genes that are normally targeted for repression (*i.e.* myeloid genes) are not appropriately repressed, implicating Gon4l in mechanisms of transcriptional repression. Biochemical analysis of the Gon4l protein demonstrates that Gon4l can form a complex with the transcription factor YY1 and the co-repressors Sin3a/HDAC1. These co-factors could provide Gon4l with DNA binding and epigenetic modification activities to mediate transcriptional repression. In this study, Gon4l was demonstrated to indeed repress the activity of a nearby promoter when tethered to DNA. The transcriptional repression activity of Gon4l correlates with its ability to form a complex with YY1 and Sin3a/HDAC1, suggesting that the Gon4l complex is biologically functional. Both YY1 and Sin3a/HDAC1 regulate a wide range of biological processes and the malfunction of these proteins is implicated in tumorigenesis. Strikingly, *Justy* mice develop a variety of tumors with very high incidence. Some of the tumors found in *Justy* mice are analogous to those developed in mouse models or in humans where the function of YY1 or Sin3a/HDAC1 is disrupted. Thus, this study provides further evidence that Gon4l can function by associating with YY1 and Sin3a/HDAC1. The data also suggest that the Gon4l is essential for not only B lymphopoiesis but also tumor suppression.

Introduction

Differentiation of hematopoietic progenitor cells to the B lineage is associated with tightly controlled activation of B lineage appropriate genes and repression of non-B lineage genes (2-4,13,26). One important mechanism to control such lineage specific

gene expression is by regulating chromatin accessibility (129-132). In eukaryotic cells, DNA is wrapped around histones, which is further packed into chromatin. The packaging of DNA into higher order structures is inhibitory to transcription, because it impedes transcriptional activators from binding to the target genes. Thus, for a gene to be transcribed, its associated chromatin needs to take on an "open" configuration, where the binding sites for transcription factors are exposed. Conversely, changing the chromatin conformation into a "compact" form can be utilized to mediate transcriptional repression (61,135). The importance of chromatin modification in the regulation of transcription became exceedingly clear after the histone modification enzymes were identified. Histone acetyltransferases (HATs) covalently add acetyl groups to histone tails which reduces the affinity of histone-DNA interaction and thus facilitates transcriptional activation (137,138). Histone deacetylases (HDACs) on the other hand remove the acetyl groups from histones and therefore enhance the interaction between histones and DNA and is associated with transcriptional repression (135).

The recruitment of HDACs to target gene promoters can be achieved via transcription factors that either directly associate with HDACs (*i.e.*YY1) or associate with the Sin3 co-repressors that contain HDACs (141-150). During B cell development, Ikaros is one of the transcription factors that restrict multipotent progenitor cells to the lymphoid lineage by repressing myeloid genes (66). The Sin3a/HDAC1 co-repressor can also be recruited by Ikaros to mediate transcriptional repression (68). Similarly, Pax5 recruits the Grg4 co-repressor that interacts with HDACs to repress gene expression (124-128), which is implicated in its repression of non-B lineage genes (113,119,154).

YY1 is a zinc finger containing transcription factor that is critical for B cell development (182,183,240,241). Previous studies show that YY1 can function as a transcriptional activator, repressor or transcription-initiator element-binding protein depending on the cellular context (168-170,173,174,182). YY1 is a highly conserved protein and has been shown to function as a Polycomb Group repressor during

development (169,171-173).The transcriptional repression mediated by YY1 is via the recruitment of HDACs (168). In addition, YY1 can interact and modulate the function of other proteins such as p53, which is independent of the transcriptional activity of YY1 (180,242). Meanwhile, other factors that interact with YY1 can modulate the function of YY1 as well. YY1AP is a human protein that interacts with YY1 and can enhance YY1 mediated transcriptional regulation (228). SAP30 (Sin3a-associated polypeptide p30) is a subunit of the Sin3a co-repressor complex that physically associates with YY1 and enhances YY1-mediated transcriptional repression (158,235). Both YY1 and Sin3a/HDAC1 are ubiquitously expressed and regulate a diversity of biological processes ranging from embryonic development to tumorigenesis (155,170). However, whether YY1 functions in concert with the Sin3 co-repressors in the regulation of transcription, especially that in B cell development is unknown.

The *Justy* mutation is a unique point mutation in the *Gon4l* gene that abrogates Gon4l protein expression. As a result, B cell development in mice harboring the *Justy* mutation (*Justy* mice) is blocked at the pre-pro-B to pro-B transition. Under normal conditions, B lineage commitment occurs at this transition stage with genes associated with other lineages greatly repressed. But in *Justy* cells, the repression of myeloid genes is severely impaired, suggesting that Gon4l plays a vital role in this process. Gon4l protein is phylogenetically conserved and contains multiple protein-interacting domains that are found in other transcriptional regulators. Those include the motif homologous to the human YY1AP protein and the PAH repeats that are found in the Sin3 co-repressors. More importantly, Gon4l forms a complex with YY1 and Sin3a/HDAC1, which strongly suggests that Gon4l could mediate transcriptional repression by employing the DNA binding activity of YY1 and the HDAC activity from the Sin3a/HDAC1 co-repressor.

In this chapter, Gon4l was shown to confer transcriptional repression when tethered to DNA, which was consistent with its association with transcriptional co-repressors. This function of Gon4l correlates with its ability to form a complex with YY1

and Sin3a/HDAC1, suggesting that the Gon4l-YY1-Sin3a/HDAC1 complex is biologically functional. On the other hand, significant increases of neoplastic growth in a variety of tissues were observed in *Justy* mice about 10-12 months old, which is analogous to tumorigenesis in mouse models or humans where the function of Sin3 or YY1 is disrupted. Thus, these data further suggested the potential biological significance of the identified Gon4l-complex and implicated it in the suppression of tumor growth.

Materials and Methods

Cell culture

Raw264.7 macrophage cells (Raw cells) were maintained in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum, 100 U/ml penicillin, 100 μg/ml streptomycin, 1 mM sodium pyruvate and 2 mM L-glutamine. M12 cells and HEK293T cells were cultured as indicated in chapter III.

Plasmids, transient transfections, immunoprecipitation and immunoblot analysis

pcDNA3.1 (Invitrogen) was used as the backbone for all Gon4l expression constructs. A DNA fragment from pCMV-BD (Stratagene) that encodes the Gal4 DNA-binding domain (DBD) was used to generate plasmids expressing Gal4-DBD-Gon4l fusion proteins. pUAS-TK-Luc was kindly provided by Dr. Christopher Glass (UCSD). CD19-Luc and pHis-Pax5 were described previously (243). The promoter regions of *Csf1r* (*c-fms*) and *C/EBPα* were PCR amplified and inserted into pGL2 and pGL3 (Promega) luciferase reporter vectors respectively. All transfections were performed using Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions. Immunoprecipitation and immunoblot analysis were carried out as described in chapter III. The anti-GAL4-DBD (Santa Cruz, sc-510) was used to probe the Gal4-DBD-Gon4l fusion proteins and its derivatives.

Luciferase reporter assays

HEK 293T cells were cultured in 24 well plates and transfected with UAS-TK-Luc 0.1ug, PRL-CMV 2ng (Promega: E2261), and increasing amount of pcDNA-Gal4-Gon4l or its derivatives. Total DNA transfected was adjusted to 0.8ug/reaction by adding empty pcDNA vectors. Cells were harvested 24 hours post-transfection. Firefly and *Renilla* luciferase activity were measured using the dual-luciferase reporter assay system (Promega: E1960) following the manufacturer's instructions. All the samples were read in triplicates. Transfection of Raw cells was performed in 24-well plate with 1 million cells/well using the jetPEI-Macrophage DNA transfection reagent (Polplus Trasnfection) according to the manufacturer's instruction. 0.04pmol pcDNA-FLAG-Gon4l, 0.13pmol c-fms-Luc, 0.5fmol pRL-CMV and increasing amount of pHis-Pax5 were used. The total DNA transfected was adjusted to 2μg/reaction by adding empty pcDNA vector. Cells were harvested and analyzed as the 293T cells. For M12 cell transfections, cells were grown in RPMI 1640. 16 hours before the transfection, cells were transferred to fresh medium. Each transfection consists of 5 million cells, 10ug C/EBPα-Luc or empty pGL3-Luc reporter vector, 20ng pRL-CMV, and increasing amount of pcDNA3.1-FLAG-Gon4l. Total DNA transfected was adjusted to 20μg by adding empty pcDNA3.1 vector. Electroporation was carried out with 1pulse at 220v and a 30 millisecond pulse width. Cells were harvested and analyzed 24 hours post transfection as the 293T cells.

Histopathology of tumor samples

Samples were collected and fixed in 10% neutral buffered formalin as described (193). The fixed tissues were then processed, stained and analyzed by the Comparative Pathology Facility at the University of Iowa.

Flow cytometric analysis

Single cell suspensions were prepared from *Justy* lymphoma spleens, stained and analyzed by FACS as described in Chapter II. The following antibodies were purchased

from eBioscience: PerCP-B220 (RA3-6B2), PE-Cy7 CD19 (1D3), and APC-CD3ε (145-2C11). FITC conjugates: CD8 (H35-17.2) and Mac-1 (M1/70). PE conjugates: CD11c (N418) and Gr-1(RB6-8C5). PE-CD4 (H129.19) was purchased from BD Pharmingen. The FITC-Ly6c antibody was provided by the Waldschmidt lab at the University of Iowa.

Electrophoretic mobility shift assay (EMSA)

EMSA was performed using the Gel Shift Assay System (Promega E3050) according to the manufacturer's instruction. Briefly, nuclear extracts were prepared as previously described (244,245). Complimentary oligonucleotides containing the YY1 consensus binding sites or the mutated binding sites (Santa Cruz: sc-2533 and -2534) were radiolabeled using T4 polynucleotide kinase and [γ-^{32}P]ATP. The labeled oligonucleotides were purified by using the Illustra MicroSpin™ G-50 Columns (GE Healthcare 27-5330-01). 9 μg of nuclear extract and 0.5 μg of poly(dI-dC) were used in each reaction. The antibodies used for supershift were the same as those used for immunoblotting in Chapter III. The reactions were run on native 6% TBE polyacrylamide gels (Invitrogen) and the gels were then dried and exposed to x-ray film at −80 °C.

Oligonucleotide-precipitation assays

5'-end biotin labeled oligonucleotides that contain consensus YY1 binding sits (5'-CGCTCCGCGGCCATCTTGGCGGCTGGT-3') or a mutated sequence (5'-CGCT CCGCGATTATCTTGGCGGCTGGT-3') were annealed with their complementary oligonucleotides as described (174). The annealed oligonucleotides (40ug) were then used to coat 100ul Streptavidin Dynabeads (Invitrogen) following the bead manufacturer's instructions. NE was prepared as described above and pre-cleared by incubating with 100μl uncoated streptavidin-dynabeads for 2-3 hours at 4 °C with rotation. 1000 μg cleared NE and 125 μg/ml poly(dI.dC) were then incubated with the oligo-coated beads at 4 °C overnight with rotation. Beads were intensively washed and

the bound proteins were eluted with the sample buffer (Invitrogen, NP0007). Proteins were separated by SDS-PAGE and visualized by immunoblotting.

Results

Gon4l represses the activity of a nearby promoter when tethered to DNA.

In the previous studies, Gon4l was shown to form a complex with YY1 and the transcriptional co-repressor Sin3a/HDAC1. By forming the complex, Gon4l will be able to possess the HDAC and DNA-binding activities to confer transcriptional repression. To address this, a reporter gene assay system was utilized, in which Gon4l was directed to a target promoter via a Gal4-DNA-binding domain to sites upstream of a promoter sequence (Figure 19A). An expression plasmid was constructed to encode a chimeric protein with the DNA-binding domain of the yeast protein Gal4 (Gal4-DBD) fused to the N-terminus of Gon4l (Gal4-Gon4l). At the same time, expression plasmids encoding only the Gal4-DBD or FLAG-Gon4l were included as controls. Immunoblot analysis of lysates from 293T cells transfected with these plasmids confirmed correct protein expression (Figure 19A).

Increasing amounts of each expression plasmid were transiently transfected into 293T cells together with a reporter containing five Gal4 binding sites upstream of a TK promoter and the firefly luciferase gene (147,246). The total amount of DNA transfected in all experiments was kept the same by the addition of empty plasmids. A plasmid that encodes *Renilla* luciferase was included as an internal control. 24 hrs post transfection, cells were harvested and analyzed for firefly and *Renilla* luciferase activities. The absolute firefly luciferase activities were normalized to that of the *Renilla* luciferase to calculate the relative luciferase activity (Figure 19 B-D). The presence of Gal4-Gon4l repressed the reporter luciferase activity in a dose dependent manner (Figure 19B). In contrast, co-transfection of the FLAG-Gon4l expression plasmid had no effect on the

reporter (Figure 19C), while co-transfection of the Gal4-DBD expression plasmid modestly increased luciferase activity (Figure 19D). These data thus indicate that, when bound to DNA, Gon4l can inhibit the activity of a nearby promoter.

The YY1AP homology region of Gon4l is required for Gon4l mediated transcriptional repression

The regions in Gon4l that mediates the repression of the reporter were determined. Expression plasmids were constructed to encode the Gon4l C-terminal truncated mutants fused to Gal4-DBD (Figure 20A). These plasmids were transiently transfected into 293T cells and the correct expression of the encoded proteins was confirmed by immunoblot. The repression of the UAS-TK-Luc reporter by these Gon4l mutants was assayed. The results indicated that co-transfection of plasmids encoding the Gon4l mutants lacking either the SANT domain (Gal4-Gon4l-ΔC1) or both the SANT and PAH domains (Gal4-Gon4l-ΔC2) substantially repressed luciferase activity similarly as did WT Gon4l (Figure 20B and C). However, when a Gal4-Gon4l mutant with a further deletion of the YY1AP homology region was used (Gal4-Gon4l-ΔC3), the repression of the reporter was greatly reduced (Figure 20D), even though the mutant protein was expressed at a relatively higher level in the cells (Figure 20A). These data suggest that the YY1AP homology domain is a functional domain that mediates the transcriptional repression. In support of this, luciferase reporter assay using the YY1AP homology fused to Gal4-DBD indicated that the YY1AP homologous domain alone was able to maintain partial transcriptional repression activity as compared to the full length Gal4-Gon4l (Figure 20E).

The YY1AP homology region of Gon4l also mediates the interaction with YY1 and Sin3a/HDAC1

To assess how the transcriptional repression activity of Gon4l is related to its ability to associate with YY1 and Sin3a/HDAC1, expression plasmids were generated to

encode FLAG epitope tagged Gon4l C-terminal truncated mutants (Figure 21A). These Gon4l mutants have the same protein sequences as the Gal4-Gon4l mutants described above, except for that of the epitope tags. Each of these constructs was transfected into 293T cells together with plasmids encoding YY1, myc tagged Sin3a and HA tagged HDAC1 (Figure 21B, right panel). Lysates were prepared from these transfected cells and IP analysis was carried out with either anti-FLAG antibodies or control IgG (Figure 21B, left panel). The deletion of amino acid sequences spanning the SANT domain and the two PAH repeats did not have an obvious effect on the interactions between Gon4l and the other proteins. However, deletion of the YY1AP homology resulted in complete loss of association between Gon4l and YY1 and Sin3a/HDAC1. Thus, in addition to mediating transcriptional repression, the YY1AP homology region in Gon4l is also required for its association with YY1 and Sin3a/HDAC1. More importantly, the dependence on the same YY1AP homology region suggests that the transcriptional repression function of Gon4l and its ability to form the complex are interrelated.

Justy mice develop tumors

YY1 and Sin3a/HDAC1 are mediators of epigenetic gene repression. Based on our data, the loss of Gon4l could potentially affect gene silencing mediated by YY1 and Sin3a/HDAC1 in *Justy* mice. This is supported by the observation that the silencing of alternative lineage genes is disrupted in *Justy* B progenitor cells (192). Because epigenetic gene silencing is tightly associated with tumorigenesis in mouse models and in humans (155,170,217,247-250), a cohort of *Justy* mice was followed for spontaneous tumor development. In contrast to the control WT mice (0 tumor incidence), 58% *Justy* mice (a total of 55 mice) developed tumors at around 10-12 months of age (Figure 22A). Necropsy analysis of these mice revealed a wide spectrum of tumors, including lymphoma, salivary gland tumors (myoepitheliomas), and pulmonary papillary adenocarcinoma, *etc* (Figure 22 B). The most predominant tumors found in the *Justy*

mice were lymphomas (~28%). These mice displayed obvious splenomegaly and lymphadenopathy, which are typical for lymphomas (Figure 22C). Hematoxylin and eosin staining (H&E staining) revealed wide-spread invasion of lymphoma cells into a variety of organs and tissues and the disruption of their infrastructure (Figure 22D and E). FACS analysis of the spleens from *Justy* lymphoma mice showed that the lymphoma cells were $CD3^+$ and $CD4^+$, indicating their T cell origin (Figure 23). The splenic $CD3^+CD8^+$ T cell population was almost completely lost in the lymphoma mice, in contrast to that of the WT or non-lymphoma *Justy* mice (compare with Figure 3C). The myeloid populations in the spleen were present and seemed to be intact (Figure 23 lower right). Together, these data suggest that the absence of Gon4l results in the development of tumors, especially lymphoma, in *Justy* mice.

