DISSERTATION

Fakultät für Medizin

Institut für Medizinische Mikrobiologie, Immunologie und Hygiene

Determining the master regulator responsible for negative selection in thymocytes

Jane Beil-Wagner

Vollständiger Abdruck der von der Fakultät für Medizin der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat)

genehmigten Dissertation.

Vorsitzende(r):

Prof. Dr. Thomas Korn

Prüfer der Dissertation:

1. apl. Prof. Dr. Clarissa Prazeres da Costa

2. Prof. Dr. Bernhard Küster

Die Dissertation wurde am 23.06.2016 bei der Technischen Universität München eingereicht und durch die Fakultät für Medizin am 12.10.2016 angenommen.

DISCLAIMER

The thesis is based upon and partly adapted from the following published manuscript:

T cell-specific inactivation of mouse CD2 by CRISPR/Cas9

J. Beil-Wagner, G. Doessinger, K. Schober, J. Vom Berg, A. Tresch, M. Grandl, P. Palle, F. Mair, M. Gerhard, B. Becher, D. Busch, T. Buch

Scientific Reports 6, Article number: 21377 (2016); doi:10.1038/srep21377

Furthermore the results of this publication were incorporated into a **patent application** which is currently pending.

STATUTORY DECLARATION

I declare that I have authored this thesis independently, that I have not used other than the declared sources / resources and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

Contents

List of figures

List of tables

List of abbreviations

AIRE ..... *autoimmune regulator*
Amp ..... *ampicillin*
AmpR ..... *ampicillin resistance*
APC ..... *Allophycocyani*
ATP ..... *adenosine triphosphate*
bp ..... *base pair*
Cas ..... *CRISPR associated proteins*
CBh ..... *chicken beta actin*
cDNA ..... *complementary DNA*
CmR ..... *chloramphenicol resistance*
Cnot6 ..... *CCR4-NOT transcription complex, subunit 6*
CNS ..... *central nervous system*
Cpf1 ..... *CRISPR from Prevotella and Francisella 1*
CRISPR ..... *clustered regularly interspaced short palindromic repeats*
crRNA ..... *crispr RNA*
cTECs ..... *cortical thymic epithelial cells*
DBD ..... *DNA binding domain*
DC ..... *dendritic cells*
dCas9 ..... *nickase form of Cas9*
DKO ..... *double knockout*
DN ..... *double negative*
DNA ..... *d*eoxyribonucleic acid
DP ..... *double positive*
DSB ..... double stand breaks
DTT ..... *Dithiothreitol*
E ..... *exon*
E. coli ..... *Escherichia coli*
E2F ..... *E2F transcription factor*
ERK ..... *extracellular-signal-regulated kinase*
ES cells ..... *embryonic stem cells*
F. novicida ..... *Francisella novicida*
FACS ..... *Fluorescence Activated Cell Sorting*
FITC ..... Fluorescein *isothiocyanate*
FM ..... *follicular mature B cells*
FO ..... *follicular B cells*
fwd ..... *forward*
G1 ..... *Gap 1*
G2 ..... *Gap 2*
GRB2 ..... *growth factor receptor-bound protein 2*
Gtf2h4 ..... *general transcription factor II H, polypeptide 4*

Gtf2ird1 ..... *general transcription factor II I repeat domain-containing 1*
H&E ..... *hematoxylin and eosin*
HCG ..... *human chorionic gonadotropin*
HCL ..... hydrogen chloride
HET ..... *heterozygous*
HIF1 ..... *hypoxia inducible factor 1*
HPV ..... *human papilloma virus*
HR ..... homologous recombination
HSA ..... *heat stable antigen*
i.p. ..... *intraperitoneal*
ITAM ..... *immunoreceptor tyrosine-based activation motifs*
IU ..... *international unit*
JNK ..... c-Jun NH2-terminal kinase
KO ..... *knockout*
LAT ..... *linker for activation of T cells*
LB ..... *lysogeny broth*
Lck ..... *lymphocyte-specific protein tyrosine kinase*
LN ..... *lymph node*
LTs ..... *labyrinth trophoblasts*
Lztfl1 ..... *leucine zipper transcription factor-like 1*
M ..... *mitosis*
MAPK ..... *mitogen-activated protein kinase*
MHC ..... *major histocompatibility complex*
MIQE ..... *Minimum Information for Publication of Quantitative Real-Time PCR Experiments*
MM ..... *Mus musculus*
mRNA ..... *messenger RNA*
MS ..... *multiple sclerosis*
mTECs ..... *medullary thymic epithelial cells*
MZ ..... *marginal zone B cells*
N. gregoryi ..... *Natronobacterium gregoryi*
N. meningitidis ..... *Neisseria meningitidis*
NER ..... *nucleotide excision repair*
NFκB ..... *nuclear factor-κB*
NHEJ ..... non-homologous end-joining
NK cell ..... *natural killer cell*
NLS ..... *nuclear localization sequence*
ORI ..... *origin of replication*
pA ..... *poly A*
PAM ..... *proto-spacer adjacent motifs*
PB ..... *pacific blue*
PBS ..... *phosphate-buffered saline*
PCR ..... *polymerase chain reaction*

PE ..... R-*Phycoerythrin*
PerCP ..... *Peridinin chlorophyll protein*
PI ..... *propidium iodide*
PLCγ1 ..... *phospholipase Cγ1*
pMHC ..... *peptide/MHC*
PMSG ..... *pregnant mare's serum gonadotropin*
PNI ..... *pronuclear injection*
PNK ..... *Polynucleotide Kinase*
Pou6f1 ..... *POU domain, class 6, transcription factor 1*
Rag ..... *recombination activation gene*
RB ..... *retinoblastoma*
RBC ..... *red blood cell*
rev ..... *reverse*
RNA ..... *ribonucleic acid*
RT-PCR ..... *Real-Time PCR*
S ..... *synthesis*
S. aureus ..... *Staphylococcus aureus*
S. pyogenes ..... *Streptococcus pyogenes*
S. thermophilius ..... *Streptococcus thermophilus*
S1 ..... *safety level 1*
SLE ..... *systemic lupus erythematosus*
SP ..... *single positive*
SPF ..... *specific pathogen free*
STs ..... *spongio throphoblasts*
T. denticola ..... *Treponema denticola*
T1 ..... *transitional T1 B cells*
T2 ..... *transitional T2 B cells*
TAE ..... *Tris acetate EDTA*
TALEN ..... *Transcription Activator-like Effector Nucleases*
Tcf12 ..... *transcription factor 12*
TCR ..... *T cell receptor*
TF ..... *transcription factor*
TFIIH ..... *Transcription factor II H*
TG ..... *transgene*
TGCs ..... *trophoblast giant cells*
TRA ..... *tissue-restricted antigen*
tracr RNA ..... *trans-activating RNA*
VEGFA ..... *vascular endothelial growth factor*
WT ..... *wildtype*
ZAP70 ..... *Zeta-chain-associated protein kinase 70*
ZFN ..... *Zinc Finger Nucleases*

Abstract

Thymocytes undergo multiple selection processes on their way to becoming mature T cells. One of these processes is called negative selection and ensures that self-reactive thymocytes are eliminated. However, the molecular pathway how the TCR signal of double positive thymocytes results in negative selection in consequence of the binding to pMHC is not yet known. Therefore this thesis deals with the identification of the transcription factor(s) (TF) that modifies the apoptotic pathway in double positive thymocytes such that the TCR signal leads to negative selection instead of proliferation and differentiation in mature T cells. To achieve this goal an expressional approach was chosen and E2F7, E2F8 and Gtf2h4 were identified as potential candidates. These were analyzed *in vivo* in more detail by specific mouse models. For E2F7 and E2F8 an already existing mouse model was used to obtain T cell-specific double deficient mice. In addition, to investigate the broad range of the impact of Gtf2h4 on T cell development two mouse models were generated, showing a deficiency for Gtf2h4 and an overexpression in CD4 positive thymocytes and T cells. Astonishingly, E2F7 and E2F8 did not have a key role in neither T cell development nor negative selection. Nevertheless, the deficiency and overexpression of Gtf2h4 led to an altered cell number with reverse effect. Additionally, different thymocytes subsets as well as peripheral T and B cell populations were affected. In summary, Gtf2h4 seems to match the hypothesis and play a role in T cell development that needs to be investigated further in the future. Moreover, because the homozygous deficiency of Gtf2h4 in mice turned out to be lethal, the CRISPR/Cas9 system was modified to achieve cell type-specific gene editing in mice. Therefore, Cas9 was expressed under the CD4 promoter. To proof the concept Exon2 of CD2 in CD4 positive cells were targeted. Although, the efficiency was very low, we could show that in principle the system of conditional gene targeting by the use of CRISPR/Cas9 can be performed in mice. Further modifications will show if the targeting efficiency can be improved in the future.

Zusammenfassung

Thymozyten durchlaufen mehrere Selektionsprozesse auf dem Weg zur reifen T-Zelle. Einer dieser Prozesse wird negative Selektion genannt und gewährleistet, dass selbstreaktive Thymozyten eliminiert werden. Allerdings ist es bisher nicht klar wie der molekulare Signalweg des T-Zellrezeptors (TZR) doppelt positiver Thymozyten aufgrund der Bindung zum MHC präsentierten Eigenpeptid zu negativer Selektion führt. Demzufolge beschäftigt sich diese Arbeit mit der Identifizierung des Transkriptionsfaktors (TF), der den Apoptosesignalweg doppelt positiver Thymozyten dahingehend lenkt, dass das Signal des TZRs zu negativer Selektion und nicht zu Proliferation und Differenzierung von reifen T-Zellen führt. Um dieses Ziel zu erreichen wurde ein Expressionsansatz gewählt und E2F7, E2F8 und Gtf2h4 als potentielle Kandidaten identifiziert. Diese wurden mit Hilfe von Mausmodellen *in vivo* genauer analysiert. Für E2F7 und E2F8 wurde ein bereits existierendes Mausmodel verwendet, um einen T-Zell spezifischen Verlust dieser TFen zu erhalten. Um außerdem den Einfluss von Gtf2h4 auf die T-Zellentwicklung zu untersuchen, wurde ein Mausmodell generiert, welches einen Verlust von Gtf2h4 zeigt, sowie ein weiteres Modell, welches Gtf2h4 in CD4 positiven Thymozyten und T-Zellen überexprimiert. Erstaunlicherweise spielten E2F7 und E2F8 weder eine Rolle in der T-Zellentwicklung noch in der negativen Selektion. Im Fall von Gtf2h4 jedoch, führen der Verlust sowie die Überexpression zu einer veränderten Zellanzahl mit entgegengesetztem Effekt. Außerdem waren verschiedene Thymozytenklassen sowie periphere T- und B-Zellpopulationen beeinflusst. Zusammenfassend erscheint es, als ob Gtf2h4 die Hypothese bestätigt und eine Rolle in der T-Zellentwicklung spielt, welche in der Zukunft noch genauer untersucht werden muss. Da sich herausgestellt hat, dass der homozygote Verlust von Gtf2h4 in Mäusen letal ist, kommt hinzu, dass eine neue Methode entwickelt wurde, um das CRISPR/Cas9 System für konditionelle Geneditierung zu verwenden. Hierfür wurde Cas9 unter dem CD4 Promoter exprimiert. Um dieses Konzept zu bestätigen, wurde das Exon2 von CD2 in allen CD4 positiven Zellen gezielt verändert. Auch wenn die Effizienz sehr gering ist, konnte gezeigt werden, dass das Prinzip, das CRISPR/Cas9 System für konditionelle Genveränderung zu verwenden, in Mäusen angewendet werden kann. In der Zukunft werden weitere Optimierungen des Systems zeigen, ob die Effizienz der Genveränderung gesteigert werden kann.

1 Introduction

1.1 T cell development

The immune system of vertebrates can be classified into the innate and adaptive immune system. T cells are part of the adaptive immune system and necessary for the cell-mediated immune response and develop in the thymus. The thymus is known as a primary lymphoid organ and divided into two main structural parts: the cortex and the medulla. T cells are characterized by the expression of the T cell receptor (TCR) and evolve from lymphoid precursor cells that migrate from the bone marrow into the thymus through the cortico-medullary junction (Lind et al., 2001). Thymocytes undergo different functional checkpoints on their way to mature T cells before they start the journey from the thymus to secondary lymphoid organs. Lymphoid precursor cells entering the thymus lack CD4 and CD8 surface marker expression and are so called double negative (DN).

This maturation stage can be further subdivided into DN1-4 by the surface marker CD44 and CD25. DN1 ($CD44^+$ $CD25^-$) cells, still able to develop into αβ T cells, γδ T cells, natural killer (NK) cells, dendritic cells (DC), macrophages and B cells, enter the cortex and start to proliferate before they begin the TCRβ, TCRγ and TCRδ rearrangement during DN2 ($CD44^+$ $CD25^+$). As soon as the TCRβ chain rearrangement is successfully completed TCRβ can associate with the invariant pre-TCRα chain as well as CD3 and form the pre-TCR-complex. This process defines the maturation towards the αβ T cells and occurs during stage DN3 ($CD44^-$ $CD25^+$). The last DN stage, DN4 ($CD44^-$ $CD25^-$), is characterized by the TCRα rearrangement that later on leads to the initiation of the expression of CD4 and CD8 (reviewed in (Shah and Zuniga-Pflucker, 2014))(Godfrey et al., 1994; Godfrey et al., 1993) (Fig. 1+2). Importantly, the TCR rearrangements are regulated in two waves by Rag-1 and Rag-2 (Wilson et al., 1994).

Another way to term and discriminate the different stages of thymocyte development is related to terms and surface marker used in the B cell development. Here the DN1 stage is called Pro-T1 and is characterized by the marker $CD17^{+++}$, $CD44^{+++}$, $CD25^-$, $CD127^-$ and $CD3^-$. DN2 is equivalent to Pro-T2 ($CD17^{+++}$, $CD44^{+++}$, $CD25^{++}$, $CD127^+$, $CD3^-$), DN3 correspond to Pre-T1 ($CD17^+$, $CD44^+$, $CD25^{++}$, $CD127^{low}$, $CD3^{low}$) and DN4 matches Pre-T2 ($CD17^{low}$, $CD44^-$, $CD25^-$, $CD127^-$, $CD3^{low}$) (reviewed in (Ceredig and Rolink, 2002)). In addition, all DN stages are

characteristic for a high CD24 (HSA, J11d) expression that disappears slowly while maturation (Crispe et al., 1987).

Because of the random nature of the α and β rearrangement the TCRs are quality controlled in terms of function, self-restriction and self-tolerance. This is necessary since only an immune competent thymocyte, which attacks “foreign” but not “self”, qualifies for further maturation. The first step towards the maturation process is the positive selection process. Double positive (DP) thymocyte that express CD4 and CD8 migrate among cortical thymic epithelial cells (cTECs) in search of the possibility to bind to self-peptides presented by the major histocompatibility complex (MHC) of those cells. αβ thymocytes with a non-functional TCR, that is not able to bind to peptide/MHC (pMHC), can undergo TCRα re-rearrangement. Cells which are still not able to bind, die by neglect. Furthermore, another process, that is called negative selection or clonal deletion, takes place in the cortex or the medulla of the thymus (reviewed in (Yates, 2014)). Thymocytes that bind with a strong avidity to tissue-restricted antigen (TRA) presented by the MHC of medullary thymic epithelial cells (mTECs) or to ubiquitous proteins presented by cTECs, die by apoptosis. The expression of the TRA is mediated by the autoimmune regulator (AIRE) and the expression takes place either directly by the mTECs itself or indirectly through the transfer of TRA to dendritic cells and the presentation by them (reviewed in (Klein et al., 2011) and (Shi and Zhu, 2013)). This checkpoint, as well as the positive selection checkpoint, where the strength of the binding is characterized by an intermediate avidity, ensures that self-tolerant and self-restricted thymocytes can survive and maturate further. These stringent checkpoints lead to a survival rate of only less than 5 % of all thymocytes (reviewed in (Yates, 2014)). Afterwards, DP thymocytes downregulate one of the surface markers and become either CD4 or CD8 single positive (SP). Mature SP T cells can finally exit the thymus and migrate to secondary lymphoid organs, like lymph nodes (LN) and spleen.

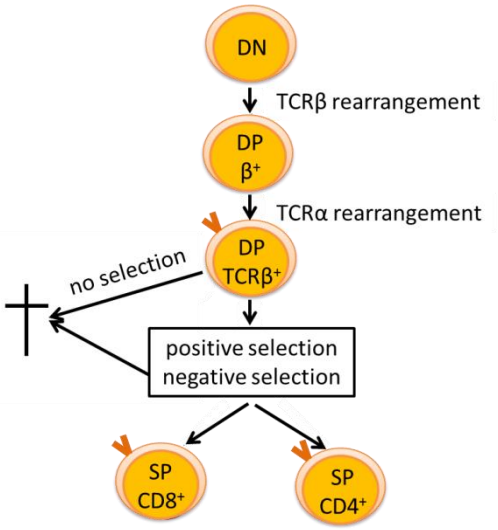


Figure 1 | Schematic illustration of T cell development in the thymus.
Lymphoid precursors enter the thymus and become double negative (DN) thymocytes. After TCRβ and TCRα rearrangement they develop into TCR expressing double positive (DP) thymocytes. Before they are fully matured they undergo positive and negative selection and finally leave the thymus into the periphery as single positive T cells that express either CD4 or CD8.

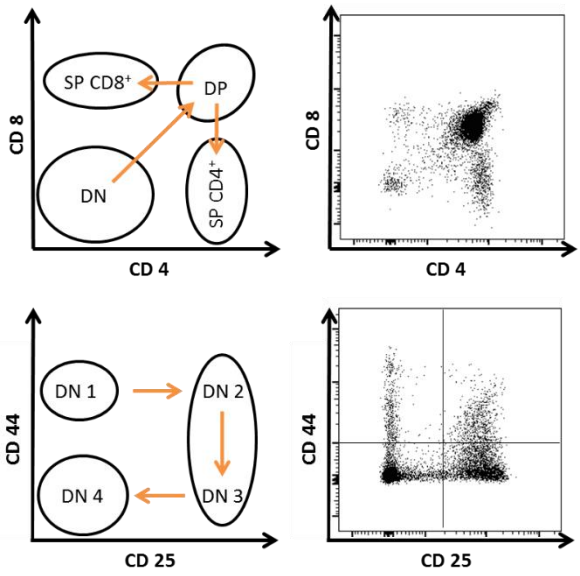


Figure 2 | Different stages of thymocyte development.
The surface marker CD4 and CD8 are used to distinguish the main stages of thymocyte development: DN ($CD4^-$ $CD8^-$), DP ($CD4^+$ $CD8^+$), SP $CD4^+$ ($CD4^+$ $CD8^-$) and SP $CD8^+$ ($CD4^-$ $CD8^+$). Depicted as schematic illustration (top left) and from flow cytometric analysis (top right). Furthermore, the DN stage can be further subdivided into DN1 ($CD44^+$ $CD25^-$), DN2 ($CD44^+$ $CD25^+$), DN3 ($CD44^-$ $CD25^+$) and DN4 ($CD44^-$ $CD25^-$). Represented as schematic illustration (bottom left) and from flow cytometric analysis (bottom right).

1.2 T cell receptor signaling and negative selection

Negative selection, which is important to eliminate self-reactive thymocytes, is controlled by the avidity of the interaction between the TCR and the pMHC. However, the exact molecular pathway(s) through which the TCR signal of DP thymocytes results in negative selection as a consequence of binding with high avidity to pMHC is not yet known. Astonishingly, in mature T cells the same TCR signal results in proliferation and differentiation. In general, the TCR binding to pMHC leads to the activation of lymphocyte-specific protein tyrosine kinase (Lck) and Fyn which are members of the Src family kinase (Ericsson and Teh, 1995; Samelson et al., 1990). Lck bound to the co-receptor CD4 or CD8 is brought into position, close to its targets. The immunoreceptor tyrosine-based activation motifs (ITAM) of TCR associated CD3γ, CD3δ, CD3ε and CD3ζ chains are phosphorylated, leading to the recruitment of the Zeta-chain-associated protein kinase 70 (Zap70) (Johnson et al., 1995; Neumeister et al., 1995; Osman et al., 1995). Subsequently, Zap70 undergoes conformational changes, due to autophosphorylation which, finally triggers its kinase activity. One of the ZAP70 target molecules is the linker for activation of T cells (LAT), where it phosphorylates four tyrosine residues which in turn activates LAT downstream signaling molecules, like phospholipase Cγ1 (PLCγ1) and growth factor receptor-bound protein 2 (GRB2). The activation of LAT downstream molecules triggers the activation of three main signaling pathways, namely the Ca^{2+}, the mitogen-activated protein kinase (MAPK) and the nuclear factor-κB (NFκB) pathways (Zhang et al., 1999a; Zhang et al., 1999b) which finally decide on the outcome of the TCR-pMHC interaction of this particular T cell (reviewed in (Brownlie and Zamoyska, 2013)).

Nevertheless, some pathway regulations and the outcomes of different stimuli are identified. It is, for instance, known, by comparing negative and positive selection, that LAT is a central molecule. The strong TCR-pMHC binding during negative selection leads to a fully phosphorylated LAT which in turn responds with a stable interaction with Grb2-SOS that provokes a strong and transient extracellular-signal-regulated kinase (ERK) activation. However, the weak interaction of the TCR to pMHC during positive selection phosphorylates LAT only partially which activates, among others, PLCγ1. This initiation leads to a recruitment of RasGRP1 to the Golgi and a slow and sustained ERK activation (reviewed in (Labrecque et al., 2011)). In addition, Zap70 plays an essential role regarding the decision of positive or negative selection. Differences in the amount of Zap70, which is regulated by the amount of

expressed TCR, result in different outcomes in which a partial inhibition of Zap70 kinase leads to a conversion of negative selection into positive selection (Mallaun et al., 2010).

Furthermore, two typical factors influencing negative selection are Bim and Nur77. Whereas the expression of Nur77 is required for negative selection (Calnan et al., 1995), it is not as clear for Bim. In the HY mouse model (Kisielow et al., 1988) it was shown that Bim is not necessary for negative selection (Hu et al., 2009). However, others investigated that TCR induced killing in thymocytes is independent of Fas pathway but requires Bim, which can be inhibited by Bcl-2 (Villunger et al., 2004). Nevertheless, it seems that there are Bim independent and dependent pathways involved in the selection process. During negative selection, Bim is required upon a strong TCR signal, whereas in case of binding to a ubiquitous self-antigen negative selection occurs independently of Bim (Suen and Baldwin, 2012).

Interestingly, for some molecules the different outcome of TCR signaling is already known. For instance, Schnurri2 (Shn2) deficiency alters the threshold of TCR induced cell death towards enhanced negative selection that can be rescued by Bim deficiency (Staton et al., 2011). Moreover, c-Jun NH2-terminal kinase (JNK) was found to be activated in MAP kinase kinase MKK7 mediated negative selection (Rincon et al., 1998). Also detailed investigations with JNK1 and JNK2 deficient mice showed that c-jun phosphorylation is reduced in $JNK1^{-/-}$ thymocytes, but not affected in mature T cells, stimulated with anti-CD3, indicating a normal activation induced T cell apoptosis but resistance in thymocyte apoptosis (Sabapathy et al., 1999; Sabapathy et al., 2001). Nevertheless, the complex differentiation of the TCR signaling during negative selection of thymocytes and the activation of mature T cells needs to be determined by further investigation.

1.3 Gtf2h4

RNA polymerases are found in all species and are required to transcribe DNA into RNA. RNA polymerases exist in three different subclasses, RNA polymerase I-III, each transcribing a different class of RNA. RNA polymerase II is essential for mRNA transcription. The first step of this process is the formation of the pre-initiation complex. This involves Polymerase II, the TATA binding protein TBP as well as the general transcription factors TFIIA, TFIIB, TFIIE, TFIIF and TFIIH. The complex then binds to the promoter and unwinds the DNA. The latter can be accomplished through TFIIH which is the only transcription factor in this complex with an

enzymatic activity (Grunberg and Hahn, 2013; Luse, 2013). TFIIH is additionally involved in the nucleotide excision repair (NER) process for DNA damage repair (Drapkin et al., 1994; Seroz et al., 1995), promoter escape (Kugel and Goodrich, 1998; Spangler et al., 2001) and cell cycle regulation (Dynlacht, 1997; Matsuno et al., 2007). It is a multi-subunit protein consisting of XPB, XPD, p52, p8, p62, p44, p34, cdk7, cyclin H and MAT1. Whereas p52, p8, p62, p44, p34 form a core complex, cdk7, cyclin H and MAT1 form a kinase complex (Marinoni et al., 1997)(reviewed in (Zhovmer et al., 2010)). XPD is known as the link between the core and kinase complex and has, in addition to XPB, an ATP-dependent helicase function (Fuss and Tainer, 2011) (Fig. 3). Furthermore, both subunits are part of the p53-mediated apoptosis pathway. They directly bind p53 at the carboxy-terminal domain which can lead to an inhibition of the TFIIH helicase activity (Wang et al., 1996; Wang et al., 1995).

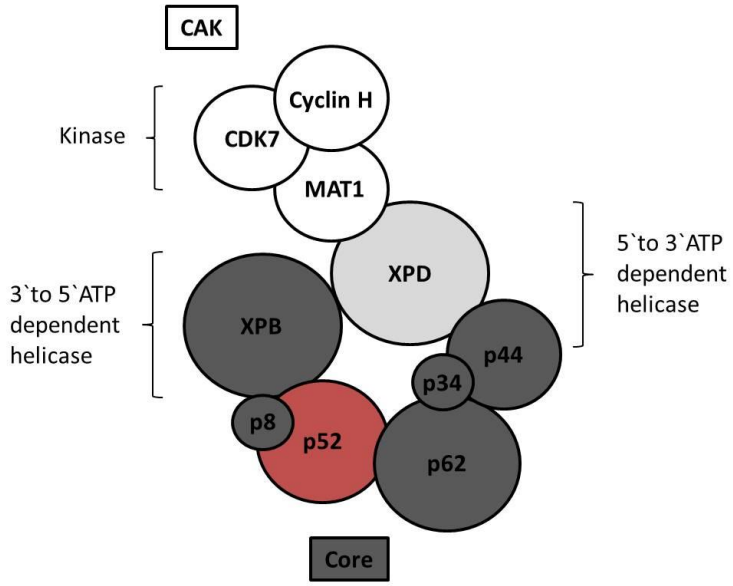


Figure 3 | Transcription factor II H (TFIIH) with its 10 subunits.
XPB, XPD, p52, p8, p62, p44, p34, cdk7, cyclin H and MAT1. cdk7, cyclin H and MAT1 form the kinase complex (CAK) (white) which is bridged via XPD (light grey) to the core complex (Core) (dark grey) consisting of p44, p34, p62, p52, p8 and XPB. p52 is also known as Gtf2h4 (red) (reviewed in (Zhovmer et al., 2010)).

The p52 subunit of TFIIH, also known as Gtf2h4, is conserved from yeast to human. Compared to the yeast counterpart TFB2, human p52 shows only a 37% sequence identity (Marinoni et al., 1997). However, the comparison of p52 from mouse and human reveals 91% nucleotide and 98% protein identity (Lanning and Lafuse, 1999). It was shown that a mutation of the C terminal region of p52 reduced the NER and transcription activity of TFIIH that led to an inhibition of the promoter opening. Additionally, p52 is anchoring XPB within the TFIIH complex, because the p34/p44/p62 subcomplex is unable to bind to XPB in

absence of p52 (Jawhari et al., 2002). Case studies and SNP analysis of Gtf2h4 showed that Gtf2h4 is upregulated in patients with Merkel-cell carcinoma, a rare eyelid tumor (Kumar et al., 2007), and that Gtf2h4 is associated with multiple sclerosis (MS) susceptibility (Briggs et al., 2010), reduced risk of lung cancer (Buch et al., 2012), HPV persistence as well as cervical cancer (Wang et al., 2010).

Interestingly, a connection between TFIIH and the E2F family, which is described in the following section, can be found. The subunit p62 is responsible for the binding of TFIIH to E2F1 and leads to its phosphorylation which is cdk7 dependent. p62 and cdk7 bind E2F1 at the retinoblastoma (RB) binding site which means that RB can inhibit and regulate the phosphorylation. Furthermore, it was shown that cdk7 expression stays the same during cell cycle but p62 specifically interact with E2F1 during the S phase and leads to its downregulation, probably through RB control (Pearson and Greenblatt, 1997; Vandel and Kouzarides, 1999). Furthermore, a link to the T cell development can be drawn by binding site analysis of the promoters of the TFIIH genes. It was shown that Ap-1 is one of the factors which have the highest binding site frequency in the sequence of Gtf2h4 of human and zebrafish (Silva et al., 2014).

1.4 E2F7 and E2F8

The E2F family of transcription factors plays an important role in apoptosis, proliferation and cell differentiation. This family consists of eight family members (E2F1-E2F8) which can be distinguished upon different characteristics. E2F1, E2F2 and E2F3a are known to be activators whereas E2F3b, E2F4, E2F5, E2F6, E2F7 and E2F8 belong to the group of repressors. Another way to describe these members is according to their structural characteristics. E2F1-6 are typical members because they consist of an N-terminally located DNA binding domain (DBD) which is followed by a dimerization domain. In contrast, E2F7 and E2F8 are atypical family members as they have a duplicated DBD, bind in a dimerization pattern and independent of RB (de Bruin et al., 2003; Di Stefano et al., 2003; Lammens et al., 2009; Logan et al., 2004; Logan et al., 2005). E2F7 and E2F8 are the most recent discovered family members and share not only structural but also functional similarities. During cell cycle they are both induced when it comes to G1 to S transition and their expression increases in the S phase of the cell cycle (Di Stefano et al., 2003; Endo-Munoz et al., 2009).

This means that an overexpression of E2F7/8 leads to an accumulation of cells in G2 phase and decreased proliferation potential (de Bruin et al., 2003; Maiti et al., 2005).

E2F7 is located on mouse chromosome 10 and human chromosome 12 and shows the highest expression in skin and thymus whereas no or less expression is detected in brain, muscle and stomach (de Bruin et al., 2003). E2F8 is also highly expressed in skin and thymus in addition to the liver and testis. It is found on mouse chromosome 7 and human chromosome 11 (Christensen et al., 2005; Maiti et al., 2005). Interestingly, E2F7 and E2F8 are able to form homo- and heterodimers, with the formation of E2F7 homodimers being preferred over E2F7/E2F8 heterodimers. The less preferred formation consists of E2F8 homodimers (de Bruin et al., 2003). Analysis of E2F7 and E2F8 knockout mice showed their importance in embryonic development and apoptosis. Double knockout (DKO) mice are lethal by embryonic day E11.5 because of massive apoptosis induced by an increased E2F1 and p53 expression (Li et al., 2008). Furthermore, these mice show vascular defects as the loss of E2F7/8 prevents the complex formation with hypoxia inducible factor 1 (HIF1) which is important for transcriptional activation of VEGFA (Weijts et al., 2012). In addition, it was shown that E2F7/8 are highly expressed in placental tissue, more precisely in three major trophoblast lineages: labyrinth trophoblasts (LTs), spongio throphoblasts (STs) and trophoblast giant cells (TGCs). The placentas of DKO mice are smaller than normal and trophoblasts are not able to enter maternal decidua. Interestingly, this effect can be reversed in E2F3a, E2F7 and E2F8 triple KO mice. A comparison of E2F3a KO and E2F7/8 DKO mice indicates that they share more than 90% of their target genes but expressed in the opposite direction. Structural analysis confirmed this antagonistic modulation because of an E2F3a binding site within the promoter of E2F7 and E2F8 (Ouseph et al., 2012).

1.5 Genetically modified mice and the CRISPR/Cas9 System

The mouse has become an essential animal model in immunological research not least because its genome is highly homologous to the human genome (Waterston et al., 2002) and easy to manipulate. The first genetically modified mice were generated in the earliest 80s by microinjections of plasmids into fertilized oocytes (Brinster et al., 1981; Gordon and Ruddle, 1981; Gordon et al., 1980; Palmiter et al., 1982) which was later on improved by site-directed gene targeting in mouse embryonic stem cells (ES cells) (Thomas and Capecchi, 1987). To study the function of one specific gene in certain cell types or organs, the methods

of choice is the generation of conditional KO mice by the use of the Cre/loxP recombination system (Schwenk et al., 1998). The Cre recombinase is a type I topoisomerase from P1 bacteriophage and can be used to recognize the 34 bp loxP sites. Cre expressed by a specific promoter leads to the deletion or inversion of the allele flanked by loxP sites in this particular cell type or tissue (Hoess et al., 1982; Schwenk et al., 1998; Sternberg and Hamilton, 1981).

The induction of a complete gene deletion has become much easier during the last years. Designer nucleases like Zinc Finger Nucleases (ZFN) (Lee et al., 2010; Meyer et al., 2010; Sollu et al., 2010), Transcription Activator-like Effector Nucleases (TALEN) (Cermak et al., 2011; Zhang et al., 2011) and CRISPR/Cas9 (Cong et al., 2013; Jinek et al., 2013; Mali et al., 2013) can be injected into fertilized oocytes and introduce specific double stand breaks (DSB) in the target sequence of the genome which is followed by a repair through either non-homologous end-joining (NHEJ) or homologous recombination (HR). HR leads to integration of a co-injected plasmid or ssDNA containing homology arms on both sites and NHEJ can lead to gene deficiencies because of deletions or insertions (Fig. 4). Among these three designer nucleases the CRISPR/Cas9 system became the method of choice because it is very easy to work with. Furthermore, it is the faster approach. Only recently, it was shown that the procedure can be simplified by using oocytes from a transgenic mouse expressing Cas9 oocyte-specific under the Zp3 promoter (Zhang et al., 2016). In addition, instead of pronuclear injection (PNI) electroporation of the oocytes with Cas9 and gRNA should be considered (Chen et al., 2016; Qin et al., 2015).

Originally, the CRISPR/Cas9 system was discovered in *E. coli* in 1987 (Ishino et al., 1987), however it is also found in 37% *eubacteria* and in 90% *archaea*. CRISPR/Cas stands for clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR associated proteins (Cas) (Koonin and Makarova, 2009). In prokaryotes it is part of the adaptive immune system and responsible for the protection against foreign genetic elements by silencing. This process involves three steps: adaptation, expression and interference. Adaptation means that short sequences of foreign DNA from viruses or plasmids are taken up and are acquired as new proto-spacers at the end of the CRISPR array. The proto-spacers that vary between 26 bp and 72 bp are separated by repeats of 21 to 48 bp. Subsequently, in case of a repeated occurrence, the CRISPR array is transcribed by RNA polymerase as a pre-crRNA and maturates successively into crRNA. Finally, during the interference stage the crRNA guides

the Cas protein to the plasmid or bacteriophage DNA to form a complex which leads to targeted cleavage of the foreign nucleic acids (Bhaya et al., 2011; Doudna and Charpentier, 2014; Garneau et al., 2010; Horvath and Barrangou, 2010). Additionally, it was shown that CRISPR3/Cas from *S. thermophilus* can be transferred to *E. coli* and leads to specific CRISPR mediated immunity against transformed plasmids. Furthermore, proto-spacer and the proto-spacer adjacent motifs (PAM) are required for a functional immunity, whereas some single nucleotide mismatches at certain positions are tolerated to achieve a functional binding (Sapranauskas et al., 2011).

Until now ten major Cas proteins has been classified (Cas1 – Cas10) whereof Cas1 and Cas2 are universal Cas proteins that can be found in all three CRISPR/Cas systems: Type I, II and III. The other Cas proteins are either unique or typical for more than one type. The Type II system is the most interesting as it allows specifically targeted gene modification. It is found only in bacteria and can be distinguished into Type IIA and Type IIB. The Cas9 protein is a representative of this type of CRISPR/Cas system and contains two endonuclease domains. The RuvC-like nuclease is located at the N-terminus and the HNH domain in the center of the protein (Bhaya et al., 2011; Doudna and Charpentier, 2014). For a successful cleavage, crispr RNA (crRNA), trans-activating RNA (tracrRNA) and Cas9 are required. The crRNA pairs with the proto-spacer DNA or target DNA, in terms of targeted genetic modification, which leads to a Cas9 binding via the tracrRNA. The PAM is necessary for the target binding that enables Cas9 to separate the DNA strand and form the R-loop. Subsequently, the complementary DNA strand is cleaved three base pairs upstream of the PAM by the HNH domain whereas the RuvC-like domain cuts the non-complementary DNA strand at one or more sites within three to eight base pairs upstream of the PAM and is successively clipped by a 3´-5´exonuclease activity. The system can be minimized by the fusion of crRNA and tracrRNA to one single guide RNA (gRNA) (Jinek et al., 2012) (Fig. 4) .

In the beginning of 2013, the first results were published which showed that CRISPR/Cas9 originating from *Streptococcus pyogenes* can be used to specifically mutate DNA in human cell lines as well as induced pluripotent stem cells in a multiplex manner. Cas9 was codon-optimized and crRNA and tracrRNA used as one single gRNA. Furthermore, it was shown that the PAM sequence NGG is indispensable for the target binding. Interestingly, a nickase form of Cas9 (Cas9D10 or dCas9) could increase the HR/NHEJ ratio (Cong et al., 2013; Jinek et al., 2013; Mali et al., 2013). Since then, the system was, among others, used to target genomic

DNA in zebrafish (Chang et al., 2013; Hwang et al., 2013), *Saccharomyces cerevisiae* (DiCarlo et al., 2013), *Caenorhabditis elegans* (Cho et al., 2013; Friedland et al., 2013; Katic and Grosshans, 2013), plants (Jiang et al., 2013; Nekrasov et al., 2013; Xie and Yang, 2013), *Drosophila melanogaster* (Bassett et al., 2013; Gratz et al., 2013; Yu et al., 2013), *Xenopus tropicalis* (Nakayama et al., 2013), rats (Hu et al., 2013; Ma et al., 2014), mice (Wang et al., 2013; Yang et al., 2013) and monkeys (Niu et al., 2014). Moreover, some adaptations were made to modulate the system for more specific applications. For instance, dCas9 is used to specifically activate or repress gene expression when fused to a transcriptional effector domain (Cheng et al., 2013; Gilbert et al., 2013; Qi et al., 2013). Furthermore, dCas9 containing an antibody can be applied for purification (Fujita and Fujii, 2013) or dCas9 fused to EGFP for imaging purpose (Chen et al., 2013). Additionally, Cas9 also exists in other species than *S. pyogenes.* Using Cas9 from another species has the advantage of different PAM recognition sites and extended variability in experimental design. Cas9 from *S. aureus* recognizes NNGRRT or NNGRR(N) (Ran et al., 2015), *N. meningitidis* binds *to* NNNNGATT (Hou et al., 2013), *S. thermophiles* targets NNAGAAW (Esvelt et al., 2013; Jinek et al., 2012) and *T. denticola* identifies the PAM NAAAAC (Esvelt et al., 2013).