The binding of the Gon4l complex to oligonucleotides containing YY1 binding sites is not detected

YY1 could recruit co-repressors to facilitate transcriptional repression (150,168,170,251). To examine whether YY1 could recruit Gon4l and Sin3a/HDAC1 to target genes via its DNA-binding activity, EMSA assays were performed using complementary oligonucleotides containing consensus YY1 binding sites (Figure 24A). The specific binding of YY1 to its binding sequences (YY1 shift) was observed with nuclear extracts (NE) prepared from M12 B cells. Addition of the anti-YY1 antibodies, but not the control non-specific IgG, to the reactions led to the super-shift of the YY1 band. However, no super-shift was observed when antibodies against Gon4l, Sin3a or HDAC1 were added (Figure 24A and data not shown).

EMSA assays using nuclear extracts from 293T cells transfected to express FLAG-Gon4l, YY1, myc-Sin3a and HA-HDAC1 were also carried out (Figure 24B). This system allowed the use of antibodies against the epitope tags fused to the proteins. However, similar to that observed in the experiments described above, the addition of

YY1 antibodies into the reactions resulted in the super-shift of the YY1 band but none of the other antibodies used led to detectable shifts of the YY1 band.

The binding of the Gon4l-YY1-Sin3a/HDAC1 complex to oligonucleotides containing consensus YY1 binding sites was also tested by oligonucleotide-precipitation assays. Oligonucleotides were biotin labeled and incubated with nuclear extracts derived from M12 cells. Proteins associated with the oligonucleotides were precipitated with streptavidin-coated beads (Figure 24C). As previously reported, YY1 was purified specifically by oligonucleotides containing the YY1 binding sites (174). Mutation of the YY1 binding sites abrogated YY1 precipitation. However, co-purification of Gon4l or Sin3a/HDAC1 was not observed. The binding of the oligonucleotides by purified Gon4l-containing complexes was tested. 293T cells were transiently transfected to express FLAG-Gon4l, YY1, myc-Sin3a and HA-HDAC1. Complexes containing FLAG-Gon4l were purified from the transfected 293T cells by anti-FLAG IPs and eluted with FLAG-peptides as in Figure 17C. Successful purification of the complexes was confirmed by western blot (data not shown). The purified complexes were then subjected to the oligonucleotide precipitation assay. However, no specific YY1 precipitation was observed (Figure 24D).

Screening of genes regulated by Gon4l

Genes that may be regulated by Gon4l were analyzed. In *Justy* pro-B cells, *Colony stimulating factor 1 receptor* gene (*Csfs1r*) is one of the myeloid genes (208) that are not appropriately repressed and are expressed at significantly higher levels compared to those in the WT cells (192). Bioinformatic analysis of the *Csfs1r* promoter revealed multiple YY1 consensus binding sequences (data not shown), suggesting that this promoter may be a direct target gene of the Gon4l-YY1-Sin3a/HDAC1 complex. To address this, a luciferase reporter containing the *Csf1r* promoter (also known as *c-fms*) was constructed (Figure 25A upper panel). Luciferase reporter assays were carried out in

the myeloid cell line Raw264.7 cells where the expression of Csf1r protein is high (data not shown) and that of Gon4l is fairly low (Figure 11C), which would facilitate the assessment of how forced expression of Gon4l could affect the *Csfs1r* promoter activity. In addition to the YY1 binding sites, the *Csfs1r* promoter is known to be repressed by Pax5 and contains functional Pax5 binding sites (122). To test whether Gon4l regulated the *Csfs1r* promoter in concert with Pax5, expression plasmids encoding Pax5 were also included in the assays. Raw 264.7 cells were transfected and harvested 24 hours post transfection. As previously reported, Pax5 repressed the promoter activity in a dose-dependent manner (Figure 25A) (122). However, transfection of the FLAG-Gon4l encoding plasmids did not lead to obvious repression of the reporter. Neither did it synergize the transcriptional repression mediated by Pax5.

Csf1r is activated by the transcription factor PU.1 and C/EBPα during myeloid differentiation (252,253). In *Justy* pro-B cells, *C/EBPα* mRNA is also abnormally elevated (192). It was thus reasoned that the elevated *Csfs1r* expression may be secondary to the up-regulation of C/EBPα. More importantly, the promoter of *C/EBPα* is known to be regulated (254) and contains multiple YY1 binding sites (data not shown). Therefore the *C/EBPα* promoter was examined to see whether it could be a target of the Gon4l complex. The *C/EBPα* promoter was cloned into a luciferase reporter and assayed in both M12 B cells and 293T cells (Figure 25B and data not shown). However, no repression of the reporter by expression of Gon4l was observed.

Genes that might be activated by Gon4l were also screened. Recent studies indicate that HDACs are not only involved in gene repression but are also involved in pathways that activate gene expression (255,256). In support of this, the binding of the Sin3a co-repressor is found in both repressed and activated genes (167). In *Justy* B cell precursors, the loss of Gon4l results in down-regulation of several genes, suggesting that Gon4l may also be involved in gene activation (Dr. Colgan unpublished data). One of the hallmarks of the pre-pro-B to early pro-B transition is the up-regulation of *CD19*. *Justy*

pre-pro-B cells are unable to up-regulate CD19 expression and cannot transit to the pro-B stage (192). The *CD19* promoter is known to be activated by Pax5 (243), which is expressed normally in *Justy* B precursor cells (192). We thus hypothesized that Gon4l may be a co-factor of Pax5 and be required for its activation of the *CD19* gene. To test that, 293T cells were transfected with the expression plasmids encoding Pax5 together with or without the FLAG-Gon4l expression plasmids. A luciferase reporter consisting of three Pax5 binding sites from the *CD19* promoter was co-transfected (Figure 25C). Although the reporter was activated by Pax5 as reported (243), no synergistic effect between Gon4l and Pax5 was observed.

Discussion

B lineage commitment depends on repression of non-B lineage genes. Previous studies of the *Justy* mice implicated Gon4l in the silencing of myeloid genes during early B lymphopoiesis. Further, Gon4l was found to associate with the co-repressors Sin3a/HDAC1 and the transcription factor YY1, which suggested a potential mechanism underlying the function of Gon4l. In this current study, Gon4l is shown to mediate repression of a nearby promoter. More importantly, this activity of Gon4l correlates with its ability to form a complex with YY1 and Sin3a/HDAC1.

Both Gon4l mediated transcriptional repression and complex formation required the YY1AP homology motif. In addition, the YY1AP homology domain from Gon4l maintained partial repression activity. Given the fact that the human protein, YY1AP, can interact with YY1 and that YY1AP homology motif is the most conserved domain in Gon4l, our data suggest that YY1 plays a critical role in the function of Gon4l. YY1 is a zinc finger protein that can bind DNA. Depending on the chromatin context, YY1 mediates either transcriptional repression or activation by recruiting co-repressors such as HDACs or co-activators such as HATs. In the context of the Gon4l-YY1-Sin3a/HDAC1 complex, YY1 could provide the complex with DNA targeting activity and direct the

complex to the loci where transcriptional repression is needed. However, the promoter screening results suggested that YY1 alone may not be sufficient to target the putative Gon4l regulated genes. Other molecules may be required in the targeting of the regulated promoters. All the promoters screened contained consensus YY1 binding sites, but they were not regulated by Gon4l in the luciferase assays. Drawing from the studies of Pax5, where concerted action of Pax5 and other transcription factors are required in some occasions (257), it is postulated that the targeting of the Gon4l-complex may require collaboration between YY1 and some unknown transcription factors. If that is the case, such factors are either associated with the Gon4l-complex or they are recruited through other ways (*i.e.* by the *cis*-regulatory elements in the Gon4l-regulated target gene loci). In the current promoter screening assays, such factors or regulatory sequences may have not been included. Alternatively, YY1 could also serve as a factor that stabilizes the Gon4l-complex. This could be used to explain why the deletion of the YY1AP homology led to the dissociation of the whole complex. Even more importantly, the fact that YY1AP homology is required for both Gon4l mediated repression and its complex formation suggests that these two functions of Gon4l are interrelated. In other words, the Gon4l forms a complex to mediate its biological functions. Lastly, YY1 is known to mediate posttranslational modifications of other factors, which would affect their functions (170,180,181). Therefore, it is possible that the Gon4l-complex could indirectly regulate transcription by posttranslationally modifying other B lineage associated transcription factors. It also needs to be noted that YY1 may play some or all aforementioned roles simultaneously in the Gon4l-complex.

The functional roles of the SANT domain and PAH repeats in Gon4l were not revealed by the current study. Yet, these domains may still be important for the function of Gon4l. The SANT domain is implicated in histone binding and maintenance of HDAC activities in some co-repressors (135,159,258). These may also be critical for the function of Gon4l. Indeed, in the study of ugly duckling (Udu), the zebrafish ortholog of Gon4l,

sequences encoding the SANT domain were found required to rescue erythropoiesis in *Udu*-mutant zebrafish (213). The PAH repeats are the major protein-interaction domains in the Sin3 proteins and can be utilized to mediate interaction with DNA-binding transcription factors and many other co-factors (141,147,158,160,230,237,259,260). If other transcription factors were recruited to the Gon4l-complex as proposed above, the PAH repeats would likely be the domain to mediate such recruitment (159,258). In support of this, some evidence indicates that Udu interacts with DNA replication regulation proteins through its PAH repeats and SANT domain (261).

A wide spectrum of tumors developed in *Justy* mice, suggesting that Gon4l is a factor critical for suppressing tumor growth. In this regard, recent studies of *Udu*, suggest that Gon4l is important for the regulation of cell division and genome stability (213,261). If Gon4l does render its effects by forming the Gon4l-YY1-Sin3a/HDAC1 complex, the function of YY1 and Sin3a/HDAC1 would be affected in the absence of Gon4l. Both YY1 and Sin3a/HDAC1 are regulators of epigenetic gene silencing, which is closely related to cancer development (217,247-249). In human, estrogen-responsive breast cancer is proposed to be a result of abnormal removal of the YY1/HDAC1 complex from the *cyclin D1* promoter, which results in the accumulation of cyclin D1 and enhanced cell cycle progression (262). In other cases, the constitutive transcriptional repression function of YY1 can be disrupted by the oncoprotein c-myc and that is implicated in the development of c-myc induced tumors such as lymphomas and plasma cell tumors (170,263,264). Moreover, restoring YY1-binding to the *c-myc* promoter has been shown in some cases reduces c-myc over-expression and reverses tumor formation (265). T cell lymphoma is the predominant type of tumor developed in *Justy* mice. Whether this is due to defective repression of oncogenes like *c-myc* by YY1 needs to be examined.

Similar to YY1, the co-repressor Sin3a is also involved in a variety of signaling pathways that are associated with tumorigenesis. These include embryonic development, homeostasis, DNA replication, cell cycle progression, and genomic stability

(151,155,230). More importantly, Sin3a/HDAC complexes are recruited by tumor suppressor proteins such as p53 and phosphorylated Rb to modulate target gene activity (160,266,267). Decreased expression of Sin3a has been correlated with the development of non-small cell lung cancer (NSCLC) in human (268). Moreover, abnormal recruitment of the Sin3a/HDAC complexes by mutated transcription factors or changes of its enzymatic activity are implicated in the pathogenesis of several human hematopoietic malignancies such as acute promyelocytic leukemia and acute myeloid leukemia (230,269-271). During B lymphopoiesis, Sin3a/HDAC1 can be directed to mediate gene repression by the transcription factor Ikaros (68). However, when the DNA-binding capability is abrogated by a mutation in *Ikaros*, mice expressing such mutant Ikaros protein all develop T cell lymphoma (272). The development of T cell lymphomas in *Justy* mice is suggestive that Gon4l, YY1 and Sin3a/HDAC1 function together, because dysfunction of any of these proteins could lead to this same type of malignancy. Lastly, tumorigenesis is associated with disrupted genomic stability which could be a result of defective epigenetic modification of the chromosomes. Thus, based on the current data, it seems reasonable to propose that the loss of Gon4l will lead to dysregulated epigenetic gene silencing and as a consequence, oncogenes are overexpressed and chromosomal stability is lost. Those effects in combination eventually contribute to the transformation of the *Justy* cells.

In summary, Gon4l is capable of mediating transcriptional repression and this function relies on the same motif required for the formation of a complex with YY1 and Sin3a/HDAC1. The loss of Gon4l expression as a result of the *Justy* mutation led to significantly increased tumor development and that may be a consequence of defective epigenetic repression mediated by pathways involving YY1 and Sin3a/HDAC1.

Figure 19. Gon4l is capable of repressing transcription when bound to DNA.

(A). Schematic (upper panel) of the luciferase reporter gene and the Gal4-Gon4l fusion protein. FLAG-Gon4l is fused at the N-terminus to the DNA-binding domain of the yeast protein Gal4. UAS-TK-Luciferase reporter contains five Gal4 binding sites in tandem upstream of the promoter region. Expression plasmids encoding FLAG-Gon4l or just the DNA biding domain of Gal4 (Gal4-DBD) were used as controls. Immunblots (IB, lower panel) confirmed correct expression of the indicated proteins. (B-D), 293T cells were co-transfected with the UAS-TK-Luciferase reporter and the indicated amounts of the expression plasmids. A plasmid expressing *Renilla* luciferase was also co-transfected to provide a normalization control. 24 hrs after the addition of DNA, cells were harvested and lysates analyzed for firefly and *Renilla* luciferase activities. Graphs show the normalized luciferase activity (y-axis) expressed from the reporter plasmid when co-transfected with the indicated amount of expression plasmid (x-axis): Gal4-Gon4l (panel B); FLAG-Gon4l (panel B) or Gal4-DBD (panel D). Values shown are the average and standard deviations of triplicate samples after normalization to *Renilla* luciferase activity. Data are representative of at least 4 independent experiments.

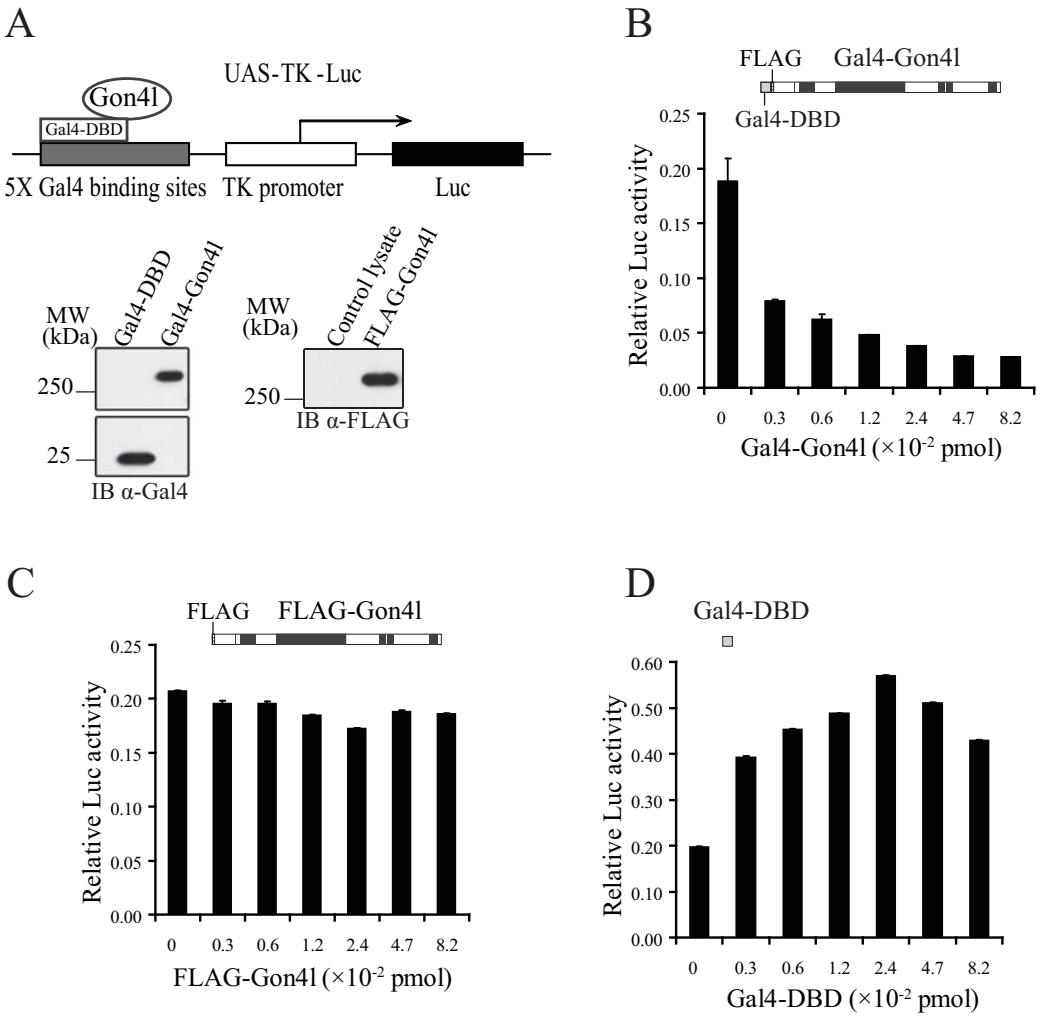

A
UAS-TK -Luc
Gon4l
Gal4-DBD
5X Gal4 binding sites
TK promoter
Luc
MW (kDa)
Gal4-DBD
Gal4-Gon4l
250
25
IB α-Gal4
MW (kDa)
Control lysate
FLAG-Gon4l
250
IB α-FLAG
B
FLAG
Gal4-Gon4l
Gal4-DBD
Relative Luc activity
0.25
0.20
0.15
0.10
0.05
0.00
0
0.3
0.6
1.2
2.4
4.7
8.2
Gal4-Gon4l (×10-2 pmol)
C
FLAG
FLAG-Gon4l
Relative Luc activity
0.25
0.20
0.15
0.10
0.05
0.00
0
0.3
0.6
1.2
2.4
4.7
8.2
FLAG-Gon4l (×10-2 pmol)
D
Gal4-DBD
Relative Luc activity
0.60
0.50
0.40
0.30
0.20
0.10
0.00
0
0.3
0.6
1.2
2.4
4.7
8.2
Gal4-DBD (×10-2 pmol)

Figure 20. The YY1AP homology domain of Gon4l mediates transcriptional repression.