Moreover, just recently two additional enzymes were published. Cpf1 (CRISPR from Prevotella and Francisella 1) is another single RNA guided endonuclease of the type II system and needs only a crRNA of ~42 bp to specifically bind target DNA. Cpf1 cuts, in contrast to Cas9 that results in blunt ends, with an 5` overhang which might be an enhancement for homologous recombination. In terms of the specific experimental design, it could be also useful that Cpf1 from *F. novicida* recognizes two different PAMs, TTN or CTA (Zetsche et al.). Additionally, besides the already mentioned RNA guided endonucleases there is also a DNA guided enzyme from *N. gregoryi,* Argonaute. Argonaute needs a 5’ phosphorylated-ssDNA of ~24 bp to specifically target DNA and cuts without the requirement of a PAM sequence that successfully leads to NHEJ and HR (Gao et al., 2016).

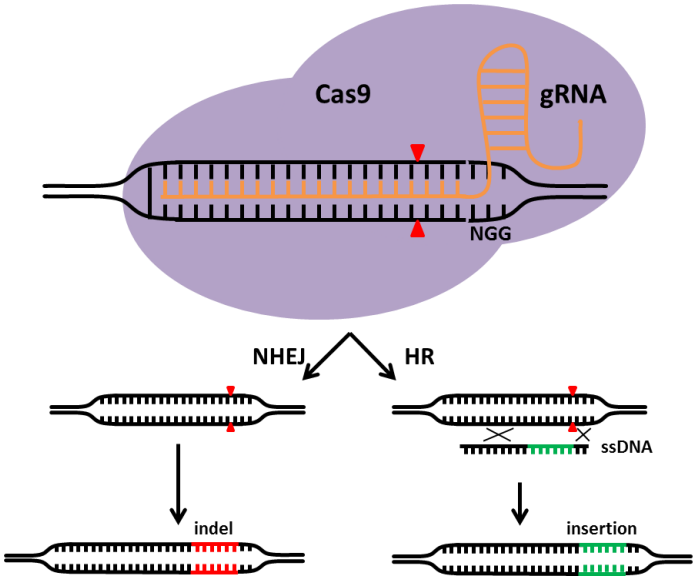


Figure 4 | Schematic illustration of gRNA and Cas9 binding to the target DNA and the possible outcomes.
The gRNA (orange) recognizes a 20 bp target sequence that leads to the binding of *S. pyogenes* Cas9 (purple). This can either induce NHEJ (left site) or HR (right site) if single-stranded DNA (ssDNA) is provided. NHEJ results in indel and HR in the insertion of the ssDNA. Red arrows indicate the cutting site.

Table 1 | Overview of newly available endonuclease tools and their characteristics.

Tool	Species	Component 1	Component 2	Recognition site	Function	Reference
CRISPR/Cas9	*S. pyogenes*	gRNA	Cas9	NGG	cutting	(Cong et al., 2013; Jinek et al., 2012; Mali et al., 2013)
CRISPR/Cas9	*S. aureus*	gRNA	Cas9	NNGRRT or NNGRR(N)	cutting	(Ran et al., 2015)
CRISPR/Cas9	*N. meningitidis*	gRNA	Cas9	NNNNGATT	cutting	(Hou et al., 2013)
CRISPR/Cas9	*S. thermophilus*	gRNA	Cas9	NNAGAAW	cutting	(Esvelt et al., 2013; Jinek et al., 2012)
CRISPR/Cas9	*T. denticola*	gRNA	Cas9	NAAAAC	cutting	(Esvelt et al., 2013)
CRISPR/Cas9	*S. pyogenes*	gRNA	dCas9	NGG	nicking	(Cong et al., 2013; Mali et al., 2013)
CRISPR/Cas9	*S. pyogenes*	gRNA	dCas9-FokI	NGG	cutting, nicking	(Guilinger et al., 2014; Tsai et al., 2014)
CRISPR/Cas9	*S. pyogenes*	gRNA	dCas9-EGFP	NGG	imaging	(Chen et al., 2013)
CRISPR/Cas9	*S. pyogenes*	gRNA	dCas9-effector domain	NGG	activation/ repression	(Cheng et al., 2013; Gilbert et al., 2013; Qi et al., 2013)
CRISPR/Cas9	*S. pyogenes*	gRNA	dCas9-CIBN	NGG	light inducible gene activation	(Nihongaki et al., 2015; Polstein and Gersbach, 2015)
CRISPR/Cas9	*S. pyogenes*	gRNA	dCas9	NGG	purification	(Fujita and Fujii, 2013)
Cpf1	*F. novicida*	crRNA	CPF1	TTN or CTA	cutting	(Zetsche et al.)
Argonaute	*N. gregoryi*	5′ phosphorylated-ssDNA	Ago	none	cutting	(Gao et al., 2016)

1.6 Specific aim of the thesis

1st Project

Determining the master regulator responsible for negative selection in thymocytes

The aim of the first project of the thesis was to identify one or more transcription factor (TF) that are able to connect the TCR signaling with the apoptotic process during T cell development in the thymus without enabling apoptosis in mature T cells. This included carrying out the following analyses:

- Identification of one or more TF highly expressed at the DP stage and less expressed at DN and SP stages
- Verify the expression profile of these TFs by RT-PCR
- Investigate the role of the selected candidates in T cell development and negative selection *in vivo* in the TF deficient mouse model
- Investigate the role of the selected candidates in T cell development and negative selection *in vivo* in the TF overexpression mouse model

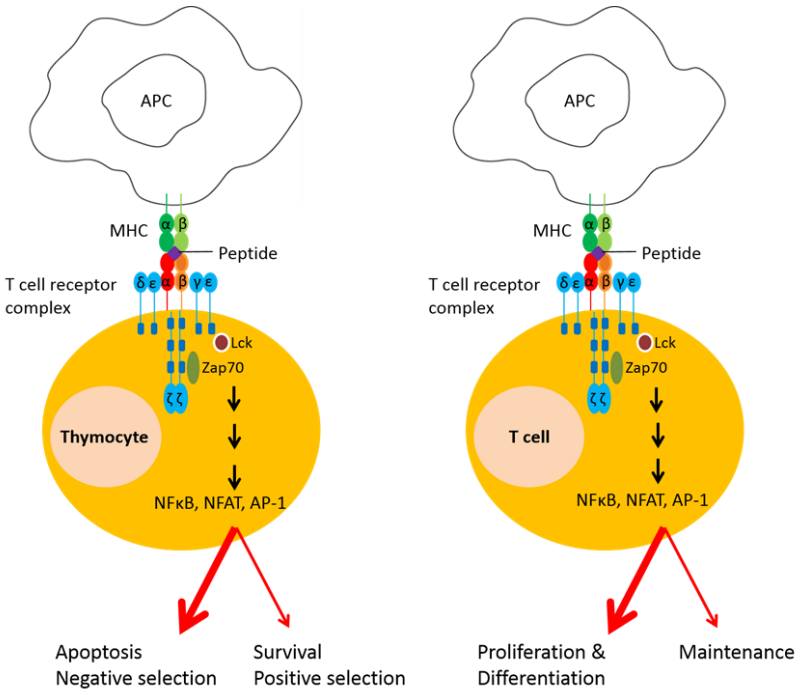


Figure 5 | Schematic illustration of the reverse outcome of the TCR signaling in thymocytes and mature T cells.
TCR signaling in thymocytes (left) and mature T cells (right) leads to reverse outcomes. In thymocytes a strong TCR (thick red arrow) signal induces apoptosis or negative selection whereas in mature T cells it leads to proliferation and differentiation. A weak TCR signal (thin red arrow) induces survival and positive selection in thymocytes and maintenance in mature T cells.

2nd Project

T cell-specific inactivation of mouse CD2 by CRISPR/Cas9

During the analyses of the first project, the Gtf2h4 deficient mouse model was found to be embryonically lethal. Consequently, we wanted to modify the CRISPR/Cas9 system to target a gene of interest in a cell type-specific manner. This has not been shown before in mice. The following steps were carried out:

- Design of a useful concept that allows the proof of concept without major changes in the phenotype of the mouse
- The generation of a double transgenic mouse that expressed Cas9 under a cell type-specific promoter as well as the gRNA specific to the target exon under the U6 promoter.
- Identification and characterization of cell-specific deficiency by single cell sorting, FACS analysis and sequencing

The methods and results of this subproject can be found in the publication Beil-Wagner et al. 2016 which is attached to this thesis.

2 Material and Methods

2.1 *In silico* analysis

The Immgen (Heng and Painter, 2008) analysis was performed by the microarray core facility of the Institute of Medical Microbiology, Immunology and Hygiene, Technische Universität München in 2011. The raw Immgen datasets (Painter et al., 2011) of the T cell stages DN2, DN3, DP, DPbl, DPsm, 4SP24, 8SP24, 4nve and 8nve were downloaded from NCBI's GEO data repository (GSE15907) (Edgar et al., 2002). The quality control box plot of the robust multi-array average (Irizarry et al., 2003) normalized data showed that the datasets are comparable and could be used without correcting for a batch-effect. Significantly expressed genes were selected by using a cut-off false discovery rate of 0 calculated by the significance analysis of microarrays (Tusher et al., 2001). The cluster analysis revealed a cluster of genes showing an up-regulation in the DP stages and a downregulation in DN2, DN3, 4SP24, 8SP24, 4nve, and 8nve. From this cluster transcription factors were selected for further analysis.

2.2 Real-Time PCR

Real Time (RT)-PCR was performed using either total organ RNA or RNA from sorted thymocytes or T cells (see 2.5.4). All organs were treated as described in 2.4.1. RNA was isolated with peqGOLD Trifast (Peqlab) or RNeasy Micro/Mini Kit (Qiagen) according to manufacturer's instructions. cDNA was transcribed with iScript cDNA Synthesis Kit (BioRad) and SsoFast EvaGreen Supermix (BioRad) was used to perform RT-PCR on the CFX384 Touch Real-Time PCR Detection System (BioRad) or QuantStudio 7 Flex Real-Time PCR System (ThermoFisher Scientific). The primer pairs (Tab. 2) were all designed to be used at an annealing temperature of 60°C and 45 cycles were performed. According to the MIQE Guidelines (Bustin et al., 2009) the RT-PCR was performed with triplets in 3 different experiments. Furthermore, 2 reference genes were used to calculate the $\Delta\Delta C_T$ and the fold change ($2^{-\Delta\Delta CT}$). The analysis was performed using Microsoft Excel and GraphPad Prism 6.

Table 2 | RT-PCR primer and their sequences.

Oligo Name	Sequence (5'-3')
RT-CD4 for	aga act ggt tcg gca tga ca
RT-CD4 rev	aag gag aac tcc gct gac
RT-CD4Gtf2h4 fwd	aga agc aga gtg aag gaa gg
RT-CD4Gtf2h4 rev	cat tgt agg tgt gct cgg t
RT-CD8 for	cgt gcc agt cct tca gaa
RT-CD8 rev	tcc cct tca ctg agc cac
RT-Cnot6 for	gga gtt gca gta ctg cta gaa ctt cga a
RT-Cnot6 rev	tgc ata tga gca tta gcc acg aga ata
RT-Cxxc1 for	ggt tgt tgc acg ggg tcc ag
RT-Cxxc1 rev	ccc cat tct cag act tgc tgt cg
RT-E2F2 for	gag cag aca gtg att gcg gtc aag
RT-E2F2 rev	tgg ggc ctt ggg tac tct tta gat aaa
RT-E2F7 for	ccg gag aga ccc tcc atc
RT-E2F7 rev	gca tct tca atg gca aaa tct
RT-E2F8 for	tca gcc caa aca aca gtg g
RT-E2F8 rev	gca gac tgc tca gcc tct aag
RT-Gtf2h4 for	ctg ggg tgc tgg acc gat tgt a
RT-Gtf2h4 rev	cca gaa aga gca tcc tca tga ccc
RT-Gtf2ird1 for	tgc ctg agc ccc atg tcc c
RT-Gtf2ird1 rev	ttc tgg atg aga ggc cca gca g
RT-Lztfl1 for	aac aag ttg ctg aat ttg aaa agg cag
RT-Lztfl1 rev	tgt tta gga gct ctg tag ttc cac ctt ca
RT-Nr1h2 for	ctt gca gtt ggg ccg gga
RT-Nr1h2 rev	acg tga tgc att ctg tct cgt ggt
RT-Pou6f1 for	tca gct gca act ccc atc cca
RT-Pou6f1 rev	tcc cgt cct cat cca gac ttg g
RT-Tcf12 for	cca gta gtt atg gca gcc t
RT-Tcf12 rev	cgg tgg aag ctg gac att
RT-Ywhaz for	gtt act tgg ccg agg ttg ctg ct
RT-Ywhaz rev	ggt gtg tcg gct gca tct cct t
RT-Zfp703 for	tcc aca ccc gtc agc cca att
RT-Zfp703 rev	ttg agc ttg gat gag ggc gg

2.3 Animal models

2.3.1 Vector cloning

2.3.1.1 Specific cloning material

Table 3 | Cloning enzyme list.

Enzyme	Supplier
AvaI	NEB
BbsI	NEB
NotI	NEB
PstI	NEB
SalI	Promega
XbaI	NEB

Table 4 | Cloning vector list.

Plasmid name	Origin/Supplier
pGEM-T Easy	Promega
CD4 Depe	gift from Marc Schmidt-Supprian
px330	gift from Feng Zhang (Addgene # 42230) (Cong et al., 2013)

2.3.1.2 px330 (gRNA Gtf2h4) vector design and cloning

The Crispr design tool at crispr.mit.edu was used to identify the protospacer specific for Gtf2h4. The gRNA was cloned into the pX330-U6-Chimeric_BB-CBh-hSpCas9 (px330) (a gift from Feng Zhang (Addgene plasmid # 42230)) (Cong et al., 2013) or px330 ccdB (see 2.3.1.3) (Fig. 6) and the resulting clones were verified for correct cloning. Finally, the plasmids were injected as circular plasmids into the pronucleus of fertilized oocytes.

2.3.1.3 px330 ccdB vector cloning

The CmR/ccdB cassette was amplified by PCR using primers suitable for Gibson cloning (Gibson et al., 2009). The px330 vector was digested and gel-purified. The PCR amplified CmR/ccdB cassette was then cloned into the px330 vector (Fig. 6). Finally, the colonies were verified for correct cloning.

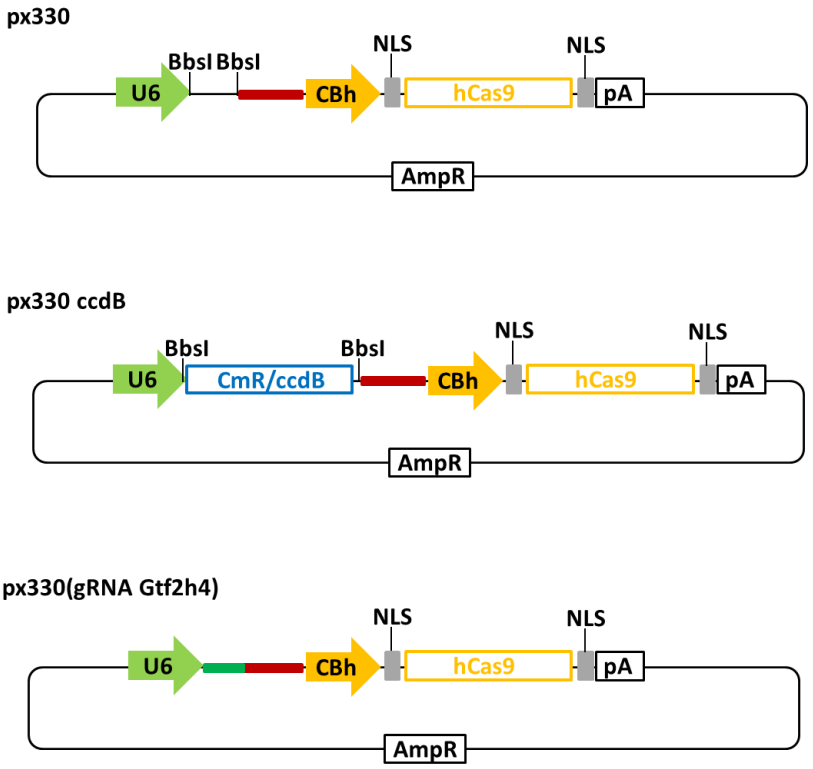


Figure 6 | Schematic illustration of the constructed vector px330, px330 ccdB and px330(gRNA Gtf2h4).
px330: U6 propmoter (green arrow), two BbsI recognition sites, chimeric guide RNA (solid red bar), chicken beta actin (CBh) promoter (yellow arrow), humanized (h) Cas9 flanked by two NLS (nuclear localization sequence) (grey rectangle), pA (poly A), AmpR (ampicillin resistance); px330 ccdB: px330 + chloramphenicol resistance and cytotoxic protein ccdB (CmR/ccdB, blue box); px330(gRNA Gtf2h4): px330 + 20 bp protospacer specific for exon 2 of Gtf2h4 (solid green bar).

2.3.1.4 CD4Gtf2h4 vector cloning

Total thymic wildtype C57Bl/6JOla RNA was transcribed into cDNA and the cDNA of Gtf2h4 was amplified using primers, which contain the Kozak sequence (Kozak, 1984) and two additional SalI digestion sites, one at each end. The PCR product was cloned into the pGEM-T Easy vector and verified by sequencing. The cDNA was then amplified from the pGEM-T Easy Gtf2h4 with primers suitable for Gibson cloning to clone the Gf2h4 PCR product into the CD4 Depe vector (gift from Marc Schmidt-Supprian). This vector consists of the CD4 promoter, the distal and the proximal enhancer, exon 1 and parts of exon 2 but lacking the intronic silencer (Sawada et al., 1994). After amplification, the appropriate band was gel-purified and cloned into the digested CD4 Depe vector (Fig. 7). The CD4Gtf2h4 plasmids were isolated from the colonies and verified by restriction digest and sequencing. To perform the pronuclear injection (PNI) the CD4Gtf2h4 vector had to be linearized.

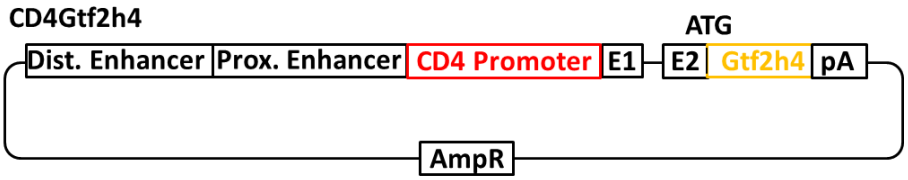


Figure 7 | Schematic illustration of the CD4Gtf2h4 vector.
The vector contains the distal and the proximal enhancer, the CD4 promoter (red), exon 1 (E1), parts of exon 2 (E2) and Gtf2h4 (yellow). AmpR (ampicillin resistance), pA (poly A), ATG = start codon

2.3.2 Transformation of competent cells

DNA was added to chemically competent DH5α or Top10 *E. coli* cells. The cells were then incubated on ice for 15 min followed by a 90 sec heat shock at 42°C and subsequent incubation on ice for 90 sec. Afterwards, 300 µl LB medium (lysogeny broth (Bertani, 1951)) were added to the cells and incubated on a shaker at 37°C for 1 h. The cells were pelleted by centrifugation at 300 g for 3 min and 200 µl supernatant were discarded. The remaining 100 µl were resuspended and plated on LB agar plates containing the appropriate selection antibiotics and incubated overnight at 37°C.

2.3.3 Plasmid preparations

For plasmid preparations 5 ml (mini) or 200 ml (midi) of LB medium containing the selecting antibiotic was inoculated with one clone or approx. 100 µl of a pre-culture. The plasmid purification was performed with the PureYield™ Plasmid Mini/Midiprep System (Promega) according to the manufacturer's instructions.

2.3.4 Restriction digest

Restriction digests were performed with 1 unit of enzyme per 1 µg of DNA in the corresponding buffer and a total volume of 20-100 µl. The digestion was performed at required temperature for 1 h – overnight, depending on the amount of DNA.

2.3.5 Agarose gel electrophoresis

The size of the DNA fragments was analyzed by agarose gel electrophoresis. Agarose was dissolved in 1x TAE buffer (40 mM Tris acetate, 0.05 M EDTA, pH 8.2-8.4) by heating in a microwave. 0.01 % GelRed (Biotium) or Roti-Safe (CarlRoth) were added. Were necessary 6x DNA loading buffer (Promega, Fermentas) was added to samples prior the loading. Gels were run for 1-2 hours at 60-100 mV. Gel images were taken with the Eagle Eye imaging system (BioRad) using the Quantity One Software (Version 4.6.1) or EXview CCD camera (Intas) and Intas GelDoc. 1kb ladder (Promega, Fermentas) was used as a size marker.

2.3.6 Pronuclear injection

The pronuclear injection (PNI) was performed by the transgenic core facility of the Institute of Medical Microbiology, Immunology and Hygiene, Technische Universität München. After superovulation of FVB/N mice with PMSG (5IU, i.p.) at 01:00 pm and HCG (5IU, i.p.) 47 h later the animals were mated. Next day the oviducts were dissected and zygotes were removed. For the generation of transgenic mice 1-2 µg/ml of the appropriate linearized plasmid were injected into the pronucleus of the fertilized oocytes. To obtain knockout (KO) mice with the CRISPR/Cas9 technique 5 ng/µl of each circular plasmid were injected into the pronucleus. The plasmids were purified with Qiaex II Gel Extraction Kit (Qiagen) or precipitated with ethanol before use and diluted in injection buffer (0.5 M EDTA, 1M HCL, 1M Tris-HCL) for injection. Afterwards, the founders were screened by PCR.

2.3.7 Mouse maintenance

Animals were housed in a barrier-SPF or S1 level animal facilities at the Institute of Medical Microbiology, Immunology and Hygiene, Technische Universität München and the Institute for Laboratory Animal Sciences, University of Zurich according to German and Swiss Animal Protection laws. Animal experiments were performed under the license number 55.2-1-54-2532-2-12 (Germany) and ZH141/15 (Switzerland).

2.3.8 Genotyping

For genotyping tail biopsies or toes were lysed in 200 µl DirectPCR Lysis Reagent Tail (Peqlab) substituted with 0.2 – 0.4 mg/ml proteinase K. The following tables show the primer sequences (Tab. 5), the components of the PCR mix (Tab. 6) and the expected PCR product sizes (Tab. 7) for all genotyping protocols. The PCRs were performed on the C1000 Touch and S1000 Thermal Cycler (BioRad).

Table 5 | Genotyping primer list with corresponding sequences and applied PCR mixes.

Oligo Name	Sequence (5'-3')	Mix
CD4-cre fwd	ccc aac caa caa gag ctc aag g	B
CD4-cre rev	ccc aga aat gcc aga tta cg	B
E2F7 fwd	cag tac ttt tgg gac aga gg	A
E2F7 rev	gtg ttt aca gct gtg caa gg	A
E2F8 fwd	aat tgg tga gcc cca ggt cc	A
E2F8 rev	gca gga aag gag gct agc ag	A
lck-cre fwd	cct cct gtg aac ttg gtg ctt gag	B
lck-cre rev	tgc atc gac cgg taa tgc ag	B

Table 6 | PCR mix A and B and the corresponding master mix components.

	Mix A (in µl)	Mix B (in µl)
GoTaq®Green Master Mix (Promega)	12.5	12.5
Primer fwd (10 mM)	0.25	0.5
Primer rev (10 mM)	0.25	0.5
H_2O	11.9	10.5
Template DNA	0.1	1

Table 7 | Expected PCR product sizes of WT and KO or TG and the corresponding PCR primer name.

PCR	Primer fwd	Primer rev	Product size WT	Product size KO/TG
E2F7	E2F7 fwd	E2F7 rev	318 bp	362 bp
E2F8	E2F8 fwd	E2F8 rev	189 bp	226 bp
Lck-cre	lck-cre fwd	lck-cre rev	-	170 bp
CD4-cre	CD4-cre fwd	CD4-cre rev	-	600 bp

2.3.8.1 Lck-cre/CD4-cre E2F7$^{fl/fl}$ and E2F8$^{fl/fl}$

To perform the genotyping of the Lck-cre/CD4-cre E2F7$^{fl/fl}$ and E2F8$^{fl/fl}$ mouse strain four different PCR had to be performed (see primer sequences, master mixes and band sizes above). The following tables show the applied temperature protocols.

Table 8 | **Temperature protocol of the genotyping PCR Lck-cre.**

Lck-cre

Step	Temp. (°C)	Time (mm:ss)
1	94	04:00
2	94	00:30
3	57	00:40
4	72	00:40
2-4	35 cycles	
5	72	10:00
6	12	∞

Table 9 | **Temperature protocol of the genotyping PCR CD4-cre.**

CD4-cre

Step	Temp. (°C)	Time (mm:ss)
1	95	03:00
2	94	00:45
3	62	00:45
4	72	01:00
2-4	35 cycles	
5	72	05:00
6	12	∞

Table 10 | Temperature protocol of the genotyping PCR E2F7.

E2F7

Step	Temp. (°C)	Time (mm:ss)
1	95	04:00
2	95	00:30
3	56.5	00:30
4	72	00:30
2-4	35 cycles	
5	72	07:00
6	12	∞

Table 11 | Temperature protocol of the genotyping PCR E2F8.

E2F8

Step	Temp. (°C)	Time (mm:ss)
1	95	04:00
2	95	00:20
3	56.4	00:30
4	72	00:30
2-4	35 cycles	
5	72	07:00
6	12	∞

2.4 Lymphocyte isolation

2.4.1 Lymphocyte isolation from lymphatic organs

Spleen, lymph nodes (LN) and thymi were removed from dead mice and transferred into ice-cold 1x PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4). The organ was crushed with the plunger of a syringe and filtered through a 70 µm cell strainer (Falcon). This was followed by a centrifugation step for 5 min at 300 g. Afterwards, the supernatant was discarded. In case of the spleen, the erythrocytes of the pellet were lysed with 1 ml 1x RBC lysis solution (4.15 g NH_4Cl, 0.55 g $KHCO_3$, 0.185 g EDTA filled up to 500 ml H_20, pH 7.4) per spleen for 3 min at room temperature. Subsequently, the tube was filled up with PBS, filtered again through a 70 µm cell strainer and centrifuged with the same conditions as described above. Finally, the pellet was resuspended in PBS or FACS buffer (PBS with 2% FCS and 0.01% NaN_3), depending on the following application.

2.4.2 Cell counting

Cell counting was conducted with either a Neubauer counting chamber or the Cellometer Auto1000 (Nexcelom Bioscience). For the former, 90 µl of the cell suspension were mixed with 10 µl trypan blue. Approximately 10 µl of the solution was pipetted under the coverslip of the counting chamber and the viable cells of four big quadrants or at least 100 cells were counted. The total cell number was calculated according to following formula: (counted cell number/numbers of big quadrant) x dilution factor x 10^4 = numbers of cells in ml. The automatic cell counting is performed with 20 µl cell solution (1 in 2 in trypan blue) loaded into the Cellometer Disposable Chambers (Nexcelom Bioscience).

2.5 Flow cytometry

2.5.1 Antibody surface staining

The pelleted cells were resuspended in 20-30 µl of the respective antibody (Tab. 12) diluted in FACS buffer and incubated for at least 20 min at 4°C in the dark. Incubation was followed by a washing step with PBS and subsequently the cells were centrifuged for 5 min at 300 g and 4°C. The supernatant was discarded and the cells were resuspended in FACS buffer. The measurement was carried out on a CANTO II (BD Biosciences) or LSR Fortessa (BD Biosciences) flow cytometer using the FACSDiva (BD Biosciences) software. Live/Dead discrimination was performed by propidium iodide (PI) or Live Dead Fixable Aqua Dead Cell Stain Kit (Life Technologies). For the analysis FlowJo 9.4 and 10.0.8 software (TreeStar) were used.

Table 12 | **Flow cytometric staining antibodies, their color, clone and corresponding supplier.**

Marker	Color	Clone	Supplier
B220	APC/Cy7	RA3-6B2	BioLegend
CD11b	PB	M1/70	BioLegend
CD11b	PB	M1/70	BioLegend
CD11c	PE/Cy7	N418	BioLegend
CD19	BP	6D5	BioLegend
CD19	FITC	6D5	BioLegend
CD19	APC/Cy7	6D5	BioLegend
CD21	APC	7E9	BioLegend
CD23	FITC	B3B4	BioLegend
CD24	PB	M1/69	BioLegend
CD25	APC/Cy7	PC61	BioLegend
CD25	PerCP	PC61	BioLegend
CD4	PE/Cy7	GK1.5	BioLegend
CD4	PerCP	GK1.5	BioLegend
CD4	PB	GK1.5	BioLegend
CD4	APC/Cy7	GK1.5	BD Biosciences
CD4	APC	GK1.5	BioLegend
CD44	FITC	IM7	BioLegend
CD62L	APC/Cy7	MEL-14	BioLegend
CD69	PE/Cy7	H1.2F3	BioLegend
CD69	FITC	H1.2F3	BioLegend
CD8	PerCP	53-6.7	BioLegend
CD8	PE/Cy7	53-6.7	BioLegend
CD8	APC	53-6.7	BioLegend
CD8	PE	53-6.7	BD Biosciences
Foxp3	Alexa Fluor 488	MF-14	BioLegend
IgM	PE	RMM-1	BioLegend
Nk1.1	PE/Cy7	PKI36	BioLegend
TCRβ	APC/Cy7	H57-597	BioLegend
TCRβ	PE	H57-597	BioLegend
TCRγδ	PerCP	GL3	eBioscience
TCRγδ	PE/Cy7	GL3	BioLegend
TCRγδ	FITC	UC7-13D5	BioLegend

2.5.2 Intracellular antibody staining

After the antibody surface staining the intracellular Foxp3 antibody staining was performed using Foxp3/Transcription Factor Staining Buffer Set (eBioscience) according to manufacturer's instructions.

2.5.3 Cell cycle analysis

The isolated cells were resuspended thoroughly in 0.5 ml PBS, added to 4.5 ml of 70 % EtOH in 15 ml Falcon tubes and placed on a rotation table overnight at 4°C in the dark. Afterwards, the cells were centrifuged for 5 min at 1000 g, the ethanol was decanted completely and the cells were resuspended in 5 ml PBS. After 1 h incubation at room temperature the cells were centrifuged for 5 min at 1000 g again and the supernatant was discarded. The pellet was resuspended in 1 ml PI staining solution (100 ml PBS containing 0,1% Triton X-100, 20 mg DNAse-free RNAse A and 2 mg PI) and the cells were incubated for 2 h at room temperature on a rotation table before acquisition. The measurement of PI had to be performed in linear acquisition of the PE-channel. The measurement was carried out on the Calibur or CANTO II (BD Biosciences) flow cytometer using CellQuest Pro or FACSDiva (BD Biosciences) software. For the analysis FlowJo 9.4 and 10.0.8 software (TreeStar) were used.

2.5.4 Cell sorting

For cell sorting, total thymus or spleen of three to five wildtype (WT) C57Bl/6JOla mice between six to nine weeks old were pooled and antibody-stained with different cocktails containing anti-CD44 FITC, anti-CD69 FITC, anti-TCRβ PE, anti-CD8 PerCP, anti-CD25 PE-Cy7, anti-CD4 APC or anit-CD24 PB (all BioLegend). PI was used to exclude dead cells. The following table shows which population of which organ was sorted using MoFlo Cell Sorter (Beckman Coulter).

Table 13 | Sorting strategy of twenty thymocyte maturation stages in thymus and spleen and the corresponding surface marker.

	CD25	CD44	CD4	CD 8	TCRβ	CD24
Thymus						
DN1	-	+	-	-	-	hi
DN2	+	+	-	-	-	hi
DN3	+	low	-	-	-	hi
DN4	-	-	-	-	-	hi
	CD25	**CD 69**	**CD4**	**CD 8**	**TCRβ**	**CD24**
DP1	-	-	+	+	-	hi
DP1 CD 69+	-	+	+	+	-	hi
DP2	-	-	+	+	+	hi
DP2 CD 69+	-	+	+	+	+	hi
SP immature CD4	-	-	+	-	+	hi
SP immature CD4 CD 69+	-	+	+	-	+	hi
SP immature CD8	-	-	-	+	+	hi
SP immature CD8 CD 69+	-	+	-	+	+	hi
SP mature CD 4	-	-	+	-	+	lo
SP mature CD 4 CD 69+	-	+	+	-	+	lo
SP mature CD 8	-	-	-	+	+	lo
SP mature CD 8 CD 69+	-	+	-	+	+	lo
Spleen						
CD4	-	-	+	-	+	
CD4 CD 69+	-	+	+	-	+	
CD8	-	-	-	+	+	
CD8 CD 69+	-	+	-	+	+	

2.6 Histology

The immunohistochemistry was performed by the group of Prof. Mathias Heikenwälder, Institute of Virology, Technische Universität München. The formalin-fixed paraffin-embedded (FFPE) thymi, spleens and lymph nodes were cut in sections of 4 µm, deparaffinized in xylene and hydrated in a descending grade of alcohol. The anti-CD3 (Clone SP7, ThermoFisher Scientific) staining was performed with Bond Primary antibody diluent (Leica) and the automatic BOND-MAX stainer (Leica) using BOND polymer refine detection solution for DAB (Leica). The staining for hematoxylin and eosin (H&E) were performed after

Harris. The data was analyzed using ImageJ (Schneider et al., 2012). The total and stained areas were calculated by adjusting the color threshold in a way that the appropriate area was marked. This area was then measured and the ratio (stained to total area) was calculated. In addition, the width and height of the T cell-rich areas were measured. Therefore, the following steps had to be performed: 1. Converting the picture into an 8-bit picture 2. Measuring the scale bar of the histology picture in pixel and set the scale to one millimeter 3. Adjusting the color threshold to the T cell-rich areas 4. Analyzing the particles with a set threshold bigger than 100 pixels (to exclude false positive areas) 5. Measuring the width and height and combining data points. The final analysis was performed using Microsoft Excel and Graphpad Prism 6 (GraphPad Software).

2.7 Statistical data analysis

Statistical analyses were performed when n>5 using unpaired students t test of Graphpad Prims 6 (GraphPad Software).

2.8 Software

FlowJo 9.4 was used on Macintosh. All other data analysis and processing was performed on Windows computers using the following software.

Table 14 | Data analysis and processing software as well as the corresponding supplier.

Software	Supplier
BD FACSDiva	BD
BioRad CFX Manager 3.1	BioRad
CELLQuest Pro	BD
CLC Main Workbench 7.6	Qiagen
FlowJo 9.4 and 10	TreeStar
Graphpad Prism 6	GraphPad Software
ImageJ 1.51a	open source, NIH
Microsoft Office 2010 +2013	Microsoft

3 Results

3.1 Immgen database analysis – 26 potential candidates

To examine the role of one or several transcription factors (TFs) that are responsible to set the molecular machinery towards connecting the TCR signaling with the apoptotic process during thymocyte development we hypothesized that this TF is highly expressed in the DP stages compared to DN and SP stages. Therefore, we screened the Immgen database and selected significantly expressed genes by using a cut-off false discovery rate of 0 calculated by the significance analysis of microarrays (Tusher et al., 2001). This cluster analysis revealed in one cluster of genes showing an up-regulation in the DP stages and a down-regulation in DN2, DN3, 4SP24, 8SP24, 4nve, and 8nve. This gene cluster contained 3325 genes with 26 TFs (Tab. 15) that were selected. The expression profiles of those TFs were re-checked on the heatmap (Fig. 8). Some factors did not show the expected profile of a higher expression (red) in DP, DPbl and DPsm compared to a low expression (green) of the other stages and were not selected. In summary, this analysis confirmed nine TFs according to our hypothesis (red marked, Fig. 8) which were subsequently verified by RT-PCR.

Table 15 | List of transcription factors included in the selected Immgen cluster of total 3325 genes.

Transcription Factors

1 MM AT-hook transcription factor (Akna)
2 MM CCR4-NOT transcription complex, subunit 6 (Cnot6)
3 MM E2F transcription factor 1 (E2f1)
4 MM E2F transcription factor 2 (E2f2)
5 MM E2F transcription factor 3 (E2f3)
6 MM E2F transcription factor 6 (E2f6), transcript variant 1
7 MM E2F transcription factor 7 (E2f7)
8 MM E2F transcription factor 8 (E2f8)
9 MM general transcription factor II H, polypeptide 4 (Gtf2h4)
10 MM general transcription factor II I repeat domain-containing 1 (Gtf2ird1), transcript variant 2
11 MM leucine zipper transcription factor-like 1 (Lztfl1)
12 MM metal response element binding transcription factor 2 (Mtf2)
13 MM POU domain, class 6, transcription factor 1 (Pou6f1)
14 MM runt related transcription factor 1 (Runx1), transcript variant 3
15 MM runt related transcription factor 3 (Runx3)
16 MM signal transducer and activator of transcription 1 (Stat1)
17 MM signal transducer and activator of transcription 4 (Stat4)
18 MM transcription elongation factor B (SIII), polypeptide 1 (Tceb1)
19 MM transcription factor 12 (Tcf12)
20 MM transcription factor 19 (Tcf19), transcript variant 1
21 MM transcription factor 4 (Tcf4), transcript variant 1
22 MM transcription factor AP4 (Tcfap4)
23 MM transcription factor Dp 1 (Tfdp1)
24 MM transcription factor Dp 2 (Tfdp2)
25 MM transcription termination factor, RNA polymerase II (Ttf2)
26 MM transcriptional adaptor 2A (Tada2a)

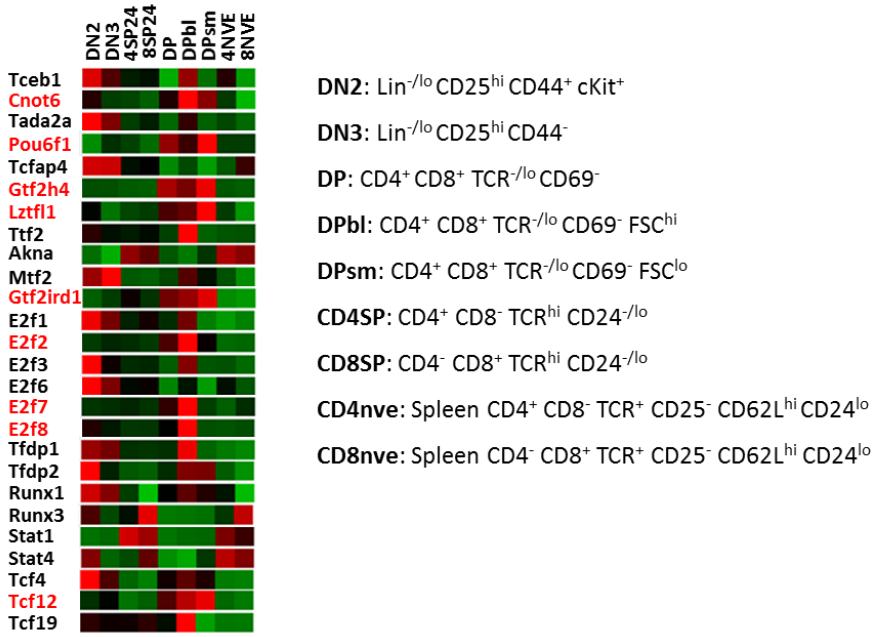


Figure 8 | **Heatmap of 26 transcription factors in nine different T cell developmental stages.** The expression of 26 different transcription factor were compared using Immgen database regarding nine different stages of T cell development (DN2, DN3, DP, DPbl, DPsm, CD4SP, CD8SP, CD4nve and CD8nve). The red marked transcription factors were used for further analysis.

3.2 RT-PCR – The mRNA expression of eleven selected candidates

The cluster analysis resulted in nine potential candidates for our hypothesis, i.e. a TF important for negative selection should be upregulated in the DP stages compared to DN and SP stages. To verify the expression profile of these nine candidate TFs (Cnot6, Pou6f1, Gtf2h4, Lztfl1, Gtf2ird1, E2F2, E2F7, E2F8 and Tcf12) the mRNA expression where measured by RT-PCR in 20 sorted stages of T cell development (for sorting strategy see 2.5.4). More specifically, this included four DN stages, four DP stages and eight SP stages in the thymus as well as four splenic stages. Furthermore, preliminary data of our group showed two additional interesting potential candidates, Nr1h2 and Zfp703, which were included into the analysis.

In the interest of correct sorting the mRNA level of CD4 and CD8 were also measured. CD4 and CD8 were already slightly expressed in the DN4 stage and had an expression peak at the DP stage (Fig. 9). In addition, CD4 was contrary expressed to CD8 at the SP CD4$^+$ stages (Fig. 9A). CD8, however, was expressed at the SP CD8$^+$ stages (Fig. 9B), indicating a correct sorting in terms of the major thymocyte subclasses.

The expression of Cnot6 was very constant in all stages (Fig. 10A), and thus Cnot6 was not included into the further analysis. E2F2 and E2F7, however, showed a peak in stage DP1 (Fig. 10B and 10C). The expression of E2F7 reduced drastically after DP1, yet in further maturated stages E2F2 was still slightly expressed. Nevertheless, this classified both TFs as candidates supporting the hypothesis. E2F8 was expressed in all stages and did not show an expression peak (Fig. 11A). The expression of Gtf2h4 increased continuously until stage DP1 $CD69^+$ and decreased drastically afterwards. Though, Gtf2h4 was weakly expressed in the mature SP stages (Fig. 11B). Although, the expression profile of Gtf2h4 did not fit exactly the criteria associated with the hypothesis, it still might be an interesting candidate because of the decimation of its expression after the DP stages. The expression of Gtf2ird1 started in DN1 and showed a peak during stage DP2 but it was also expressed in the immature SP stages of the thymus as well as splenic SP stages (Fig. 11C) assuming another role than having influence on negative selection. Moreover, Lztfl1 showed a continuous increasing expression until stage DP2 $CD69^+$ and decreased continuously afterwards (Fig. 12A), showing the perfect hypothesized expression profile, yet slightly shifted towards the further maturated DP cells. In contrast, Nr1h2 did not show any expression during the DP stages and its expression started only in mature thymic stages. In addition, it was expressed in the splenic SP stages (Fig. 12B), suggesting a different function and excluding it for further analysis. The expression of Pou6f1 started in DN1 and had a peak in DP1 $CD69^+$. Afterwards the expression was decreased in immature stages but increased in mature stages again (Fig. 12C). Like Gtf2h4, Pou6f1 did not exactly demonstrate the proposed expression profile, but it still might be an interesting TF. The expression of Tcf12 started in DN1 and had a peak in the stages from DN3 until DP1. Afterwards the expression was decreasing and almost lacking from the SP stages onwards (Fig. 13A), indicating a well matching expression profile. Finally, Zfp703 was expressed in most stages except for SP immature CD8 $CD69^+$ as well as the mature SP stages in the thymus (Fig. 13B), which led to the exclusion for further analysis.

In summary, the mRNA expression profile of Tcf12, E2F2, E2F7 and Ltzfl1 matched our hypothesis. Nevertheless, because Pou6f1 and Gtf2h4 were extremely downregulated after the DP stages they were still interesting candidates. A closer look into the literature led to the selection of E2F7, E2F8 and Gtf2h4 for further detailed analysis regarding their role in T cell development and negative selection.

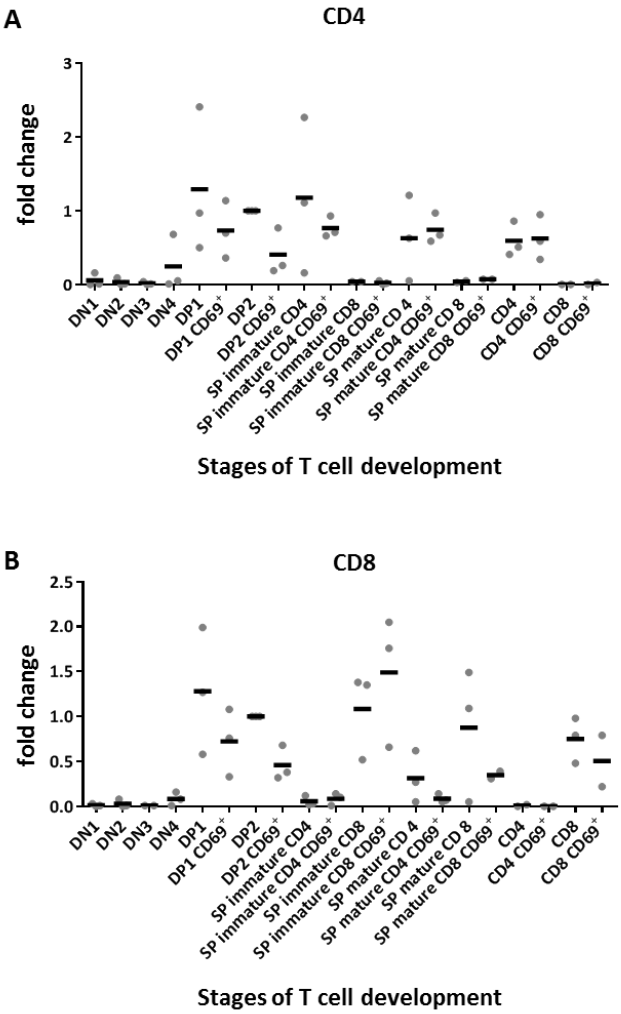


Figure 9 | mRNA expression profile of CD4 and CD8 in 20 different stages of T cell development.
Shown are the single values and means of the calculated fold change of three independent sorts per stage, normalized to two reference genes (Ywhaz and Cxxc1) and stage DP2, of CD4 (A) and CD8 (B).

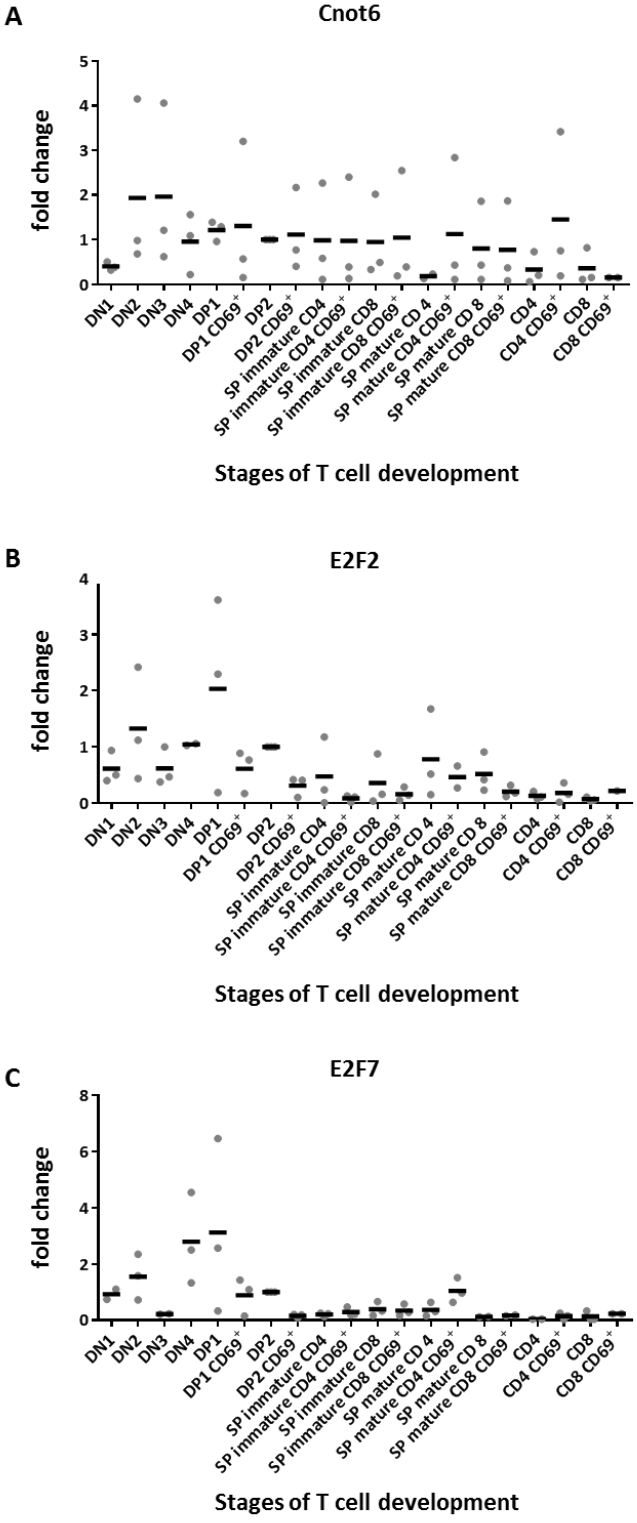


Figure 10 | **mRNA expression profile of Cnot6, E2F2 and E2F7 in 20 different stages of T cell development.**
Shown are the single values and means of the calculated fold change of three independent sorts per stage, normalized to two reference genes (Ywhaz and Cxxc1) and stage DP2, of Cnot6 (A), E2F2 (B) and E2F7 (C).

Figure 11 | mRNA expression profile of E2F8, Gtf2h4 and Gtf2ird1 in 20 different stages of T cell development.

Shown are the single values and means of the calculated fold change of three independent sorts per stage, normalized to two reference genes (Ywhaz and Cxxc1) and stage DP2, of E2F8 (A), Gtf2h4 (B) and Gtf2ird1 (C).

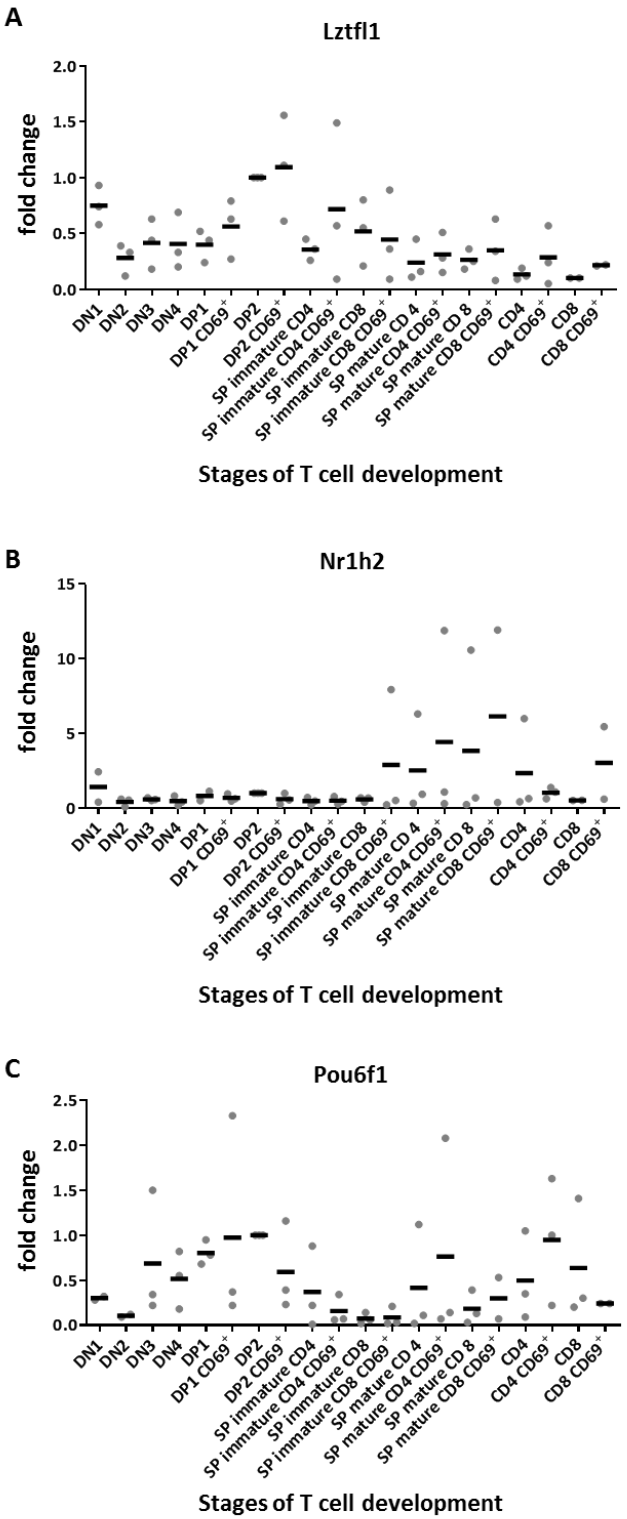


Figure 12 | mRNA expression profile of Lztfl1, Nr1h2 and Pou6f1 in 20 different stages of T cell development.

Shown are the single values and means of the calculated fold change of three independent sorts per stage, normalized to two reference genes (Ywhaz and Cxxc1) and stage DP2, of Lztfl1 (A), Nr1h2 (B) and Pou6f1 (C).

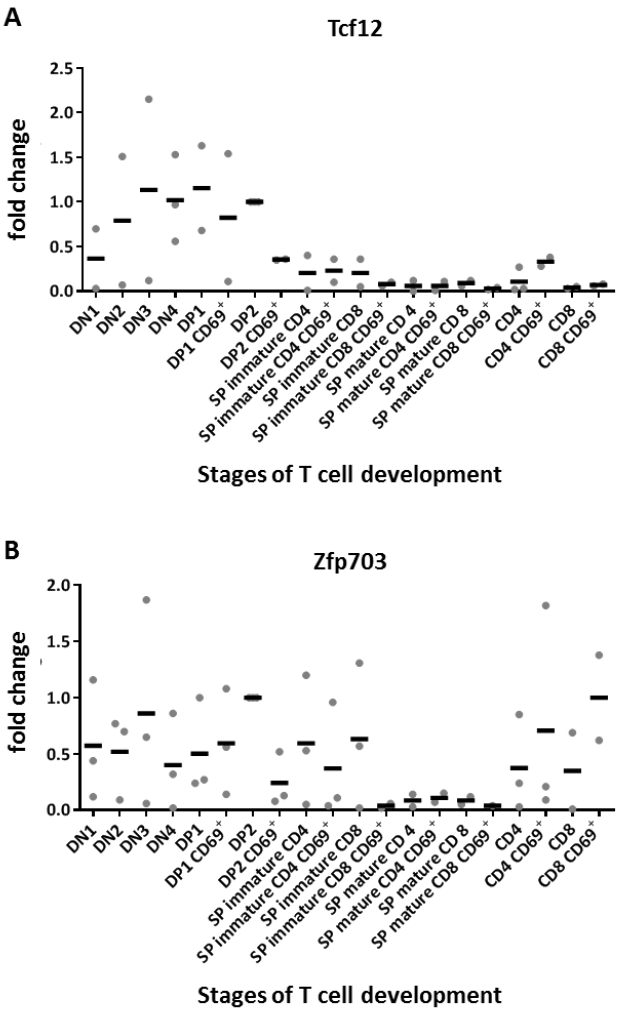


Figure 13 | mRNA expression profile of Tcf12 and Zfp703 in 20 different stages of T cell development.
Shown are the single values and means of the calculated fold change of three independent sorts per stage, normalized to two reference genes (Ywhaz and Cxxc1) and stage DP2, of Tcf12 (A) and Zfp703 (B).

3.3 Thymocyte-specific E2F7 and E2F8 deficiencies do not lead to altered T cell development or cell cycle

To investigate the influence of E2F7 and E2F8 in T cell development and negative selection $E2F7^{fl/fl}$ and $E2F8^{fl/fl}$ mice (Li et al., 2008) were crossed with Lck-cre and CD4-cre mice (Lee et al., 2001; Zhang et al., 2005) to achieve thymocyte-specific KO mice beginning at stage DN3 (Lck-cre) and DP (CD4-cre). Different thymocyte, T cell as well as B cell population of thymus (Fig. 14A+B) and spleen (Fig. 14C-E) were analyzed by flow cytometry. If not otherwise stated, all cell populations are pre-gated on lymphocytes, single cells and live cells (Fig. 14A). Preliminary data showed no differences between the different genotypes of all control mice.

Therefore, the control mice could have one of the following genotypes: E2F7$^{fl/fl}$ E2F8$^{fl/fl}$, Lck-cre/CD4-cre E2F7$^{fl/fl}$ E2F8$^{fl/+}$, Lck-cre/CD4-cre E2F7$^{fl/fl}$ E2F8$^{+/+}$, Lck-cre/CD4-cre E2F7$^{fl/+}$ E2F8$^{fl/fl}$, Lck-cre/CD4-cre E2F7$^{+/+}$ E2F8$^{fl/fl}$, Lck-cre/CD4-cre E2F7$^{fl/+}$ E2F8$^{fl/+}$).

The analysis of the Lck-cre E2F7$^{fl/fl}$ E2F8$^{fl/fl}$ mice was performed with one KO (Lck-cre E2F7$^{fl/fl}$ E2F8$^{fl/fl}$) mouse and three control mice, except for CD4$^+$ and CD8$^+$ cells of the spleen were two KO mice and six control mice were investigated. In the thymus DN, SP CD4$^+$, SP CD8$^+$ and DP as well as all four DN stages (DN1-4) did not show any difference in the KO mouse compared to the control mice. The same was true for the different cell populations of the spleen. No differences in the relative cell number of CD4$^+$, CD8$^+$, TCRβ$^+$, CD19$^+$, TCRγδ$^+$, Treg as well as Tem, Tcm, and TNaive cells of CD4$^+$ and CD8$^+$ pre-gated T cells could be detected (Fig. 15).

The analysis of the thymus of CD4-cre E2F7$^{fl/fl}$ E2F8$^{fl/fl}$ mice was performed on nine KO and nine control mice. Equally to the analysis of Lck-cre E2F7$^{fl/fl}$ E2F8$^{fl/fl}$ no differences between the relative cell number of KO and control mice could be found. For the analysis of the spleen three to five KO and six to eleven control mice were used. Additionally to the cell subsets analyzed for the Lck-cre strain, CD4$^+$ CD69$^+$ and CD8$^+$ CD69$^+$ cells were investigated. However, none of the cell subsets showed any differences in the relative cell number of KO and control mice (Fig. 16).

Furthermore, by the use of propidium iodide the three different phases of the cell cycle, which can be distinguished into G_0/G_1, S and G_2/M phase, were analyzed by flow cytometry. The thymi of two KO mice and six control mice were compared regarding their cell cycle phases but no difference could be observed in both Lck-cre E2F7$^{fl/fl}$ E2F8$^{fl/fl}$ and CD4-cre E2F7$^{fl/fl}$ E2F8$^{fl/fl}$ mice compared to the controls (Fig. 17). In summary, conditional E2F7 and E2F8 double deficiency in thymocytes does not influence the T cell development.

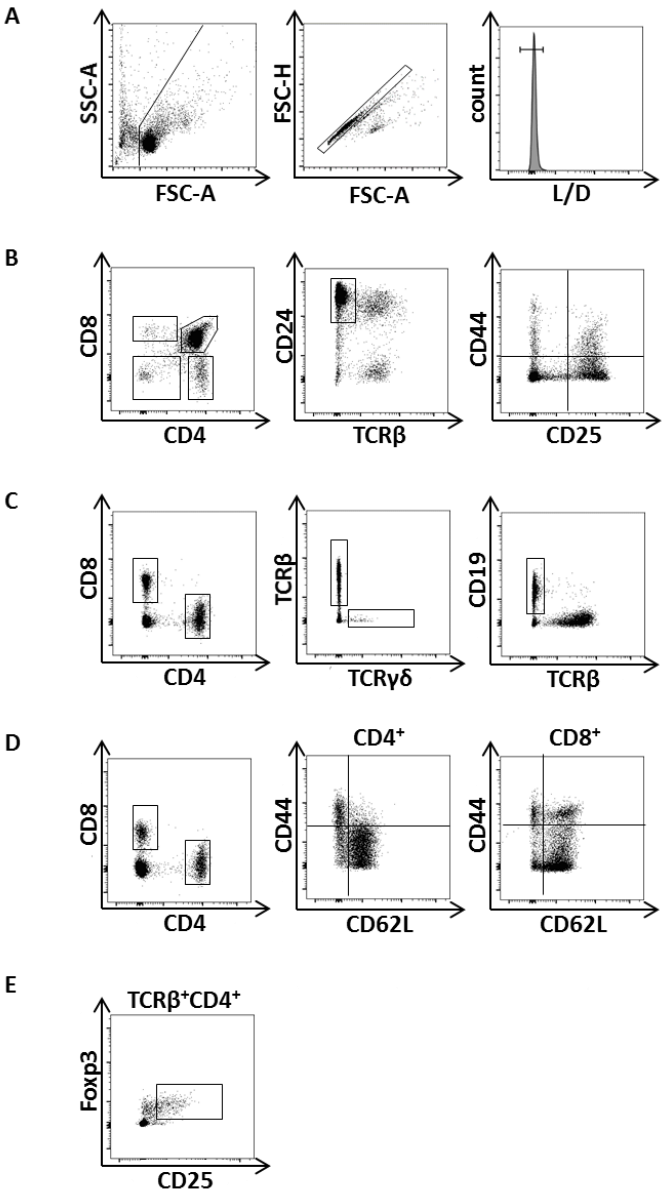


Figure 14 | **Gating strategy of the analyzed cell populations.**
One representative sample of thymus and spleen are shown. (A) lymphocyte, singlet and live/dead gating (from left to right) of one thymus sample. (B) $CD4^-$ $CD8^-$ (DN), $CD4^+$ $CD8^+$ (DP) as well as $CD4^+$ $CD8^-$ and $CD8^+$ $CD4^-$ (left), $CD24^{hi}$ $TCR\beta^-$ (pre-gated on DN) (middle), $CD44^+$ $CD25^-$ (DN1), $CD44^+$ $CD25^+$ (DN2), $CD44^-$ $CD25^+$ (DN3) and $CD44^-$ $CD25^-$ (DN4) (all pre-gated on DN and $CD24^{hi}$ $TCR\beta^-$) (right) of one thymus sample. (C) $CD4^+$ and $CD8^+$ (left), $TCR\beta^+$ and $TCR\gamma\delta^+$ (middle) as well as $CD19^+$ (right) of one spleen sample. (D) $CD4^+$ and $CD8^+$ (left) as well as $CD44^+$ $CD62L^-$ (Tem), $CD44^+$ $CD62L^+$ (Tcm) and $CD44^-$ $CD62L^+$ (TNaive) on $CD4^+$ (middle) and $CD8^+$ (right) pre-gated cells, shown on one spleen sample. (E) $Foxp3^+$ $CD25^+$ cells of $TCR\beta^+$ and $CD4^+$ pre-gated cells of one spleen sample. All populations are pre-gated on lymphocytes, singlets and live cells (except for A).

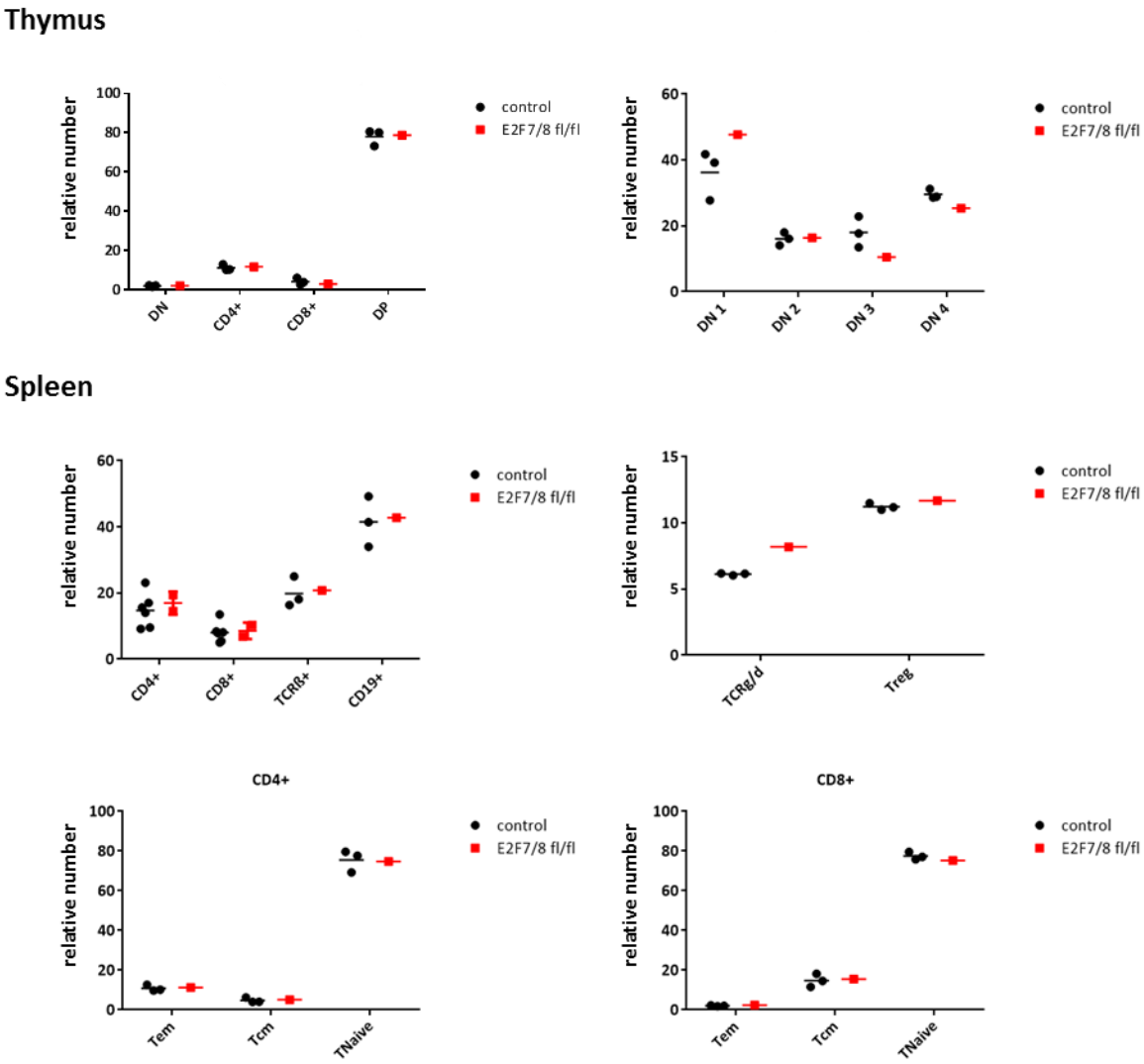


Figure 15 | Relative cell numbers of different cell populations of T cell development of Lck-cre E2F7$^{fl/fl}$ E2F8$^{fl/fl}$ and control mice.

Shown are the relative cell numbers of different cell populations (top) DN, CD4$^+$, CD8$^+$ and DP as well as DN1, DN2, DN3 and DN4 in the thymus of one experiment with Lck-cre E2F7$^{fl/fl}$ E2F8$^{fl/fl}$ (E2F7/8 fl/fl) (n=1) and control mice (n=3). (middle) CD4$^+$, CD8$^+$, TCRβ$^+$, CD19$^+$, TCRγδ$^+$ and Treg cells in the spleen of one to two independent experiments with Lck-cre E2F7$^{fl/fl}$ E2F8$^{fl/fl}$ (E2F7/8 fl/fl) (n=1) and control mice (n=3). (bottom) Tem, Tcm and TNaive on CD4$^+$ (left) and CD8$^+$ (right) pre-gated cells in the spleen of one experiment with Lck-cre E2F7$^{fl/fl}$ E2F8$^{fl/fl}$ (E2F7/8 fl/fl) (n=1) and control mice (n=3). Bars indicate the mean of each group.

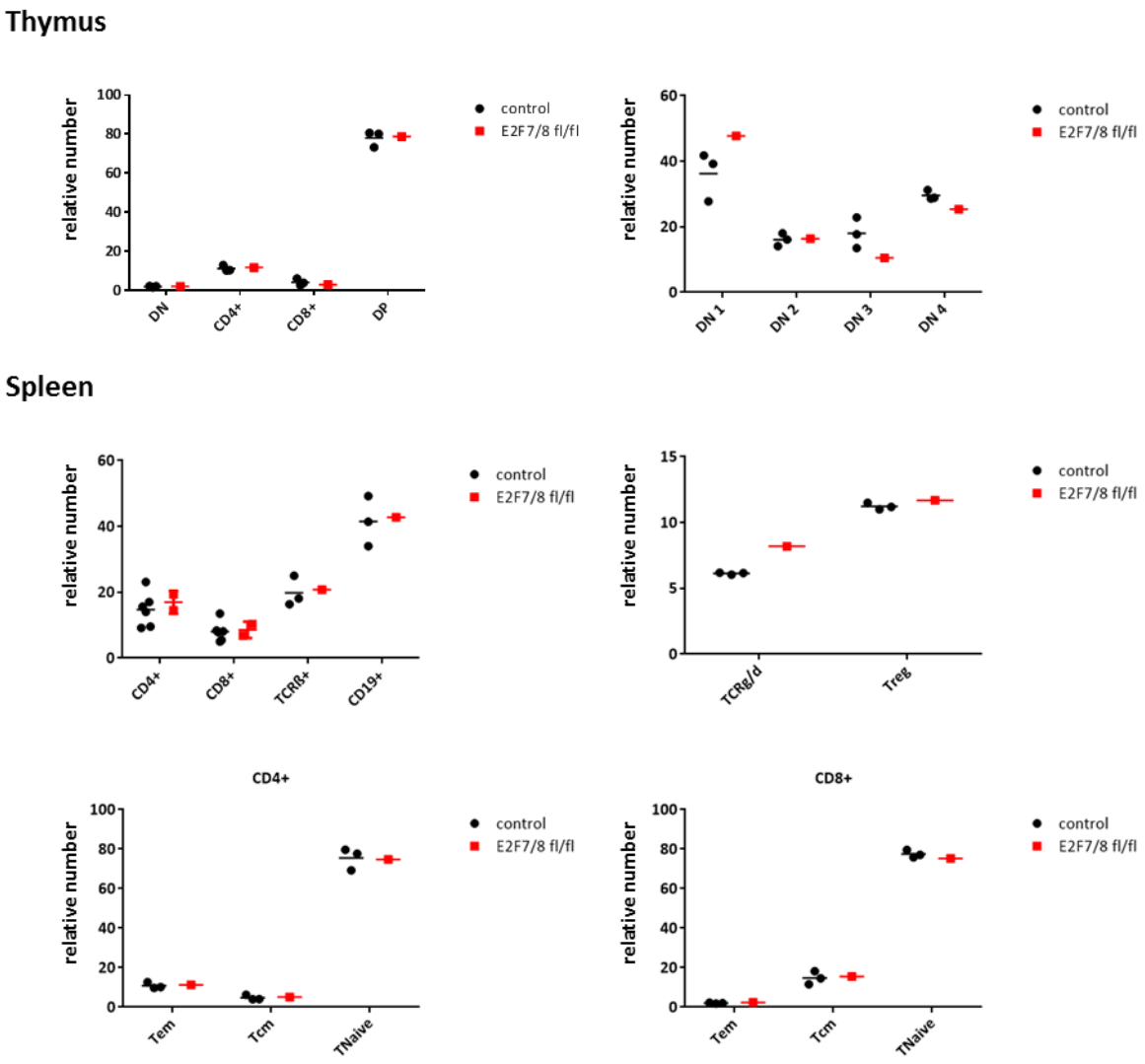


Figure 16 | Relative cell numbers of different cell populations of T cell development of CD4-cre E2F7$^{fl/fl}$ E2F8$^{fl/fl}$ and control mice.

Shown are the relative cell numbers of different cell populations (top) DN, CD4^{+}, CD8^{+} and DP as well as DN1, DN2, DN3 and DN4 in the thymus of four independent experiments with CD4-cre E2F7$^{fl/fl}$ E2F8$^{fl/fl}$ (E2F7/8 fl/fl) (n=9) and control mice (n=9). (middle) CD4^{+}, CD8^{+}, TCRβ^{+} CD19^{+}, TCRγδ$^{+,}$ Treg cells in the spleen of two to four independent experiments with CD4-cre E2F7$^{fl/fl}$ E2F8$^{fl/fl}$ (E2F7/8 fl/fl) (n=3-5) and control mice (n=6-11). (bottom) Tem, Tcm and TNaive on CD4^{+} (left) and CD8^{+} (right) pre-gated cells in the spleen of three independent experiments with CD4-cre E2F7$^{fl/fl}$ E2F8$^{fl/fl}$ (E2F7/8 fl/fl) (n=3) and control mice (n=8). Bars indicate the mean of each group.

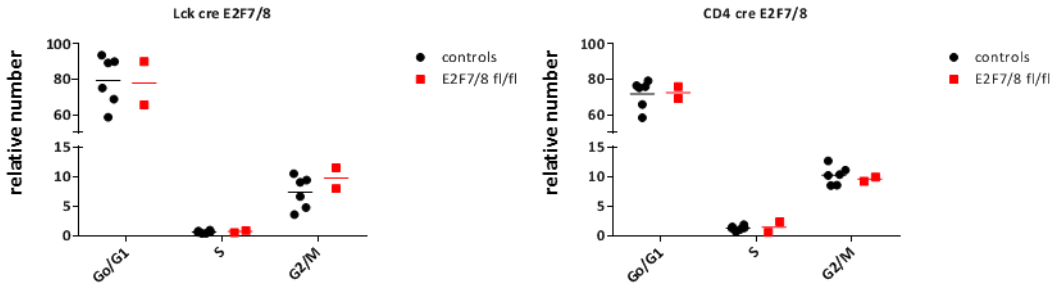


Figure 17 | Relative cell numbers of different stages during the cell cycle of Lck-cre E2F7$^{fl/fl}$ E2F8$^{fl/fl}$ and CD4-cre E2F7$^{fl/fl}$ E2F8$^{fl/fl}$ as well as control mice.
The relative cell numbers and the corresponding mean of G_0/G_1, S and G_2/M stages of the cell cycle analysis of two independent experiments of Lck-cre E2F7$^{fl/fl}$ E2F8$^{fl/fl}$ (left) (n=2) and CD4-cre E2F7$^{fl/fl}$ E2F8$^{fl/fl}$ (right) (n=2) as well as control mice (n=6 each).

3.4 Generation of Gtf2h4 deficient mice with CRISPR/Cas9

3.4.1 *In silico* design for px330(gRNA Gtf2h4)

The KO mouse strain of Gtf2h4 was generated by use of the CRISPR/Cas9 system from *S. pyogenes*. The protospacer was designed using Crispr design tool at crispr.mit.edu. Exon 2 of Gtf2h4 was selected for targeting to ideally target all of the different transcripts. Gtf2h4 has twelve splice variants of which five are protein-coding. The following sequence was used for the *in silico* analysis:

5`-gtg atg gag atc acc ccg gcg agg ggt gga ctg aac cga gca cac cta caa tgc agg aat ctc cag gag ttc tta gga ggc ctg agc cct ggg gtg ctg gac cga ttg tat ggg cac cct gcc act tgt ctg gct gtc ttc ag-3`

27 gRNAs with scores from 95-40 were suggested by the program. The score is calculated with an assumption of 100% on target affinity subtracted by the number and affinity to off-targets. Only gRNAs with a score above 50 are recommended to be considered by crispr.mit.edu (Hsu et al., 2013). Two gRNAs were chosen to target exon 2 of Gtf2h4, namely gRNA Gtf2h4 fwd (5`-ggg gtg ctg gac cga ttg ta-3`) with a score of 91 and gRNA Gtf2h4 rev (5`-ggt cca gca ccc cag ggc tc-3`) with a score of 51. The gRNAs were named after the forward and reverse strand that they target. Those two gRNAs were chosen because of the score above 50 and the 5`-GG-$N_{18}$-NGG-3` sequence which was suggested at the time when the mice were generated (Gagnon et al., 2014).