(A) Immunoblot analysis of whole cell lysates from 293T cells transfected with the indicated Gal4-Gon4l expression plasmids. Anti-Gal4 antibodies were used as a probe. The correct expression of Gal4-YY1AP domain is shown in panel E. Schematics showing the structures of Gal4-Gon4l fusion proteins used in each experiment are indicated in panels (B-E). (B-E) 293T cells were co-transfected with the UAS-TK-Luciferase reporter and the indicated amounts Gal4-Gon4l expression plasmids or its derivatives. A plasmid expressing *Renilla* luciferase was also co-transfected to provide a normalization control. Graphs show the normalized luciferase activity (y-axis) expressed from the reporter plasmid when co-transfected with the indicated amount of Gal4-Gon4l or its derivative expression plasmids (x-axis). All data are representative of at least 3 independent experiments.

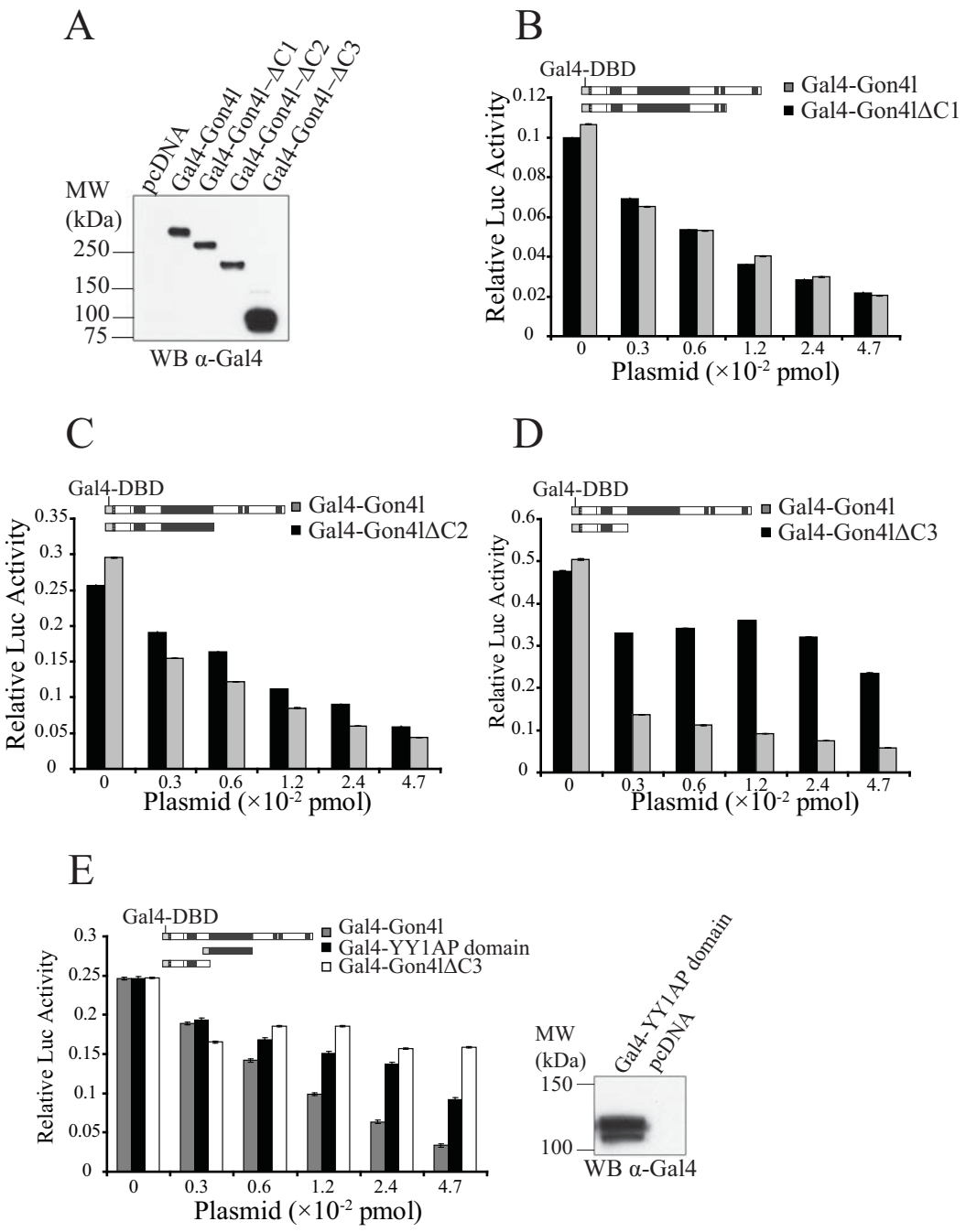
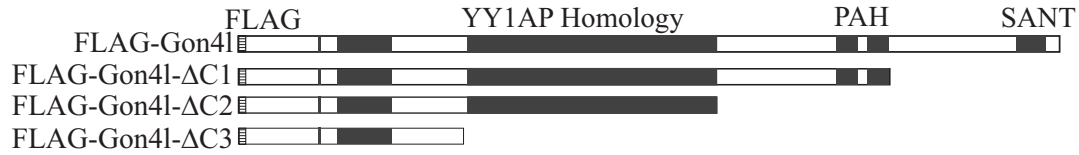

A
pcDNA
Gal4-Gon4l
Gal4-Gon4l−ΔC1
Gal4-Gon4l−ΔC2
Gal4-Gon4l−ΔC3
MW
(kDa)
250
150
100
75
WB α-Gal4
B
Gal4-DBD
Gal4-Gon4l
Gal4-Gon4lΔC1
Relative Luc Activity
Plasmid (×10-2 pmol)
C
Gal4-DBD
Gal4-Gon4l
Gal4-Gon4lΔC2
Relative Luc Activity
Plasmid (×10-2 pmol)
D
Gal4-DBD
Gal4-Gon4l
Gal4-Gon4lΔC3
Relative Luc Activity
Plasmid (×10-2 pmol)
E
Gal4-DBD
Gal4-Gon4l
Gal4-YY1AP domain
Gal4-Gon4lΔC3
Relative Luc Activity
Plasmid (×10-2 pmol)
MW
(kDa)
Gal4-YY1AP domain
pcDNA
150
100
WB α-Gal4

Figure 21. The YY1AP homology domain of Gon4l mediates its association with YY1 and Sin3a/HDAC1.

(A) Schematic showing the structures of FLAG-Gon4l C-terminal truncated mutants used for the experiments shown in panel (B). (B) Deletion of the YY1AP homology abrogates the association between Gon4l and its co-factors. 293T cells were co-transfected with the indicated FLAG-Gon4l or the mutant expression plasmids together with plasmids encoding YY1, myc-Sin3a and HA-HDAC1. Transfections with plasmids lacking a cDNA insert were used as a control (ctrl). Cells were harvested 48 hours after transfection, lysed and subjected to anti-FLAG-Gon4l IP. Nonspecific IgG (Ig) was used to perform control IPs in parallel. The immunoprecipitates were analyzed by immunoblotting using the antibodies indicated (left panel). Whole cell lysate (WCL) were similarly probed in the right panel. The YY1 signal in the “ctrl” lane in the right panel is from the endogenous YY1 protein in 293T cells. Data are representative of 2 independent experiments.

A


FLAG
YY1AP Homology
PAH
SANT
FLAG-Gon4l
FLAG-Gon4l-ΔC1
FLAG-Gon4l-ΔC2
FLAG-Gon4l-ΔC3

B

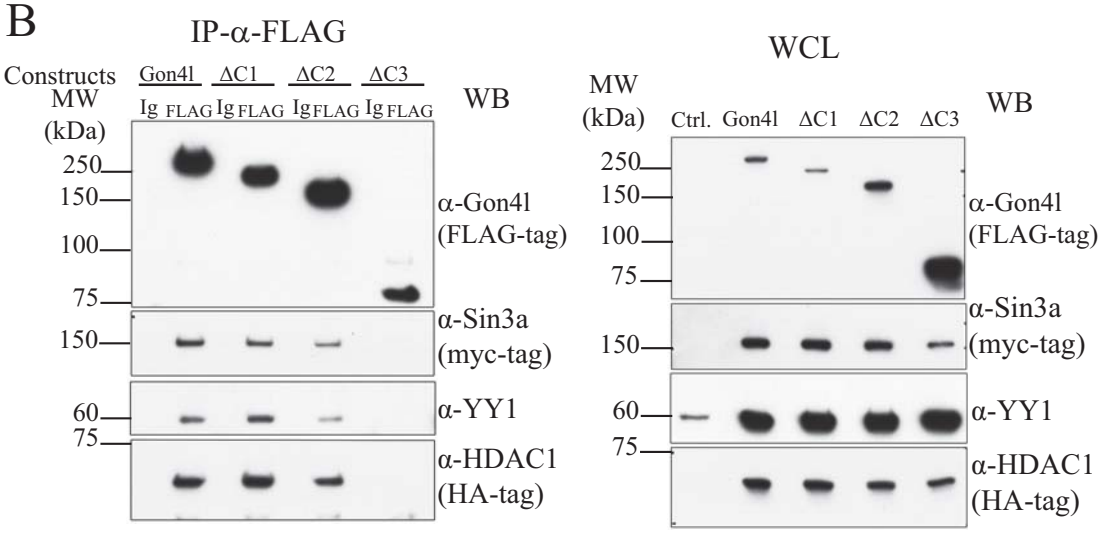

IP-α-FLAG
Constructs
Gon4l
ΔC1
ΔC2
ΔC3
Ig FLAG Ig FLAG Ig FLAG Ig FLAG
MW (kDa)
250
150
100
75
150
60
75
WB
α-Gon4l (FLAG-tag)
α-Sin3a (myc-tag)
α-YY1
α-HDAC1 (HA-tag)
WCL
MW (kDa)
Ctrl.
Gon4l
ΔC1
ΔC2
ΔC3
250
150
100
75
150
60
75
WB
α-Gon4l (FLAG-tag)
α-Sin3a (myc-tag)
α-YY1
α-HDAC1 (HA-tag)

Figure 22. *Justy* mice develop tumors.

(A) Tumor development in WT and *Justy* mice. A cohort of 55 *Justy* mice was followed. The percentage of mice with tumors was calculated by dividing the number of tumor mice by the total number of mice followed and was indicated on the Y axis (B) Distribution of histological tumor types found in *Justy* mice. The percentage of a given tumor type (Y axis) was calculated by dividing the number of mice carrying that type of tumor (indicated by the X axis) by the number of all mice with tumors. (C) Representative pictures of splenomegaly and lymphadenopathy in *Justy* mice with T cell lymphoma (indicated by yellow arrows). (D) Hematoxylin and eosion (H&E) staining of spleen sections from WT, tumor free *Justy* mice, and *Justy* mice with T cell lymphoma. The growth of lymphoma is indicated by the expansion of white pulp, which is not seen in the WT and *Justy* tumor-free mice (upper panel) (indicated by arrows). The histological splenic structure and the white pulp hypoplasia seen in *Justy* spleen is absent in *Justy* lymphoma spleen (lower panel) (indicated by arrows). (E) Lymphoma in thymus and metastatic lymphoma in the muscle, kidney and lung. H&E staining of infiltrating lymphoma cells in thymus and kidney (indicated by arrows). The T cell origin of these tumor cells was indicated by anti-CD3 staining (left panel, brown staining, indicated by arrows). H&E staining of lymphoma cells metastasized to the lung and muscles adjacent to the thymus (right panel).

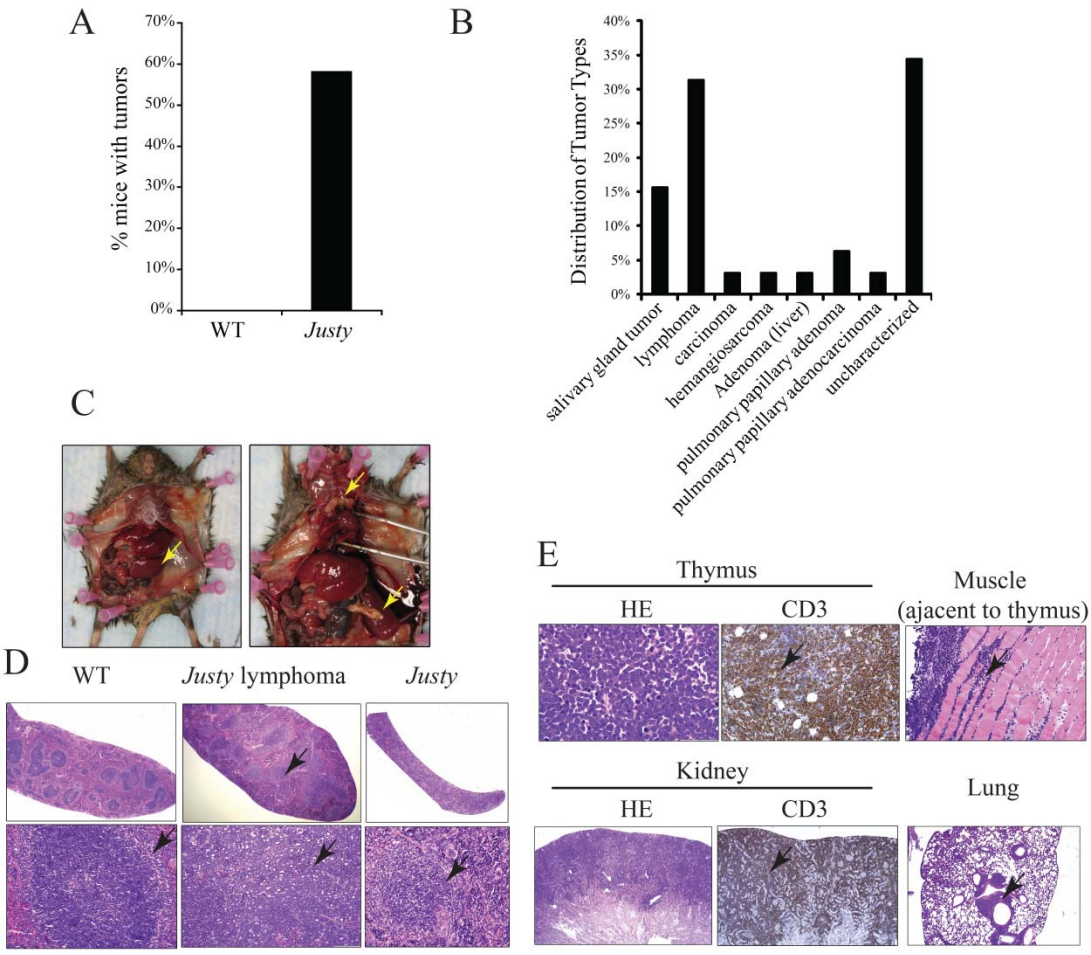

A
% mice with tumors
70%
60%
50%
40%
30%
20%
10%
0%
WT
Justy
B
Distribution of Tumor Types
40%
35%
30%
25%
20%
15%
10%
5%
0%
salivary gland tumor
lymphoma
carcinoma
hemangiosarcoma
Adenoma (liver)
pulmonary papillary adenoma
pulmonary papillary adenocarcinoma
uncharacterized
C
D
WT
Justy lymphoma
Justy
E
Thymus
HE
CD3
Muscle
(ajacent to thymus)
Kidney
HE
CD3
Lung

Figure 23. Flow cytometric analysis of splenocytes from *Justy* mice with lymphoma.

Total splenocytes were stained with antibodies detecting the indicated markers. CD3 and CD19 were used to differentiate T ($CD3^+$), B ($CD19^+$) and myeloid cell populations ($CD19^-CD3^-$) (middle column). The $CD3^+$T cell population was further analyzed by expression of CD4 and CD8 (upper left panel). The quadruple negative myeloid population was analyzed for the DC markers CD11c and Ly6C; the macrophage marker Mac-1; and the granulocyte marker Gr-1(lower left two panels). The $CD19^+$ population was further analyzed for B220 expression (upper right panel) or the expression of myeloid markers (lower right two panels). A small population of $CD19^+$ cells was seen but those were positive for DC and macrophage makers and negative for B220. Data are representative of 2 independent experiments.