3.4.2 Cloning of the px330(gRNA Gtf2h4 fwd/rev) vector

After the protospacer identification for Gtf2h4 the cloning into the px330 vector was performed according to the protocol provided by the Zhang Lab (Cong et al., 2013) with some adaptations. 1 µl of each specific oligonucleotide (100 µM each) was mixed with 1 µl 10x T4 PNK Buffer (NEB) and 7 µl H_20 and annealed with following temperature protocol: 30 minutes at 37°C followed by 95°C ramped down to 25°C at -1.5 °C/min. Afterwards, the annealed oligonucleotides were diluted 1:250 in nuclease free water. Subsequently, the digestion ligation reaction was set up as followed: 4 µl px330 (25 ng/µl), 2 µl annealed oligonucleotides (1:250 dilution), 2 µl 10x NEB 2.1 Buffer, 1 µl DTT (500 mM) (Fermentas), 1 µl ATP (10 mM) (Fermentas), BbsI 1 µl (NEB), 0.5 µl T7 ligase (NEB), 8.5 µl H_2O. The reaction was incubated in the C1000 Touch or S1000 Thermal Cycler (BioRad) using following program: 10 min at 37°C followed by 5 min at 23°C; both steps were repeated for 6 cycles. Finally, 2 µl of the digestion ligation product were used for transformation of *E. coli* DH5α cells. To verify the colonies for correct cloning, a colony screening PCR with subsequent restriction digest was designed and established. The PCR was performed using Taq PCR Master Mix (Qiagen) and the px330 screen fwd (5`-ggc cta ttt ccc atg att-3`) and rev (5`-ata caa aat tgg ggg tgg-3`) primer pair. 3 µl of this cell lysate (one picked colony mixed in 50 µl H_2O) was added to the PCR mixture, which contained 12.5 µl Taq PCR Master Mix (Qiagen), 1 µl of each primer (10 µM) and 8.5 µl H_2O. The following temperature protocol was used: 1. initial denaturation: 95°C for 3 min, 2. denaturation 95°C for 30 sec, 3. annealing 58°C for 30 sec, 4. elongation 72°C for 50 sec (2-4: 35 cycles), 5. final elongation 72°C 5 min. After PCR amplification 15 µl of the amplified mix were digested by adding 0.2 µl AvaII (NEB) and 4.8 µl H_2O and incubated at 37°C for 1 h. Finally, the digestion products were analyzed on a 1% agarose gel. This protocol was used to construct px330 (gRNAGtf2h4 fwd) that results in a PCR band of 793 bp, subsequently digested into a 255 bp and 538 bp product (Fig. 18). Only digested PCR products resulted from the correct cloning. The cloning of px330(gRNA Gtf2h4 rev) was performed with px330 ccdB (see 3.4.3) without necessity of verification. Both circular plasmids were injected in parallel into fertilized oocytes.

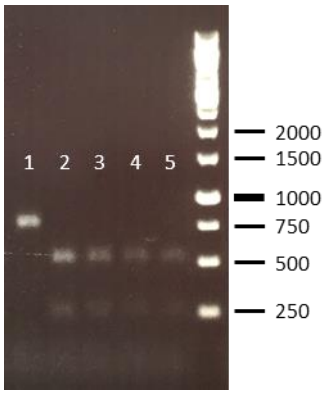


Figure 18 | Clone verification of px330(gRNA Gtf2h4 fwd).
Agarose gel picture of the colony PCR with the px330 screen fwd and rev primer pair after the digestion with AvaII. Lane 1 shows an undigested PCR product of 793 bp. Lane 2-5 show the digested PCR products with two bands, one with 255 bp and the other with 538 bp. 1kb ladder, assigned in bp (right end)

3.4.3 Increasing the cloning efficiency of the px330 vector

The cloning strategy with the px330 vector provided by Cong et al. (Cong et al., 2013) with some adaptions worked quite efficiently (Fig. 19). Nevertheless, to ensure the right cloning, colonies should be verified. To simplify this, the chloramphenicol (CmR)/ccdB cassette was amplified by PCR from an available Gateway destination vector in our plasmid stock using following primer Gibson CmR/ccdB fwd (5`- tgt gga aag gac gaa aca ccg ggt ctt ccg gcc gca tta ggc acc cca g-3`) and rev (5`- gct att tct agc tct aaa aca ggt ctt cgt cga cct gca gac tgg ctg t-3`). An overhang of 20 bp to each primer was added to perform Gibson Assembly Cloning. Additionally, restriction sites for BbsI were added on both sites with the intention of using the same established cloning protocol as in 3.4.2. The PCR was performed using Herculase II Fusion Enzyme with dNTPs Combo (Agilent). The px330 was digested with BbsI and gel-purified with QIAEX II Gel Extraction Kit (Qiagen). The PCR amplified CmR/ccdB cassette was then cloned into the vector between of the U6 promoter and the chimeric guide RNA using the Gibson assembly cloning kit (NEB) according to manufacturer's instructions. The ccdB resistant strain *E. coli* ccdB Survival™ 2 T1R (ThermoFisher Scientific) was transformed by electroporation (1700 V, 5 ms, 1 mm cuvette), subsequently plated onto LB plates containing chloramphenicol and ampicillin and incubated overnight at 37°C. The plasmids of single colonies were purified with PureYield™ Plasmid Miniprep System (Promega) and a restriction digest with BbsI (NEB) was performed. Plasmids carrying the CmR/ccdB cassette should show a band at 8488 bp and 1478 bp. One positive clone was purified using PureYield™ Plasmid

Midiprep System (Promega) and as a negative control the same plasmid was used to transform *E. coli* DH5α cells and plated on a LB plate with ampicillin resistance. With this newly generated vector px330 ccdB only colonies that had inserted the 20 bp protospacer successfully could survive in the DH5α or Top10 *E. coli* strains which led to an increase of the cloning efficiency to 100% compared to 83% positive clones with the former vector px330 (Fig. 19).

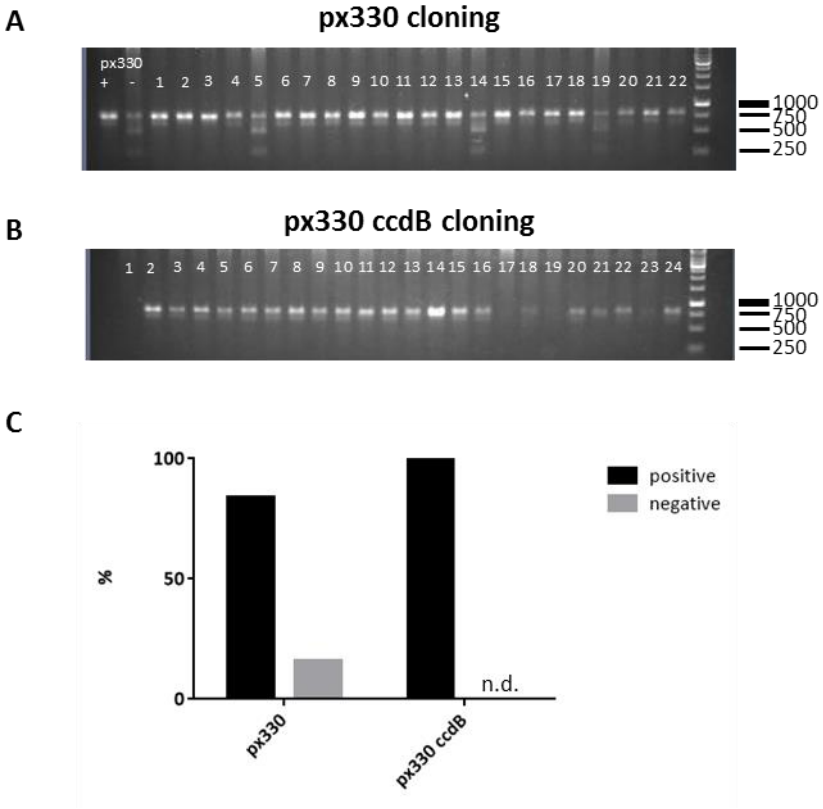


Figure 19 | **Higher cloning efficiency by the use of the px330 ccdB vector.**
(A) Agarose gel picture after BbsI digestion of a PCR product (750 bp) spanning the cloning site after the cloning of the protospacer into the conventional px330 vector. Lane 5, 14 and 19 showed digested PCR products (250 bp and 500 bp); px330+ = positive control, vector with correctly inserted gRNA; px330- = negative control, empty px330 vector; 1 kb ladder, assigned in bp (right end). (B) Agarose gel picture after BbsI digestion of a PCR product spanning the cloning site after the cloning of the protospacer into the px330 ccdB vector. No digested PCR products are visible; 1 kb ladder, assigned in bp (right end). (C) Summary of all screened colonies depicted in percentage; positive = correct gRNA insertion, no digest; negative = no gRNA insertion, PCR product digested.

3.4.4 Genotyping of the Gtf2h4 deficient mice

After PNI of the circular px330(gRNA Gtf2h4 fwd/rev) plasmids twelve founder mice were born. To verify their genotype the below-mentioned PCR was performed (Tab. 16+17) and the PCR product was purified using Wizard® SV Gel and PCR Clean-Up System (Promega) and sent for sequencing using the primer Gtf2h4 KO fwd (5`- gag ggg ttt ggg agg aat ag-3`). One heterozygous mouse could be detected with a 7 bp deletion in exon 2 (Fig. 20A) that leads to a stop codon in exon 3 of Gtf2h4. This mutation introduced a recognition site for the restriction enzyme MwoI which was used for genotyping. The genotyping PCR was established with the primer pair Gtf2h4 KO fwd (5`- gag ggg ttt ggg agg aat ag-3`) and rev (5`- acg gtt gac att agc cct tg-3`) and the Taq PCR Master Mix (Qiagen) using following mix and temperature protocol.

Table 16 | **PCR Mix for the genotyping of the Gtf2h4 deficient mouse strain.**

Mix components	µl
Taq PCR Master Mix (Qiagen)	12.5
Primer fwd (10 mM)	0.5
Primer rev (10 mM)	0.5
H_2O	10.5
Template DNA	1

Table 17 | **Temperature protocol for the PCR of the Gtf2h4 deficient mouse strain.**

Step	Temp. (°C)	Time (mm:ss)
1	94	03:00
2	94	00:30
3	57	00:30
4	72	01:00
2-4	35 cycles	
5	72	10:00
6	12	∞

The PCR led to a 602 bp (WT) or 595 bp (KO) product which was subsequently digested with MwoI. The mix of 3 µl CutSmart Buffer (NEB), 0.5 µl MwoI (Neb) and 1.5 µl H_20 was added directly to each 25 µl PCR reaction. After incubation at 37°C for 1 hour the digestion product was analyzed on a 2% agarose gel. In case of a mutated allele the product was digested by

MwoI into three bands of 318 bp, 213 bp and 64 bp in which the latter band was not visible on the gel. In contrast, the digestion of the WT PCR product led to two bands with 284 bp and 318 bp (Fig. 20B). Interestingly, heterozygous breeding resulted in 27% WT, 73% HET and no KO mice (Fig. 20C), indicating a non-mendalian ratio that implies a lethality of Gtf2h4 KO mice. The newly generated mouse strain is named B6FVBN-Gtf2h4$^{131\text{-}137del}$ThBu.

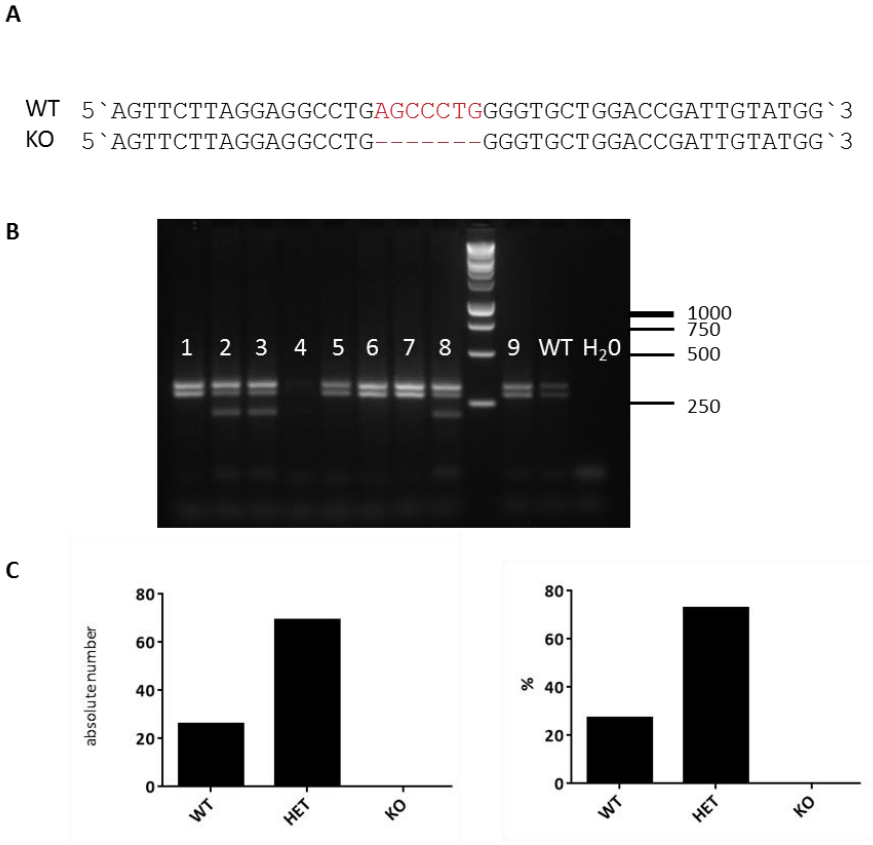


Figure 20 | Sequence of Gtf2h4 mutation generated by CRISPR/Cas9, agarose gel picture of the genotyping and distribution of genotypes.
(A) Sequence section of Gtf2h4 exon 2 of the WT sequence and KO sequence with 7 bp deletion. (B) Agarose gel of MwoI digested PCR products; lane 1,5,6,7 and 9 showed WT bands of 284 bp and 318 bp; lane 2, 3 and 8 showed heterozygous bands of 318 bp and 213 bp (64 bp was not detected); in lane 4 no PCR product was amplified; positive control = WT, negative control = H_20, ladder = 1 kb, assigned in bp. (C) Absolut number of genotyped mice and the percentage sorted into WT, HET and KO.

3.5 Embryonic lethality in mice homozygous for Gtf2h4 deficiency

To analyze the time point of lethality of the embryos, the embryos of heterozygous breeding pairs were removed at day E12.5. Out of 46 embryos eleven were WT (24%) and 35 HET (76%) (Fig. 21A+B), which was almost equal to the number of pups of heterozygous breeding pairs. Furthermore, the morphology of the embryos was investigated with a stereo

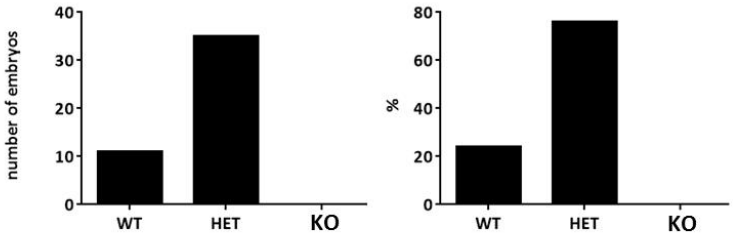
microscope (Wild M3, Heerbrugg) without obvious differences (Fig. 21C). Additionally, one experiment was conducted at an earlier time point at day E8.5. Among the eight isolated embryos none of them had a homozygous genotype. Furthermore, fertilized oocytes were extracted and intended to be cultured to the blastocyst stage. Astonishingly, in two experiments only two oocytes each developed into blastocysts, which did not allow any conclusion.

A

	Number of Embryos
WT	11
HET	35
KO	0

	%
WT	23,9
HET	76,1
KO	0

B



C

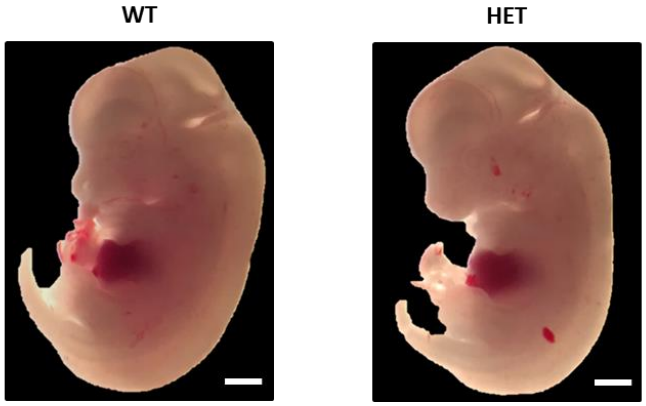


Figure 21 | **Analysis of E12.5 embryos from Gtf2h4 HET breeding pairs.**
Absolut number of genotyped embryos and the percentage sorted into WT, HET and KO depicted in a table (A) and as bar graphs (B). (C) Stereo microscopic picture of one embryo of WT and HET each. Scale bar = 1mm.

3.6 Gtf2h4 heterozygous deficient mice showed a reduced number of cells in the spleen and LN and an increased number in the thymus

Subsequent to the generation of the Gtf2h4 KO mouse we investigated the influence of Gtf2h4 on T cell development and negative selection. Therefore, Gtf2h4 HET mice were analyzed by flow cytometry and compared to WT littermates. First, the total cell numbers of spleen, LN and thymus of Gtf2h4 HET (n=3) mice and WT littermates (n=4) were compared. We observed a tendency for a decreased number of cells in the spleen and LN as well as an increased number in the thymus of HET mice (Fig. 22). Furthermore, the different cell population of spleen, LN and thymus were analyzed by flow cytometry. In addition, to the already mentioned cell populations in 3.3 B cells were further analyzed as followed: $B220^+$, $B220^+$ $CD23^{hi}$ $CD21^{lo}$ (follicular B cells (FO)), $B220^+$ IgM^+ $CD21^-$ (transitional T1 B cells (T1)), $B220^+$ IgM^+ $CD21^+$ (transitional T2 – marginal zone B cells (T2-MZ)), $B220^+$ $CD23^{lo/-}$ $CD21^{hi}$ (marginal zone B cells (MZ)) and $B220^+$ IgM^{lo} $CD21^{mid}$ (follicular mature B cells (FM)). Moreover, $CD11b^+$, $CD11c^+$ and $Nk1.1^+$ cells were investigated (Fig. 23). Flow cytometry analysis of the thymus revealed a tendency of an increased cell number of especially $DN3^+$, $DN4^+$, $CD4^+$, DP as well as Treg cells. However, the number of $CD19^+$ seemed to be slightly downregulated (Fig. 24). In contrast, $CD4^+$, $CD8^+$, $TCR\beta^+$ and $CD19^+$ cell numbers in the lymph nodes decreased, which was not observed in the spleen (Fig. 25A). Comparison of $TCR\gamma\delta^+$ as well as $CD11b^+$ and $CD11c^+$ of spleen and lymph nodes of HET and WT mice resulted in no differences. Nevertheless, Treg cells and $Nk1.1^+$ cells are less present in both peripheral organs of HET mice (Fig. 25B). The same was true for $CD4^+$ Tcm. For $CD4^+$ TNaive as well as $CD8^+$ Tcm and $CD8^+$ TNaive this lower abundance was only observed in lymph nodes but not in spleen (Fig. 25C+D). With regard to B cells, a reduction in the cell numbers was detected, especially in $B220^+$, FO, MZ and FM of lymph nodes, whereas in the spleen no differences of the B cell populations were found (Fig. 25E). In summary, several cell populations in the Gtf2h4 HET mice were downregulated in the peripheral lymphoid organs but upregulated in the thymus.

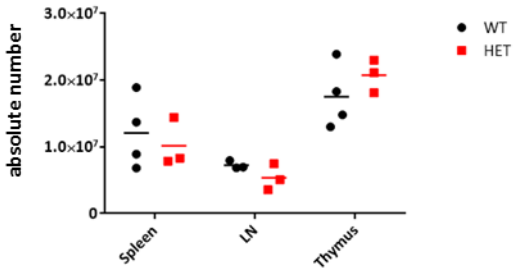


Figure 22 | Comparison of cell numbers of various organs of WT and Gtf2h4 HET mice.
Depicted are the cell numbers of spleen, LN and thymus of WT (n=4) and Gtf2h4 HET (n=3) mice. Bars indicate the mean of each group.

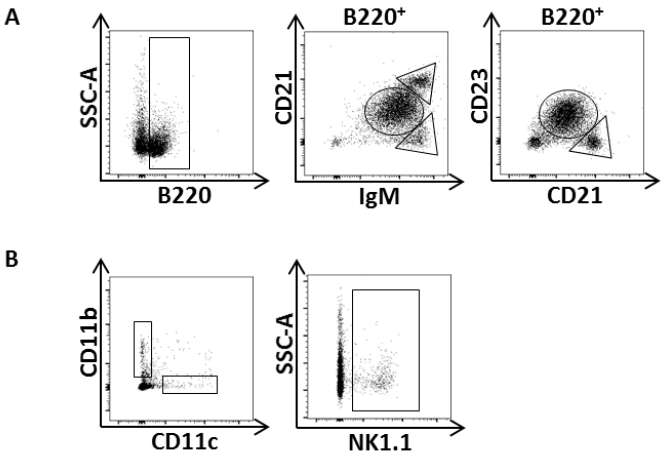


Figure 23 | Gating strategy of additionally analyzed cell population shown on one representative sample of the spleen.
(A) $B220^{+}$, $B220^{+}$ $CD23^{hi}$ $CD21^{lo}$ (FO), $B220^{+}$ IgM^{+} $CD21^{-}$ (T1), $B220^{+}$ IgM^{+} $CD21^{+}$ (T2-MZ), $B220^{+}$ $CD23^{lo/-}$ $CD21^{hi}$ (MZ) and $B220^{+}$ IgM^{lo} $CD21^{mid}$ (FM) of one spleen sample. (B) $CD11b^{+}$, $CD11c^{+}$ and $Nk1.1^{+}$ of one spleen sample. All populations are pre-gated on lymphocytes, singlets and live cells.

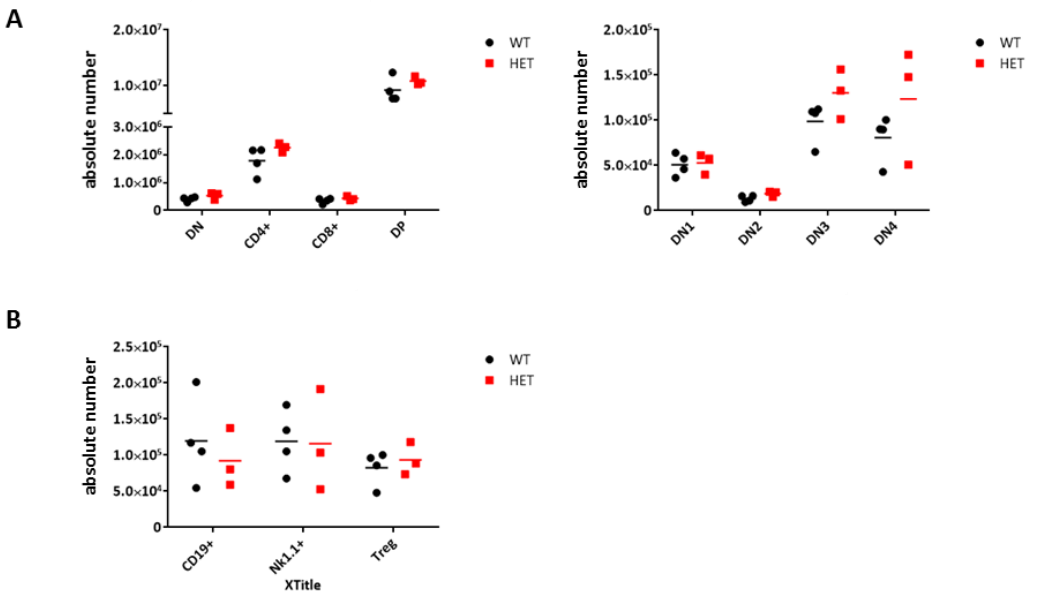


Figure 24 | **Cell numbers of different cell populations in the thymus of Gtf2h4 HET and control mice.**

Shown are the cell numbers of different cell populations (A) DN, $CD4^+$, $CD8^+$ and DP (left) as well as DN1, DN2, DN3 and DN4 (right) in the thymus of two independent experiments with Gtf2h4 HET (n=3) and control WT mice (n=4). (B) $CD19^+$, $Nk1.1^+$ and Treg cells in the thymus of two independent experiments with Gtf2h4 HET (n=3) and control WT mice (n=4). Bars indicate the mean of each group.

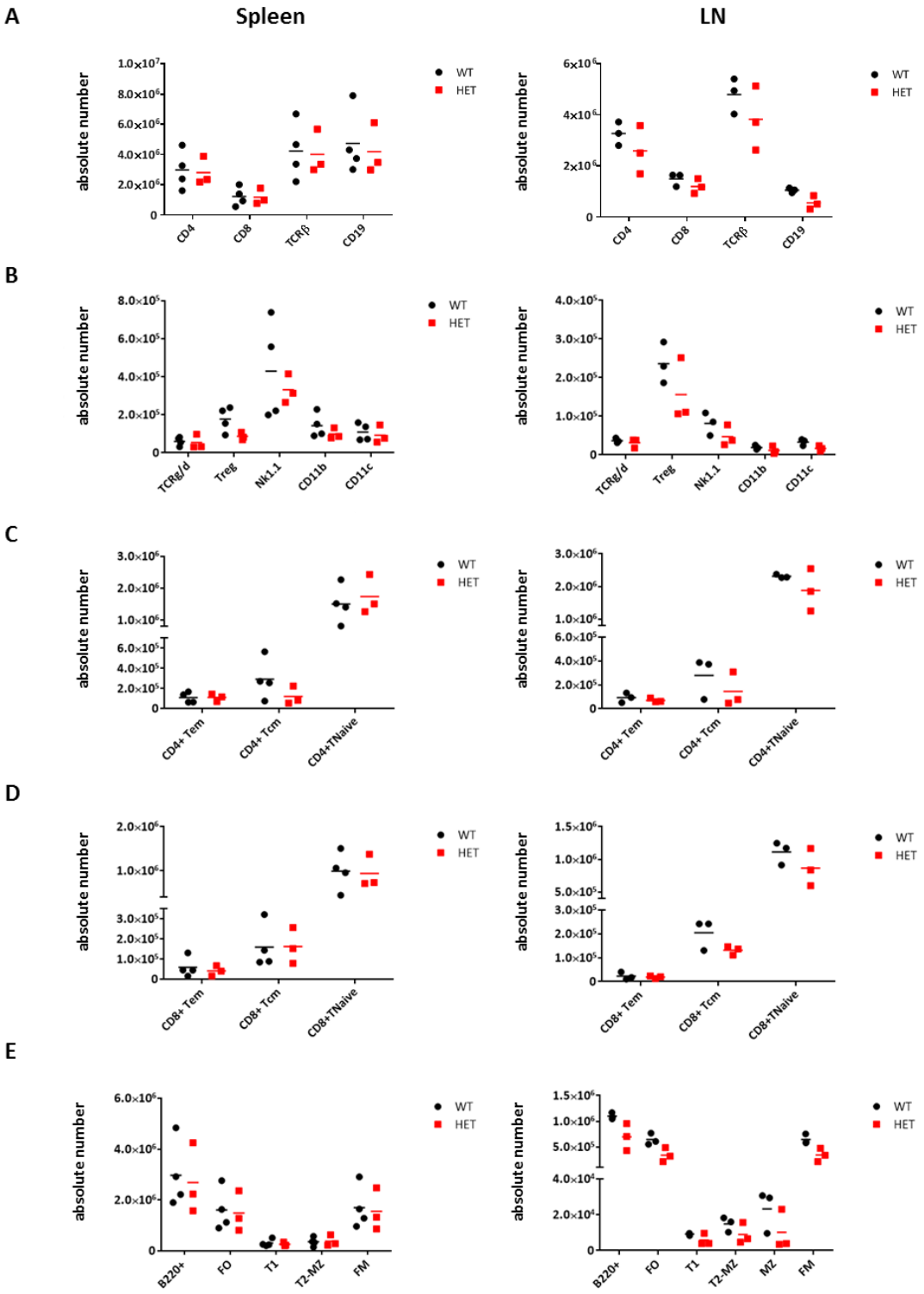


Figure 25 | Cell numbers of different cell populations in the spleen and LN of Gtf2h4 HET and control mice.

Shown are the cell numbers of different cell populations (A) $CD4^+$, $CD8^+$, $TCR\beta^+$ and $CD19^+$ in the spleen (left) and LN (right) of two independent experiments with Gtf2h4 HET (n=3) and control WT mice (n=3-4). (B) $TCR\gamma\delta^+$, Treg, $Nk1.1^+$, $CD11b^+$ and $CD11c^+$ in the spleen (left) and LN (right) of two independent experiments with Gtf2h4 HET (n=3) and control WT mice (n=3-4). Tem, Tcm and TNaive of $CD4^+$ (C) and $CD8^+$ (D) pre-gated cells in the spleen (left) and LN (right) of two independent experiments with Gtf2h4 HET (n=3) and control WT mice (n=3-4). (E) $B220^+$, FO, T1, T2-MZ, MZ and FM B cells in the spleen (left) and LN (right) of two independent experiments with Gtf2h4 HET (n=3) and control WT mice (n=3-4). Bars indicate the mean of each group.

3.7 The generation of the transgenic mouse model overexpressing Gtf2h4 under the CD4 promoter

3.7.1 Cloning of the CD4Gtf2h4 vector

To investigate the influence of a TF it is a standard approach to consider KO as well as overexpression mouse models. To broaden the knowledge about the role of Gtf2h4 in T cell development and negative selection we therefore overexpressed Gtf2h4 in T cells and generated a transgenic mouse model which expressed Gtf2h4 under the CD4 promoter. Total thymic wildtype C57Bl/6JOla RNA was transcribed into cDNA with iScript cDNA Synthesis Kit (BioRad). cDNA of the transcript variant Gtf2h4-001 was amplified with USB Fideli Taq PCR Master Mix (Affymetrix) and primer ORF gtf2h4 fwd (5`-agg tcg acg cca cca tgg aga tca ccc cgg cga-3`) and rev (5`- ctg tcg act cag gag ctg tgc ttc tgc cgc ttc-3`), which contained the Kozak sequence (Kozak, 1984) and two additional SalI digestion sites, one at each end. The PCR product was cloned into the pGEM-T Easy vector using the pGEM-T Easy Vector System (Promega) and verified by sequencing. To clone the Gtf2h4 PCR product into a construct consisting of the CD4 promoter the cDNA was amplified from the pGEM-T Easy Gtf2h4 with USB Fideli Taq PCR Master Mix (Affymetrix) and the primer Gibson Gtf2h4 fwd (5`- acc aac aag agc tca agg agt cga cgc cac cat gga gat c-3`) and rev (5`- atg atc cat gga atc gtc gac tca gga gct gtg ctt ctg c -3`) containing a 20 bp overhang suitable for Gibson cloning. The Gibson cloning simplifies the cloning procedure and ensures correct orientation of the insert. After amplification, the 1444 bp band was gel-purified using QIAEX II Gel Extraction Kit (Qiagen) and cloned into the SalI digested CD4 Depe vector by Gibson Assembly Cloning Kit (NEB). Transformed *E. coli*, provided by the kit, were plated on LB Amp plates and next day single colonies were inoculated in 5 ml LB Amp media. The CD4Gtf2h4 plasmids were isolated by PureYield™ Plasmid Miniprep System (Promega) according to manufacturer's instructions. Subsequently, colonies were picked and verified by enzymatic digestions. The digestion with SalI resulted into two bands, one with 1408 bp and the other with 7529 bp (Fig. 26A). Furthermore, three more enzymatic digestions were set up to confirm the correct cloning. The digestion with XbaI yields two bands (1812 bp and 7125 bp), PstI to six bands (13 bp, 313 bp, 651 bp, 852 bp, 3064 bp and 4060 bp) and AvaI to four bands (619 bp, 964 bp, 2734 bp and 4628 bp) (Fig. 26B).

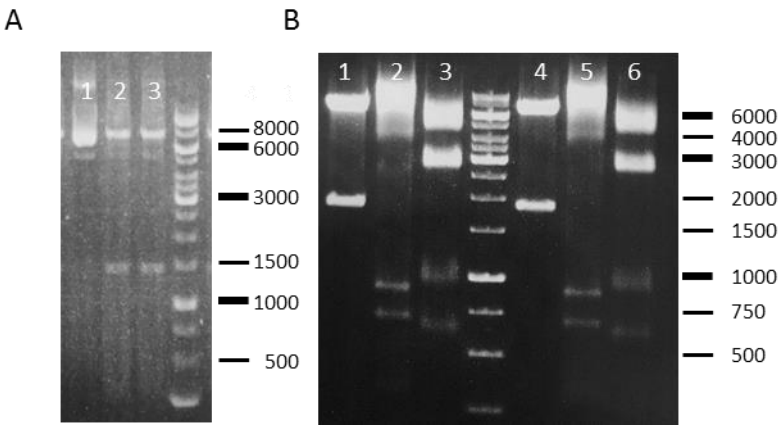


Figure 26 | **Agarose gel picture of the verification digest of CD4Gtf2h4.**
After cloning the cDNA of Gtf2h4 into the CD4 Depe vector the correct cloning was verified by four different digestions: (A) SalI (1408 bp and 7529 bp), (B) XbaI (1812 bp and 7125 bp) (lane 1 and 4), PstI (13 bp, 313 bp, 651 bp, 852 bp, 3064 bp and 4060 bp) (lane 2 and 5) and AvaI (619 bp, 964 bp, 2734 bp and 4628 bp) (lane 3 and 6). For PstI the 313 bp band was very thin. Shown are two representative colonies for A and B. Additionally, lane 1 of picture A showed the undigested vector. 1kb ladder, assigned in bp, was used for both gels (right end each).

Finally, to exclude mutations caused by the PCR, the insert was fully sequenced using the primer Gtf2h4 1a (5`-tct ctg tca tgc act gct-3`), Gtf2h4 1b (5`-aac tcc ttc ttc acc cac-3`), Gtf2h4 2a (5`-gga ctt tga act gct gct-3`) and Gtf2h4 2b (5`-acg gtg aaa acc tct gac-3`). One positive clone was inoculated overnight in 200 ml LB media and PureYield™ Plasmid Midiprep System (Promega) was performed the following day. For PNI, 10 µg of the CD4Gtf2h4 vector was digested overnight with NotI (NEB). The 6284 bp band was gel-purified with QIAEX II Gel Extraction Kit (Qiagen) and used for injection (Fig. 27).

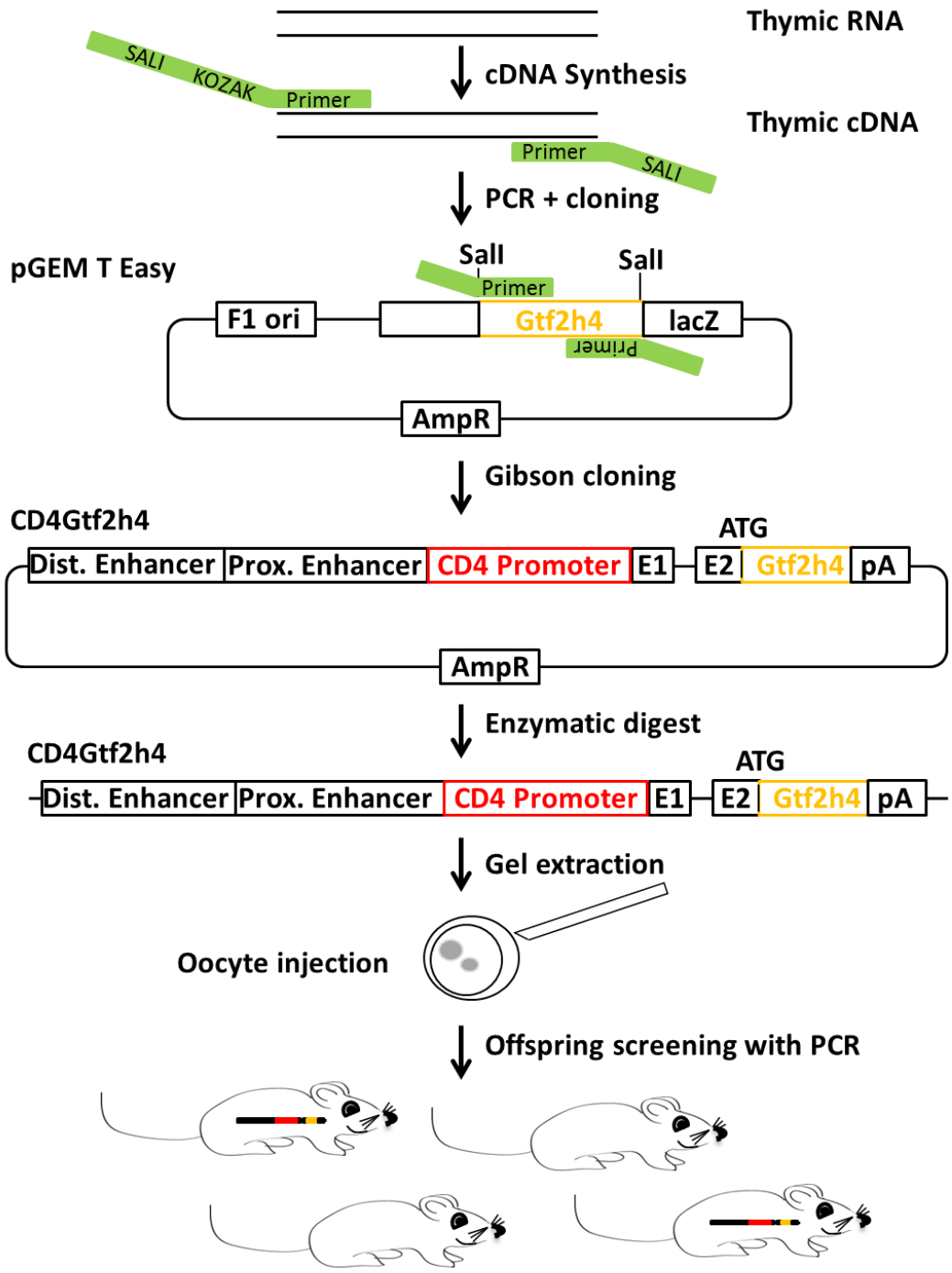


Figure 27 | Schematic illustration of the generation of CD4Gtf2h4 transgenic mice.
Total thymic wildtype C57Bl/6JOla RNA were used to transcribe thymic cDNA. Subsequently, the Gtf2h4 transcript (yellow) was amplified using a primer pair (green) that contains the Kozak sequence and two additional SalI digestion sites, one at each end. The PCR product was cloned into the pGEM-T Easy vector. Then the cDNA was amplified from the pGEM-T Easy Gtf2h4 with primers suitable for Gibson cloning and cloned into the linearized CD4 Depe vector containing the distal and the proximal enhancer, the CD4 promoter (red), exon 1 (E1) and parts of exon 2 (E2) but lacking the intronic silencer. CD4Gtf2h4 vector was digested with NotI. Next, the 6284 bp band was gel-purified and used for injection into the pronucleus of fertilized oocytes. Finally, the offspring were screened by PCR. F1 ori = phaged derived origin of replication, AmpR (ampicillin resistance), pA (poly A), ATG = start codon

3.7.2 Genotyping of CD4Gtf2h4 transgenic mice

After injection of the CD4Gtf2h4 vector into fertilized FVB/N oocytes the founder were genotyped for transgenesis by PCR using the primer pair Gtf2h4 TG fwd (5`-tgg tgg aga ttc tct cct tcc t-3`) and rev (5`-cag gcg tgt ggg gta gta ac-3`) and the following PCR mix as well as the temperature protocol.