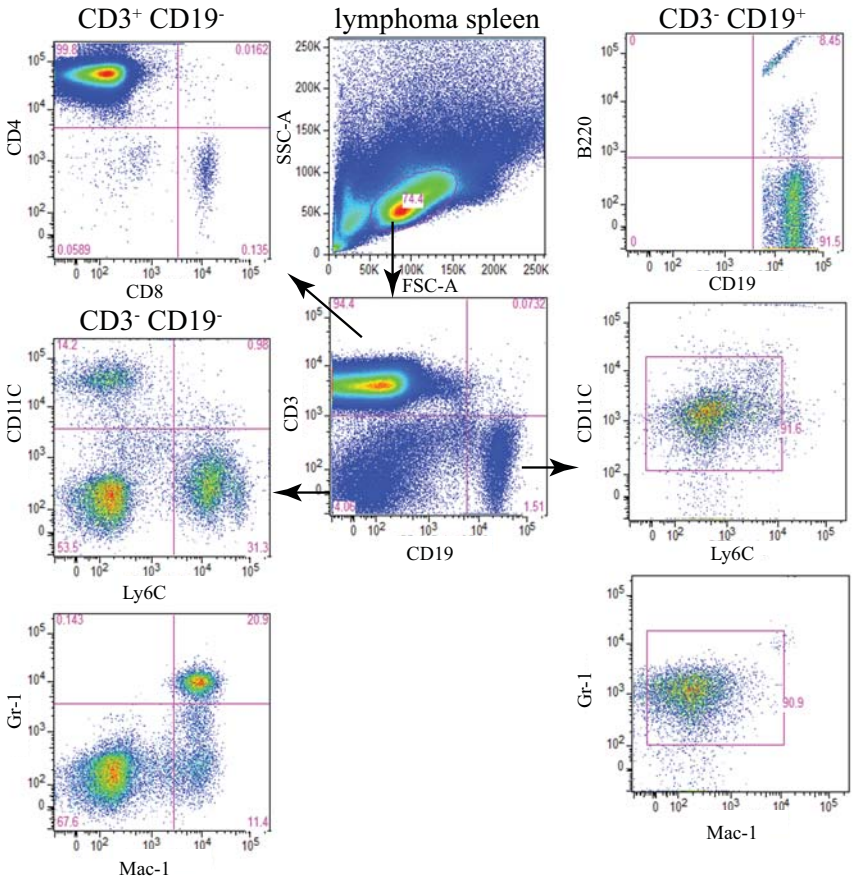

CD3+ CD19-
lymphoma spleen
CD3- CD19+
CD4
CD8
SSC-A
FSC-A
B220
CD19
CD3- CD19-
CD11C
Ly6C
CD3
CD19
CD11C
Ly6C
Gr-1
Mac-1
Gr-1
Mac-1

Figure 24. Gon4l or Sin3a/HDAC1 does not associate with oligonucleotides containing consensus YY1-binding sites.

(A-B) YY1, but not Gon4l or Sin3a/HDAC1, binds the complementary oligonucleotides with YY1 binding sites. EMSA analysis was performed using probes containing consensus YY1-binding sites and nuclear extract derived from M12 B cells (A) or 293T cells transfected to express FLAG-Gon4l, YY1, myc-Sin3a and HA-HDAC1 (B). "-NE" denotes no addition of nuclear extracts to the reaction. Cold YY1 probes or cold mutant YY1 probes containing mutated YY1 binding sites were added as indicated. Antibodies against YY1 (α-YY1), Gon4l, and the epitope tags, were added as indicated. Normal Ig from corresponding species was included as negative controls. Arrows denote YY1 shift, super shift and free probes. (C-D) Gon4l, YY1 and Sin3a/HDAC1 do not form a complex on the probes containing consensus YY1-binding sites. (C).The nuclear extracts of M12 B cells were incubated with streptavidin-magnetic beads coated with annealed WT form of biotinylated YY1 probes (WT). For control experiments, the nuclear extracts were incubated with beads coated with mutant probes (Mu). After incubation, the beads were washed and the bound fractions were analyzed by immunoblotting with antibodies indicated. (D) Similar oligo pull-down assays as in (C) were performed on purified FLAG-Gon4l complexes. WCL was prepared from 293T cells transfected to express FLAG-Gon4l, myc-Sin3a, HA-HDAC1 and YY1. FLAG-Gon4l containing complexes were purified by anti-FLAG IP as in Figure 17 (A). YY1 oligo-precipitation assay was then performed with streptavidin-beads coated with WT, or a mutant (Mu) YY1 probe, uncoated beads were also included as a control (-oligo). WCL prepared from untransfected 293T cells was included as a negative control for immunoblotting. Data are representative of at least 2 independent experiments.

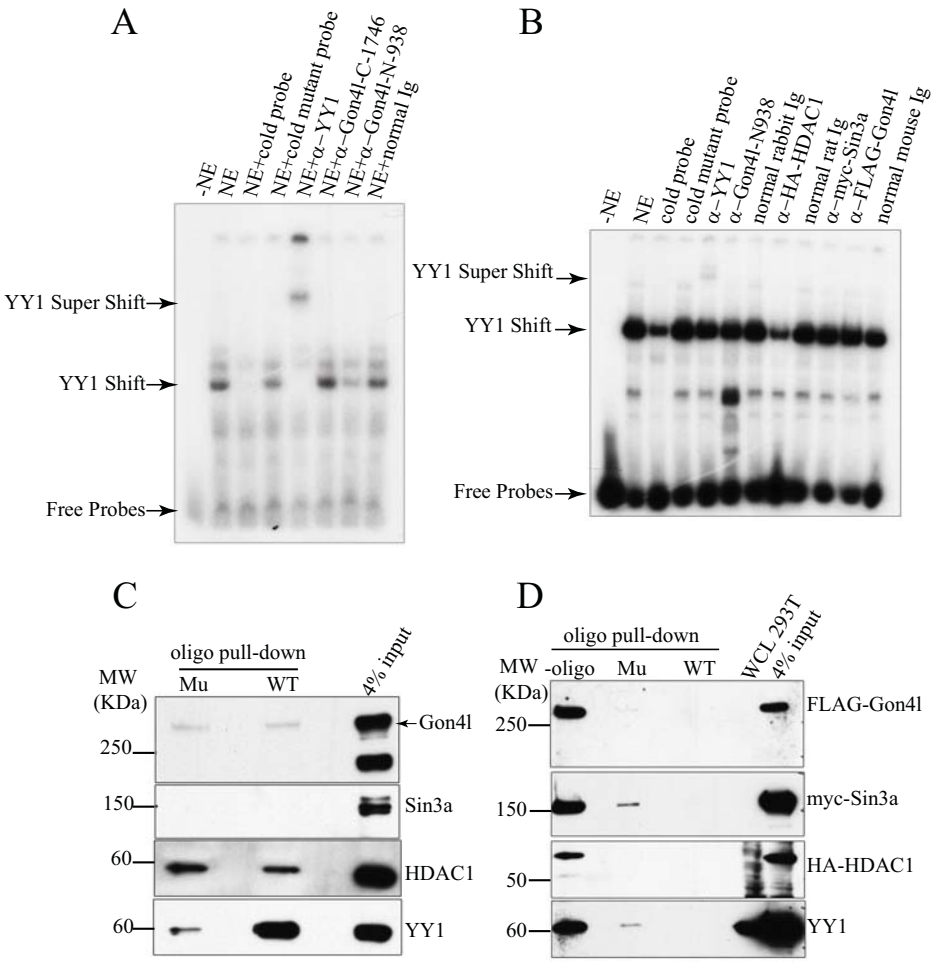
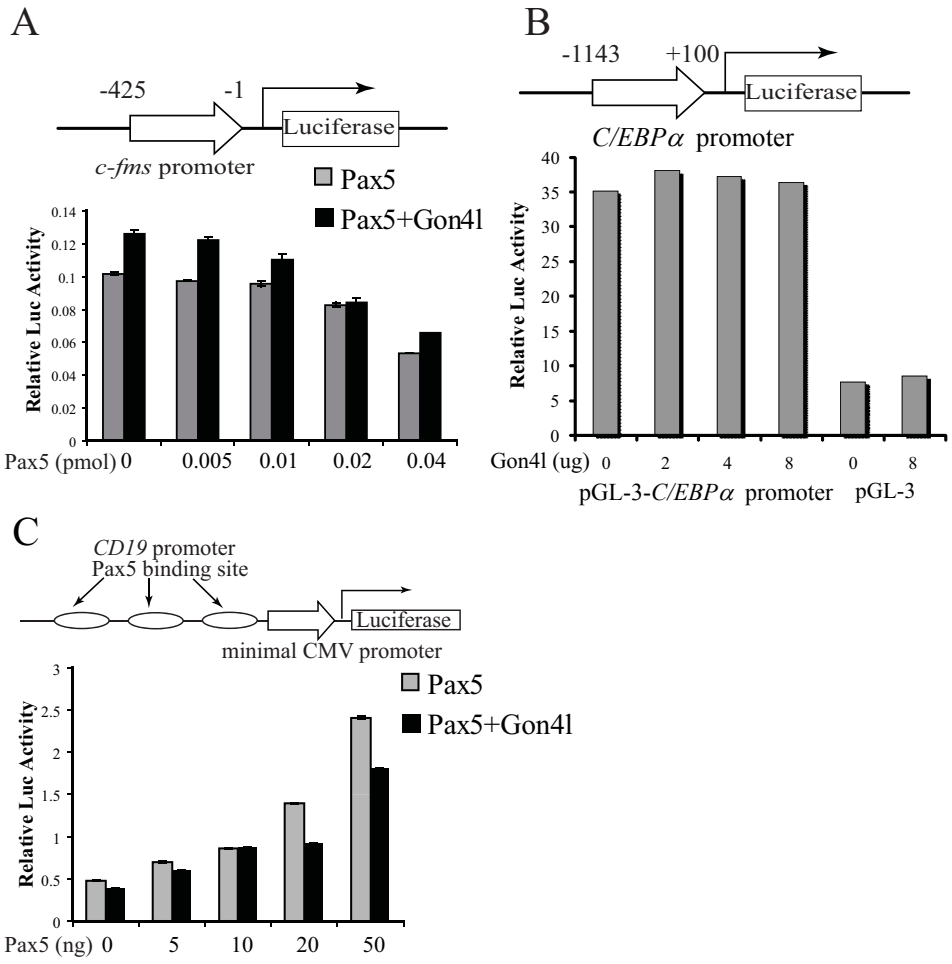

A
-NE
NE
NE+cold probe
NE+cold mutant probe
NE+α−YY1
NE+α−Gon4l-C-1746
NE+α−Gon4l-N-938
NE+normal Ig
YY1 Super Shift
YY1 Shift
Free Probes
B
-NE
NE
cold probe
cold mutant probe
α−YY1
α−Gon4l-N938
normal rabbit Ig
α−HA-HDAC1
normal rat Ig
α−myc-Sin3a
α−FLAG-Gon4l
normal mouse Ig
YY1 Super Shift
YY1 Shift
Free Probes
C
oligo pull-down
Mu
WT
4% input
MW (KDa)
250
150
60
60
Gon4l
Sin3a
HDAC1
YY1
D
oligo pull-down
MW (KDa)
-oligo
Mu
WT
WCL 293T
4% input
250
150
50
60
FLAG-Gon4l
myc-Sin3a
HA-HDAC1
YY1

Figure 25. Screening of target genes regulated by Gon4l.

(A). Gon4l does not enhance Pax5 mediated repression on *Csf1r* (also known as *c-fms*) promoter. The schematic (upper panel) indicates the luciferase reporter gene used (*c-fms*-Luc). The promoter contains 425 base pairs upstream the transcription starting site. Plasmids encoding FLAG-Gon4l and/or Pax5 were transfected into Raw 264.7 macrophage cells together with the reporter. Plasmids encoding *Renilla* luciferase were also co-transfected to serve as a normalization control. Cells were harvested 24 hrs after the addition of DNA, and analyzed for firefly and *Renilla* luciferase activities as in Figure 19. Graphs show the normalized luciferase activity (y-axis) expressed from the reporter plasmid when co-transfected with the indicated amount of expression plasmid (x-axis). Values shown are the average and standard deviations of triplicate samples after normalization to *Renilla* luciferase activity. (B) Gon4l does not regulate the *C/EBPα* promoter. The schematic (upper panel) indicates the structure of the luciferase reporter (pGL-3-C/EBPα promoter-Luc). The reporter contains the promoter region of *CEBP/α* from 1143 base pairs upstream to 100 base pair downstream the transcription starting site. Plasmids encoding FLAG-Gon4l were transfected into M12 B cells together with the reporter or a control empty pGL-3 reporter. Plasmids encoding *Renilla* luciferase were also co-transfected to serve as a normalization control. Luciferase activity was analyzed as in (A). (C). Gon4l does not enhance Pax5 mediated activation of the *CD19* promoter. The schematic (upper panel) indicates the structure of the luciferase reporter (CD19-Luc). The reporter contains three Pax5 binding sites from the *CD19* promoter. Plasmids encoding FLAG-Gon4l and/or Pax5 were transfected into 293T cells together with the reporter. Plasmids encoding *Renilla* luciferase were also co-transfected. Luciferase activity was analyzed as in (A). All data are representative of 3 independent experiments for (A) and (C); and 2 experiments for (B).


A
-425
-1
Luciferase
c-fms promoter
Pax5
Pax5+Gon4l
Relative Luc Activity
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
Pax5 (pmol) 0
0.005
0.01
0.02
0.04
B
-1143
+100
Luciferase
C/EBPα promoter
Relative Luc Activity
40
35
30
25
20
15
10
5
0
Gon4l (ug) 0
2
4
8
0
8
pGL-3-C/EBPα promoter
pGL-3
C
CD19 promoter
Pax5 binding site
Luciferase
minimal CMV promoter
Pax5
Pax5+Gon4l
Relative Luc Activity
3
2.5
2
1.5
1
0.5
0
Pax5 (ng) 0
5
10
20
50

CHAPTER V

GENERAL DISCUSSION

The studies detailed in this dissertation have contributed to the understanding of a phylogenetically conserved protein, Gon4-like, in the development of B lymphocytes. Further, a potential molecular mechanism with which Gon4l may function was described. In chapter II, the Gon4l protein was shown to be required for early B lymphopoiesis. In *Justy* pro-B cells, loss of Gon4l protein resulted in ineffective repression of myeloid genes, suggesting that Gon4l is part of the pathway that represses myeloid genes. In support of this, Gon4l expression was excluded or at very low level in myeloid cell lines. In chapter III, the identification of a high molecular weight complex containing Gon4l and the transcriptional regulators, YY1 and Sin3a/HDAC1 was described. This finding suggested that Gon4l could function as a mediator of transcriptional repression. In chapter IV, data were presented to demonstrate that Gon4l could function to repress transcription when tethered to DNA. More importantly, the repression mediated by Gon4l correlated with its ability to form a complex with its co-factors, indicating the biological significance of the Gon4l-complex. Lastly, a high incidence of tumors was observed in *Justy* mice, which may be due to disrupted transcriptional regulation mediated by Gon4l and/or its co-factors.

The data that are presented in this dissertation have implications for the understanding of the protein network that guides B cell development in bone marrow, especially the roles that transcriptional co-repressors play in the repression of alternative lineage genes. The results suggest that Gon4l is a new factor in this network. In addition, the data tied together three other well-documented transcriptional regulators, YY1, Sin3a and HDAC1, in the context of B cell development. All these proteins are ubiquitously expressed. Therefore the collaboration between these proteins and Gon4l may represent a general mechanism for gene repression. The presented data also suggest that Gon4l may

function as a tumor suppressor and the loss of it resulted in high incidences of hematopoietic tumors. Collectively, these data provide insight into the regulation of B lymphopoiesis and suggested a potential target for the prevention of tumor development.

Gon4l is an essential factor for B lymphopoiesis

In Chapter II of this dissertation, Gon4l is shown to be another molecule required for early B lymphopoiesis. B lymphopoiesis is a highly regulated process. In humans, defect in genes encoding critical proteins involved in this pathway lead to primary B cell deficiency (273). The current treatment involves adaptive antibody transfer, and bone marrow or hematopoietic stem cell transplantation (274,275). Recent studies revealed that IPS (induced pluripotent stem cell) may be a new way of generating transplantable hematopoietic stem cells (276). While a few transcriptional factors have been demonstrated to induce hematopoietic IPS (277-284), how to guide these cells to effectively differentiate along the B lineage pathway requires a better understanding of the key factors and mechanisms that instruct B lymphopoiesis. In addition, dysregulated B lymphopoiesis may lead to hematopoietic neoplastic growth. Thus a better understanding of B lymphopoiesis will also provide information for establishing molecular targets for cancer therapies (1). The current paradigm of the regulatory network contains only a handful of proteins (2), whether and how other proteins specifically contribute to this process is unknown. *Gon4l* is a conserved gene and has been shown to be important for hematopoiesis, erythropoiesis in particular, in zebrafish (213). The analysis of the *Justy* mutation revealed, for the first time in mammals, that Gon4l is an indispensible protein for early B lymphopoiesis. More importantly, genes encoding key transcription factors of B lymphopoiesis such as E2A, EBF, and Pax5 are expressed normally in *Justy* early pro-B cells, and that further demonstrates that Gon4l is another key factor to be added into the current paradigm.