Table 18 | **PCR Mix for the genotyping of the CD4Gtf2h4 transgenic mouse strain.**

Mix components	**µl**
GoTaq®Green Master Mix (Promega)	12.5
Primer fwd (10 mM)	0.5
Primer rev (10 mM)	0.5
H2O	10.5
Template DNA	1

Table 19 | **Temperature protocol for the PCR of the CD4Gtf2h4 transgenic mouse strain.**

Step	Temp. (°C)	Time (mm:ss)
1	95	04:00
2	95	00:30
3	63	00:30
4	72	00:40
2-4	35 cycles	
5	72	07:00
6	12	∞

The transgenic PCR product has a length of 170 bp compared to WT PCR product of 514 bp (Fig. 28). Five transgenic founders could be identified by PCR which were subsequently bred for further analysis. The newly generated mouse strain expressing Gtf2h4 under the CD4 promoter was named B6FVBN-Tg(CD4-Gtf2h4)ThBu or CD4Gtf2h4 Tg.

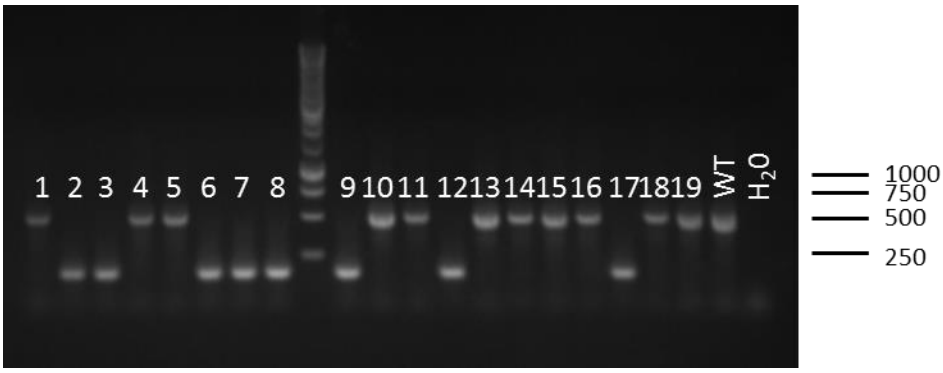


Figure 28 | **Agarose gel picture of the CD4Gtf2h4 TG genotyping.**
The PCR product of the TG with 170 bp was depicted in lane 2, 3, 6, 7, 8, 9, 12 and 17. The WT band with a size of 514 bp was shown in lane 1, 4, 5, 10, 11, 13, 14, 15, 16, 18 and 19. positive control = WT, negative control = H_20, ladder = 1kb, assigned in bp

3.8 Two transgenic CD4Gtf2h4 mouse strains with different expression profile

The five transgenic founders that had a TG genotype were subsequently bred for further analysis to verify the amount of the CD4Gtf2h4 mRNA expression. To ensure that CD4Gtf2h4 is specifically expressed in T cell containing organs nine different organs, namely, thymus, spleen, lymph nodes, liver, lung, heart, kidney, stomach and intestine were compared. As a control CD4 and CD8 mRNA expression was also assessed (Fig. 29A+B). The pups of two founders showed a moderate expression of CD4Gtf2h4 (Fig. 29C). CD4Gtf2h4 is highly expressed in spleen and LN of mice of the transgenic line 2 (TG 2) and lower expressed in the thymus. In contrast, CD4Gtf2h4 is less expressed in spleen and LN and it is slightly more expressed in the thymus of mice of the transgenic line 3 (TG 3). The other three lines did not show any CD4Gtf2h4 mRNA expression. Comparing the WT/TG genotype ratio of the pups of TG 2 and TG 3, TG 3 has a higher ratio than TG 2 (Fig. 30), indicating an unequal distribution of WT and TG 3 mice and assuming an advantage of WT mice over TG mice. In conclusion, out of five transgenic founders two were expressing CD4Gtf2h4, yet with a different expression profile.

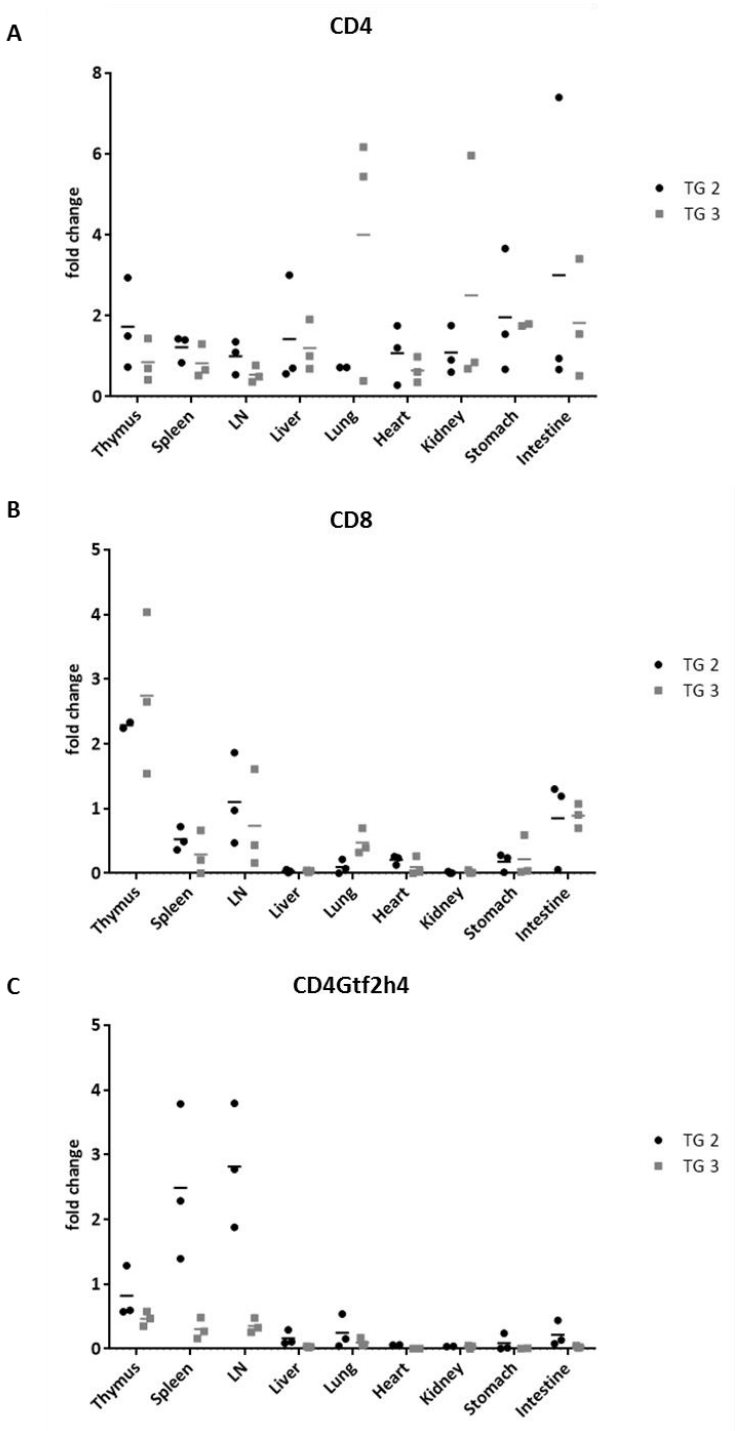


Figure 29 | mRNA expression profile of CD4, CD8 and CD4Gtf2h4 in nine different organs of two different transgenic lines.
Shown is the calculated fold change and mean of three independent experiments per organ, normalized to two reference genes (Ywhaz and Cxxc1) and WT littermates, of CD4 (A), CD8 (B) and CD4Gtf2h4 (C). Nine different organs were compared (thymus, spleen, lymph nodes (LN), liver, lung, heart, kidney stomach and intestine) in two different CD4Gtf2h4 transgenic line (TG 2 and TG 3).

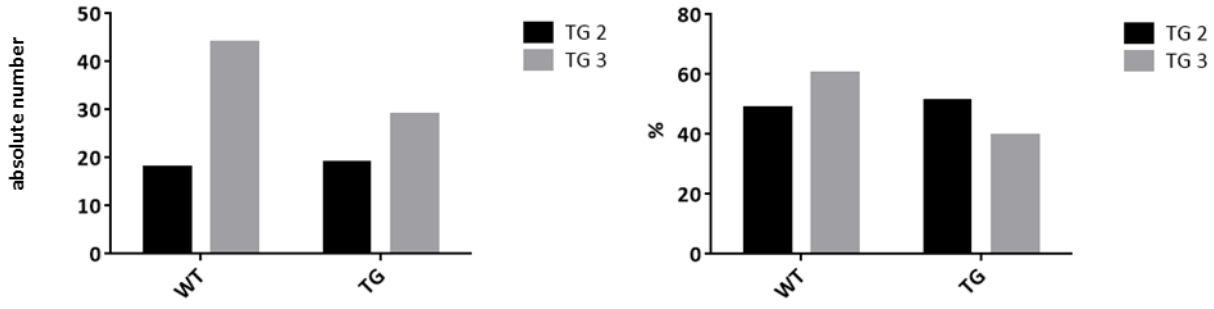


Figure 30 | **Number of mice per genotype of both TG mouse strains**
Number of mice per genotype and the percentage of both transgenic lines and control mice. The absolut number (left) of mice and the percentage (right) of both transgenic lines (TG2 and TG 3) compared to wildtype littermates is depicted in bar graphs.

3.9 Reverse cell numbers of CD4Gtf2h4 TG mice compared to Gtf2h4 HET deficient mice

To investigate the influence of the overexpression of Gtf2h4 in T cells regarding their development the number of cells in thymus, spleen and LN were counted first. The total cell number of spleen and LN did not differ among both transgenic lines. In contrast, the thymus of TG 2 showed a tendency of a higher number of cells compared to wild type littermates, whereas the opposite is the case for TG 3 (Fig. 31). To subdivide the cells further, spleen, LN and thymus of TG 2 and TG 3 were analyzed by flow cytometry with the same strategy as in 3.3 and 3.6. When analyzing the spleen of TG 2 and TG 3 none of the populations were found to be altered compared to wildtype littermates except for $Nk1.1^+$ and $CD11b^+$ cells in TG 3 (Fig. 32). Those two populations showed a slight decrease in the cell numbers compared to the cell populations of wildtype mice (Fig. 32A). In lymph nodes of TG 2 no change in the cell number of any cell population was observed (Fig. 33). However, in the LN of TG 3 $CD4^+$, $CD8^+$, $TCR\beta^+$, $CD19^+$, Tregs and $Nk1.1^+$ were slightly upregulated (Fig. 33A+B). In addition, the cell number of $CD4^+$ TNaive as well as $CD8^+$ TNaive was upregulated (Fig. 33C+D). With regard to the B cell numbers, a higher number of $B220^+$, FO, T1 and FM cells were detected (Fig. 33E). In the thymus, the absolute cell number of all populations of TG 2 was not altered compared to the wildtype except for $CD4^+$ cells, which are slightly more in the transgene (Fig. 34A). In the TG 3 mice DN, $CD4^+$, $CD8^+$, DP, DN4 as well as $CD19^+$, $Nk1.1^+$ and Treg cells are less present compared to the wildtype (Fig. 34).

In conclusion, several cell populations in the LN showed an upregulation in the TG 3 mice compared to the WT mice. In contrast, most of the investigated populations in the thymus were downregulated in terms of numbers. These alterations are reverse to those found in the Gtf2h4 HET mice.

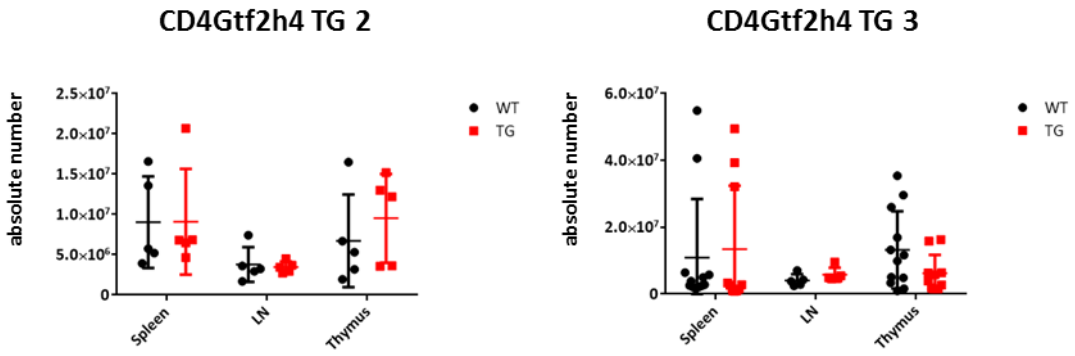


Figure 31 | **Comparison of the cell numbers of various organs of WT and TG mice.**
Depicted are the cell numbers, the calculated mean and SD of spleen, LN and thymus of TG 2 (left) and TG 3 (right) compared to wildtype control mice. TG 2: two to four independent experiments with WT (n=5-12) and TG (n=5-10) mice. TG 3: two independent experiments with WT (n=5) and TG (n=5).

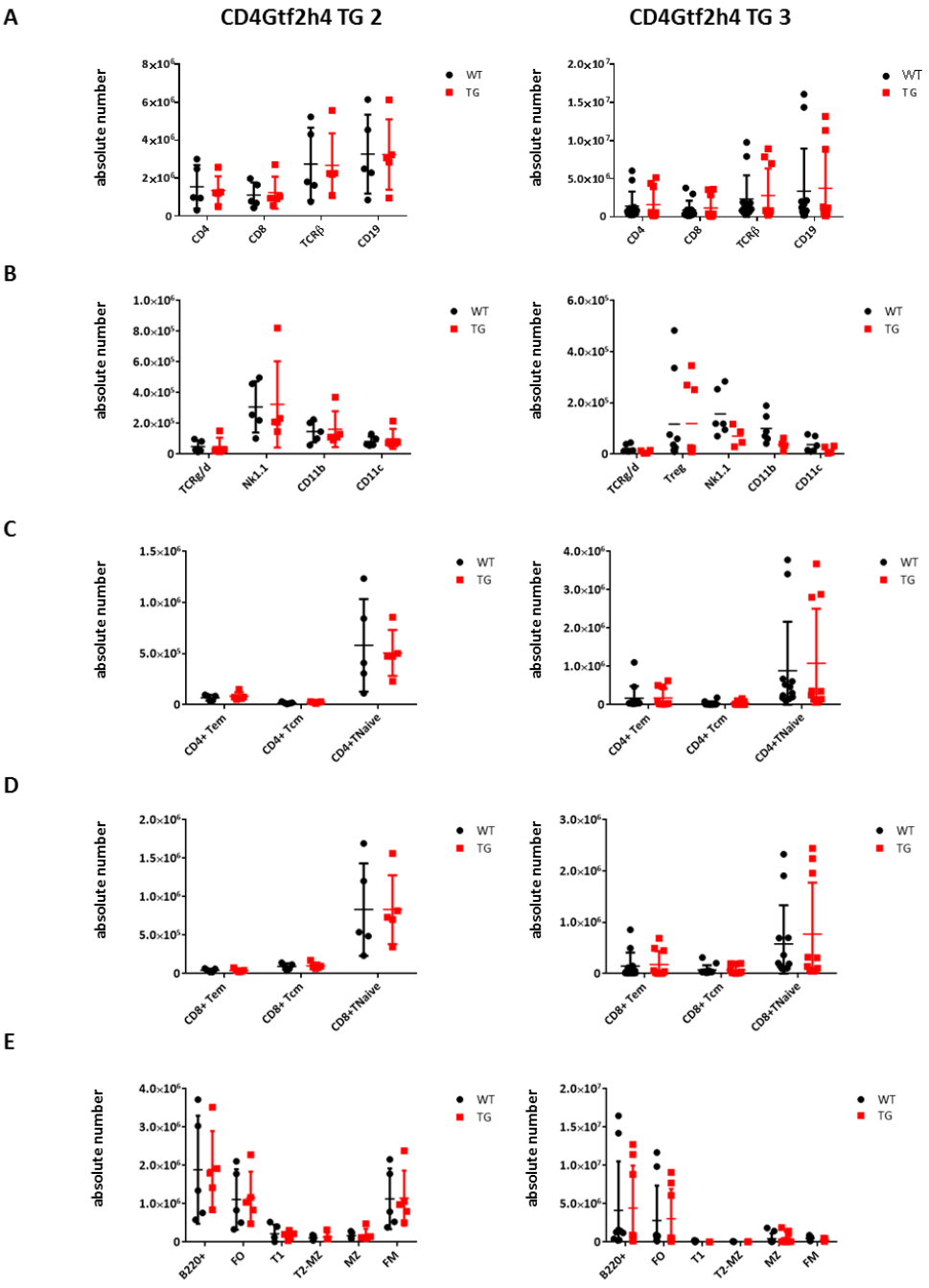


Figure 32 | **Cell numbers of different cell populations in the spleen of CD4Gtf2h4 TG 2 and TG 3 as well as wildtype control mice.**

Shown are the cell numbers of different cell populations (A) $CD4^+$, $CD8^+$, $TCR\beta^+$ and $CD19^+$ in the spleen of TG 2 (left) and TG 3 (right) as well as wildtype control mice. TG 2: two independent experiments with WT (n=5) and TG (n=5) mice. TG 3: four independent experiments with WT (n=12) and TG (n=10) mice. (B) $TCR\gamma\delta^+$, Tregs (only TG 3), $Nk1.1^+$, $CD11b^+$ and $CD11c^+$ in the spleen of TG 2 (left) and TG 3 (right) as well as wildtype control mice. TG 2: two independent experiments with WT (n=5) and TG (n=5) mice. TG 3: two independent experiments with WT (n=6) and TG (n=4) mice. Tem, Tcm and TNaive of $CD4^+$ (C) and $CD8^+$ (D) pre-gated cells in the spleen of TG 2 (left) and TG 3 (right) as well as wildtype control mice. TG 2: two independent experiments with WT (n=5) and TG (n=5) mice. TG 3: four independent experiments with WT (n=12) and TG (n=10) mice. (E) $B220^+$, FO, T1, T2-MZ, MZ and FM B cells in the spleen of TG 2 (left) and TG 3 (right) as well as wildtype control mice. TG 2: two independent experiments with WT (n=5) and TG (n=5) mice. TG 3: two to three independent experiments with WT (n=7-9) and TG (n=5-8) mice. Bars indicate the mean of each group, in case of n>5 SD is depicted.

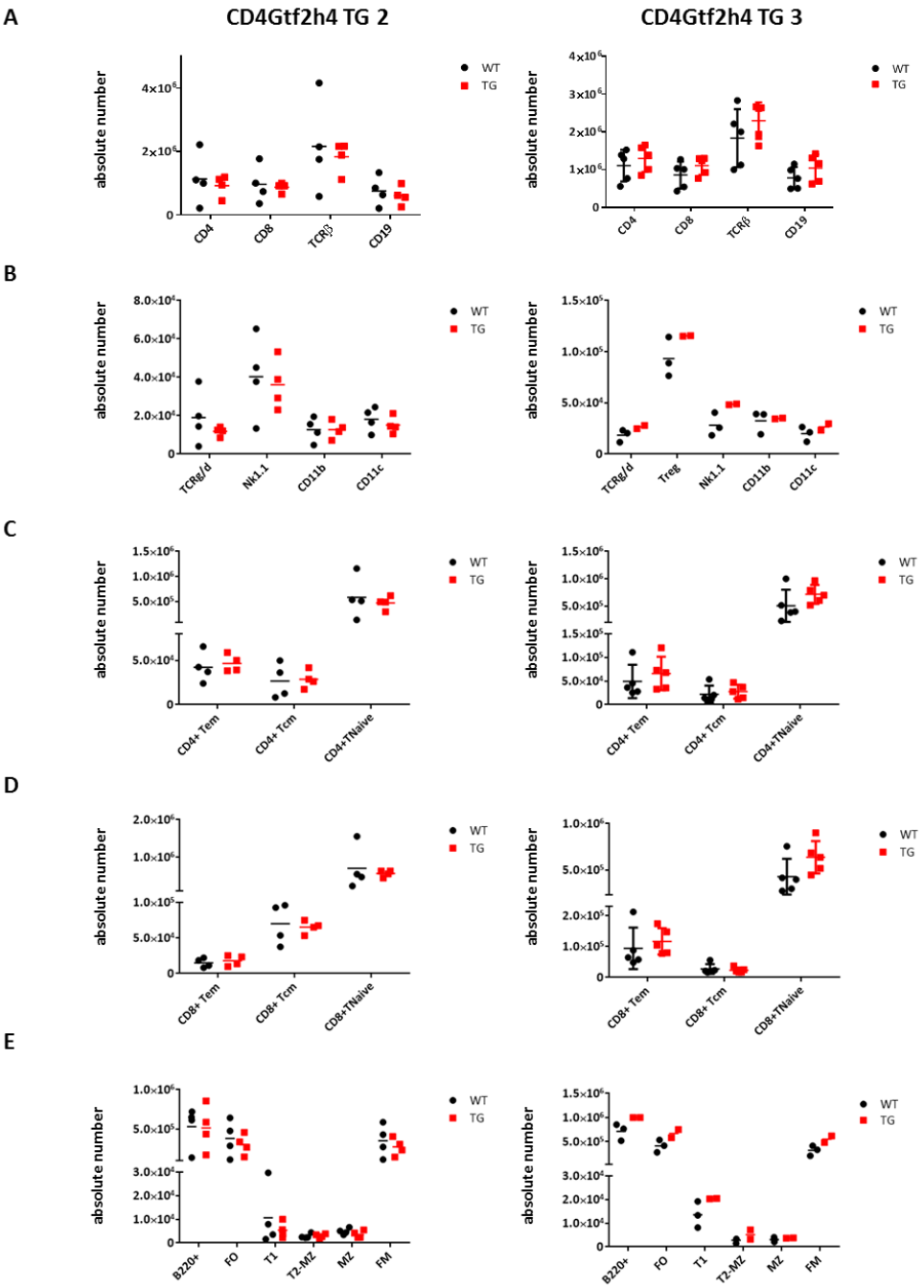


Figure 33 | Cell numbers of different cell populations in the lymph nodes of CD4Gtf2h4 TG 2 and TG 3 as well as wildtype control mice.

Shown are the cell numbers of different cell populations (A) $CD4^+$, $CD8^+$, $TCR\beta^+$ and $CD19^+$ in the LN of TG 2 (left) and TG 3 (right) as well as wildtype control mice. TG 2: two independent experiments with WT (n=4) and TG (n=4) mice. TG 3: two independent experiments with WT (n=5) and TG (n=5) mice. (B) $TCR\gamma\delta^+$, Tregs (only TG 3), $Nk1.1^+$, $CD11b^+$ and $CD11c^+$ in the LN of TG 2 (left) and TG 3 (right) as well as wildtype control mice. TG 2: two independent experiments with WT (n=4) and TG (n=4) mice. TG 3: one experiments with WT (n=3) and TG (n=2) mice. Tem, Tcm and TNaive of $CD4^+$ (C) and $CD8^+$ (D) pre-gated cells in the LN of TG 2 (left) and TG 3 (right) as well as wildtype control mice. TG 2: two independent experiments with WT (n=4) and TG (n=4) mice. TG 3: two independent experiments with WT (n=5) and TG (n=5) mice. (E) $B220^+$, FO, T1, T2-MZ, MZ and FM B cells in the LN of TG 2 (left) and TG 3 (right) as well as wildtype control mice. TG 2: two independent experiments with WT (n=4) and TG (n=4) mice. TG 3: one experiments with WT (n=3) and TG (n=2) mice. Bars indicate the mean of each group, in case of n>5 SD is depicted.

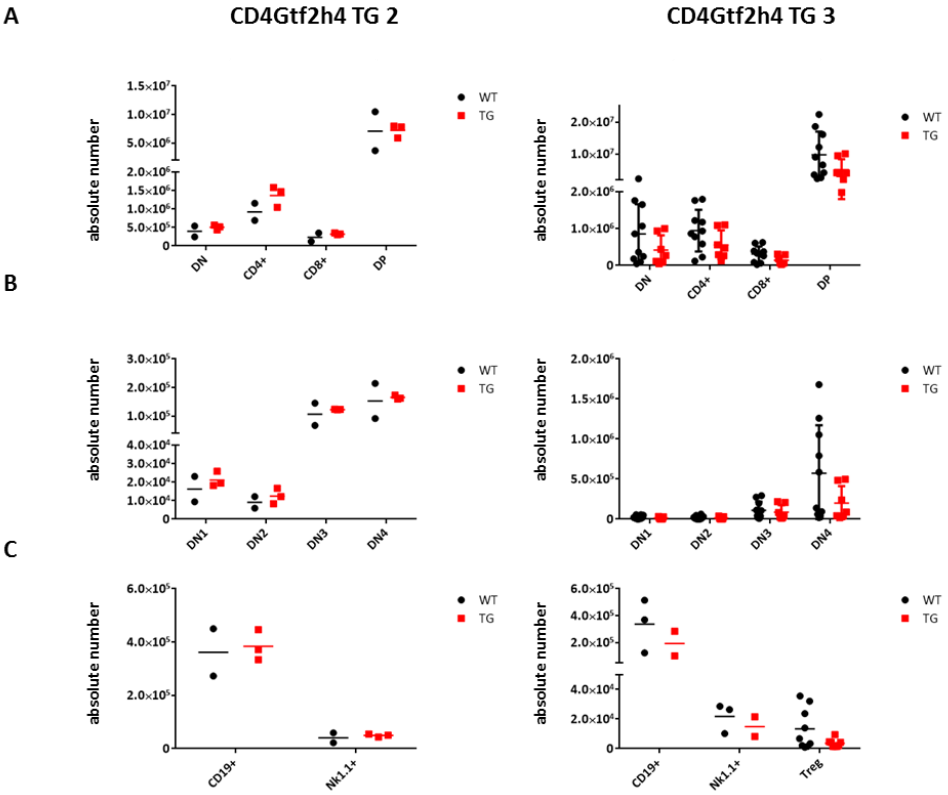


Figure 34 | **Cell numbers of different cell populations in the thymus of CD4Gtf2h4 TG 2 and TG 3 as well as wildtype control mice.**
Shown are the cell numbers of different cell populations (A) DN, $CD4^+$, $CD8^+$ and DP in the thymus of TG 2 (left) and TG 3 (right) as well as wildtype control mice. TG 2: one experiment with WT (n=2) and TG (n=3) mice. TG 3: three independent experiments with WT (n=10) and TG (n=7) mice. (B) DN1, DN2, DN3 and DN4 in the thymus of TG 2 (left) and TG 3 (right) as well as wildtype control mice. TG 2: one experiment with WT (n=2) and TG (n=3) mice. TG 3: three independent experiments with WT (n=10) and TG (n=7) mice. (C) $CD19^+$, $Nk1.1^+$, Tregs (only TG 3) in the thymus of TG 2 (left) and TG 3 (right) as well as wildtype control mice. TG 2: one experiment with WT (n=2) and TG (n=3) mice. TG 3: one to three independent experiment with WT (n=3-9) and TG (n=2-8) mice. Bars indicate the mean of each group, in case of n>5 SD is depicted.

3.10 Reduced T cell-rich areas in the histology of spleen and LN of CD4Gtf2h4 TG mice

To confirm the flow cytometric data and to get a closer look into the organ structure histology was performed. Thymus, spleen and lymph nodes of TG 2 and TG 3 as well as WT littermates were collected for histology and stained with H&E (blue) and anti-CD3 (brown) (Fig. 35+36). The experiment was repeated three times with one mouse each and one representative picture is shown for each TG line and control mice. For both transgenes a reduction of anti-CD3-stained area in the spleen and LN of the TG was observed compared to the WT littermate, which was not apparent in the thymus. By measuring the stained areas

and total area with ImageJ and calculating their ratio it was shown that for both transgenic lines the ratio of stained vs total area was lower in spleen and lymph nodes compared to the wildtype, indicating a reduced T cell number in the TG mice. In the thymus the ratio was lower in TG 2 but higher in TG 3 (Fig. 37). Analyzing the spleen in more detail it was observed that the T cell-rich areas in the transgenic mice were smaller than in wildtype littermates. Therefore, the diameter of these areas was measured using ImageJ and compared. Smaller T cell regions with regard to the diameter in the spleen of both transgenic lines compared to wildtype mice were found (Fig. 38). This observation was expressed as mean value of the diameter (n=3) and was smaller in both transgenic lines compared to control mice (Fig. 38B). In summary, overexpression of Gtf2h4 in T cells led in both TG lines to a decreased number of cells in spleen and lymph nodes and to a reduced diameter of T cell-rich areas in the spleen. However, the outcome in the thymus is different. In TG 2 it resulted in a lower number of cells compared to a higher number in the TG 3 line.

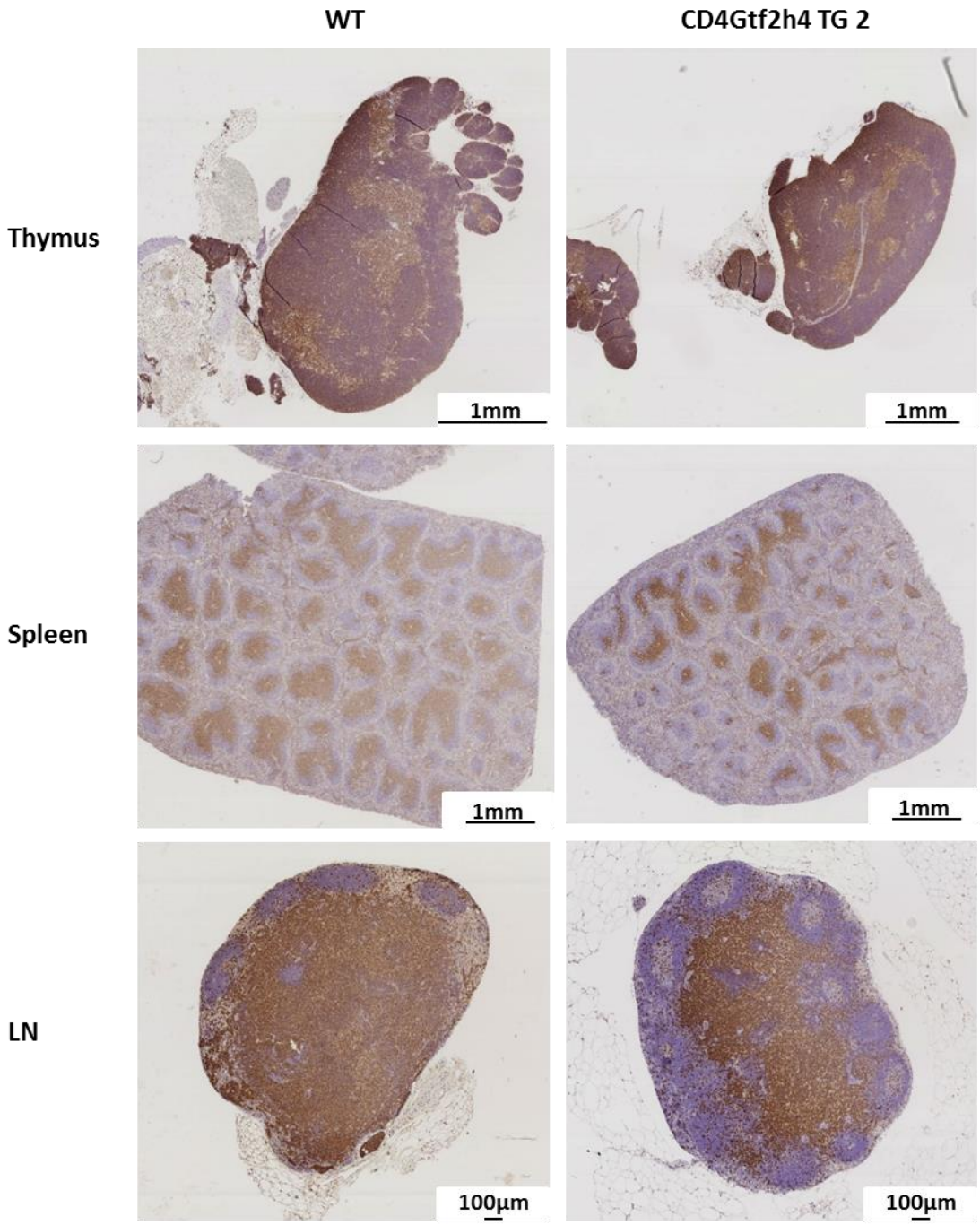


Figure 35 | Histology of thymus, spleen and LN of WT and TG 2 mice.
Thymus, spleen and LN of WT and TG 2 mice are stained with H&E (blue) and anti-CD3 (brown). Shown is one representative of two to three independent experiments each with n=1. The length of scale bare is depicted in each individual picture.

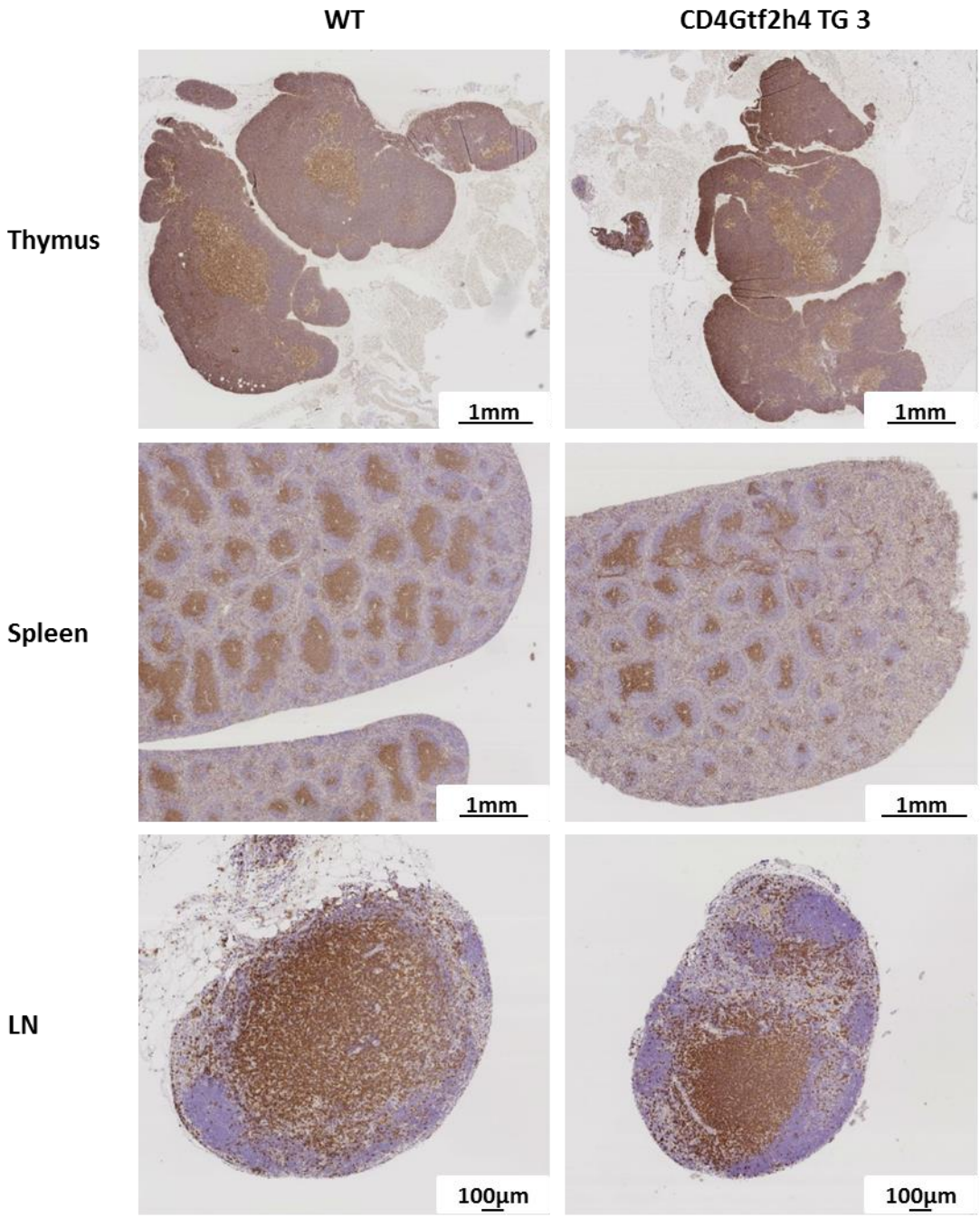


Figure 36 | Histology of thymus, spleen and LN of WT and TG 3 mice.
Thymus, spleen and LN of WT and TG 3 mice are stained with H&E (blue) and anti-CD3 (brown). Shown is one representative of two to three independent experiments each with n=1. The length of scale bare is depicted in each individual picture.