A Gon4l parolog is located 5' of the *Gon4l* gene. This paralog is likely generated by segmental duplication of *Gon4l* as suggested by studies showing that the *YY1AP* gene is a partial duplication of *GON4L* in the human genome (201). The *Gon4l* paralog encodes an mRNA consisting of 17 exons, which correspond to exons 1-3 and 12-25 of *Gon4l*. If the mRNA of the *Gon4l* paralog is translated, the resultant protein would lack the PAH repeats and SANT domain. Whether and how this *Gon4l* paralog can influence the function of *Gon4l* remains to be investigated.

Gon4l and transcriptional regulation during B lymphopoiesis

Genes associated with different lineages are accessible and transcribed at base line levels in multipotent progenitor cells (4). When the precursor cells become committed to the B lineage, genes associated with other lineages need to be turned off (4). The mechanisms that regulate such repression are not fully understood. Epigenetic modification controls transcription by modulating chromatin accessibility to transcription factors. In B cell development, Pax5 is responsible for continuous repression of non-B lineage genes (2,4). Although Pax5 has been found to interact with the Grg4 co-repressor to mediate transcriptional repression (124,125), the biological relevance of this interaction has only been tested in cell lines but not in the context of B cell development. Whether this is the mechanism that truly regulates B lineage commitment is unknown. More importantly, whether there are other pathways that regulate this process also needs to be evaluated. B cell development in *Justy* mice is blocked at the lineage commitment phase, and the analysis shows that when the expression of Gon4l protein is reduced, the repression of the myeloid genes is defective, suggesting that Gon4l is required for repressing alternative lineage genes.

In chapter III, the data presented suggest that Gon4l possess characteristic domains associated with classical transcriptional regulators. Furthermore, Gon4l can

associate with the DNA-binding transcription factor YY1 and the core co-repressor complex Sin3a/HDAC1. These findings suggest that Gon4l may function to regulate gene expression by recruiting these factors. Following that, in chapter IV, Gon4l was shown to indeed mediate repression of a nearby promoter and such repression was associated with its ability to form a complex with the co-factors YY1 and Sin3a/HDAC1. These findings suggest a new mechanism that could potentially regulate repression of alternative lineage genes in addition to the Pax5 pathway.

Sin3a and HDAC1 are highly conserved and ubiquitously expressed proteins, and are the core of the many Sin3a co-repressor complexes (158). This is consistent with the fact that Sin3a/HDAC1 is a global regulator of transcription (146,158). Inactivation of *Sin3a* results in profound defects in cells (155). Sin3a/HDAC1 does not have a DNA binding domain. Thus the recruitment of the Sin3a/HDAC1 complex by DNA binding transcription factors is the key in determining which genes are to be repressed. Whether or how Sin3a/HDAC1 participates in the transcriptional regulation during B lymphopoiesis is unknown. The data presented in this dissertation show that Sin3a/HDAC1 can interact with Gon4l, thus connecting Sin3a/HDAC1 to pathways that regulate transcription during B lymphopoiesis. Similarly, YY1 is linked to B lymphopoiesis by its interaction with Gon4l. YY1 is highly conserved and ubiquitously expressed, and can mediate both transcriptional activation and repression (175). The repression mediated by YY1 is through its interaction with HDACs. However, the mechanisms by which YY1 regulates transcription in B lymphopoiesis are unknown.

In general, somatic cells all contain the same genetic material of a given host, even though they may be of different cell types in different tissues. Thus, genes that are not associated with the appropriate lineage or cell types should not be expressed. Recruiting Sin3a/HDAC complexes may be a common mechanism to mediate repression of alternative lineage genes. In the lymphoid system, Ikaros is critical in early lymphoid specification and commitment by inhibiting myeloid genes. Ikaros recruits Sin3/HDAC

and mediates transcriptional repression (68). Similarly, non-neuron cells constitutively repress neuron associated genes by recruiting Sin3a/HDAC complexes with the help of the transcription factor, REST (161,285,286). Because Gon4l, YY1 and Sin3a/HDAC1 are all conserved proteins and are expressed in a variety of tissues, the association of the four proteins may represent a core co-repressor complex that can be recruited by different factors to mediate lineage gene repression in different settings. If this is the case, other B lineage specific factors may be involved in Gon4l mediated transcriptional repression.

In the attempt to screen for Gon4l target genes, it was found that although the putative genes screened all have YY1 binding sites in their promoter region, they did not seem to be regulated by Gon4l in transient transfection assays (Figure 25). Several reasons may contribute to that. First, YY1 may not serve as the targeting factor in the Gon4l complex, instead it is a co-factor required for stabilizing the complex, or for post-translational modification of other proteins as previously reported (181). This would be consistent with the fact that the Gon4l-complex did not bind DNA-oligonucleotides containing YY1 consensus binding sites (Figure 24). In that case, another B cell specific transcription factor may be involved. *E2A* deficient mice have a similar B cell development block as that of the *Justy* mice. In addition, the AD1 domain of E2A can function as a repressor domain by interacting with the ETO co-repressor, which in turn recruits the N-COR and Sin3a/HDAC1 to mediate transcriptional repression (224,287). Thus, E2A would be a potential factor that co-regulates gene expression together with the Gon4l-complex. The second reason for not seeing repression of the target genes in the screening is that factors that contribute to the regulation of the putative Gon4l target genes may not be present in the cell lines used for the assays. Third, the *cis*-elements that are regulated by the Gon4l complex were not included in the luciferase reporters used. Fourth, the reporter gene transfected into the cells may not recapitulate the natural chromatin context required for the repression. In this case, making a stable line expressing the reporter gene may help to address this question.

Gon4l and tumorigenesis in *Justy* mice

In chapter IV, the *Justy* mutation is shown to lead to significantly increased tumorigenesis in mice. In the *Justy* mouse cohort followed, 58% of the mice developed tumors at around 10-12 months of age. The dominant type of tumor is T cell lymphoma, which accounts for about 30% of the tumors identified. However, tumors from other tissues were also occasionally observed. The wide distribution of histological tumor types suggests that Gon4l may function as a general tumor suppressor like p53 or Rb, where disruption of these proteins is directly related to the progression into a tumorigenic state (288,289). This hypothesis would also be consistent with the expression of Gon4l in multiple tissues and its association with the general transcriptional regulators, YY1 and Sin3a/HDAC1. On the other hand, disruption of the function of YY1 or Sin3a is also closely related to tumorigenesis in human and mice (See the Discussion in chapter IV).

The predominance of T cell lymphomas suggests that T cells are most susceptible to the loss of Gon4l in terms of transformation. This may be due to the relatively high proliferation rate of T cells compared to other cell types. Such a high proliferation rate may facilitate more mutations to occur (290). The loss of Gon4l in *Justy* cells may also provide a growth advantage so as to increase the number of cells available for mutations. Lastly, the loss of Gon4l may directly contribute to the malignant transformation by some unknown mechanisms.

In the tissues examined in *Justy* mice with lymphoma, malignant lymphoma cells replaced the normal cell populations of the affected organs and disrupted the normal infrastructure of the tissue. The T cell origin of these tumor cells were confirmed by immunohistology and FACS analysis. Although clonal expansion of the tumors cells has not been examined, all the tumor cells are CD4 single positive, suggesting that the T cells were transformed after they reached the later stages of T cell development. It is unknown whether the transformation took place in the thymus or in the periphery. This could be determined by comparing VDJ recombination status of the TCR locus (T cell receptor)

among different lymphoid organs at different stages of the lymphoma as previously described (272). In addition, the autonomous growth and tumorigenic properties of these malignant T cells could be tested by adoptive transfer into immunodeficiency mice such as the nude mice (272).

The zebrafish Gon4l ortholog, Udu, is implicated in the maintenance of genome stability (261). Thus, it will be interesting to analyze the karyotype of *Justy* lymphoma cells to see if there are any abnormalities in chromosomal morphology. In the study of Ikaros, radiation or chemical induced murine lymphomas have been linked to homozygous deletions or mutations of the *Ikarso* gene (272). Whether such lymphomagenesis also requires mutations in *Gon4l* would be direct evidence to support the notion that Gon4l is a tumor suppressor.

The *Justy* mutation causes aberrant splicing of the *Gon4l* gene which in turn abrogates the expression of the encoded protein. Whether the splicing pattern is changed after transformation is unknown. Thus, analysis of Gon4l expression in the lymphoma cells may help to determine if restoration of Gon4l expression could potentially reverse the tumorigenic state of the lymphoma cells. Using the same system, the domain/s of Gon4l that mediates tumor suppression could be determined.

Salivary gland myoepithelioma is another major type of tumors that arose in the *Justy* mice (~16%). Myoepithelioma is a benign but rare salivary gland neoplasm, composed entirely of myoepithelial cells (291,292). In some even rarer cases, malignant myoepitheliomas, also known as myoepithelial carcinoma, can develop from benign myoepithelioma (293,294). Although studying salivary gland myoepitheliomas can provide valuable insight into myoepithelial differentiation and transformation, no good mouse model has been established for such studies (Personal communication with Drs. Colgan and Meyerholz). Given its relatively high myoepithelioma incidence rate, the *Justy* mice may serve as a very good model for such studies. Moreover, our preliminary

analysis indicated that some of the *Justy* myoepitheliomas may have undergone malignant transformation and could metastasize into adjacent tissues.

Why the *Justy* mice develop myoepithelioma in the salivary gland with such a high incidence rate is unknown. Studies found in the literature suggest that myoepithelioma is associated with chromosomal translocations (291,295). Non-random chromosomal translocations are known to contribute to a diversity of cancers, but the mechanisms that cause such translocations are not fully understood. In general, processes that cause DNA breaks or fragile sites in the genome are thought to play a role. Those include V(D)J recombination, class switching recombination, and homologous recombination during cell replication. In addition, defective DNA-repair pathways are also thought to cause chromosomal translocations in humans, mice and yeast (296). Whether chromosomal translocation occurs in *Justy* myoepitheliomas can be evaluated by karyotype analysis. Further, the zebrafish Gon4l ortholog, Udu, regulates genome integrity and associates with heterochromatin during DNA replication (261).

The proposed model

In summary, the data shown in this dissertation suggest that the Gon4l protein is required for the normal progression of early B lymphopoiesis (Figure 26). The earliest manifestation of the loss of Gon4l protein as a consequence of the *Justy* mutation is the block of B cell development at the B lineage commitment phase. Whether Gon4l is required for later stages of B lymphopoiesis is unknown. But based on the fact that Gon4l is expressed in B lineage cells after the pro-B stage, it is very likely that Gon4l plays a role in those stages (Figure 27). In addition, Gon4l is required for maintaining the biological integrity of differentiated T cells (Figure 26). The loss of Gon4l leads to development of T cell lymphoma from differentiated T cells in *Justy* mice around 10 to 12 months old, although T cell development was normal in those mice early in their lives.

On the molecular level, Gon4l may associate with the general transcriptional repression mediators YY1 and Sin3a/HDAC1 in a high molecular weight protein complex. More importantly, deletion of the YY1AP homology domain from Gon4l abrogates the formation of the Gon4l-complex and, at the same time, the Gon4l-mediated transcriptional repression, indicating the biological significance of this complex. Consistently, correct silencing of myeloid genes is lost in *Justy* pro-B cells, suggesting that Gon4l is required for transcriptional repression. The data therefore provided a potential mechanism as how Gon4l could function (Figure 27).

The presented data also suggest that the interaction between YY1 and Gon4l may be critical for the stabilization of the Gon4l-complex as well as the function of the complex, because both require the YY1-interacting motif, YY1AP. Whether YY1 directs the Gon4l-complex to DNA targets is unknown (Figure 25). It is likely that YY1 may require cooperation with other transcription factors as suggested by the study of Pax5 mediated transcriptional repression (124). Moreover, the possibility that YY1 functions in ways other than targeting DNA cannot be excluded, since YY1 plays more roles than just to regulate transcription (170,181). It is also worth noting that HDACs have other substrates in addition to histones (297). Whether the B lineage transcription factors are posttranslationally modified (*i.e.* deacetylation) is unknown, but such modifications will likely affect the activities of these proteins or their subcellular distribution. Thus, the Gon4l-complex could achieve the regulation of transcription by mediating posttranscriptional modification of key B lineage transcription factors.

Together, these data indicate that Gon4l is an essential factor for B lymphopoiesis by participating in pathways that regulate alternative lineage gene repression. This function of Gon4l is likely rendered through its association with YY1 and Sin3a/HDAC1. The same mechanism may also keep mature T cells and other types of cells from malignant transformation.

Figure 26. Loss of Gon4l in the hematopoietic system reveals distinct roles of the protein.

A simplified B and T cell developmental pathway is illustrated. The *Justy* mutation leads to a developmental block at the pre-pro-B to pro-B transition, indicating Gon4l is critical for early B cell development at the B lineage commitment phase. The *Justy* mutation also leads to malignant transformation of fully differentiated T cells, suggesting that Gon4l is important for maintaining the biological integrity of these cells. The letter "A" and "B" in the circles stand for fraction A and B according to the Hardy fractionation nomenclature. Abbreviations: HSC, hematopoietic stem cells; LMPP, lymphoid-primed progenitor; CLP, common lymphoid progenitor; ETP, early T cell progenitors, DN, double negative; DP, double positive.

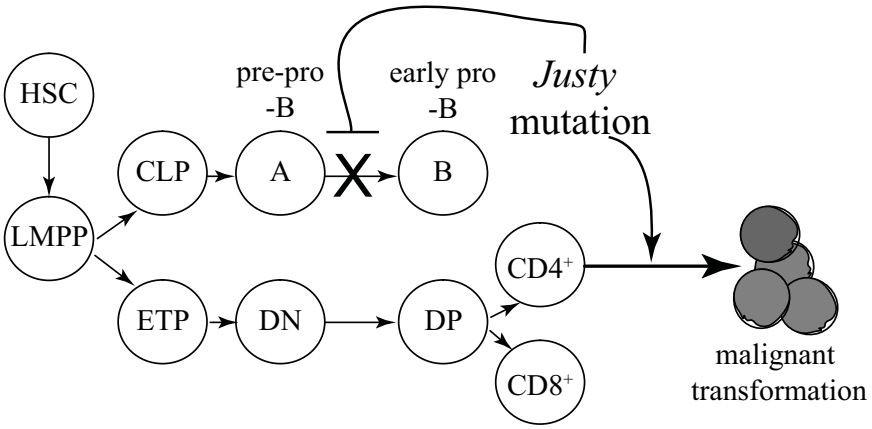
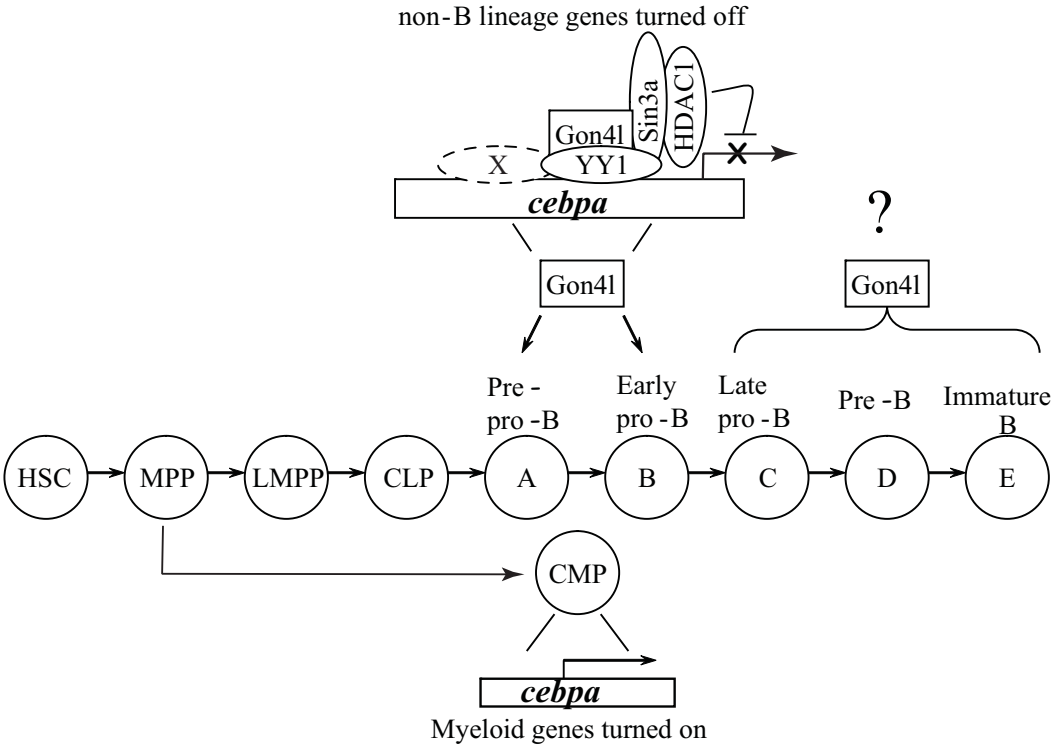

HSC
LMPP
CLP
ETP
pre-pro
-B
A
early pro
-B
B
DN
DP
CD4+
CD8+
Justy
mutation
malignant
transformation

Figure 27. The model of Gon4l-mediated transcriptional repression during early B lymphopoiesis.