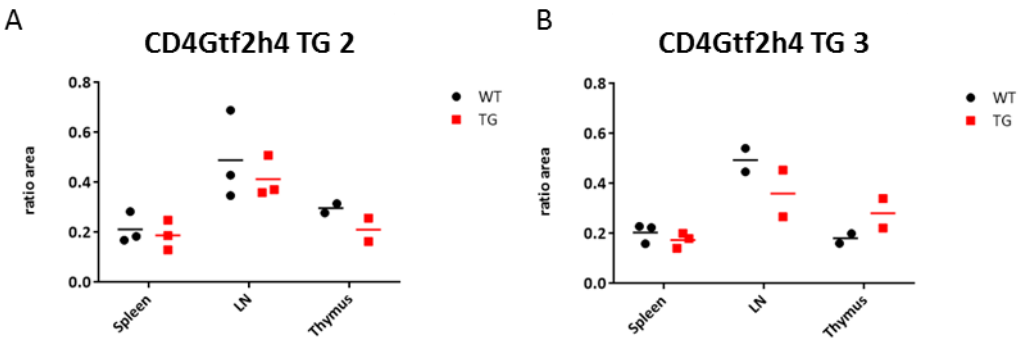


Figure 37 | Ratio of anti-CD3 stained vs total area of the TG mice compared to WT littermates.
Shown is the ratio of anti-CD3 stained area versus total area calculated with ImageJ of the single organ histology of TG 2 (left) and TG 3 (right) compared to WT littermates of spleen, LN and thymus. Two to three indepent experiments with n=1 each are depicted. Bars indicate the mean of each group.

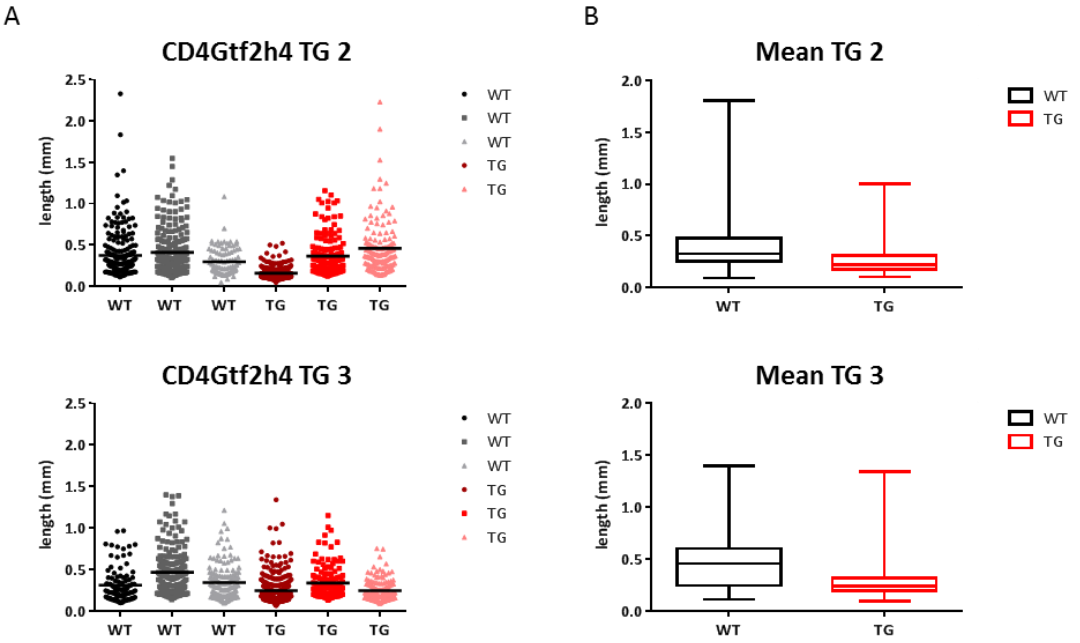


Figure 38 | Analysis of the diameter of the T cell-rich areas in the spleen of TG mice compared to WT littermates.
(A) The diameter of TG 2 (top) and TG 3 (bottom) as well as littermates WT mice were measured using ImageJ. Each mouse is depicted individually. Bars indicate the mean of each group. (B) Calculated mean and the min. and max. value of the diamter of TG 2 (top) and TG 3 (botton) as well as littermates WT mice is shown as a summary over all three experiments.

4 Discussion

4.1 Three promising candidates that could be involved in negative selection

To identify the transcription factor (TF) that is necessary to put the TCR signaling machine of DP thymocytes towards negative selection instead of proliferation that is induced in peripheral T cells, the Immgen database was screened. We hypothesized that this particular TF is highly expressed in the DP stage compared to the DN and SP immature and mature stages. The analysis revealed nine TF meeting the criteria, to which we added another two identified in a previous study of our group. All of them were subsequently verified by RT-PCR. With the purpose of selecting a TF for further analysis of it potential function in T cell development and negative selection a careful literature search for all eleven TFs was performed. Despite the finding that many of the initially selected TFs were not related to T cells, some single nucleotide polymorphism (SNP) analysis studies revealed their association to autoimmune diseases.

Nr1h2 also known as LXRβ for example is associated with late onset Alzheimer´s disease (Adighibe et al., 2006) and connected to impaired insulin secretion which may lead to type 2 diabetes (Ketterer et al., 2011). In addition, SNPs in the human Nr1h2 may predispose to preeclampsia (Mouzat et al., 2011). Furthermore, hypergylcemia in mice suppressed macrophage activity and aneurysmal degeneration by the activation of Nr1h2 (Tanaka et al., 2016). Nevertheless, since in our analysis the mRNA expression was not upregulated in the DP stages we decided to omit Nr1h2 from further analysis.

Another TF, Zfp703, has a function in promoting metastasis by the repression of E-cadherin expression (Slorach et al., 2011) and inhibits the Wnt/β-catenin activity. Therefore it prevents the mesoderm differentiation of embryonic stem cells (Kumar et al., 2016). Since Zfp703 was expressed in the DP and immature SP stages the expression profile did not correspond to the predicted profile and consequently it was not selected for further analysis.

Cnot6 (CCR4-NOT 6), a less studied TF, belongs to a huge transcriptional regulatory complex (Dupressoir et al., 2001) and may play a role in the differentiation of neuronal stem cells (Chen et al., 2011). However, its expression remained elevated throughout all stages and was not selected for deeper analysis.

Another potential TF that could influence the outcome of the TCR signaling was Gtf2ird1. Mice deficient in Gtf2ird1 show decreased fear and aggression potential and increased social

behaviour that was associated with the Williams-Beuren syndrom as a result of enhanced serotonin levels in several regions of the brain (Young et al., 2008). Nevertheless, no *in vivo* neuronal targets of Gtf2ird1 have been identified (O'Leary and Osborne, 2011). Furthermore, Gtf2ird1 expression regulates normal retinal function by binding to the promoter region of rhodopsin, M- and S-opsin genes in the mouse (Masuda et al., 2014). SNP analysis of a Chinese Han population showed a correlation to systemic lupus erythematosus (SLE) (Li et al., 2015). Although there might be a connection to SLE, Gtf2ird1, when verified by RT-PCR, Gtf2ird1 did not show the hypotesized mRNA expression profile since it was also expressed in some SP immature thymocyte stages.

In summary, four out of eleven TF (Nr1h2, Zfp703, Cnot6 and Gtf2ird1) that were initially selected based on the Immgen database screening could not be verified by RT-PCR. Supported by the absence of literature indicating a potential role of these TF in T cell development they were ommitted from further analysis.

The mRNA expression profile of the remaining selected eleven TFs showed a peak expression in the DP stages compared to less expression in DN and SP stages, some more distinctly than the other and were therfore in line with our hypothesis.

Leucine zipper transcription factor-like 1 (Lztfl1), was suggested as a prognostic marker for gastric cancer. It is higly expressed in epithelial cells of normal tissue, but downregulated in gastric cancer, with expression levels correlated to survival (Wei et al., 2010). In addition, Lztfl1 can bind to β-catenin in the cytoplasm, inhibiting its nuclear translocation in gastric cancer cells, and leading to supression of cell migration (Wang et al., 2014). Interestingly, in human CD4 T cells treated with all-trans retinoic acids Lztfl1 is upregulated and concentrated on the APC contact site while immunological synapse formation takes place. Furthermore, its overexpression enhanced the NFAT signaling induced by the TCR (Jiang et al., 2016). Lztfl1 showed the hypotesized expression profile of a high expression in the DP stage compared to a less expression in the DN and SP stages, however we excluded Lztfl1 from further analysis since a collegue was already studying the involment of Lztfl1 in T cell development.

Another TF is Tcf12, which is also known as HEB and has several known functions in T and B cell development and is mostly functional as a hetero- or homodimer. HEB, together with E2-2, is necessary for normal pro-B cell development (Zhuang et al., 1996), while together with E12 it is able to block the CD4 enhancer activity of the CD4 promoter (Sawada and Littman, 1993). Experiments with T cell-specific transgenic mice expressing Id1, the inhibitor for HEB

and other E-proteins, showed an enhanced apoptosis of differentiated thymocytes with rearranged TCRs (Kim et al., 1999). Furthermore, two E-box elements were found in the pTα promoter allowing the HEB and E2A heterodimer binding resulting in the activation of the promoter (Takeuchi et al., 2001; Tremblay et al., 2003). Moreover, HEB is necessary for the rearrangement of the TCRA, TCRD and TCRG but not TCRB locus (D'Cruz et al., 2010; Langerak et al., 2001). Thymocytes deficient for HEB presented an impaired survival and development of the invariant NKT cells, as well as a block in the TCRα rearrangement (Barndt et al., 1999; D'Cruz et al., 2010). In conclusion, although a large body of evidence exists for the role of Tcf12 (HEB) in B and T cell development, no fundamental role in negative selection was shown. Consequently, Tcf12 was not selected for further investigations.

Pou6f1, also known as Brn-5 or TCFβ_1, has a role in the developing and adult central nervous system (CNS) of the mouse (Cui and Bulleit, 1998; Okamoto et al., 1993). Furthermore, it is able to bind the TCRβ enhancer (Messier et al., 1993) and can be phosphorylated by JNK (Kasibhatla et al., 1999). Pou6f1 showed a high mRNA expression at the DP stages and was subsequently downregulated, but expressed again in mature SP stages. Because of the drastic reduction it was chosen for further investigation regarding its function in T cell development and negative selection. However, the group of Chance J. Luckey, Harvard Medical School, studied the role of Pou6f1 due to its high expression in DP thymocytes. Contrary to the expectations, Pou6f1 had no fundamental role in T cell development nor negative selection (Carrasco-Alfonso et al., 2013).

E2F2 that belongs together with E2F7 and E2F8 to the E2F gene family, plays a fundamental role in a normal T cell development. E2F2$^{-/-}$ mice developed late onset autoimmune disease because of an increased proliferation triggered through a TCR signaling with a lower activation threshold (Murga et al., 2001). Our data support the role of E2F2, as the mRNA expression was high in the DP stage and low in the other stages. E2F2 was not selected because of its known function.

In conclusion, Cnot6, Gtf2ird1, Nr1h2 and Zfp703 did not show the mRNA expression profile corresponding to our hypothesis i.e. a low expression in the DN stages, an expression peak in the DP stages and a transiently reduction in the SP stages of the thymus and spleen and therefore were not selected. E2F2, Lztfl1, Pou6f1 and Tcf12 showed the hypothesized mRNA expression profile, but were excluded from further studies because of their already

discussed function in T cell development. Therefore E27, E2F8 and Gtf2h4 were chosen for further studies.

4.2 E2F7 and E2F8 do not play a role in T cell development and negative selection

Because of the fundamental role of the E2F family of transcription factors in apoptosis, proliferation and cell differentiation, E2F7 and E2F8 were two encouraging candidates for being involved in the modification of the apoptotic pathway in DP thymocytes, initiated by the TCR signal leading to negative selection instead of proliferation and differentiation in mature T cells. Other members of this gene family even have already been described for their function in T cell development. For example, $E2F1^{-/-}$ mice developed normally, but showed an expansion of mature $CD4^{+}$ and $CD8^{+}$ thymocytes as well as peripheral T cells resulting from a thymic hyperproliferation that takes place mainly in the cortex. These findings led to the assumption of an important role of E2F1 in promoting apoptosis and suppressing proliferation (Field et al., 1996) as well as an involvement in the p53 pathway (Pan et al., 1998). Investigating the role of E2F1 in negative selection showed that thymocytes of $E2F1^{-/-}$ mice had a defect in the TCR-mediated negative selection (Garcia et al., 2000). As already mentioned, E2F2 also plays a role in the maintenance of normal T cell proliferation through the influence of the TCR activation threshold (Murga et al., 2001). The effect of E2F1 and E2F2 on the TCR activation threshold in single KO mice could be even increased as shown in the double KO mice (Zhu et al., 2001).

However, E2F7 and E2F8, two other E2F family members, have not been shown to have a role in T cell development. Nevertheless, they have a critical function in cell survival and embryonic development as it was shown in the corresponding KO mice. The deficiency of E2F7 and E2F8 led to an increased expression of E2F1 and p53 resulting in major apoptosis (Li et al., 2008; Zalmas et al., 2008). Those findings suggest that E2F7 and E2F8 are important for T cell development and negative selection and both were potential candidates resulting from our Immgen database screen. The RT-PCR analysis revealed that these two E2F gene family members were differently expressed. While E2F7 showed an expression peak in the DP1 stages, the mRNA expression of E2F8 remained elevated to some extent throughout all stages. These results suggest E2F7 to be a perfect candidate for further analysis in line with our hypothesis that a TF that is important for negative selection should be highly expressed

in DP thymocytes and less expressed in the DN stages as well SP immature and mature stages of the thymus and peripheral secondary organs. The mRNA E2F8 expression profile did not perfectly correspond to our hypothesis because it was expressed in all stages and did not show an expression peak. Nevertheless, because of its unknown role regarding T cell development we decided to include both TF for further analysis.

Astonishingly, neither Lck-cre E2F7$^{fl/fl}$ E2F8$^{fl/fl}$ nor CD4-cre E2F7$^{fl/fl}$ E2F8$^{fl/fl}$ had a differentially regulated T cell population in thymus or spleen compared to control mice. In addition, also the cell cycle analysis revealed no differences. Therefore, it can be concluded that E2F7 and E2F8 do not play a fundamental role in T cell development and negative selection.

4.3 Gtf2h4 – a small but important switch downstream of the TCR signaling

Another promising candidate according to our hypothesis that a TF which is important for negative selection in thymocytes should be highly expressed in the DP stages but less expressed in the DN and SP stages was Gtf2h4 (p52). It belongs to the TFIIH complex that is involved in NER (Drapkin et al., 1994; Seroz et al., 1995), promoter escape (Kugel and Goodrich, 1998; Spangler et al., 2001) and cell cycle regulation (Dynlacht, 1997; Matsuno et al., 2007). So far nothing is known about the influence of Gtf2h4 on the T cell development. Interestingly, human SNP analyses of 1343 Multiple Sclerosis (MS) cases and 1379 healthy volunteers resulted in a significant association of one SNP of Gtf2h4, located in intron 11, with MS susceptibility (Briggs et al., 2010).

Our mRNA expression analysis of Gtf2h4 throughout all stages of T cell development resulted in a peak expression at the DP stage that was drastically reduced thereafter. However, Gtf2h4 was expressed again in more mature SP thymocytes as well as peripheral T cells. Similarly to Pou6f1, it was included for further analysis because of its drop in expression after the DP stages.

To analyze the influence of Gtf2h4 on T cell development and especially negative selection, two mouse models were generated, i.e. one deficient for Gtf2h4 and one overexpressing it under the control of the CD4 promoter. We hypothesized that a transcription factor, which is highly expressed in the DP stage of thymocytes, could be necessary to modify the apoptotic pathway in double positive thymocytes such that the TCR signal leads to negative selection instead of proliferation and differentiation as it does in mature T cells. In theory that means that the KO mouse for this particular TF (TF$^{-/-}$) should show less negative selection

(apoptosis) in thymocytes that leads to more DP thymocytes and a higher total number of cells in the thymus. In contrast, mature peripheral T cells of the theoretical $TF^{-/-}$ mouse model, engaged with the same TCR signaling, should induce more apoptosis and should have a lower proliferation rate leading to a lower number of total cells in the secondary lymphoid organs. In case of the overexpression of the TF under the control of the CD4 promoter one would expect the opposite effect with enhanced negative selection in the thymus and therefore less DP thymocytes and less total number of cells. Furthermore, the TCR signal of peripheral T cells would then lead to a reduction of apoptosis and an increase in proliferation and total cell number (Tab. 20).

Table 20 | **The theoretical outcome of the TCR signal in thymocytes and peripheral T cells in the KO and overexpression mouse model.** ↑ upregulation ↓ downregulation

		$TF^{-/-}$	CD4-TF
Thymocytes	negative selection	↓	↑
	DP cells	↑	↓
	total cell number	↑	↓
mature peripheral T cells	apoptosis	↑	↓
	proliferation	↓	↑
	total cell number	↓	↑

Mice deficient for Gtf2h4 showed a reduction of the total cell number in spleen and LN as well as an increase in the thymus, matching our hypothesis that negative selection is reduced and apoptosis of peripheral T cells is increased. The investigation of the DN subpopulations revealed that the number of cells of DN3 and DN4 were higher in the KO compared to WT, whereas the cell numbers of DN1 and DN2 were equal in the KO and WT controls. This finding could be due to the impaired start of negative selection in these stages in the KO mouse model. Furthermore, the DP and SP $CD4^{+}$ populations were slightly increased which is a further indication of less negative selection matching the hypothesis. Interestingly, also the Treg population appeared to be marginally upregulated. This might be explained by two different events implemented in the main process of negative selection, first the clonal deletion and second the upregulation of Foxp3 in self-reactive T cells (Hu et al., 2016). If in Gtf2h4-deficient mice negative selection occurs less efficiently, the number of

self-reactive thymocytes would increase that may lead to more thymocytes converting into Treg cells.

In contrast to the increasing number of DN3, DN4, DP and SP $CD4^+$ cells, the $CD19^+$ population slightly decreased. However, it was shown before that there might be a reverse effect of thymic B cells influencing thymocytes and negative selection (Fujihara et al., 2014; Tokoro et al., 1998).

Comparing the absolute cell number of several investigated T cell populations of the thymus with those of spleen and LN the data showed reciprocal cell counts in the lymph nodes while being unaffected in the spleen. This could be due to the priming of naïve T cells. During the development of autoimmune diseases it was shown that the priming of naïve T cells takes place in the lymph nodes but not spleen (Gagnerault et al., 2002; Phillips et al., 1997).

In addition to the Gtf2h4-deficient mouse model, we also overexpressed Gtf2h4 under the control of the CD4 promoter. Two transgenic lines were obtained (TG 2 and TG 3) which showed different amounts of CD4Gtf2h4 mRNA expressed in various organs. Due to the random integration of the transgene altered levels of expression of the transgene are known. Therefore, it is not surprising that the analysis of the cell populations in two Gtf2h4 overexpressing mouse strains resulted in different outcomes. While the total amount of some populations of LN and thymus of TG 3 had altered numbers compared to WT littermates, the cell numbers of various investigated cell population of TG 2 did not show differences. The cell counts of subpopulations in the thymus DN1-DN3 of TG 3 were not affected. However, those of DN4 showed a trend towards reduction in the thymus of the overexpression model compared to the WT control mice. DN4 is the stage were the CD4 mRNA expression started as it could be shown with the sorting experiment comparing 20 stages of T cell development. Furthermore, the number of cells of DP, SP $CD4^+$ and $CD8^+$ was reduced as well which might be due to enhanced negative selection. Surprisingly, the total cell count of $CD19^+$ cell was also decreased.

Furthermore, most of the T and B cell populations in the LN of TG 3 showed a slightly increased cell count, providing evidence for the presence of mature T cells with a reduced apoptotic capability. This should be associated with a higher number of total cells, however it was not the case in either spleen or LN. In addition, the total number of cells in the thymus, which should be reduced according to the hypothesis, remained unaltered. However, the immunohistochemistry showed a reduction of the ratio of stained T cell area

versus total area in spleen and LN and an increase in the thymus which contradicts the flow cytometric data. Furthermore, also the size of the T cell-rich areas in the spleen decreased in the transgenic mouse models compared to WT littermates.

In summary, although preliminary, the data indicate that Gtf2h4 may play a role in the outcome of the TCR signaling. Most of the results support the original hypothesis that Gtf2h4 modifies the apoptotic pathway in double positive thymocytes such that the TCR signal leads to negative selection instead of proliferation and differentiation in mature T cells. The most obvious indication is the opposite outcome within one particular mouse model comparing thymus and peripheral lymphoid organs. Nevertheless, further experiments have to be performed to obtain detailed information about the influence of Gtf2h4 on the outcome of the TCR signaling in thymocytes and mature T cells. To investigate the negative selection in more detail, flow cytometric analysis of the DN and DP thymocytes has been developed to discriminate the state of the cells by the use of antibodies against CD69, PD-1, Helios and CCR7 (Baldwin and Hogquist, 2007; Daley et al., 2013; Hu et al., 2016). Moreover, both mouse models should be crossed to mouse models where changes in negative selection are easily detectable. For instance, the TCR transgenic HY mouse model could be used, in which T cells-specific for the male antigen (H-Y) are deleted in males (Kisielow et al., 1988). In addition, apoptosis and proliferation assays of thymocytes and peripheral T cells should be performed. Since the Gtf2h4 KO mice are lethal it would be useful to study the loss of Gtf2h4 in a cell type-specific manner. Finally, a reporter luciferase assay should be considered to investigate if Ap-1 can regulate Gtf2h4 in thymocytes and peripheral T cells in mice. AP-1 is expressed in DP thymocytes and plays a role in thymic selection processes (Rincon and Flavell, 1996). It was shown that in human and zebrafish Ap-1 is one of the factors which have the highest binding site frequency in the sequence of Gtf2h4 (Silva et al., 2014).

4.4 Conditional gene targeting - CRISPR/Cas9 the all-round solution?

Over the last decades the method of choice with regard to conditional gene targeting was the Cre/loxP recombination system. Cre, when expressed by a specific promoter, leads to a cell type- or tissue-specific deletion or inversion of the allele flanked by loxP sites (Hoess et al., 1982; Schwenk et al., 1998; Sternberg and Hamilton, 1981). Flanking the gene of interest with two loxP sites can be achieved through transfection of a targeting vector into ES cells

that is very time consuming. Furthermore, a high number of mice is necessary because of the back crossing into the background strain of interest as well as crossing with the particular Cre mouse strain. In the case of Gtf2h4, which is embryonic lethal in homozygous mice, it would be of great interest to investigate its function in a T cell-specific manner. In this case Lck-cre and CD4-cre mice (Lee et al., 2001; Zhang et al., 2005) would have to be crossed with mice having one or several exons of Gtf2h4 flanked with loxP. This would result in T cell-specific KO mice beginning at stage DN3 (Lck-cre) and DP (CD4-cre). To finally obtain the cell type-specific mouse of interest a lot of mice are only used for breeding purpose. According to the 3Rs principles (replace, reduce and refine) (Russell and Burch, 1959) it should be, among other things, the responsibility of the researchers to minimize the mouse numbers. Therefore, we modified the available CRISPR/Cas9 system (Cong et al., 2013; Jinek et al., 2013; Mali et al., 2013) to obtain T cell-specific KO mice.

To prove this concept, Cas9 was cloned under the control of the CD4 promoter and the gRNA, under the expression of U6, was designed to target exon 2 of CD2. CD2 was selected because it was already known that mutations in this gene do not influence cell viability (Killeen et al., 1992). Furthermore, CD2 is easy to detect by flow cytometry. Surprisingly, this conditional Cas9 approach targeting CD2 resulted in only 1% loss of CD2 in $CD4^+$ cells and 0.6% in $CD8^+$ cells of spleen and LN that could be verified by sequencing (Beil-Wagner et al., 2016). However, by targeting one or several genes in drosophila and zebrafish others have shown that a mutation rate of up to 90% can be achieved (Xue et al., 2014; Yin et al., 2015). Furthermore, just recently it has been shown that oocyte-specific expression of Cas9 under the Zp3 promoter works efficiently (Zhang et al., 2016).

In the experiments presented here, we showed Cas9 expression on mRNA level but we were not able to detect Cas9 protein by western blot. Therefore, this might be the explanation for the low mutation efficiency of CD2 in CD4 positive T cells. To improve the T cell-specific CRISPR/Cas9 system we devised a two-step plan. Firstly, screen thoroughly by RT-PCR and western blot the founder mice immediately after injection of only CD4Cas9, to choose the founder with the highest CD4Cas9 expression on protein level. Furthermore, a GFP will be fused to Cas9 with the purpose of simplifying the possibility to track the cells with flow cytometry and to measure their mutation rate easier. Secondly, instead of using one gRNA to target CD2 three different gRNAs, targeting the same exon, will be cloned on one plasmid to increase the mutation potential (Beil-Wagner et al., 2016). Most likely, this will optimize

the system that it can be used for cell type-specific targeting of the gene of interest which saves time and reduces mouse numbers because of the direct injection of the targeting vector into fertilized oocytes of the strain of interest. Even if the system cannot be improved to achieve efficiencies like those of the Cre/loxP system, it could be applied to address specific questions in, for instance, cancer research, where a conditional KO with a minor efficiency represents a much more physiological model.

Importantly, when working with the CRISPR/Cas9 system the off-target activity should not be disregarded since off-targets with up to five mismatches compared to the on-target are tolerated (Fu et al., 2013; Lin et al., 2014). While using the system to target a gene of interest in the complete mouse any potentially off-target is most likely crossed out in the F1 generation. In the transgenic conditional approach off-targets should be considered in theory. Nevertheless, in the presented T cell-specific CRISPR/Cas9 system off-target effects were not considered since the mutation efficiency in our system was very low. In case the system can be optimized off-targets should be taken into account. But, it was shown that by the use of the inactive form of Cas9 (dCas9) (Shen et al., 2014) or by a high-fidelity variant of Cas9 (SpCas9-HF1) (Kleinstiver et al., 2016) they can be minimized.

Taken together, the CRISPR/Cas9 system revolutionized the way of generating genetically modified mice and made their generation much easier and faster. Furthermore, it appears that the CRIPSR/Cas9 system is very easy to modify to address specific research questions as it was shown by our conditional approach targeting CD2 in CD4 positive T cells. Moreover, using the CRIPSR/Cas9 system supports the reduction of mouse numbers to obtain the genetically modified mouse of interest.

5 References

Adighibe, O., Arepalli, S., Duckworth, J., Hardy, J., Wavrant-De Vrieze, F., 2006. Genetic variability at the LXR gene (NR1H2) may contribute to the risk of Alzheimer's disease. Neurobiol Aging 27, 1431-1434.

Baldwin, T.A., Hogquist, K.A., 2007. Transcriptional analysis of clonal deletion in vivo. Journal of immunology 179, 837-844.

Barndt, R., Dai, M.F., Zhuang, Y., 1999. A novel role for HEB downstream or parallel to the pre-TCR signaling pathway during alpha beta thymopoiesis. Journal of immunology 163, 3331-3343.

Bassett, A.R., Tibbit, C., Ponting, C.P., Liu, J.L., 2013. Highly efficient targeted mutagenesis of Drosophila with the CRISPR/Cas9 system. Cell Rep 4, 220-228.

Beil-Wagner, J., Dossinger, G., Schober, K., Vom Berg, J., Tresch, A., Grandl, M., Palle, P., Mair, F., Gerhard, M., Becher, B., Busch, D.H., Buch, T., 2016. T cell-specific inactivation of mouse CD2 by CRISPR/Cas9. Scientific reports 6, 21377.

Bertani, G., 1951. STUDIES ON LYSOGENESIS I. : The Mode of Phage Liberation by Lysogenic Escherichia coli. Journal of bacteriology 62, 293-300.

Bhaya, D., Davison, M., Barrangou, R., 2011. CRISPR-Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. Annual review of genetics 45, 273-297.

Briggs, F.B., Goldstein, B.A., McCauley, J.L., Zuvich, R.L., De Jager, P.L., Rioux, J.D., Ivinson, A.J., Compston, A., Hafler, D.A., Hauser, S.L., Oksenberg, J.R., Sawcer, S.J., Pericak-Vance, M.A., Haines, J.L., Barcellos, L.F., International Multiple Sclerosis Genetics, C., 2010. Variation within DNA repair pathway genes and risk of multiple sclerosis. American journal of epidemiology 172, 217-224.

Brinster, R.L., Chen, H.Y., Trumbauer, M., Senear, A.W., Warren, R., Palmiter, R.D., 1981. Somatic expression of herpes thymidine kinase in mice following injection of a fusion gene into eggs. Cell 27, 223-231.

Brownlie, R.J., Zamoyska, R., 2013. T cell receptor signalling networks: branched, diversified and bounded. Nature reviews. Immunology 13, 257-269.

Buch, S.C., Diergaarde, B., Nukui, T., Day, R.S., Siegfried, J.M., Romkes, M., Weissfeld, J.L., 2012. Genetic variability in DNA repair and cell cycle control pathway genes and risk of smoking-related lung cancer. Molecular carcinogenesis 51 Suppl 1, E11-20.

Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L., Vandesompele, J., Wittwer, C.T., 2009. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clin Chem 55, 611-622.

Calnan, B.J., Szychowski, S., Chan, F.K., Cado, D., Winoto, A., 1995. A role for the orphan steroid receptor Nur77 in apoptosis accompanying antigen-induced negative selection. Immunity 3, 273-282.

Carrasco-Alfonso, M.J., Jiang, W., Luckey, J., 2013. Pou6f1 is dispensable for mouse T cell development and virus-specific memory CD8+ T cell generation. Frontiers in immunology.

Ceredig, R., Rolink, T., 2002. A positive look at double-negative thymocytes. Nature reviews. Immunology 2, 888-897.

Cermak, T., Doyle, E.L., Christian, M., Wang, L., Zhang, Y., Schmidt, C., Baller, J.A., Somia, N.V., Bogdanove, A.J., Voytas, D.F., 2011. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic acids research 39, e82.

Chang, N., Sun, C., Gao, L., Zhu, D., Xu, X., Zhu, X., Xiong, J.W., Xi, J.J., 2013. Genome editing with RNA-guided Cas9 nuclease in zebrafish embryos. Cell research 23, 465-472.

Chen, B., Gilbert, L.A., Cimini, B.A., Schnitzbauer, J., Zhang, W., Li, G.W., Park, J., Blackburn, E.H., Weissman, J.S., Qi, L.S., Huang, B., 2013. Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. Cell 155, 1479-1491.

Chen, C., Ito, K., Takahashi, A., Wang, G., Suzuki, T., Nakazawa, T., Yamamoto, T., Yokoyama, K., 2011. Distinct expression patterns of the subunits of the CCR4-NOT deadenylase complex during neural development. Biochemical and biophysical research communications 411, 360-364.

Chen, S., Lee, B., Lee, A.Y., Modzelewski, A.J., He, L., 2016. Highly Efficient Mouse Genome Editing by CRISPR Ribonucleoprotein Electroporation of Zygotes. The Journal of biological chemistry.

Cheng, A.W., Wang, H., Yang, H., Shi, L., Katz, Y., Theunissen, T.W., Rangarajan, S., Shivalila, C.S., Dadon, D.B., Jaenisch, R., 2013. Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. Cell research 23, 1163-1171.

Cho, S.W., Lee, J., Carroll, D., Kim, J.S., Lee, J., 2013. Heritable gene knockout in Caenorhabditis elegans by direct injection of Cas9-sgRNA ribonucleoproteins. Genetics 195, 1177-1180.

Christensen, J., Cloos, P., Toftegaard, U., Klinkenberg, D., Bracken, A.P., Trinh, E., Heeran, M., Di Stefano, L., Helin, K., 2005. Characterization of E2F8, a novel E2F-like cell-cycle regulated repressor of E2F-activated transcription. Nucleic acids research 33, 5458-5470.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., Zhang, F., 2013. Multiplex genome engineering using CRISPR/Cas systems. Science 339, 819-823.

Crispe, I.N., Moore, M.W., Husmann, L.A., Smith, L., Bevan, M.J., Shimonkevitz, R.P., 1987. Differentiation potential of subsets of CD4-8- thymocytes. Nature 329, 336-339.

Cui, H., Bulleit, R.F., 1998. Expression of the POU transcription factor Brn-5 is an early event in the terminal differentiation of CNS neurons. J Neurosci Res 52, 625-632.

D'Cruz, L.M., Knell, J., Fujimoto, J.K., Goldrath, A.W., 2010. An essential role for the transcription factor HEB in thymocyte survival, Tcra rearrangement and the development of natural killer T cells. Nature immunology 11, 240-249.

Daley, S.R., Hu, D.Y., Goodnow, C.C., 2013. Helios marks strongly autoreactive CD4+ T cells in two major waves of thymic deletion distinguished by induction of PD-1 or NF-kappaB. The Journal of experimental medicine 210, 269-285.

de Bruin, A., Maiti, B., Jakoi, L., Timmers, C., Buerki, R., Leone, G., 2003. Identification and characterization of E2F7, a novel mammalian E2F family member capable of blocking cellular proliferation. The Journal of biological chemistry 278, 42041-42049.

Di Stefano, L., Jensen, M.R., Helin, K., 2003. E2F7, a novel E2F featuring DP-independent repression of a subset of E2F-regulated genes. The EMBO journal 22, 6289-6298.

DiCarlo, J.E., Norville, J.E., Mali, P., Rios, X., Aach, J., Church, G.M., 2013. Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. Nucleic acids research 41, 4336-4343.

Doudna, J.A., Charpentier, E., 2014. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science 346, 1258096.

Drapkin, R., Reardon, J.T., Ansari, A., Huang, J.C., Zawel, L., Ahn, K., Sancar, A., Reinberg, D., 1994. Dual role of TFIIH in DNA excision repair and in transcription by RNA polymerase II. Nature 368, 769-772.

Dupressoir, A., Morel, A.P., Barbot, W., Loireau, M.P., Corbo, L., Heidmann, T., 2001. Identification of four families of yCCR4- and Mg2+-dependent endonuclease-related proteins in higher eukaryotes, and characterization of orthologs of yCCR4 with a conserved leucine-rich repeat essential for hCAF1/hPOP2 binding. BMC genomics 2, 9.

Dynlacht, B.D., 1997. Regulation of transcription by proteins that control the cell cycle. Nature 389, 149-152.

Edgar, R., Domrachev, M., Lash, A.E., 2002. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic acids research 30, 207-210.

Endo-Munoz, L., Dahler, A., Teakle, N., Rickwood, D., Hazar-Rethinam, M., Abdul-Jabbar, I., Sommerville, S., Dickinson, I., Kaur, P., Paquet-Fifield, S., Saunders, N., 2009. E2F7 can regulate proliferation, differentiation, and apoptotic responses in human keratinocytes: implications for cutaneous squamous cell carcinoma formation. Cancer research 69, 1800-1808.

Ericsson, P.O., Teh, H.S., 1995. The protein tyrosine kinase p56lck regulates TCR expression and T cell selection. International immunology 7, 617-624.

Esvelt, K.M., Mali, P., Braff, J.L., Moosburner, M., Yaung, S.J., Church, G.M., 2013. Orthogonal Cas9 Proteins for RNA-Guided Gene Regulation and Editing. Nature methods 10, 1116-1121.

Field, S.J., Tsai, F.Y., Kuo, F., Zubiaga, A.M., Kaelin, W.G., Jr., Livingston, D.M., Orkin, S.H., Greenberg, M.E., 1996. E2F-1 functions in mice to promote apoptosis and suppress proliferation. Cell 85, 549-561.

Friedland, A.E., Tzur, Y.B., Esvelt, K.M., Colaiacovo, M.P., Church, G.M., Calarco, J.A., 2013. Heritable genome editing in C. elegans via a CRISPR-Cas9 system. Nature methods 10, 741-743.

Fu, Y., Foden, J.A., Khayter, C., Maeder, M.L., Reyon, D., Joung, J.K., Sander, J.D., 2013. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Nature biotechnology 31, 822-826.

Fujihara, C., Williams, J.A., Watanabe, M., Jeon, H., Sharrow, S.O., Hodes, R.J., 2014. T cell-B cell thymic cross-talk: maintenance and function of thymic B cells requires cognate CD40-CD40 ligand interaction. Journal of immunology 193, 5534-5544.

Fujita, T., Fujii, H., 2013. Efficient isolation of specific genomic regions and identification of associated proteins by engineered DNA-binding molecule-mediated chromatin immunoprecipitation (enChIP) using CRISPR. Biochemical and biophysical research communications 439, 132-136.

Fuss, J.O., Tainer, J.A., 2011. XPB and XPD helicases in TFIIH orchestrate DNA duplex opening and damage verification to coordinate repair with transcription and cell cycle via CAK kinase. DNA repair 10, 697-713.

Gagnerault, M.-C., Luan, J.J., Lotton, C., Lepault, F., 2002. Pancreatic Lymph Nodes Are Required for Priming of β Cell Reactive T Cells in NOD Mice. The Journal of experimental medicine 196, 369-377.

Gagnon, J.A., Valen, E., Thyme, S.B., Huang, P., Ahkmetova, L., Pauli, A., Montague, T.G., Zimmerman, S., Richter, C., Schier, A.F., 2014. Efficient Mutagenesis by Cas9 Protein-Mediated Oligonucleotide Insertion and Large-Scale Assessment of Single-Guide RNAs. PloS one 9, e98186.

Gao, F., Shen, X.Z., Jiang, F., Wu, Y., Han, C., 2016. DNA-guided genome editing using the Natronobacterium gregoryi Argonaute. Nat Biotech advance online publication.

Garcia, I., Murga, M., Vicario, A., Field, S.J., Zubiaga, A.M., 2000. A role for E2F1 in the induction of apoptosis during thymic negative selection. Cell Growth Differ 11, 91-98.