A simplified B lymphopoiesis pathway and the assembly of the identified Gon4l complex on a putative Gon4l-regulated gene *cebpa* are illustrated. The "X" in the oval with dashed line denotes protein X, an unidentified factor that facilitates the Gon4l complex in the regulation of its target genes. Protein X can either directly associate with the Gon4l complex or be recruited separately to the target gene loci. The role of Gon4l in B lineage cells after the pro-B stage is unknown. Abbreviations: HSC, hematopoietic stem cells; MPP: multipotent progenitor; LMPP, lymphoid-primed progenitor; CLP, common lymphoid progenitor; CMP, common myeloid progenitor; *cebpa*, CCAAT/enhancer binding protein (C/EBP), alpha.


non-B lineage genes turned off
Sin3a
HDAC1
Gon4l
X
YY1
cebpa
?
Gon4l
Gon4l
Pre -
pro -B
Early
pro -B
Late
pro -B
Pre -B
Immature
B
HSC
MPP
LMPP
CLP
A
B
C
D
E
CMP
cebpa
Myeloid genes turned on

REFERENCES

1. LeBien, T. W., and Tedder, T. F. (2008) *Blood* **112**, 1570-1580

2. Busslinger, M. (2004) *Annu Rev Immunol* **22**, 55-79

3. Hardy, R. R., Kincade, P. W., and Dorshkind, K. (2007) *Immunity* **26**, 703-714

4. Nutt, S. L., and Kee, B. L. (2007) *Immunity* **26**, 715-725

5. Charles Janeway, P. T., Mark Walport, Mark Shlomchik (2004) *Immunobiology*, 6th Edition ed., Garland Science

6. Kondo, M., Scherer, D. C., King, A. G., Manz, M. G., and Weissman, I. L. (2001) *Curr Opin Genet Dev* **11**, 520-526

7. Weissman, I. L. (2000) *Science* **287**, 1442-1446

8. Christensen, J. L., and Weissman, I. L. (2001) *Proc Natl Acad Sci U S A* **98**, 14541-14546

9. Weissman, I. L., Anderson, D. J., and Gage, F. (2001) *Annu Rev Cell Dev Biol* **17**, 387-403

10. Lai, A. Y., Lin, S. M., and Kondo, M. (2005) *J Immunol* **175**, 5016-5023

11. Lai, A. Y., and Kondo, M. (2008) *Semin Immunol* **20**, 207-212

12. Lai, A. Y., Watanabe, A., O'Brien, T., and Kondo, M. (2009) *Blood* **113**, 5757-5764

13. Adolfsson, J., Borge, O. J., Bryder, D., Theilgaard-Monch, K., Astrand-Grundstrom, I., Sitnicka, E., Sasaki, Y., and Jacobsen, S. E. (2001) *Immunity* **15**, 659-669

14. Traver, D., and Akashi, K. (2004) *Adv Immunol* **83**, 1-54

15. Ye, M., Iwasaki, H., Laiosa, C. V., Stadtfeld, M., Xie, H., Heck, S., Clausen, B., Akashi, K., and Graf, T. (2003) *Immunity* **19**, 689-699

16. Adolfsson, J., Mansson, R., Buza-Vidas, N., Hultquist, A., Liuba, K., Jensen, C. T., Bryder, D., Yang, L., Borge, O. J., Thoren, L. A., Anderson, K., Sitnicka, E., Sasaki, Y., Sigvardsson, M., and Jacobsen, S. E. (2005) *Cell* **121**, 295-306

17. Mansson, R., Hultquist, A., Luc, S., Yang, L., Anderson, K., Kharazi, S., Al-Hashmi, S., Liuba, K., Thoren, L., Adolfsson, J., Buza-Vidas, N., Qian, H., Soneji, S., Enver, T., Sigvardsson, M., and Jacobsen, S. E. (2007) *Immunity* **26**, 407-419

18. Igarashi, H., Gregory, S. C., Yokota, T., Sakaguchi, N., and Kincade, P. W. (2002) *Immunity* **17**, 117-130

19. Medina, K. L., Garrett, K. P., Thompson, L. F., Rossi, M. I., Payne, K. J., and Kincade, P. W. (2001) *Nat Immunol* **2**, 718-724

20. Fuxa, M., and Skok, J. A. (2007) *Curr Opin Immunol* **19**, 129-136

21. Balciunaite, G., Ceredig, R., Massa, S., and Rolink, A. G. (2005) *Eur J Immunol* **35**, 2019-2030

22. Allman, D., Sambandam, A., Kim, S., Miller, J. P., Pagan, A., Well, D., Meraz, A., and Bhandoola, A. (2003) *Nat Immunol* **4**, 168-174

23. Harman, B. C., Miller, J. P., Nikbakht, N., Gerstein, R., and Allman, D. (2006) *Blood* **108**, 878-885

24. Kondo, M., Weissman, I. L., and Akashi, K. (1997) *Cell* **91**, 661-672

25. Bhandoola, A., von Boehmer, H., Petrie, H. T., and Zuniga-Pflucker, J. C. (2007) *Immunity* **26**, 678-689

26. Rumfelt, L. L., Zhou, Y., Rowley, B. M., Shinton, S. A., and Hardy, R. R. (2006) *J Exp Med* **203**, 675-687

27. Gounari, F., Aifantis, I., Martin, C., Fehling, H. J., Hoeflinger, S., Leder, P., von Boehmer, H., and Reizis, B. (2002) *Nat Immunol* **3**, 489-496

28. Jung, D., Giallourakis, C., Mostoslavsky, R., and Alt, F. W. (2006) *Annu Rev Immunol* **24**, 541-570

29. Rolink, A. G., Schaniel, C., Busslinger, M., Nutt, S. L., and Melchers, F. (2000) *Immunol Rev* **175**, 104-111

30. Holmes, M. L., Pridans, C., and Nutt, S. L. (2008) *Immunol Cell Biol* **86**, 47-53

31. Geier, J. K., and Schlissel, M. S. (2006) *Semin Immunol* **18**, 31-39

32. Herzog, S., Reth, M., and Jumaa, H. (2009) *Nat Rev Immunol* **9**, 195-205

33. Kitamura, D., Roes, J., Kuhn, R., and Rajewsky, K. (1991) *Nature* **350**, 423-426

34. Kitamura, D., Kudo, A., Schaal, S., Muller, W., Melchers, F., and Rajewsky, K. (1992) *Cell* **69**, 823-831

35. Reth, M., Petrac, E., Wiese, P., Lobel, L., and Alt, F. W. (1987) *Embo J* **6**, 3299-3305

36. Chakraborty, T., Chowdhury, D., Keyes, A., Jani, A., Subrahmanyam, R., Ivanova, I., and Sen, R. (2007) *Mol Cell* **27**, 842-850

37. Chowdhury, D., and Sen, R. (2001) *Embo J* **20**, 6394-6403

38. Chowdhury, D., and Sen, R. (2004) *Immunol Rev* **200**, 182-196

39. Morshead, K. B., Ciccone, D. N., Taverna, S. D., Allis, C. D., and Oettinger, M. A. (2003) *Proc Natl Acad Sci U S A* **100**, 11577-11582

40. Sayegh, C. E., Jhunjhunwala, S., Riblet, R., and Murre, C. (2005) *Genes Dev* **19**, 322-327

41. Roldan, E., Fuxa, M., Chong, W., Martinez, D., Novatchkova, M., Busslinger, M., and Skok, J. A. (2005) *Nat Immunol* **6**, 31-41

42. Fuxa, M., Skok, J., Souabni, A., Salvagiotto, G., Roldan, E., and Busslinger, M. (2004) *Genes Dev* **18**, 411-422

43. Nagasawa, T. (2006) *Nat Rev Immunol* **6**, 107-116

44. Kang, J., and Der, S. D. (2004) *Curr Opin Immunol* **16**, 180-190

45. Domen, J., and Weissman, I. L. (2000) *J Exp Med* **192**, 1707-1718

46. Waskow, C., Paul, S., Haller, C., Gassmann, M., and Rodewald, H. R. (2002) *Immunity* **17**, 277-288

47. McKenna, H. J., Stocking, K. L., Miller, R. E., Brasel, K., De Smedt, T., Maraskovsky, E., Maliszewski, C. R., Lynch, D. H., Smith, J., Pulendran, B., Roux, E. R., Teepe, M., Lyman, S. D., and Peschon, J. J. (2000) *Blood* **95**, 3489-3497

48. Sitnicka, E., Bryder, D., Theilgaard-Monch, K., Buza-Vidas, N., Adolfsson, J., and Jacobsen, S. E. (2002) *Immunity* **17**, 463-472

49. Mackarehtschian, K., Hardin, J. D., Moore, K. A., Boast, S., Goff, S. P., and Lemischka, I. R. (1995) *Immunity* **3**, 147-161

50. Lyman, S. D., James, L., Vanden Bos, T., de Vries, P., Brasel, K., Gliniak, B., Hollingsworth, L. T., Picha, K. S., McKenna, H. J., Splett, R. R., and et al. (1993) *Cell* **75**, 1157-1167

51. Ceredig, R., Rauch, M., Balciunaite, G., and Rolink, A. G. (2006) *Blood* **108**, 1216-1222

52. Hunte, B. E., Hudak, S., Campbell, D., Xu, Y., and Rennick, D. (1996) *J Immunol* **156**, 489-496

53. Sitnicka, E., Brakebusch, C., Martensson, I. L., Svensson, M., Agace, W. W., Sigvardsson, M., Buza-Vidas, N., Bryder, D., Cilio, C. M., Ahlenius, H., Maraskovsky, E., Peschon, J. J., and Jacobsen, S. E. (2003) *J Exp Med* **198**, 1495-1506

54. Fry, T. J., and Mackall, C. L. (2002) *Blood* **99**, 3892-3904

55. Hofmeister, R., Khaled, A. R., Benbernou, N., Rajnavolgyi, E., Muegge, K., and Durum, S. K. (1999) *Cytokine Growth Factor Rev* **10**, 41-60

56. Milne, C. D., and Paige, C. J. (2006) *Semin Immunol* **18**, 20-30

57. Dias, S., Silva, H., Jr., Cumano, A., and Vieira, P. (2005) *J Exp Med* **201**, 971-979

58. Kikuchi, K., Lai, A. Y., Hsu, C. L., and Kondo, M. (2005) *J Exp Med* **201**, 1197-1203

59. Corcoran, A. E., Riddell, A., Krooshoop, D., and Venkitaraman, A. R. (1998) *Nature* **391**, 904-907

60. Chowdhury, D., and Sen, R. (2003) *Immunity* **18**, 229-241

61. Georgopoulos, K. (2002) *Nat Rev Immunol* **2**, 162-174

62. Georgopoulos, K., Bigby, M., Wang, J. H., Molnar, A., Wu, P., Winandy, S., and Sharpe, A. (1994) *Cell* **79**, 143-156

63. Wang, J. H., Nichogiannopoulou, A., Wu, L., Sun, L., Sharpe, A. H., Bigby, M., and Georgopoulos, K. (1996) *Immunity* **5**, 537-549

64. Wu, L., Nichogiannopoulou, A., Shortman, K., and Georgopoulos, K. (1997) *Immunity* **7**, 483-492

65. Lopez, R. A., Schoetz, S., DeAngelis, K., O'Neill, D., and Bank, A. (2002) *Proc Natl Acad Sci U S A* **99**, 602-607

66. Nichogiannopoulou, A., Trevisan, M., Neben, S., Friedrich, C., and Georgopoulos, K. (1999) *J Exp Med* **190**, 1201-1214

67. Morgan, B., Sun, L., Avitahl, N., Andrikopoulos, K., Ikeda, T., Gonzales, E., Wu, P., Neben, S., and Georgopoulos, K. (1997) *Embo J* **16**, 2004-2013

68. Koipally, J., Renold, A., Kim, J., and Georgopoulos, K. (1999) *Embo J* **18**, 3090-3100

69. Kim, J., Sif, S., Jones, B., Jackson, A., Koipally, J., Heller, E., Winandy, S., Viel, A., Sawyer, A., Ikeda, T., Kingston, R., and Georgopoulos, K. (1999) *Immunity* **10**, 345-355

70. Koipally, J., and Georgopoulos, K. (2000) *J Biol Chem* **275**, 19594-19602

71. Ng, S. Y., Yoshida, T., and Georgopoulos, K. (2007) *Curr Opin Immunol* **19**, 116-122

72. Nutt, S. L., Metcalf, D., D'Amico, A., Polli, M., and Wu, L. (2005) *J Exp Med* **201**, 221-231

73. Akashi, K., Traver, D., Miyamoto, T., and Weissman, I. L. (2000) *Nature* **404**, 193-197

74. Hromas, R., Orazi, A., Neiman, R. S., Maki, R., Van Beveran, C., Moore, J., and Klemsz, M. (1993) *Blood* **82**, 2998-3004

75. Anderson, M. K., Hernandez-Hoyos, G., Diamond, R. A., and Rothenberg, E. V. (1999) *Development* **126**, 3131-3148

76. DeKoter, R. P., Lee, H. J., and Singh, H. (2002) *Immunity* **16**, 297-309

77. Scott, E. W., Simon, M. C., Anastasi, J., and Singh, H. (1994) *Science* **265**, 1573-1577

78. Scott, E. W., Fisher, R. C., Olson, M. C., Kehrli, E. W., Simon, M. C., and Singh, H. (1997) *Immunity* **6**, 437-447

79. Dakic, A., Metcalf, D., Di Rago, L., Mifsud, S., Wu, L., and Nutt, S. L. (2005) *J Exp Med* **201**, 1487-1502

80. Polli, M., Dakic, A., Light, A., Wu, L., Tarlinton, D. M., and Nutt, S. L. (2005) *Blood* **106**, 2083-2090

81. Iwasaki, H., Somoza, C., Shigematsu, H., Duprez, E. A., Iwasaki-Arai, J., Mizuno, S., Arinobu, Y., Geary, K., Zhang, P., Dayaram, T., Fenyus, M. L., Elf, S., Chan, S., Kastner, P., Huettner, C. S., Murray, R., Tenen, D. G., and Akashi, K. (2005) *Blood* **106**, 1590-1600

82. Shen, C. P., and Kadesch, T. (1995) *Mol Cell Biol* **15**, 4518-4524

83. Bain, G., Maandag, E. C., Izon, D. J., Amsen, D., Kruisbeek, A. M., Weintraub, B. C., Krop, I., Schlissel, M. S., Feeney, A. J., van Roon, M., and et al. (1994) *Cell* **79**, 885-892

84. Bain, G., Robanus Maandag, E. C., te Riele, H. P., Feeney, A. J., Sheehy, A., Schlissel, M., Shinton, S. A., Hardy, R. R., and Murre, C. (1997) *Immunity* **6**, 145-154

85. Zhuang, Y., Soriano, P., and Weintraub, H. (1994) *Cell* **79**, 875-884

86. Borghesi, L., Aites, J., Nelson, S., Lefterov, P., James, P., and Gerstein, R. (2005) *J Exp Med* **202**, 1669-1677

87. Greenbaum, S., and Zhuang, Y. (2002) *Proc Natl Acad Sci U S A* **99**, 15030-15035

88. Kee, B. L., and Murre, C. (1998) *J Exp Med* **188**, 699-713

89. Smith, E. M., Gisler, R., and Sigvardsson, M. (2002) *J Immunol* **169**, 261-270

90. Zhuang, Y., Barndt, R. J., Pan, L., Kelley, R., and Dai, M. (1998) *Mol Cell Biol* **18**, 3340-3349

91. OMIM. INHIBITOR OF DNA BINDING 1. NCBI

92. Sun, X. H. (1994) *Cell* **79**, 893-900

93. Jaleco, A. C., Stegmann, A. P., Heemskerk, M. H., Couwenberg, F., Bakker, A. Q., Weijer, K., and Spits, H. (1999) *Blood* **94**, 2637-2646

94. Spits, H., Couwenberg, F., Bakker, A. Q., Weijer, K., and Uittenbogaart, C. H. (2000) *J Exp Med* **192**, 1775-1784

95. Hacker, C., Kirsch, R. D., Ju, X. S., Hieronymus, T., Gust, T. C., Kuhl, C., Jorgas, T., Kurz, S. M., Rose-John, S., Yokota, Y., and Zenke, M. (2003) *Nat Immunol* **4**, 380-386

96. Kusunoki, T., Sugai, M., Katakai, T., Omatsu, Y., Iyoda, T., Inaba, K., Nakahata, T., Shimizu, A., and Yokota, Y. (2003) *J Allergy Clin Immunol* **111**, 136-142

97. Hagman, J., Belanger, C., Travis, A., Turck, C. W., and Grosschedl, R. (1993) *Genes Dev* **7**, 760-773

98. Lin, H., and Grosschedl, R. (1995) *Nature* **376**, 263-267

99. Zhang, Z., Cotta, C. V., Stephan, R. P., deGuzman, C. G., and Klug, C. A. (2003) *Embo J* **22**, 4759-4769

100. Seet, C. S., Brumbaugh, R. L., and Kee, B. L. (2004) *J Exp Med* **199**, 1689-1700

101. Medina, K. L., Pongubala, J. M., Reddy, K. L., Lancki, D. W., Dekoter, R., Kieslinger, M., Grosschedl, R., and Singh, H. (2004) *Dev Cell* **7**, 607-617