Garneau, J.E., Dupuis, M.E., Villion, M., Romero, D.A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadan, A.H., Moineau, S., 2010. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature 468, 67-71.

Gibson, D.G., Young, L., Chuang, R.Y., Venter, J.C., Hutchison, C.A., 3rd, Smith, H.O., 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nature methods 6, 343-345.

Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A., Lim, W.A., Weissman, J.S., Qi, L.S., 2013.

CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. Cell 154, 442-451.

Godfrey, D.I., Kennedy, J., Mombaerts, P., Tonegawa, S., Zlotnik, A., 1994. Onset of TCR-beta gene rearrangement and role of TCR-beta expression during CD3-CD4-CD8- thymocyte differentiation. Journal of immunology 152, 4783-4792.

Godfrey, D.I., Kennedy, J., Suda, T., Zlotnik, A., 1993. A developmental pathway involving four phenotypically and functionally distinct subsets of CD3-CD4-CD8- triple-negative adult mouse thymocytes defined by CD44 and CD25 expression. Journal of immunology 150, 4244-4252.

Gordon, J.W., Ruddle, F.H., 1981. Integration and stable germ line transmission of genes injected into mouse pronuclei. Science 214, 1244-1246.

Gordon, J.W., Scangos, G.A., Plotkin, D.J., Barbosa, J.A., Ruddle, F.H., 1980. Genetic transformation of mouse embryos by microinjection of purified DNA. Proceedings of the National Academy of Sciences of the United States of America 77, 7380-7384.

Gratz, S.J., Cummings, A.M., Nguyen, J.N., Hamm, D.C., Donohue, L.K., Harrison, M.M., Wildonger, J., O'Connor-Giles, K.M., 2013. Genome engineering of Drosophila with the CRISPR RNA-guided Cas9 nuclease. Genetics 194, 1029-1035.

Grunberg, S., Hahn, S., 2013. Structural insights into transcription initiation by RNA polymerase II. Trends in biochemical sciences 38, 603-611.

Guilinger, J.P., Thompson, D.B., Liu, D.R., 2014. Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. Nature biotechnology 32, 577-582.

Heng, T.S., Painter, M.W., 2008. The Immunological Genome Project: networks of gene expression in immune cells. Nature immunology 9, 1091-1094.

Hoess, R.H., Ziese, M., Sternberg, N., 1982. P1 site-specific recombination: nucleotide sequence of the recombining sites. Proceedings of the National Academy of Sciences of the United States of America 79, 3398-3402.

Horvath, P., Barrangou, R., 2010. CRISPR/Cas, the immune system of bacteria and archaea. Science 327, 167-170.

Hou, Z., Zhang, Y., Propson, N.E., Howden, S.E., Chu, L.-F., Sontheimer, E.J., Thomson, J.A., 2013. Efficient genome engineering in human pluripotent stem cells using Cas9 from Neisseria meningitidis. Proceedings of the National Academy of Sciences of the United States of America 110, 15644-15649.

Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O., Cradick, T.J., Marraffini, L.A., Bao, G., Zhang, F., 2013. DNA targeting specificity of RNA-guided Cas9 nucleases. Nat Biotech 31, 827-832.

Hu, D.Y., Yap, J.Y., Wirasinha, R.C., Howard, D.R., Goodnow, C.C., Daley, S.R., 2016. A timeline demarcating two waves of clonal deletion and Foxp3 upregulation during thymocyte development. Immunol Cell Biol 94, 357-366.

Hu, Q., Sader, A., Parkman, J.C., Baldwin, T.A., 2009. Bim-mediated apoptosis is not necessary for thymic negative selection to ubiquitous self-antigens. Journal of immunology 183, 7761-7767.

Hu, X., Chang, N., Wang, X., Zhou, F., Zhou, X., Zhu, X., Xiong, J.W., 2013. Heritable gene-targeting with gRNA/Cas9 in rats. Cell research 23, 1322-1325.

Hwang, W.Y., Fu, Y., Reyon, D., Maeder, M.L., Tsai, S.Q., Sander, J.D., Peterson, R.T., Yeh, J.R., Joung, J.K., 2013. Efficient genome editing in zebrafish using a CRISPR-Cas system. Nature biotechnology 31, 227-229.

Irizarry, R.A., Hobbs, B., Collin, F., Beazer-Barclay, Y.D., Antonellis, K.J., Scherf, U., Speed, T.P., 2003. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 4, 249-264.

Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., Nakata, A., 1987. Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. Journal of bacteriology 169, 5429-5433.

Jawhari, A., Laine, J.P., Dubaele, S., Lamour, V., Poterszman, A., Coin, F., Moras, D., Egly, J.M., 2002. p52 Mediates XPB function within the transcription/repair factor TFIIH. The Journal of biological chemistry 277, 31761-31767.

Jiang, H., Promchan, K., Lin, B.R., Lockett, S., Chen, D., Marshall, H., Badralmaa, Y., Natarajan, V., 2016. LZTFL1 Upregulated by All-Trans Retinoic Acid during CD4+ T Cell Activation Enhances IL-5 Production. Journal of immunology 196, 1081-1090.

Jiang, W., Zhou, H., Bi, H., Fromm, M., Yang, B., Weeks, D.P., 2013. Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice. Nucleic acids research 41, e188.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., Charpentier, E., 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337, 816-821.

Jinek, M., East, A., Cheng, A., Lin, S., Ma, E., Doudna, J., 2013. RNA-programmed genome editing in human cells. Elife 2, e00471.

Johnson, S.A., Pleiman, C.M., Pao, L., Schneringer, J., Hippen, K., Cambier, J.C., 1995. Phosphorylated immunoreceptor signaling motifs (ITAMs) exhibit unique abilities to bind and activate Lyn and Syk tyrosine kinases. Journal of immunology 155, 4596-4603.

Kasibhatla, S., Tailor, P., Bonefoy-Berard, N., Mustelin, T., Altman, A., Fotedar, A., 1999. Jun kinase phosphorylates and regulates the DNA binding activity of an octamer binding protein, T-cell factor beta1. Molecular and cellular biology 19, 2021-2031.

Katic, I., Grosshans, H., 2013. Targeted heritable mutation and gene conversion by Cas9-CRISPR in Caenorhabditis elegans. Genetics 195, 1173-1176.

Ketterer, C., Mussig, K., Machicao, F., Stefan, N., Fritsche, A., Haring, H.U., Staiger, H., 2011. Genetic variation within the NR1H2 gene encoding liver X receptor beta associates with insulin secretion in subjects at increased risk for type 2 diabetes. J Mol Med (Berl) 89, 75-81.

Killeen, N., Stuart, S.G., Littman, D.R., 1992. Development and function of T cells in mice with a disrupted CD2 gene. The EMBO journal 11, 4329-4336.

Kim, D., Peng, X.C., Sun, X.H., 1999. Massive apoptosis of thymocytes in T-cell-deficient Id1 transgenic mice. Molecular and cellular biology 19, 8240-8253.

Kisielow, P., Bluthmann, H., Staerz, U.D., Steinmetz, M., von Boehmer, H., 1988. Tolerance in T-cell-receptor transgenic mice involves deletion of nonmature CD4+8+ thymocytes. Nature 333, 742-746.

Klein, L., Hinterberger, M., von Rohrscheidt, J., Aichinger, M., 2011. Autonomous versus dendritic cell-dependent contributions of medullary thymic epithelial cells to central tolerance. Trends in immunology 32, 188-193.

Kleinstiver, B.P., Pattanayak, V., Prew, M.S., Tsai, S.Q., Nguyen, N.T., Zheng, Z., Keith Joung, J., 2016. High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects. Nature advance online publication.

Koonin, E.V., Makarova, K.S., 2009. CRISPR-Cas: an adaptive immunity system in prokaryotes. F1000 biology reports 1, 95.

Kozak, M., 1984. Point mutations close to the AUG initiator codon affect the efficiency of translation of rat preproinsulin in vivo. Nature 308, 241-246.

Kugel, J.F., Goodrich, J.A., 1998. Promoter escape limits the rate of RNA polymerase II transcription and is enhanced by TFIIE, TFIIH, and ATP on negatively supercoiled DNA. Proceedings of the National Academy of Sciences of the United States of America 95, 9232-9237.

Kumar, A., Chalamalasetty, R.B., Kennedy, M.W., Thomas, S., Inala, S.N., Garriock, R.J., Yamaguchi, T.P., 2016. Zfp703 Is a Wnt/beta-Catenin Feedback Suppressor Targeting the beta-Catenin/Tcf1 Complex. Molecular and cellular biology 36, 1793-1802.

Kumar, A., Kumar Dorairaj, S., Prabhakaran, V.C., Prakash, D.R., Chakraborty, S., 2007. Identification of genes associated with tumorigenesis of meibomian cell carcinoma by microarray analysis. Genomics 90, 559-566.

Labrecque, N., Baldwin, T., Lesage, S., 2011. Molecular and genetic parameters defining T-cell clonal selection. Immunol Cell Biol 89, 16-26.

Lammens, T., Li, J., Leone, G., De Veylder, L., 2009. Atypical E2Fs: new players in the E2F transcription factor family. Trends in cell biology 19, 111-118.

Langerak, A.W., Wolvers-Tettero, I.L., van Gastel-Mol, E.J., Oud, M.E., van Dongen, J.J., 2001. Basic helix-loop-helix proteins E2A and HEB induce immature T-cell receptor rearrangements in nonlymphoid cells. Blood 98, 2456-2465.

Lanning, D., Lafuse, W.P., 1999. The mouse p52 subunit of the transcription/DNA repair factor TFIIH is located in the class III region of the H2 complex: cloning and sequence polymorphism. Immunogenetics 49, 498-504.

Lee, H.J., Kim, E., Kim, J.S., 2010. Targeted chromosomal deletions in human cells using zinc finger nucleases. Genome research 20, 81-89.

Lee, P.P., Fitzpatrick, D.R., Beard, C., Jessup, H.K., Lehar, S., Makar, K.W., Pérez-Melgosa, M., Sweetser, M.T., Schlissel, M.S., Nguyen, S., Cherry, S.R., Tsai, J.H., Tucker, S.M., Weaver, W.M., Kelso, A., Jaenisch, R., Wilson, C.B., 2001. A Critical Role for Dnmt1 and DNA Methylation in T Cell Development, Function, and Survival. Immunity 15, 763-774.

Li, J., Ran, C., Li, E., Gordon, F., Comstock, G., Siddiqui, H., Cleghorn, W., Chen, H.Z., Kornacker, K., Liu, C.G., Pandit, S.K., Khanizadeh, M., Weinstein, M., Leone, G., de Bruin, A., 2008. Synergistic function of E2F7 and E2F8 is essential for cell survival and embryonic development. Developmental cell 14, 62-75.

Li, Y., Li, P., Chen, S., Wu, Z., Li, J., Zhang, S., Cao, C., Wang, L., Liu, B., Zhang, F., Li, Y.Z., 2015. Association of GTF2I and GTF2IRD1 polymorphisms with systemic lupus erythematosus in a Chinese Han population. Clin Exp Rheumatol 33, 632-638.

Lin, Y., Cradick, T.J., Brown, M.T., Deshmukh, H., Ranjan, P., Sarode, N., Wile, B.M., Vertino, P.M., Stewart, F.J., Bao, G., 2014. CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences. Nucleic acids research.

Lind, E.F., Prockop, S.E., Porritt, H.E., Petrie, H.T., 2001. Mapping precursor movement through the postnatal thymus reveals specific microenvironments supporting defined stages of early lymphoid development. The Journal of experimental medicine 194, 127-134.

Logan, N., Delavaine, L., Graham, A., Reilly, C., Wilson, J., Brummelkamp, T.R., Hijmans, E.M., Bernards, R., La Thangue, N.B., 2004. E2F-7: a distinctive E2F family member with an unusual organization of DNA-binding domains. Oncogene 23, 5138-5150.

Logan, N., Graham, A., Zhao, X., Fisher, R., Maiti, B., Leone, G., La Thangue, N.B., 2005. E2F-8: an E2F family member with a similar organization of DNA-binding domains to E2F-7. Oncogene 24, 5000-5004.

Luse, D.S., 2013. The RNA polymerase II preinitiation complex: Through what pathway is the complex assembled? Transcription 5.

Ma, Y., Zhang, X., Shen, B., Lu, Y., Chen, W., Ma, J., Bai, L., Huang, X., Zhang, L., 2014. Generating rats with conditional alleles using CRISPR/Cas9. Cell research 24, 122-125.

Maiti, B., Li, J., de Bruin, A., Gordon, F., Timmers, C., Opavsky, R., Patil, K., Tuttle, J., Cleghorn, W., Leone, G., 2005. Cloning and characterization of mouse E2F8, a novel

mammalian E2F family member capable of blocking cellular proliferation. The Journal of biological chemistry 280, 18211-18220.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., Church, G.M., 2013. RNA-guided human genome engineering via Cas9. Science 339, 823-826.

Mallaun, M., Zenke, G., Palmer, E., 2010. A discrete affinity-driven elevation of ZAP-70 kinase activity initiates negative selection. Journal of receptor and signal transduction research 30, 430-443.

Marinoni, J.C., Roy, R., Vermeulen, W., Miniou, P., Lutz, Y., Weeda, G., Seroz, T., Gomez, D.M., Hoeijmakers, J.H., Egly, J.M., 1997. Cloning and characterization of p52, the fifth subunit of the core of the transcription/DNA repair factor TFIIH. The EMBO journal 16, 1093-1102.

Masuda, T., Zhang, X., Berlinicke, C., Wan, J., Yerrabelli, A., Conner, E.A., Kjellstrom, S., Bush, R., Thorgeirsson, S.S., Swaroop, A., Chen, S., Zack, D.J., 2014. The transcription factor GTF2IRD1 regulates the topology and function of photoreceptors by modulating photoreceptor gene expression across the retina. J Neurosci 34, 15356-15368.

Matsuno, M., Kose, H., Okabe, M., Hiromi, Y., 2007. TFIIH controls developmentally-regulated cell cycle progression as a holocomplex. Genes to cells : devoted to molecular & cellular mechanisms 12, 1289-1300.

Messier, H., Brickner, H., Gaikwad, J., Fotedar, A., 1993. A novel POU domain protein which binds to the T-cell receptor beta enhancer. Molecular and cellular biology 13, 5450-5460.

Meyer, M., de Angelis, M.H., Wurst, W., Kuhn, R., 2010. Gene targeting by homologous recombination in mouse zygotes mediated by zinc-finger nucleases. Proceedings of the National Academy of Sciences of the United States of America 107, 15022-15026.

Mouzat, K., Mercier, E., Polge, A., Evrard, A., Baron, S., Balducchi, J.P., Brouillet, J.P., Lumbroso, S., Gris, J.C., 2011. A common polymorphism in NR1H2 (LXRbeta) is associated with preeclampsia. BMC Med Genet 12, 145.

Murga, M., Fernandez-Capetillo, O., Field, S.J., Moreno, B., Borlado, L.R., Fujiwara, Y., Balomenos, D., Vicario, A., Carrera, A.C., Orkin, S.H., Greenberg, M.E., Zubiaga, A.M., 2001. Mutation of E2F2 in mice causes enhanced T lymphocyte proliferation, leading to the development of autoimmunity. Immunity 15, 959-970.

Nakayama, T., Fish, M.B., Fisher, M., Oomen-Hajagos, J., Thomsen, G.H., Grainger, R.M., 2013. Simple and efficient CRISPR/Cas9-mediated targeted mutagenesis in Xenopus tropicalis. Genesis 51, 835-843.

Nekrasov, V., Staskawicz, B., Weigel, D., Jones, J.D., Kamoun, S., 2013. Targeted mutagenesis in the model plant Nicotiana benthamiana using Cas9 RNA-guided endonuclease. Nature biotechnology 31, 691-693.

Neumeister, E.N., Zhu, Y., Richard, S., Terhorst, C., Chan, A.C., Shaw, A.S., 1995. Binding of ZAP-70 to phosphorylated T-cell receptor zeta and eta enhances its autophosphorylation and

generates specific binding sites for SH2 domain-containing proteins. Molecular and cellular biology 15, 3171-3178.

Nihongaki, Y., Yamamoto, S., Kawano, F., Suzuki, H., Sato, M., 2015. CRISPR-Cas9-based photoactivatable transcription system. Chemistry & biology 22, 169-174.

Niu, Y., Shen, B., Cui, Y., Chen, Y., Wang, J., Wang, L., Kang, Y., Zhao, X., Si, W., Li, W., Xiang, A.P., Zhou, J., Guo, X., Bi, Y., Si, C., Hu, B., Dong, G., Wang, H., Zhou, Z., Li, T., Tan, T., Pu, X., Wang, F., Ji, S., Zhou, Q., Huang, X., Ji, W., Sha, J., 2014. Generation of gene-modified cynomolgus monkey via Cas9/RNA-mediated gene targeting in one-cell embryos. Cell 156, 836-843.

O'Leary, J., Osborne, L.R., 2011. Global analysis of gene expression in the developing brain of Gtf2ird1 knockout mice. PloS one 6, e23868.

Okamoto, K., Wakamiya, M., Noji, S., Koyama, E., Taniguchi, S., Takemura, R., Copeland, N.G., Gilbert, D.J., Jenkins, N.A., Muramatsu, M., et al., 1993. A novel class of murine POU gene predominantly expressed in central nervous system. The Journal of biological chemistry 268, 7449-7457.

Osman, N., Lucas, S., Cantrell, D., 1995. The role of tyrosine phosphorylation in the interaction of cellular tyrosine kinases with the T cell receptor zeta chain tyrosine-based activation motif. European journal of immunology 25, 2863-2869.

Ouseph, M.M., Li, J., Chen, H.Z., Pecot, T., Wenzel, P., Thompson, J.C., Comstock, G., Chokshi, V., Byrne, M., Forde, B., Chong, J.L., Huang, K., Machiraju, R., de Bruin, A., Leone, G., 2012. Atypical E2F repressors and activators coordinate placental development. Developmental cell 22, 849-862.

Painter, M.W., Davis, S., Hardy, R.R., Mathis, D., Benoist, C., 2011. Transcriptomes of the B and T lineages compared by multiplatform microarray profiling. Journal of immunology 186, 3047-3057.

Palmiter, R.D., Brinster, R.L., Hammer, R.E., Trumbauer, M.E., Rosenfeld, M.G., Birnberg, N.C., Evans, R.M., 1982. Dramatic growth of mice that develop from eggs microinjected with metallothionein-growth hormone fusion genes. Nature 300, 611-615.

Pan, H., Yin, C., Dyson, N.J., Harlow, E., Yamasaki, L., Van Dyke, T., 1998. Key roles for E2F1 in signaling p53-dependent apoptosis and in cell division within developing tumors. Molecular cell 2, 283-292.

Pearson, A., Greenblatt, J., 1997. Modular organization of the E2F1 activation domain and its interaction with general transcription factors TBP and TFIIH. Oncogene 15, 2643-2658.

Phillips, M.J., Needham, M., Weller, R.O., 1997. Role of cervical lymph nodes in autoimmune encephalomyelitis in the Lewis rat. J Pathol 182, 457-464.

Polstein, L.R., Gersbach, C.A., 2015. A light-inducible CRISPR-Cas9 system for control of endogenous gene activation. Nat Chem Biol 11, 198-200.

Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., Lim, W.A., 2013. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell 152, 1173-1183.

Qin, W., Dion, S.L., Kutny, P.M., Zhang, Y., Cheng, A., Jillette, N.L., Malhotra, A., Geurts, A.M., Chen, Y.G., Wang, H., 2015. Efficient CRISPR/Cas9-Mediated Genome Editing in Mice by Zygote Electroporation of Nuclease. Genetics.

Ran, F.A., Cong, L., Yan, W.X., Scott, D.A., Gootenberg, J.S., Kriz, A.J., Zetsche, B., Shalem, O., Wu, X., Makarova, K.S., Koonin, E.V., Sharp, P.A., Zhang, F., 2015. In vivo genome editing using Staphylococcus aureus Cas9. Nature 520, 186-191.

Rincon, M., Flavell, R.A., 1996. Regulation of AP-1 and NFAT transcription factors during thymic selection of T cells. Molecular and cellular biology 16, 1074-1084.

Rincon, M., Whitmarsh, A., Yang, D.D., Weiss, L., Derijard, B., Jayaraj, P., Davis, R.J., Flavell, R.A., 1998. The JNK pathway regulates the In vivo deletion of immature CD4(+)CD8(+) thymocytes. The Journal of experimental medicine 188, 1817-1830.

Russell, W.M.S., Burch, R.L., 1959. The principles of humane experimental technique. Methuen, London.

Sabapathy, K., Hu, Y., Kallunki, T., Schreiber, M., David, J.P., Jochum, W., Wagner, E.F., Karin, M., 1999. JNK2 is required for efficient T-cell activation and apoptosis but not for normal lymphocyte development. Curr Biol 9, 116-125.

Sabapathy, K., Kallunki, T., David, J.P., Graef, I., Karin, M., Wagner, E.F., 2001. c-Jun NH2-terminal kinase (JNK)1 and JNK2 have similar and stage-dependent roles in regulating T cell apoptosis and proliferation. The Journal of experimental medicine 193, 317-328.

Samelson, L.E., Phillips, A.F., Luong, E.T., Klausner, R.D., 1990. Association of the fyn protein-tyrosine kinase with the T-cell antigen receptor. Proceedings of the National Academy of Sciences of the United States of America 87, 4358-4362.

Sapranauskas, R., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P., Siksnys, V., 2011. The Streptococcus thermophilus CRISPR/Cas system provides immunity in Escherichia coli. Nucleic acids research 39, 9275-9282.

Sawada, S., Littman, D.R., 1993. A heterodimer of HEB and an E12-related protein interacts with the CD4 enhancer and regulates its activity in T-cell lines. Molecular and cellular biology 13, 5620-5628.

Sawada, S., Scarborough, J.D., Killeen, N., Littman, D.R., 1994. A lineage-specific transcriptional silencer regulates CD4 gene expression during T lymphocyte development. Cell 77, 917-929.

Schneider, C.A., Rasband, W.S., Eliceiri, K.W., 2012. NIH Image to ImageJ: 25 years of image analysis. Nat Meth 9, 671-675.

Schwenk, F., Kuhn, R., Angrand, P.O., Rajewsky, K., Stewart, A.F., 1998. Temporally and spatially regulated somatic mutagenesis in mice. Nucleic acids research 26, 1427-1432.

Seroz, T., Hwang, J.R., Moncollin, V., Egly, J.M., 1995. TFIIH: a link between transcription, DNA repair and cell cycle regulation. Current opinion in genetics & development 5, 217-221.

Shah, D.K., Zuniga-Pflucker, J.C., 2014. An overview of the intrathymic intricacies of T cell development. Journal of immunology 192, 4017-4023.

Shen, B., Zhang, W., Zhang, J., Zhou, J., Wang, J., Chen, L., Wang, L., Hodgkins, A., Iyer, V., Huang, X., Skarnes, W.C., 2014. Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects. Nat Meth 11, 399-402.

Shi, Y., Zhu, M., 2013. Medullary thymic epithelial cells, the indispensable player in central tolerance. Science China. Life sciences 56, 392-398.

Silva, I.A., Cox, C.J., Leite, R.B., Cancela, M.L., Conceicao, N., 2014. Evolutionary conservation of TFIIH subunits: implications for the use of zebrafish as a model to study TFIIH function and regulation. Comparative biochemistry and physiology. Part B, Biochemistry & molecular biology 172-173, 9-20.

Slorach, E.M., Chou, J., Werb, Z., 2011. Zeppo1 is a novel metastasis promoter that represses E-cadherin expression and regulates p120-catenin isoform expression and localization. Genes & development 25, 471-484.

Sollu, C., Pars, K., Cornu, T.I., Thibodeau-Beganny, S., Maeder, M.L., Joung, J.K., Heilbronn, R., Cathomen, T., 2010. Autonomous zinc-finger nuclease pairs for targeted chromosomal deletion. Nucleic acids research 38, 8269-8276.

Spangler, L., Wang, X., Conaway, J.W., Conaway, R.C., Dvir, A., 2001. TFIIH action in transcription initiation and promoter escape requires distinct regions of downstream promoter DNA. Proceedings of the National Academy of Sciences of the United States of America 98, 5544-5549.

Staton, T.L., Lazarevic, V., Jones, D.C., Lanser, A.J., Takagi, T., Ishii, S., Glimcher, L.H., 2011. Dampening of death pathways by Schnurri 2 is essential for T cell development. Nature 472, 105-109.

Sternberg, N., Hamilton, D., 1981. Bacteriophage P1 site-specific recombination. I. Recombination between loxP sites. Journal of molecular biology 150, 467-486.

Suen, A.Y., Baldwin, T.A., 2012. Proapoptotic protein Bim is differentially required during thymic clonal deletion to ubiquitous versus tissue-restricted antigens. Proceedings of the National Academy of Sciences of the United States of America 109, 893-898.

Takeuchi, A., Yamasaki, S., Takase, K., Nakatsu, F., Arase, H., Onodera, M., Saito, T., 2001. E2A and HEB activate the pre-TCR alpha promoter during immature T cell development. Journal of immunology 167, 2157-2163.

Tanaka, T., Takei, Y., Yamanouchi, D., 2016. Hyperglycemia Suppresses Calcium Phosphate-Induced Aneurysm Formation Through Inhibition of Macrophage Activation. J Am Heart Assoc 4, e003062.

Thomas, K.R., Capecchi, M.R., 1987. Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. Cell 51, 503-512.

Tokoro, Y., Sugawara, T., Yaginuma, H., Nakauchi, H., Terhorst, C., Wang, B., Takahama, Y., 1998. A mouse carrying genetic defect in the choice between T and B lymphocytes. Journal of immunology 161, 4591-4598.

Tremblay, M., Herblot, S., Lecuyer, E., Hoang, T., 2003. Regulation of pT alpha gene expression by a dosage of E2A, HEB, and SCL. The Journal of biological chemistry 278, 12680-12687.

Tsai, S.Q., Wyvekens, N., Khayter, C., Foden, J.A., Thapar, V., Reyon, D., Goodwin, M.J., Aryee, M.J., Joung, J.K., 2014. Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. Nature biotechnology 32, 569-576.

Tusher, V.G., Tibshirani, R., Chu, G., 2001. Significance analysis of microarrays applied to the ionizing radiation response. Proceedings of the National Academy of Sciences of the United States of America 98, 5116-5121.

Vandel, L., Kouzarides, T., 1999. Residues phosphorylated by TFIIH are required for E2F-1 degradation during S-phase. The EMBO journal 18, 4280-4291.

Villunger, A., Marsden, V.S., Zhan, Y., Erlacher, M., Lew, A.M., Bouillet, P., Berzins, S., Godfrey, D.I., Heath, W.R., Strasser, A., 2004. Negative selection of semimature CD4(+)8(-)HSA+ thymocytes requires the BH3-only protein Bim but is independent of death receptor signaling. Proceedings of the National Academy of Sciences of the United States of America 101, 7052-7057.

Wang, H., Yang, H., Shivalila, C.S., Dawlaty, M.M., Cheng, A.W., Zhang, F., Jaenisch, R., 2013. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. Cell 153, 910-918.

Wang, L., Guo, J., Wang, Q., Zhou, J., Xu, C., Teng, R., Chen, Y., Wei, Q., Liu, Z.P., 2014. LZTFL1 suppresses gastric cancer cell migration and invasion through regulating nuclear translocation of beta-catenin. J Cancer Res Clin Oncol 140, 1997-2008.

Wang, S.S., Gonzalez, P., Yu, K., Porras, C., Li, Q., Safaeian, M., Rodriguez, A.C., Sherman, M.E., Bratti, C., Schiffman, M., Wacholder, S., Burk, R.D., Herrero, R., Chanock, S.J., Hildesheim, A., 2010. Common genetic variants and risk for HPV persistence and progression to cervical cancer. PloS one 5, e8667.

Wang, X.W., Vermeulen, W., Coursen, J.D., Gibson, M., Lupold, S.E., Forrester, K., Xu, G., Elmore, L., Yeh, H., Hoeijmakers, J.H., Harris, C.C., 1996. The XPB and XPD DNA helicases are components of the p53-mediated apoptosis pathway. Genes & development 10, 1219-1232.

Wang, X.W., Yeh, H., Schaeffer, L., Roy, R., Moncollin, V., Egly, J.M., Wang, Z., Freidberg, E.C., Evans, M.K., Taffe, B.G., et al., 1995. p53 modulation of TFIIH-associated nucleotide excision repair activity. Nature genetics 10, 188-195.

Waterston, R.H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J.F., Agarwal, P., Agarwala, R., Ainscough, R., Alexandersson, M., An, P., Antonarakis, S.E., Attwood, J., Baertsch, R., Bailey, J., Barlow, K., Beck, S., Berry, E., Birren, B., Bloom, T., Bork, P., Botcherby, M., Bray, N., Brent, M.R., Brown, D.G., Brown, S.D., Bult, C., Burton, J., Butler, J., Campbell, R.D., Carninci, P., Cawley, S., Chiaromonte, F., Chinwalla, A.T., Church, D.M., Clamp, M., Clee, C., Collins, F.S., Cook, L.L., Copley, R.R., Coulson, A., Couronne, O., Cuff, J., Curwen, V., Cutts, T., Daly, M., David, R., Davies, J., Delehaunty, K.D., Deri, J., Dermitzakis, E.T., Dewey, C., Dickens, N.J., Diekhans, M., Dodge, S., Dubchak, I., Dunn, D.M., Eddy, S.R., Elnitski, L., Emes, R.D., Eswara, P., Eyras, E., Felsenfeld, A., Fewell, G.A., Flicek, P., Foley, K., Frankel, W.N., Fulton, L.A., Fulton, R.S., Furey, T.S., Gage, D., Gibbs, R.A., Glusman, G., Gnerre, S., Goldman, N., Goodstadt, L., Grafham, D., Graves, T.A., Green, E.D., Gregory, S., Guigo, R., Guyer, M., Hardison, R.C., Haussler, D., Hayashizaki, Y., Hillier, L.W., Hinrichs, A., Hlavina, W., Holzer, T., Hsu, F., Hua, A., Hubbard, T., Hunt, A., Jackson, I., Jaffe, D.B., Johnson, L.S., Jones, M., Jones, T.A., Joy, A., Kamal, M., Karlsson, E.K., Karolchik, D., Kasprzyk, A., Kawai, J., Keibler, E., Kells, C., Kent, W.J., Kirby, A., Kolbe, D.L., Korf, I., Kucherlapati, R.S., Kulbokas, E.J., Kulp, D., Landers, T., Leger, J.P., Leonard, S., Letunic, I., Levine, R., Li, J., Li, M., Lloyd, C., Lucas, S., Ma, B., Maglott, D.R., Mardis, E.R., Matthews, L., Mauceli, E., Mayer, J.H., McCarthy, M., McCombie, W.R., McLaren, S., McLay, K., McPherson, J.D., Meldrim, J., Meredith, B., Mesirov, J.P., Miller, W., Miner, T.L., Mongin, E., Montgomery, K.T., Morgan, M., Mott, R., Mullikin, J.C., Muzny, D.M., Nash, W.E., Nelson, J.O., Nhan, M.N., Nicol, R., Ning, Z., Nusbaum, C., O'Connor, M.J., Okazaki, Y., Oliver, K., Overton-Larty, E., Pachter, L., Parra, G., Pepin, K.H., Peterson, J., Pevzner, P., Plumb, R., Pohl, C.S., Poliakov, A., Ponce, T.C., Ponting, C.P., Potter, S., Quail, M., Reymond, A., Roe, B.A., Roskin, K.M., Rubin, E.M., Rust, A.G., Santos, R., Sapojnikov, V., Schultz, B., Schultz, J., Schwartz, M.S., Schwartz, S., Scott, C., Seaman, S., Searle, S., Sharpe, T., Sheridan, A., Shownkeen, R., Sims, S., Singer, J.B., Slater, G., Smit, A., Smith, D.R., Spencer, B., Stabenau, A., Stange-Thomann, N., Sugnet, C., Suyama, M., Tesler, G., Thompson, J., Torrents, D., Trevaskis, E., Tromp, J., Ucla, C., Ureta-Vidal, A., Vinson, J.P., Von Niederhausern, A.C., Wade, C.M., Wall, M., Weber, R.J., Weiss, R.B., Wendl, M.C., West, A.P., Wetterstrand, K., Wheeler, R., Whelan, S., Wierzbowski, J., Willey, D., Williams, S., Wilson, R.K., Winter, E., Worley, K.C., Wyman, D., Yang, S., Yang, S.P., Zdobnov, E.M., Zody, M.C., Lander, E.S., 2002. Initial sequencing and comparative analysis of the mouse genome. Nature 420, 520-562.

Wei, Q., Zhou, W., Wang, W., Gao, B., Wang, L., Cao, J., Liu, Z.P., 2010. Tumor-suppressive functions of leucine zipper transcription factor-like 1. Cancer research 70, 2942-2950.

Weijts, B.G., Bakker, W.J., Cornelissen, P.W., Liang, K.H., Schaftenaar, F.H., Westendorp, B., de Wolf, C.A., Paciejewska, M., Scheele, C.L., Kent, L., Leone, G., Schulte-Merker, S., de Bruin, A., 2012. E2F7 and E2F8 promote angiogenesis through transcriptional activation of VEGFA in cooperation with HIF1. The EMBO journal 31, 3871-3884.

Wilson, A., Held, W., MacDonald, H.R., 1994. Two waves of recombinase gene expression in developing thymocytes. The Journal of experimental medicine 179, 1355-1360.

Xie, K., Yang, Y., 2013. RNA-guided genome editing in plants using a CRISPR-Cas system. Molecular plant 6, 1975-1983.

Xue, Z., Wu, M., Wen, K., Ren, M., Long, L., Zhang, X., Gao, G., 2014. CRISPR/Cas9 mediates efficient conditional mutagenesis in Drosophila. G3 4, 2167-2173.

Yang, H., Wang, H., Shivalila, C.S., Cheng, A.W., Shi, L., Jaenisch, R., 2013. One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. Cell 154, 1370-1379.

Yates, A.J., 2014. Theories and quantification of thymic selection. Frontiers in immunology 5, 13.

Yin, L., Maddison, L.A., Li, M., Kara, N., LaFave, M.C., Varshney, G.K., Burgess, S.M., Patton, J.G., Chen, W., 2015. Multiplex Conditional Mutagenesis Using Transgenic Expression of Cas9 and sgRNAs. Genetics.

Young, E.J., Lipina, T., Tam, E., Mandel, A., Clapcote, S.J., Bechard, A.R., Chambers, J., Mount, H.T., Fletcher, P.J., Roder, J.C., Osborne, L.R., 2008. Reduced fear and aggression and altered serotonin metabolism in Gtf2ird1-targeted mice. Genes Brain Behav 7, 224-234.

Yu, Z., Ren, M., Wang, Z., Zhang, B., Rong, Y.S., Jiao, R., Gao, G., 2013. Highly efficient genome modifications mediated by CRISPR/Cas9 in Drosophila. Genetics 195, 289-291.

Zalmas, L.P., Zhao, X., Graham, A.L., Fisher, R., Reilly, C., Coutts, A.S., La Thangue, N.B., 2008. DNA-damage response control of E2F7 and E2F8. EMBO reports 9, 252-259.

Zetsche, B., Gootenberg, Jonathan S., Abudayyeh, Omar O., Slaymaker, Ian M., Makarova, Kira S., Essletzbichler, P., Volz, Sara E., Joung, J., van der Oost, J., Regev, A., Koonin, Eugene V., Zhang, F., Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. Cell.

Zhang, D.J., Wang, Q., Wei, J., Baimukanova, G., Buchholz, F., Stewart, A.F., Mao, X., Killeen, N., 2005. Selective expression of the Cre recombinase in late-stage thymocytes using the distal promoter of the Lck gene. Journal of immunology 174, 6725-6731.

Zhang, F., Cong, L., Lodato, S., Kosuri, S., Church, G.M., Arlotta, P., 2011. Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. Nat Biotech 29, 149-153.

Zhang, L., Zhou, J., Han, J., Hu, B., Hou, N., Shi, Y., Huang, X., Lou, X., 2016. Generation of an Oocyte-Specific Cas9 Transgenic Mouse for Genome Editing. PloS one 11, e0154364.

Zhang, W., Irvin, B.J., Trible, R.P., Abraham, R.T., Samelson, L.E., 1999a. Functional analysis of LAT in TCR-mediated signaling pathways using a LAT-deficient Jurkat cell line. International immunology 11, 943-950.

Zhang, W., Sommers, C.L., Burshtyn, D.N., Stebbins, C.C., DeJarnette, J.B., Trible, R.P., Grinberg, A., Tsay, H.C., Jacobs, H.M., Kessler, C.M., Long, E.O., Love, P.E., Samelson, L.E., 1999b. Essential role of LAT in T cell development. Immunity 10, 323-332.

Zhovmer, A., Oksenych, V., Coin, F., 2010. Two sides of the same coin: TFIIH complexes in transcription and DNA repair. TheScientificWorldJournal 10, 633-643.