102. Sigvardsson, M., O'Riordan, M., and Grosschedl, R. (1997) *Immunity* **7**, 25-36

103. Sigvardsson, M. (2000) *Mol Cell Biol* **20**, 3640-3654

104. Gisler, R., and Sigvardsson, M. (2002) *J Immunol* **168**, 5130-5138

105. Sigvardsson, M., Clark, D. R., Fitzsimmons, D., Doyle, M., Akerblad, P., Breslin, T., Bilke, S., Li, R., Yeamans, C., Zhang, G., and Hagman, J. (2002) *Mol Cell Biol* **22**, 8539-8551

106. O'Riordan, M., and Grosschedl, R. (1999) *Immunity* **11**, 21-31

107. Chen, J., Young, F., Bottaro, A., Stewart, V., Smith, R. K., and Alt, F. W. (1993) *Embo J* **12**, 4635-4645

108. Serwe, M., and Sablitzky, F. (1993) *Embo J* **12**, 2321-2327

109. Schlissel, M., Voronova, A., and Baltimore, D. (1991) *Genes Dev* **5**, 1367-1376

110. Romanow, W. J., Langerak, A. W., Goebel, P., Wolvers-Tettero, I. L., van Dongen, J. J., Feeney, A. J., and Murre, C. (2000) *Mol Cell* **5**, 343-353

111. Goebel, P., Janney, N., Valenzuela, J. R., Romanow, W. J., Murre, C., and Feeney, A. J. (2001) *J Exp Med* **194**, 645-656

112. Massari, M. E., Grant, P. A., Pray-Grant, M. G., Berger, S. L., Workman, J. L., and Murre, C. (1999) *Mol Cell* **4**, 63-73

113. Cobaleda, C., Schebesta, A., Delogu, A., and Busslinger, M. (2007) *Nat Immunol* **8**, 463-470

114. Fuxa, M., and Busslinger, M. (2007) *J Immunol* **178**, 3031-3037

115. Urbanek, P., Wang, Z. Q., Fetka, I., Wagner, E. F., and Busslinger, M. (1994) *Cell* **79**, 901-912

116. Nutt, S. L., Urbanek, P., Rolink, A., and Busslinger, M. (1997) *Genes Dev* **11**, 476-491

117. Nutt, S. L., Heavey, B., Rolink, A. G., and Busslinger, M. (1999) *Nature* **401**, 556-562

118. Rolink, A. G., Nutt, S. L., Melchers, F., and Busslinger, M. (1999) *Nature* **401**, 603-606

119. Delogu, A., Schebesta, A., Sun, Q., Aschenbrenner, K., Perlot, T., and Busslinger, M. (2006) *Immunity* **24**, 269-281

120. Nutt, S. L., Morrison, A. M., Dorfler, P., Rolink, A., and Busslinger, M. (1998) *Embo J* **17**, 2319-2333

121. Souabni, A., Cobaleda, C., Schebesta, M., and Busslinger, M. (2002) *Immunity* **17**, 781-793

122. Tagoh, H., Ingram, R., Wilson, N., Salvagiotto, G., Warren, A. J., Clarke, D., Busslinger, M., and Bonifer, C. (2006) *Embo J* **25**, 1070-1080

123. Mikkola, I., Heavey, B., Horcher, M., and Busslinger, M. (2002) *Science* **297**, 110-113

124. Eberhard, D., Jimenez, G., Heavey, B., and Busslinger, M. (2000) *Embo J* **19**, 2292-2303

125. Linderson, Y., Eberhard, D., Malin, S., Johansson, A., Busslinger, M., and Pettersson, S. (2004) *EMBO Rep* **5**, 291-296

126. Choi, C. Y., Kim, Y. H., Kwon, H. J., and Kim, Y. (1999) *J Biol Chem* **274**, 33194-33197

127. Chen, G., Fernandez, J., Mische, S., and Courey, A. J. (1999) *Genes Dev* **13**, 2218-2230

128. Haberland, M., Montgomery, R. L., and Olson, E. N. (2009) *Nature reviews* **10**, 32-42

129. Farkas, G., Leibovitch, B. A., and Elgin, S. C. (2000) *Gene* **253**, 117-136

130. Jaenisch, R., and Bird, A. (2003) *Nat Genet* **33 Suppl**, 245-254

131. Tagoh, H., Melnik, S., Lefevre, P., Chong, S., Riggs, A. D., and Bonifer, C. (2004) *Blood* **103**, 2950-2955

132. Kouzarides, T. (2007) *Cell* **128**, 693-705

133. Li, B., Carey, M., and Workman, J. L. (2007) *Cell* **128**, 707-719

134. Lessard, J., and Crabtree, G. R. (2010) *Annu Rev Cell Dev Biol*

135. Shahbazian, M. D., and Grunstein, M. (2007) *Annu Rev Biochem* **76**, 75-100

136. Juliandi, B., Abematsu, M., and Nakashima, K. (2010) *Curr Opin Neurobiol*

137. Kurdistani, S. K., Tavazoie, S., and Grunstein, M. (2004) *Cell* **117**, 721-733

138. Pokholok, D. K., Harbison, C. T., Levine, S., Cole, M., Hannett, N. M., Lee, T. I., Bell, G. W., Walker, K., Rolfe, P. A., Herbolsheimer, E., Zeitlinger, J., Lewitter, F., Gifford, D. K., and Young, R. A. (2005) *Cell* **122**, 517-527

139. Clapier, C. R., and Cairns, B. R. (2009) *Annu Rev Biochem* **78**, 273-304

140. O'Carroll, D., Mecklenbrauker, I., Das, P. P., Santana, A., Koenig, U., Enright, A. J., Miska, E. A., and Tarakhovsky, A. (2007) *Genes Dev* **21**, 1999-2004

141. Ayer, D. E., Lawrence, Q. A., and Eisenman, R. N. (1995) *Cell* **80**, 767-776

142. Vermeulen, M., Walter, W., Le Guezennec, X., Kim, J., Edayathumangalam, R. S., Lasonder, E., Luger, K., Roeder, R. G., Logie, C., Berger, S. L., and Stunnenberg, H. G. (2006) *Mol Cell Biol* **26**, 5226-5236

143. Alland, L., Muhle, R., Hou, H., Jr., Potes, J., Chin, L., Schreiber-Agus, N., and DePinho, R. A. (1997) *Nature* **387**, 49-55

144. Nagy, L., Kao, H. Y., Chakravarti, D., Lin, R. J., Hassig, C. A., Ayer, D. E., Schreiber, S. L., and Evans, R. M. (1997) *Cell* **89**, 373-380

145. Schreiber-Agus, N., and DePinho, R. A. (1998) *Bioessays* **20**, 808-818

146. Hassig, C. A., Fleischer, T. C., Billin, A. N., Schreiber, S. L., and Ayer, D. E. (1997) *Cell* **89**, 341-347

147. Heinzel, T., Lavinsky, R. M., Mullen, T. M., Soderstrom, M., Laherty, C. D., Torchia, J., Yang, W. M., Brard, G., Ngo, S. D., Davie, J. R., Seto, E., Eisenman, R. N., Rose, D. W., Glass, C. K., and Rosenfeld, M. G. (1997) *Nature* **387**, 43-48

148. Laherty, C. D., Yang, W. M., Sun, J. M., Davie, J. R., Seto, E., and Eisenman, R. N. (1997) *Cell* **89**, 349-356

149. Xue, Y., Wong, J., Moreno, G. T., Young, M. K., Cote, J., and Wang, W. (1998) *Mol Cell* **2**, 851-861

150. Yang, W. M., Inouye, C., Zeng, Y., Bearss, D., and Seto, E. (1996) *Proc Natl Acad Sci U S A* **93**, 12845-12850

151. Perissi, V., Jepsen, K., Glass, C. K., and Rosenfeld, M. G. *Nat Rev Genet* **11**, 109-123

152. Rave-Harel, N., Miller, N. L., Givens, M. L., and Mellon, P. L. (2005) *J Biol Chem* **280**, 30975-30983

153. Dasen, J. S., Barbera, J. P., Herman, T. S., Connell, S. O., Olson, L., Ju, B., Tollkuhn, J., Baek, S. H., Rose, D. W., and Rosenfeld, M. G. (2001) *Genes Dev* **15**, 3193-3207

154. Pongubala, J. M., Northrup, D. L., Lancki, D. W., Medina, K. L., Treiber, T., Bertolino, E., Thomas, M., Grosschedl, R., Allman, D., and Singh, H. (2008) *Nat Immunol* **9**, 203-215

155. Dannenberg, J. H., David, G., Zhong, S., van der Torre, J., Wong, W. H., and Depinho, R. A. (2005) *Genes Dev* **19**, 1581-1595

156. Schreiber-Agus, N., Chin, L., Chen, K., Torres, R., Rao, G., Guida, P., Skoultchi, A. I., and DePinho, R. A. (1995) *Cell* **80**, 777-786

157. Ayer, D. E. (1999) *Trends Cell Biol* **9**, 193-198

158. Silverstein, R. A., and Ekwall, K. (2005) *Curr Genet* **47**, 1-17

159. Guenther, M. G., Barak, O., and Lazar, M. A. (2001) *Mol Cell Biol* **21**, 6091-6101

160. Murphy, M., Ahn, J., Walker, K. K., Hoffman, W. H., Evans, R. M., Levine, A. J., and George, D. L. (1999) *Genes Dev* **13**, 2490-2501

161. Huang, Y., Myers, S. J., and Dingledine, R. (1999) *Nat Neurosci* **2**, 867-872

162. Rayman, J. B., Takahashi, Y., Indjeian, V. B., Dannenberg, J. H., Catchpole, S., Watson, R. J., te Riele, H., and Dynlacht, B. D. (2002) *Genes Dev* **16**, 933-947

163. Zhao, X., Jankovic, V., Gural, A., Huang, G., Pardanani, A., Menendez, S., Zhang, J., Dunne, R., Xiao, A., Erdjument-Bromage, H., Allis, C. D., Tempst, P., and Nimer, S. D. (2008) *Genes Dev* **22**, 640-653

164. Yang, L., Mei, Q., Zielinska-Kwiatkowska, A., Matsui, Y., Blackburn, M. L., Benedetti, D., Krumm, A. A., Taborsky, G. J., Jr., and Chansky, H. A. (2003) *Biochem J* **369**, 651-657

165. Wysocka, J., Myers, M. P., Laherty, C. D., Eisenman, R. N., and Herr, W. (2003) *Genes Dev* **17**, 896-911

166. De Nadal, E., Zapater, M., Alepuz, P. M., Sumoy, L., Mas, G., and Posas, F. (2004) *Nature* **427**, 370-374

167. van Oevelen, C., Wang, J., Asp, P., Yan, Q., Kaelin, W. G., Jr., Kluger, Y., and Dynlacht, B. D. (2008) *Mol Cell* **32**, 359-370

168. Thomas, M. J., and Seto, E. (1999) *Gene* **236**, 197-208

169. Atchison, L., Ghias, A., Wilkinson, F., Bonini, N., and Atchison, M. L. (2003) *Embo J* **22**, 1347-1358

170. Gordon, S., Akopyan, G., Garban, H., and Bonavida, B. (2006) *Oncogene* **25**, 1125-1142

171. Brown, J. L., Mucci, D., Whiteley, M., Dirksen, M. L., and Kassis, J. A. (1998) *Mol Cell* **1**, 1057-1064

172. Brown, J. L., Fritsch, C., Mueller, J., and Kassis, J. A. (2003) *Development* **130**, 285-294

173. Srinivasan, L., and Atchison, M. L. (2004) *Genes Dev* **18**, 2596-2601

174. Mai, R. T., Yeh, T. S., Kao, C. F., Sun, S. K., Huang, H. H., and Wu Lee, Y. H. (2006) *Oncogene* **25**, 448-462

175. Shi, Y., Lee, J. S., and Galvin, K. M. (1997) *Biochim Biophys Acta* **1332**, F49-66

176. Kang, K., Chung, J. H., and Kim, J. (2009) *Nucleic Acids Res* **37**, 2003-2013

177. Kim, J. D., Hinz, A. K., Bergmann, A., Huang, J. M., Ovcharenko, I., Stubbs, L., and Kim, J. (2006) *Genome Res* **16**, 901-911

178. Donohoe, M. E., Zhang, X., McGinnis, L., Biggers, J., Li, E., and Shi, Y. (1999) *Mol Cell Biol* **19**, 7237-7244

179. Castellano, G., Torrisi, E., Ligresti, G., Malaponte, G., Militello, L., Russo, A. E., McCubrey, J. A., Canevari, S., and Libra, M. (2009) *Cell Cycle* **8**, 1367-1372

180. Gronroos, E., Terentiev, A. A., Punga, T., and Ericsson, J. (2004) *Proc Natl Acad Sci U S A* **101**, 12165-12170

181. Sui, G., Affar el, B., Shi, Y., Brignone, C., Wall, N. R., Yin, P., Donohoe, M., Luke, M. P., Calvo, D., Grossman, S. R., and Shi, Y. (2004) *Cell* **117**, 859-872

182. Liu, H., Schmidt-Supprian, M., Shi, Y., Hobeika, E., Barteneva, N., Jumaa, H., Pelanda, R., Reth, M., Skok, J., Rajewsky, K., and Shi, Y. (2007) *Genes Dev* **21**, 1179-1189

183. Gordon, S. J., Saleque, S., and Birshtein, B. K. (2003) *J Immunol* **170**, 5549-5557

184. Maier, H., Ostraat, R., Gao, H., Fields, S., Shinton, S. A., Medina, K. L., Ikawa, T., Murre, C., Singh, H., Hardy, R. R., and Hagman, J. (2004) *Nat Immunol* **5**, 1069-1077

185. Jimenez, G., Griffiths, S. D., Ford, A. M., Greaves, M. F., and Enver, T. (1992) *Proc Natl Acad Sci U S A* **89**, 10618-10622

186. Enver, T., and Greaves, M. (1998) *Cell* **94**, 9-12

187. Kontaraki, J., Chen, H. H., Riggs, A., and Bonifer, C. (2000) *Genes Dev* **14**, 2106-2122

188. Hayashi, K., Yamamoto, M., Nojima, T., Goitsuka, R., and Kitamura, D. (2003) *Immunity* **18**, 825-836

189. O'Neil, J., and Look, A. T. (2007) *Oncogene* **26**, 6838-6849

190. Rosenbauer, F., and Tenen, D. G. (2007) *Nat Rev Immunol* **7**, 105-117

191. Boyer, L. A., Langer, M. R., Crowley, K. A., Tan, S., Denu, J. M., and Peterson, C. L. (2002) *Mol Cell* **10**, 935-942

192. Lu, P., Hankel, I. L., Knisz, J., Marquardt, A., Chiang, M. Y., Grosse, J., Constien, R., Meyer, T., Schroeder, A., Zeitlmann, L., Al-Alem, U., Friedman, A. D., Elliott, E. I., Meyerholz, D. K., Waldschmidt, T. J., Rothman, P. B., and Colgan, J. D. (2010) *J Exp Med*

193. Parra, D., Manils, J., Castellana, B., Vina-Vilaseca, A., Moran-Salvador, E., Vazquez-Villoldo, N., Tarancon, G., Borras, M., Sancho, S., Benito, C., Ortega, S., and Soler, C. (2009) *Cancer Res* **69**, 6676-6684

194. Livak, K. J., and Schmittgen, T. D. (2001) *Methods* **25**, 402-408

195. Rolink, A., ten Boekel, E., Melchers, F., Fearon, D. T., Krop, I., and Andersson, J. (1996) *J Exp Med* **183**, 187-194

196. Tudor, K. S., Payne, K. J., Yamashita, Y., and Kincade, P. W. (2000) *Immunity* **12**, 335-345

197. Mansson, R., Zandi, S., Welinder, E., Tsapogas, P., Sakaguchi, N., Bryder, D., and Sigvardsson, M. *Blood* **115**, 2601-2609

198. Inlay, M. A., Bhattacharya, D., Sahoo, D., Serwold, T., Seita, J., Karsunky, H., Plevritis, S. K., Dill, D. L., and Weissman, I. L. (2009) *Genes Dev* **23**, 2376-2381

199. Bruce Alberts; Alexander Johnson, J. L., Martin Raff, Keith Roberts, Peter Walter. (2002) *Molecular Biology of the Cell, Fourth Edition* Garland Science

200. Chang, Y. F., Imam, J. S., and Wilkinson, M. F. (2007) *Annu Rev Biochem* **76**, 51-74

201. Kuryshev, V. Y., Vorobyov, E., Zink, D., Schmitz, J., Rozhdestvensky, T. S., Munstermann, E., Ernst, U., Wellenreuther, R., Moosmayer, P., Bechtel, S., Schupp, I., Horst, J., Korn, B., Poustka, A., and Wiemann, S. (2006) *Genomics* **88**, 143-151

202. Ohtomo, T., Horii, T., Nomizu, M., Suga, T., and Yamada, J. (2007) *Amino acids* **33**, 645-652

203. Ohtomo, T., Horii, T., Nomizu, M., Suga, T., and Yamada, J. (2008) *Amino acids* **34**, 155-161

204. Zhang, C., Frias, M. A., Mele, A., Ruggiu, M., Eom, T., Marney, C. B., Wang, H., Licatalosi, D. D., Fak, J. J., and Darnell, R. B. (2010) *Science*

205. Zhang, C., Zhang, Z., Castle, J., Sun, S., Johnson, J., Krainer, A. R., and Zhang, M. Q. (2008) *Genes Dev* **22**, 2550-2563

206. Li, Q., Lee, J. A., and Black, D. L. (2007) *Nat Rev Neurosci* **8**, 819-831

207. Pollard, J. W. (2009) *Nat Rev Immunol* **9**, 259-270

208. Dai, X. M., Ryan, G. R., Hapel, A. J., Dominguez, M. G., Russell, R. G., Kapp, S., Sylvestre, V., and Stanley, E. R. (2002) *Blood* **99**, 111-120

209. Heavey, B., Charalambous, C., Cobaleda, C., and Busslinger, M. (2003) *Embo J* **22**, 3887-3897

210. Hsu, C. L., King-Fleischman, A. G., Lai, A. Y., Matsumoto, Y., Weissman, I. L., and Kondo, M. (2006) *Proc Natl Acad Sci U S A* **103**, 672-677

211. Xie, H., Ye, M., Feng, R., and Graf, T. (2004) *Cell* **117**, 663-676

212. Hammerschmidt, M., Pelegri, F., Mullins, M. C., Kane, D. A., Brand, M., van Eeden, F. J., Furutani-Seiki, M., Granato, M., Haffter, P., Heisenberg, C. P.,

Jiang, Y. J., Kelsh, R. N., Odenthal, J., Warga, R. M., and Nusslein-Volhard, C. (1996) *Development* **123**, 143-151

213. Liu, Y., Du, L., Osato, M., Teo, E. H., Qian, F., Jin, H., Zhen, F., Xu, J., Guo, L., Huang, H., Chen, J., Geisler, R., Jiang, Y. J., Peng, J., and Wen, Z. (2007) *Blood* **110**, 99-106

214. Hoffmann, R., Seidl, T., Neeb, M., Rolink, A., and Melchers, F. (2002) *Genome Res* **12**, 98-111

215. Hystad, M. E., Myklebust, J. H., Bo, T. H., Sivertsen, E. A., Rian, E., Forfang, L., Munthe, E., Rosenwald, A., Chiorazzi, M., Jonassen, I., Staudt, L. M., and Smeland, E. B. (2007) *J Immunol* **179**, 3662-3671

216. Laiosa, C. V., Stadtfeld, M., and Graf, T. (2006) *Annu Rev Immunol* **24**, 705-738

217. Baylin, S. B., and Ohm, J. E. (2006) *Nat Rev Cancer* **6**, 107-116

218. Ahringer, J. (2000) *Trends Genet* **16**, 351-356

219. Jepsen, K., and Rosenfeld, M. G. (2002) *Journal of cell science* **115**, 689-698

220. Cunliffe, V. T. (2008) *Current opinion in genetics & development* **18**, 404-410

221. McDonel, P., Costello, I., and Hendrich, B. (2009) *Int J Biochem Cell Biol* **41**, 108-116

222. Zhong, J., Zhang, H., Stanyon, C. A., Tromp, G., and Finley, R. L., Jr. (2003) *Genome Res* **13**, 2691-2699

223. Li, Y., Lu, J., and Prochownik, E. V. (2007) *J Biol Chem* **282**, 13363-13371

224. Lutterbach, B., Westendorf, J. J., Linggi, B., Patten, A., Moniwa, M., Davie, J. R., Huynh, K. D., Bardwell, V. J., Lavinsky, R. M., Rosenfeld, M. G., Glass, C., Seto, E., and Hiebert, S. W. (1998) *Mol Cell Biol* **18**, 7176-7184

225. Frehlick, L. J., Eirin-Lopez, J. M., Jeffery, E. D., Hunt, D. F., and Ausio, J. (2006) *BMC Genomics* **7**, 99

226. Laskey, R. A., Mills, A. D., Philpott, A., Leno, G. H., Dilworth, S. M., and Dingwall, C. (1993) *Philos Trans R Soc Lond B Biol Sci* **339**, 263-269; discussion 268-269

227. Frehlick, L. J., Eirin-Lopez, J. M., and Ausio, J. (2007) *Bioessays* **29**, 49-59

228. Wang, C. Y., Liang, Y. J., Lin, Y. S., Shih, H. M., Jou, Y. S., and Yu, W. C. (2004) *J Biol Chem* **279**, 17750-17755

229. Lu, S., Deng, P., Liu, X., Luo, J., Han, R., Gu, X., Liang, S., Wang, X., Li, F., Lozanov, V., Patthy, A., and Pongor, S. (1999) *J Biol Chem* **274**, 20473-20478

230. Grzenda, A., Lomberk, G., Zhang, J. S., and Urrutia, R. (2009) *Biochim Biophys Acta* **1789**, 443-450

231. Boyer, L. A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L. A., Lee, T. I., Levine, S. S., Wernig, M., Tajonar, A., Ray, M. K., Bell, G. W., Otte, A. P.,

Vidal, M., Gifford, D. K., Young, R. A., and Jaenisch, R. (2006) *Nature* **441**, 349-353

232. Mo, X., Kowenz-Leutz, E., Laumonnier, Y., Xu, H., and Leutz, A. (2005) *Genes Dev* **19**, 2447-2457

233. Palko, L., Bass, H. W., Beyrouthy, M. J., and Hurt, M. M. (2004) *J Cell Sci* **117**, 465-476

234. NCBI. http://www.ncbi.nlm.nih.gov/gene/54856#geneGeneral protein info.

235. Huang, N. E., Lin, C. H., Lin, Y. S., and Yu, W. C. (2003) *Biochem Biophys Res Commun* **306**, 267-275

236. Kuhn, R., Schwenk, F., Aguet, M., and Rajewsky, K. (1995) *Science* **269**, 1427-1429

237. Alland, L., David, G., Shen-Li, H., Potes, J., Muhle, R., Lee, H. C., Hou, H., Jr., Chen, K., and DePinho, R. A. (2002) *Mol Cell Biol* **22**, 2743-2750

238. Erkeland, S. J., Valkhof, M., Heijmans-Antonissen, C., Delwel, R., Valk, P. J., Hermans, M. H., and Touw, I. P. (2003) *Blood* **101**, 1111-1117

239. Sen, R., and Oltz, E. (2006) *Curr Opin Immunol* **18**, 237-242

240. Patrone, L., Henson, S. E., Wall, R., and Malone, C. S. (2004) *Mol Biol Rep* **31**, 1-11

241. Su, K., Li, X., Edberg, J. C., Wu, J., Ferguson, P., and Kimberly, R. P. (2004) *J Immunol* **172**, 7192-7199

242. Kurisaki, K., Kurisaki, A., Valcourt, U., Terentiev, A. A., Pardali, K., Ten Dijke, P., Heldin, C. H., Ericsson, J., and Moustakas, A. (2003) *Mol Cell Biol* **23**, 4494-4510

243. Zhang, Z., Espinoza, C. R., Yu, Z., Stephan, R., He, T., Williams, G. S., Burrows, P. D., Hagman, J., Feeney, A. J., and Cooper, M. D. (2006) *Nat Immunol* **7**, 616-624

244. Schreiber, E., Matthias, P., Muller, M. M., and Schaffner, W. (1989) *Nucleic Acids Res* **17**, 6419

245. Saur, D., Seidler, B., Paehge, H., Schusdziarra, V., and Allescher, H. D. (2002) *J Biol Chem* **277**, 25798-25814

246. Chen, J. D., and Evans, R. M. (1995) *Nature* **377**, 454-457

247. Ropero, S., Ballestar, E., Alaminos, M., Arango, D., Schwartz, S., Jr., and Esteller, M. (2008) *Oncogene* **27**, 4008-4012

248. Jones, P. A., and Baylin, S. B. (2007) *Cell* **128**, 683-692

249. Wang, G. G., Allis, C. D., and Chi, P. (2007) *Trends Mol Med* **13**, 363-372

250. Sparmann, A., and van Lohuizen, M. (2006) *Nat Rev Cancer* **6**, 846-856

251. Ye, J., Ghosh, P., Cippitelli, M., Subleski, J., Hardy, K. J., Ortaldo, J. R., and Young, H. A. (1994) *J Biol Chem* **269**, 25728-25734

252. Tagoh, H., Himes, R., Clarke, D., Leenen, P. J., Riggs, A. D., Hume, D., and Bonifer, C. (2002) *Genes Dev* **16**, 1721-1737

253. Zhang, D. E., Hetherington, C. J., Meyers, S., Rhoades, K. L., Larson, C. J., Chen, H. M., Hiebert, S. W., and Tenen, D. G. (1996) *Mol Cell Biol* **16**, 1231-1240

254. Legraverend, C., Antonson, P., Flodby, P., and Xanthopoulos, K. G. (1993) *Nucleic Acids Res* **21**, 1735-1742

255. Smith, C. L. (2008) *Bioessays* **30**, 15-24

256. Vidal, M., Strich, R., Esposito, R. E., and Gaber, R. F. (1991) *Mol Cell Biol* **11**, 6306-6316

257. Singh, M., and Birshtein, B. K. (1996) *Proc Natl Acad Sci U S A* **93**, 4392-4397

258. de la Cruz, X., Lois, S., Sanchez-Molina, S., and Martinez-Balbas, M. A. (2005) *Bioessays* **27**, 164-175

259. Fleischer, T. C., Yun, U. J., and Ayer, D. E. (2003) *Mol Cell Biol* **23**, 3456-3467

260. Wong, C. W., and Privalsky, M. L. (1998) *Mol Cell Biol* **18**, 5500-5510

261. Lim, C. H., Chong, S. W., and Jiang, Y. J. (2009) *Mol Biol Cell* **20**, 4183-4193

262. Cicatiello, L., Addeo, R., Sasso, A., Altucci, L., Petrizzi, V. B., Borgo, R., Cancemi, M., Caporali, S., Caristi, S., Scafoglio, C., Teti, D., Bresciani, F., Perillo, B., and Weisz, A. (2004) *Mol Cell Biol* **24**, 7260-7274

263. Shrivastava, A., Yu, J., Artandi, S., and Calame, K. (1996) *Proc Natl Acad Sci U S A* **93**, 10638-10641

264. Shrivastava, A., Saleque, S., Kalpana, G. V., Artandi, S., Goff, S. P., and Calame, K. (1993) *Science* **262**, 1889-1892

265. Parija, T., and Das, B. R. (2003) *Mol Biol Rep* **30**, 41-46

266. Lai, A., Kennedy, B. K., Barbie, D. A., Bertos, N. R., Yang, X. J., Theberge, M. C., Tsai, S. C., Seto, E., Zhang, Y., Kuzmichev, A., Lane, W. S., Reinberg, D., Harlow, E., and Branton, P. E. (2001) *Mol Cell Biol* **21**, 2918-2932

267. Brehm, A., Miska, E. A., McCance, D. J., Reid, J. L., Bannister, A. J., and Kouzarides, T. (1998) *Nature* **391**, 597-601

268. Suzuki, H., Ouchida, M., Yamamoto, H., Yano, M., Toyooka, S., Aoe, M., Shimizu, N., Date, H., and Shimizu, K. (2008) *Lung Cancer* **59**, 24-31

269. Melnick, A., and Licht, J. D. (2002) *Curr Opin Hematol* **9**, 322-332

270. Khan, M. M., Nomura, T., Kim, H., Kaul, S. C., Wadhwa, R., Shinagawa, T., Ichikawa-Iwata, E., Zhong, S., Pandolfi, P. P., and Ishii, S. (2001) *Mol Cell* **7**, 1233-1243

271. Khan, M. M., Nomura, T., Kim, H., Kaul, S. C., Wadhwa, R., Zhong, S., Pandolfi, P. P., and Ishii, S. (2001) *J Biol Chem* **276**, 43491-43494

272. Winandy, S., Wu, P., and Georgopoulos, K. (1995) *Cell* **83**, 289-299

273. Conley, M. E., Dobbs, A. K., Farmer, D. M., Kilic, S., Paris, K., Grigoriadou, S., Coustan-Smith, E., Howard, V., and Campana, D. (2009) *Annu Rev Immunol* **27**, 199-227

274. Griffith, L. M., Cowan, M. J., Notarangelo, L. D., Puck, J. M., Buckley, R. H., Candotti, F., Conley, M. E., Fleisher, T. A., Gaspar, H. B., Kohn, D. B., Ochs, H. D., O'Reilly, R. J., Rizzo, J. D., Roifman, C. M., Small, T. N., and Shearer, W. T. (2009) *J Allergy Clin Immunol* **124**, 1152-1160 e1112

275. Shehata, N., Palda, V., Bowen, T., Haddad, E., Issekutz, T. B., Mazer, B., Schellenberg, R., Warrington, R., Easton, D., Anderson, D., and Hume, H. (2010) *Transfus Med Rev* **24 Suppl 1**, S28-50

276. Park, I. H., Arora, N., Huo, H., Maherali, N., Ahfeldt, T., Shimamura, A., Lensch, M. W., Cowan, C., Hochedlinger, K., and Daley, G. Q. (2008) *Cell* **134**, 877-886

277. Kim, D., Kim, C. H., Moon, J. I., Chung, Y. G., Chang, M. Y., Han, B. S., Ko, S., Yang, E., Cha, K. Y., Lanza, R., and Kim, K. S. (2009) *Cell Stem Cell* **4**, 472-476

278. Meissner, A., Wernig, M., and Jaenisch, R. (2007) *Nat Biotechnol* **25**, 1177-1181

279. Giorgetti, A., Montserrat, N., Aasen, T., Gonzalez, F., Rodriguez-Piza, I., Vassena, R., Raya, A., Boue, S., Barrero, M. J., Corbella, B. A., Torrabadella, M., Veiga, A., and Izpisua Belmonte, J. C. (2009) *Cell Stem Cell* **5**, 353-357

280. Haase, A., Olmer, R., Schwanke, K., Wunderlich, S., Merkert, S., Hess, C., Zweigerdt, R., Gruh, I., Meyer, J., Wagner, S., Maier, L. S., Han, D. W., Glage, S., Miller, K., Fischer, P., Scholer, H. R., and Martin, U. (2009) *Cell Stem Cell* **5**, 434-441

281. Loh, Y. H., Agarwal, S., Park, I. H., Urbach, A., Huo, H., Heffner, G. C., Kim, K., Miller, J. D., Ng, K., and Daley, G. Q. (2009) *Blood* **113**, 5476-5479

282. Giorgetti, A., Montserrat, N., Rodriguez-Piza, I., Azqueta, C., Veiga, A., and Izpisua Belmonte, J. C. (2010) *Nat Protoc* **5**, 811-820

283. Kunisato, A., Wakatsuki, M., Shinba, H., Ota, T., Ishida, I., and Nagao, K. (2010) *Stem Cells Dev*

284. Staerk, J., Dawlaty, M. M., Gao, Q., Maetzel, D., Hanna, J., Sommer, C. A., Mostoslavsky, G., and Jaenisch, R. (2010) *Cell Stem Cell* **7**, 20-24

285. Grimes, J. A., Nielsen, S. J., Battaglioli, E., Miska, E. A., Speh, J. C., Berry, D. L., Atouf, F., Holdener, B. C., Mandel, G., and Kouzarides, T. (2000) *J Biol Chem* **275**, 9461-9467

286. Roopra, A., Sharling, L., Wood, I. C., Briggs, T., Bachfischer, U., Paquette, A. J., and Buckley, N. J. (2000) *Mol Cell Biol* **20**, 2147-2157

287. Zhang, J., Kalkum, M., Yamamura, S., Chait, B. T., and Roeder, R. G. (2004) *Science* **305**, 1286-1289

288. Hainaut, P., and Wiman, K. G. (2009) *Lancet Oncol* **10**, 913-919

289. Hassler, M., Singh, S., Yue, W. W., Luczynski, M., Lakbir, R., Sanchez-Sanchez, F., Bader, T., Pearl, L. H., and Mittnacht, S. (2007) *Mol Cell* **28**, 371-385

290. Wodarz, D. (2007) *J Theor Biol* **245**, 449-458

291. el-Naggar, A. K., Lovell, M., Callender, D. L., Ordonez, N. G., and Killary, A. M. (1999) *Cancer Genet Cytogenet* **113**, 49-53

292. Cheuk, W., and Chan, J. K. (2007) *Histopathology* **51**, 1-20

293. Di Palma, S., and Guzzo, M. (1993) *Virchows Arch A Pathol Anat Histopathol* **423**, 389-396

294. Bombi, J. A., Alos, L., Rey, M. J., Mallofre, C., Cuchi, A., Trasserra, J., and Cardesa, A. (1996) *Ultrastruct Pathol* **20**, 145-154

295. Stenman, G. (2005) *Semin Cancer Biol* **15**, 224-235

296. Aplan, P. D. (2006) *Trends Genet* **22**, 46-55

297. Yang, X. J., and Seto, E. (2007) *Oncogene* **26**, 5310-5318