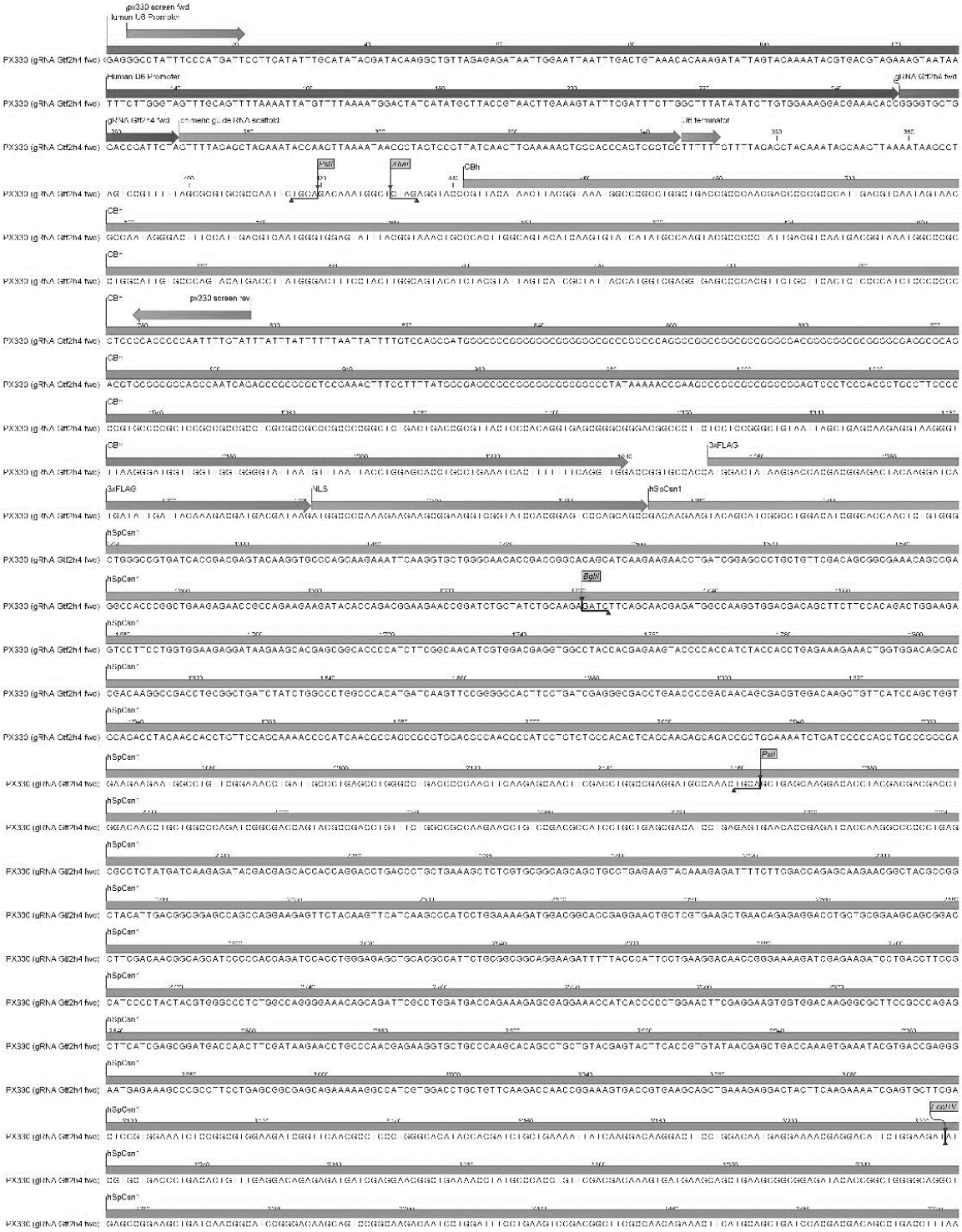
Zhu, J.W., Field, S.J., Gore, L., Thompson, M., Yang, H., Fujiwara, Y., Cardiff, R.D., Greenberg, M., Orkin, S.H., DeGregori, J., 2001. E2F1 and E2F2 determine thresholds for antigen-induced

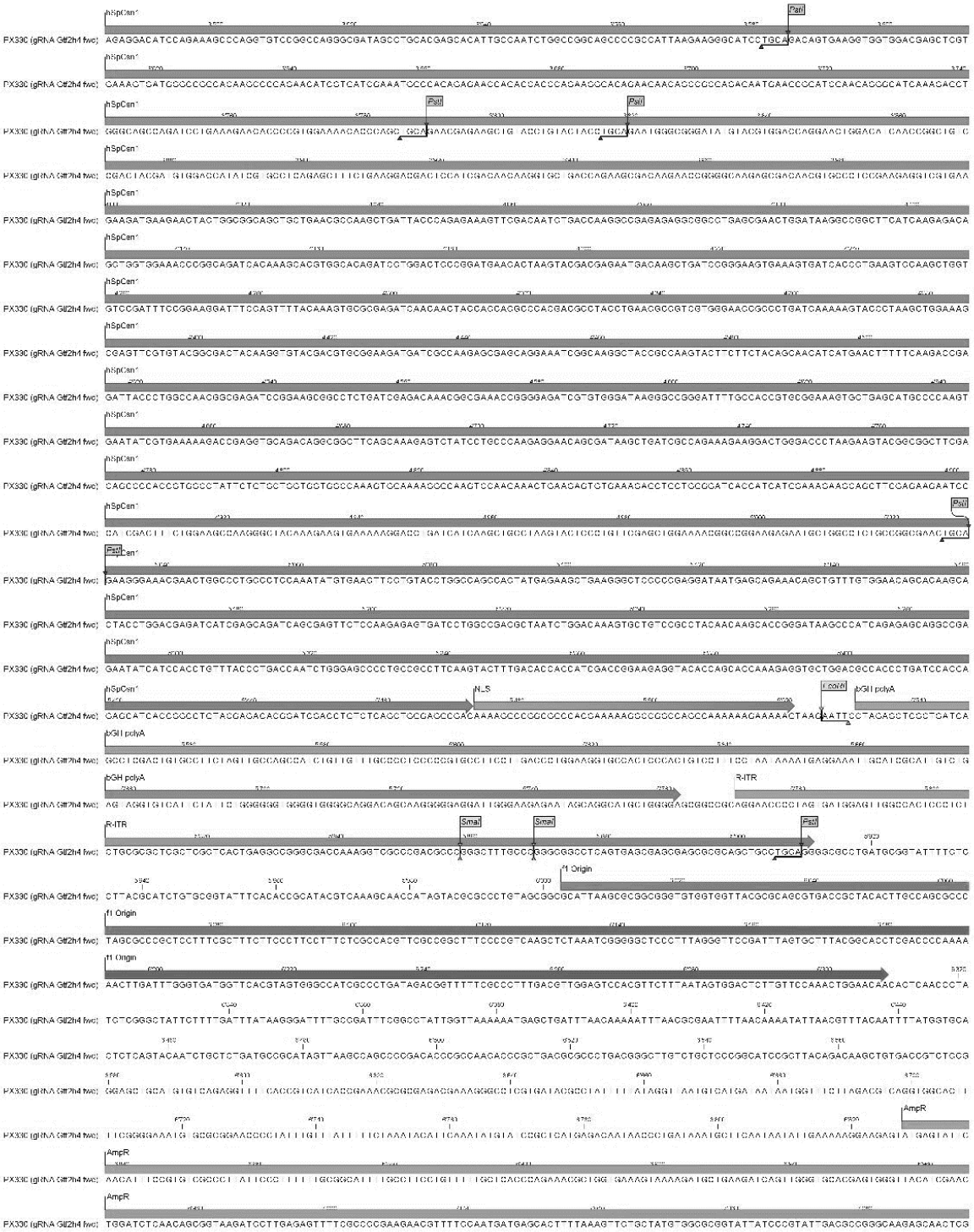
T-cell proliferation and suppress tumorigenesis. Molecular and cellular biology 21, 8547-8564.

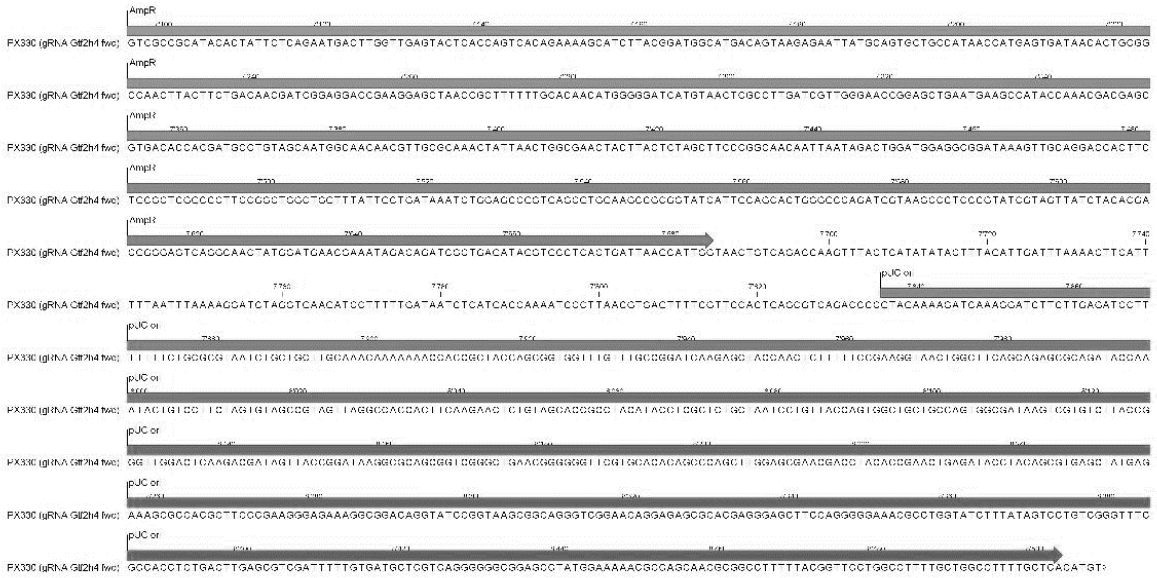
Zhuang, Y., Cheng, P., Weintraub, H., 1996. B-lymphocyte development is regulated by the combined dosage of three basic helix-loop-helix genes, E2A, E2-2, and HEB. Molecular and cellular biology 16, 2898-2905.

6 Appendix

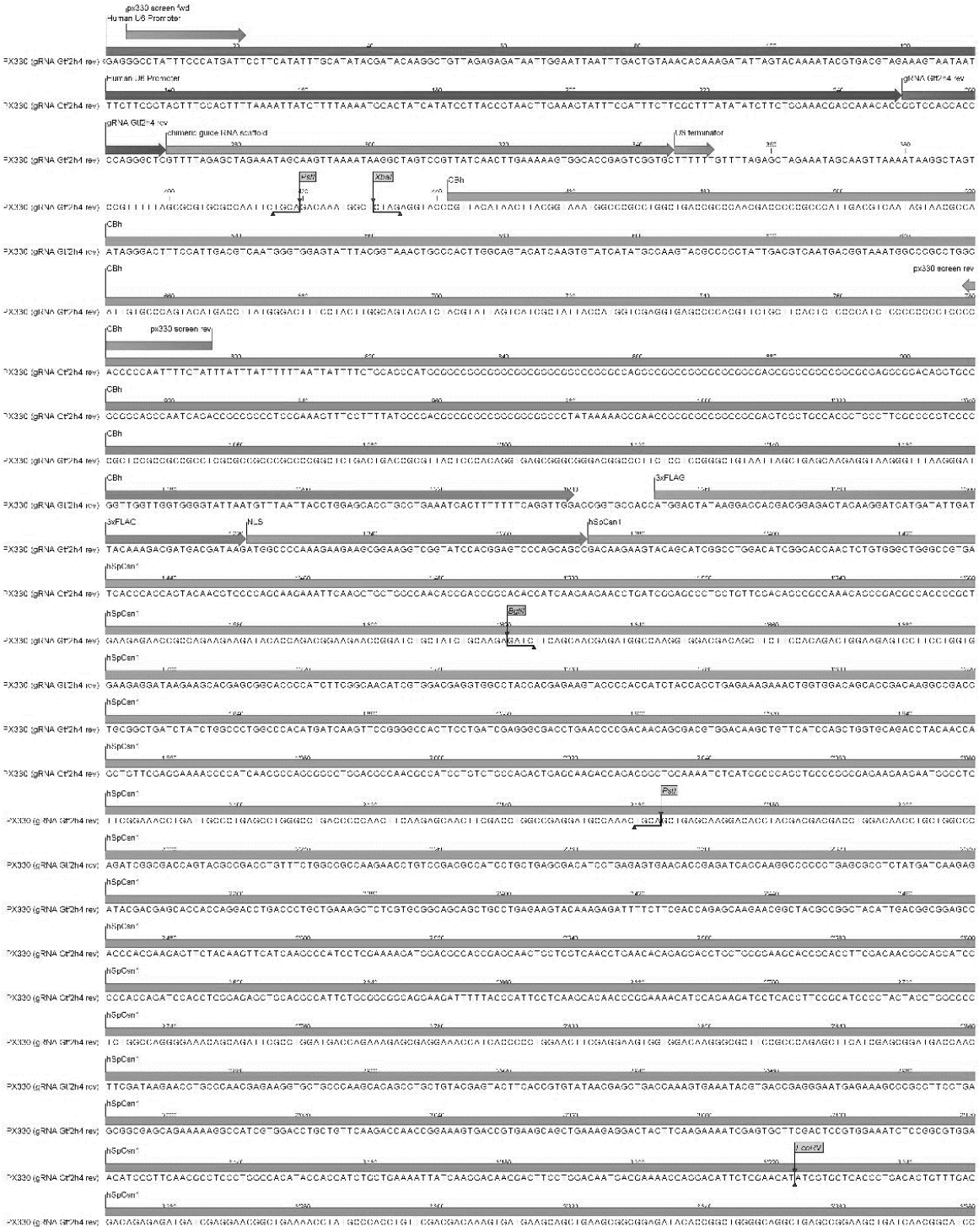
6.1 Plasmid map of px330(gRNA Gtf2h4 fwd)

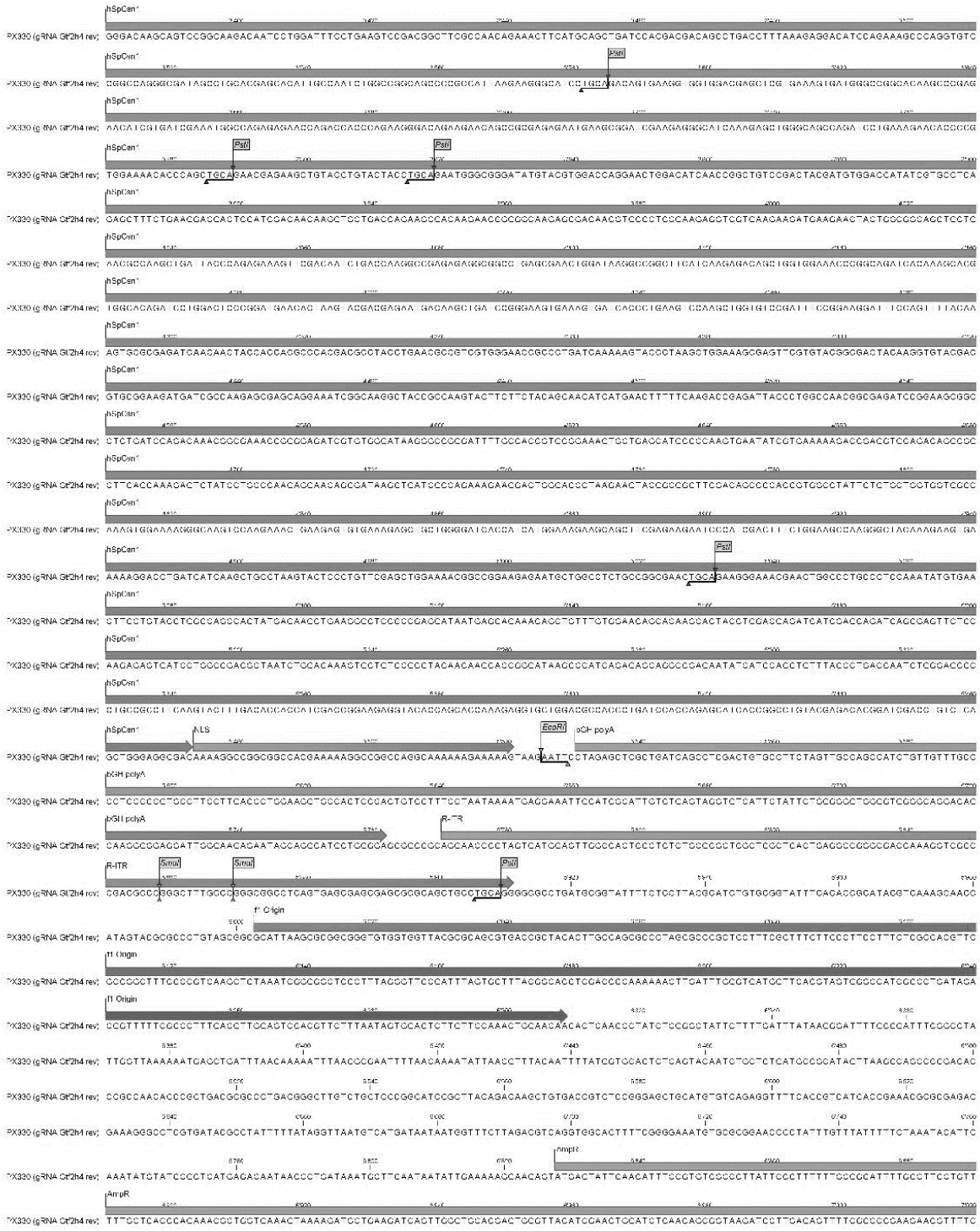


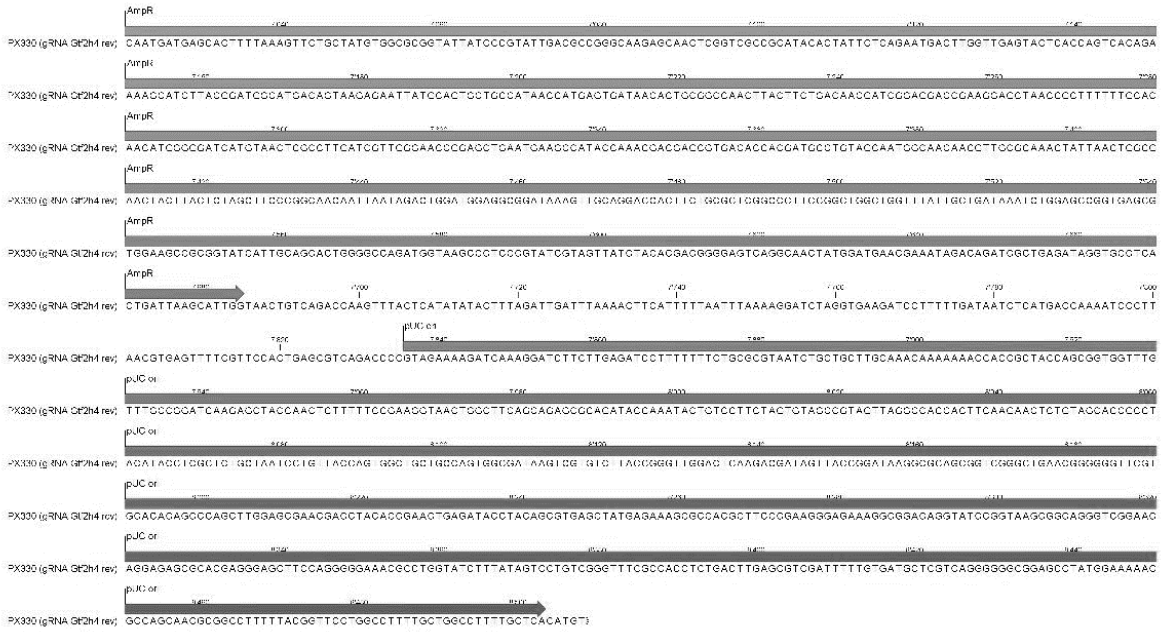

PX330 (gRNA Glt2h4 fwd)
hSpCas9
PstI
PgiI
N_S
bGH polyA
R-ITR
SmaI
f1 Origin
AmpR


AmpR
pUC ori
PX330 (gRNA Gtf2h4 fwd)

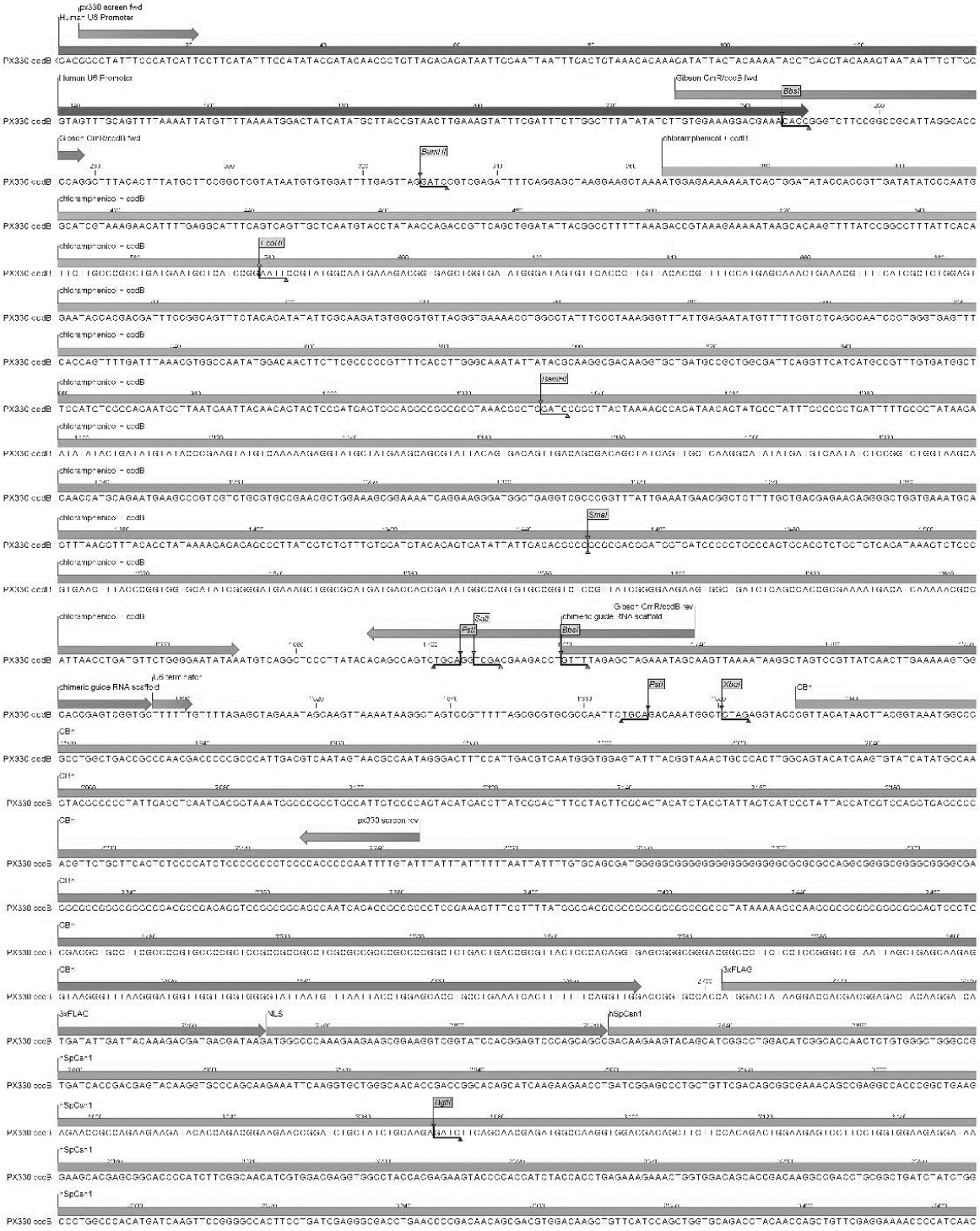
6.2 Plasmid map of px330(gRNA Gtf2h4 rev)

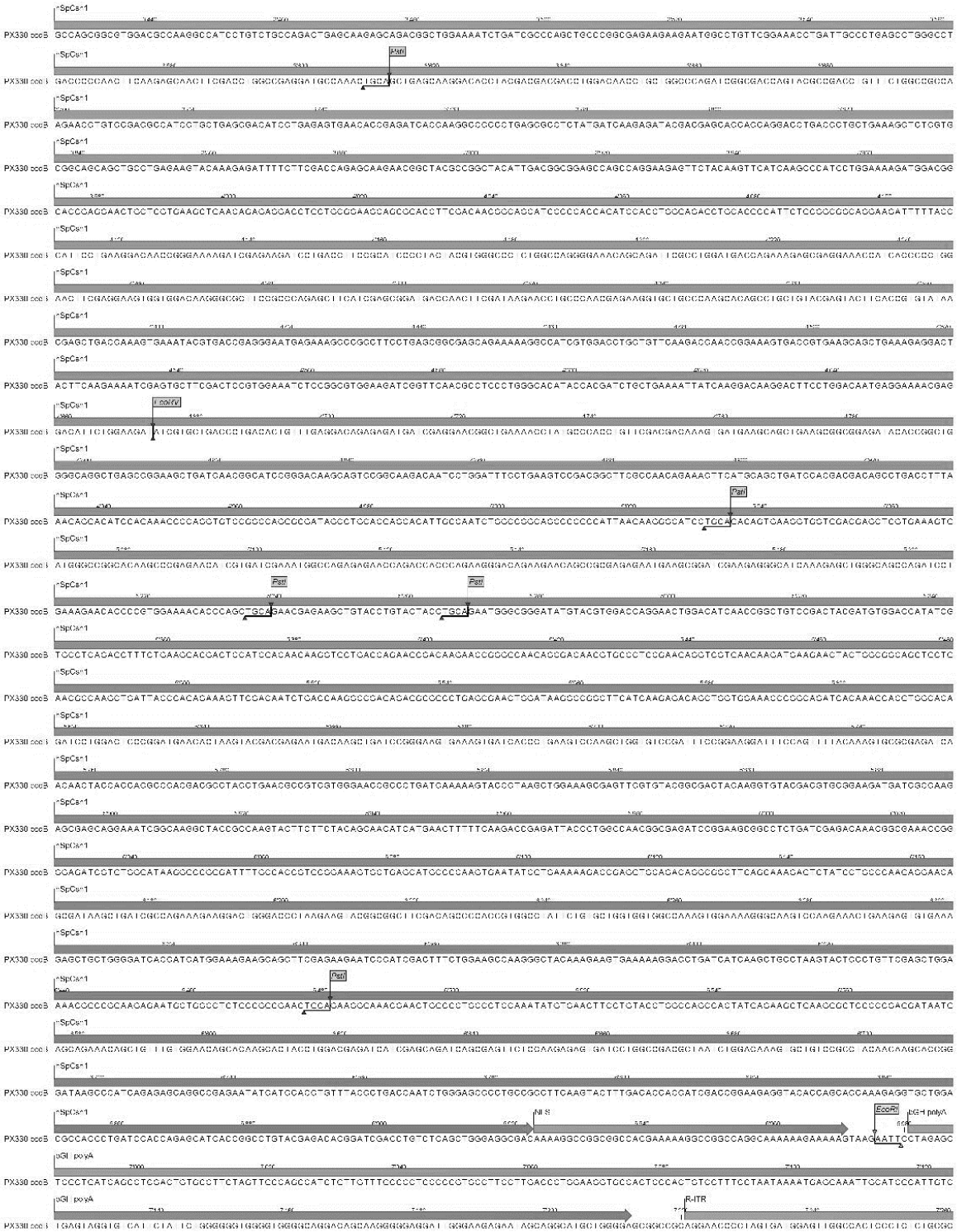


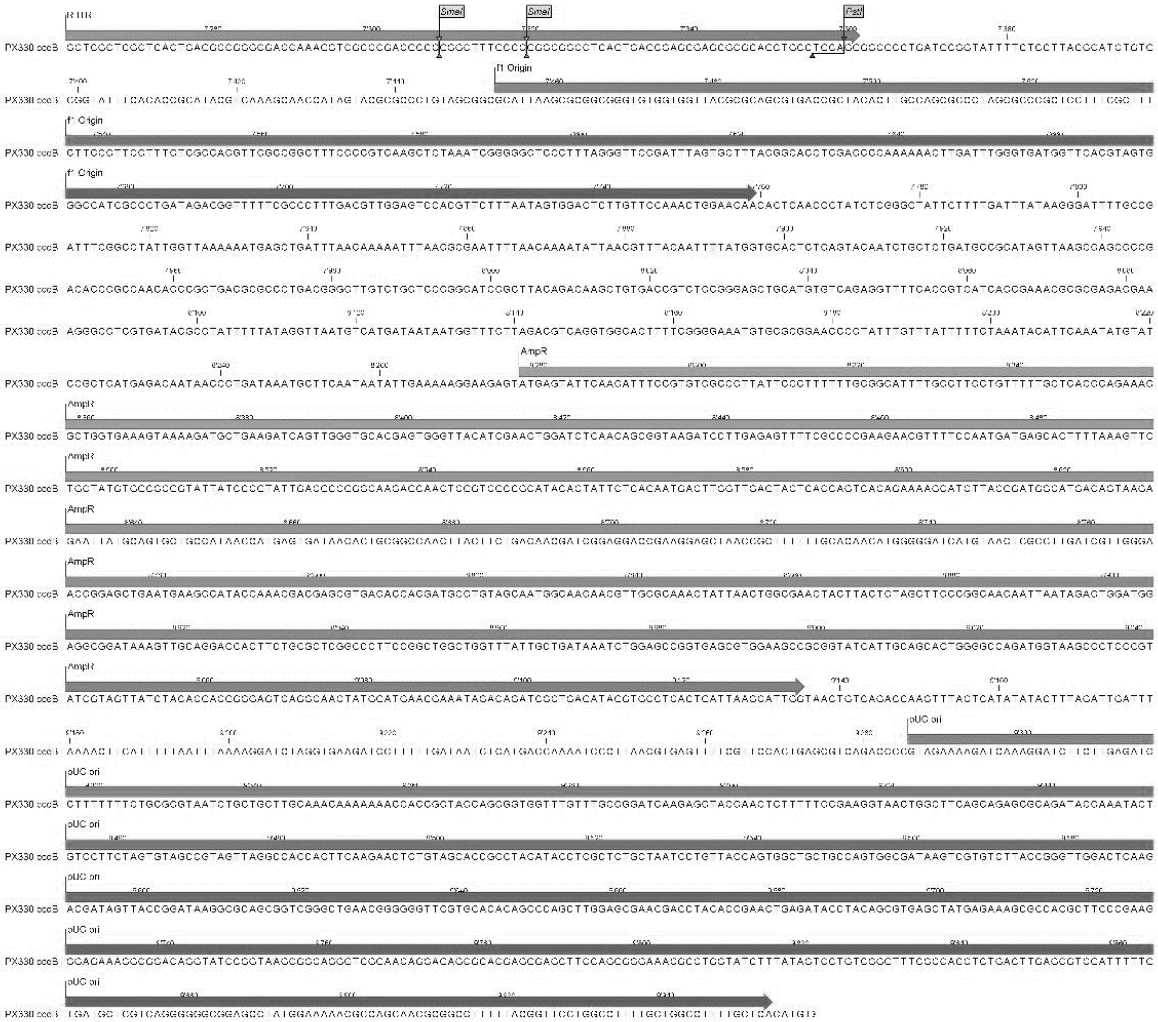



AmpR
pUC ori
pX330 (gRNA Gt2h4 rev)

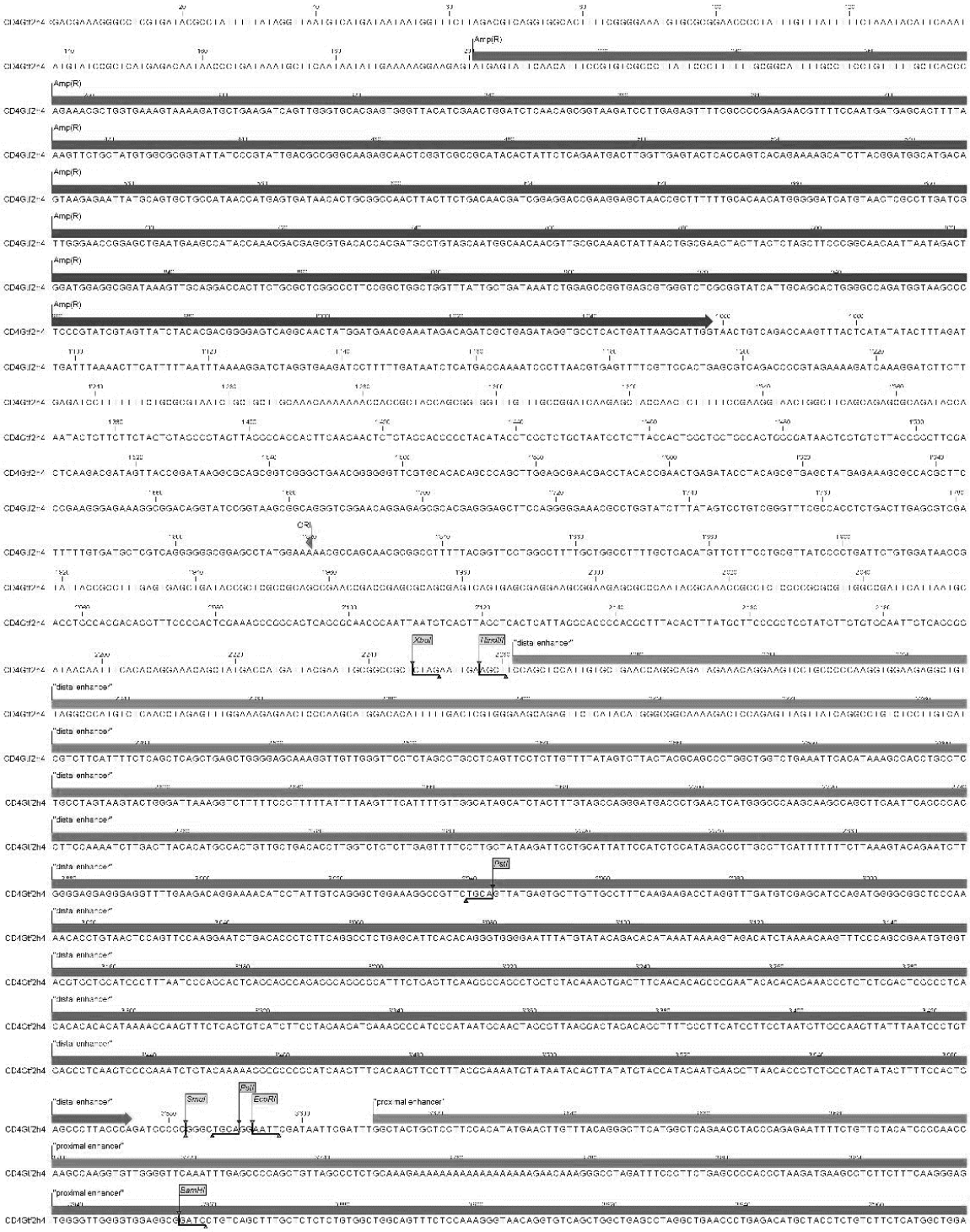
6.3 Plasmid map of px330 ccdB

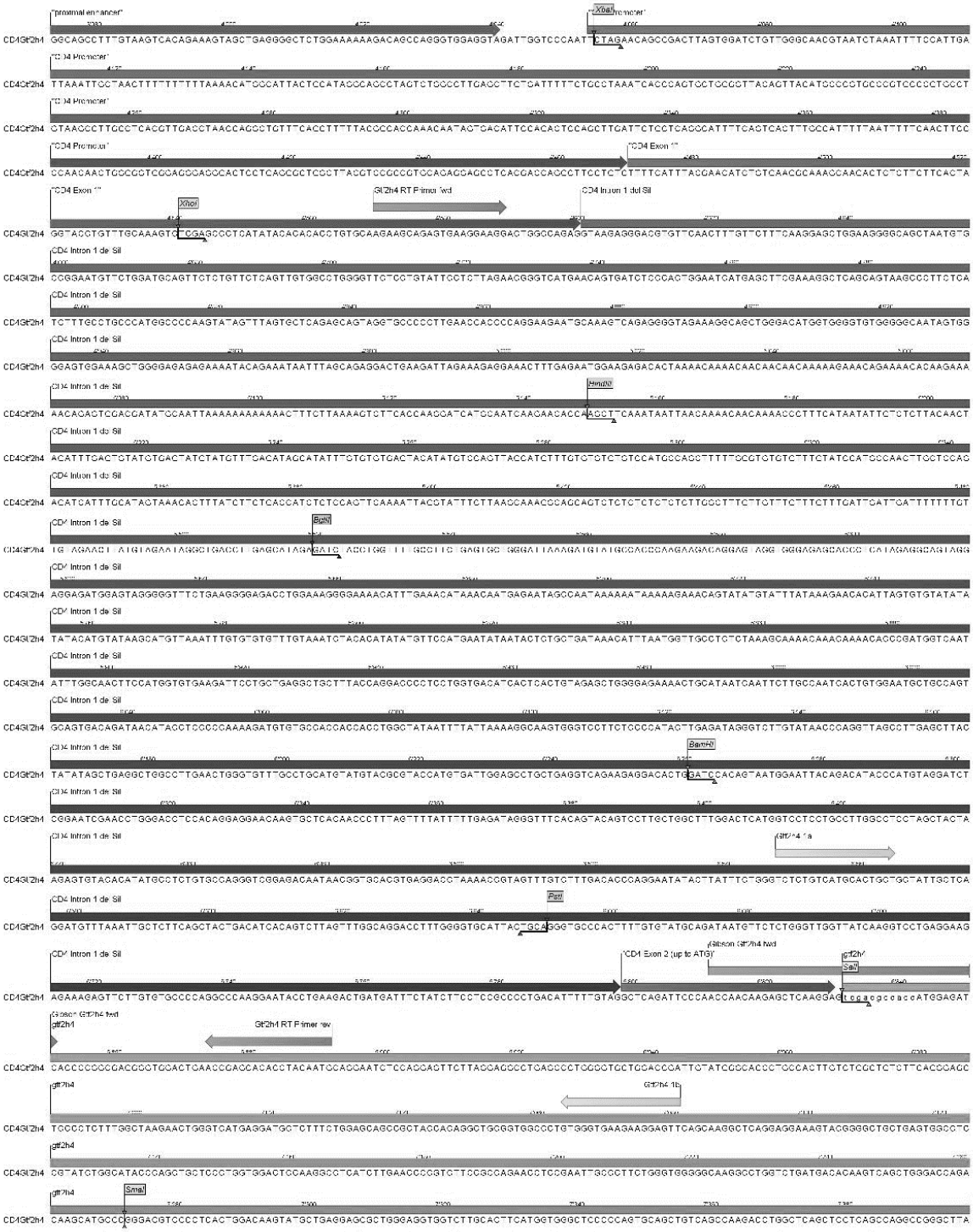


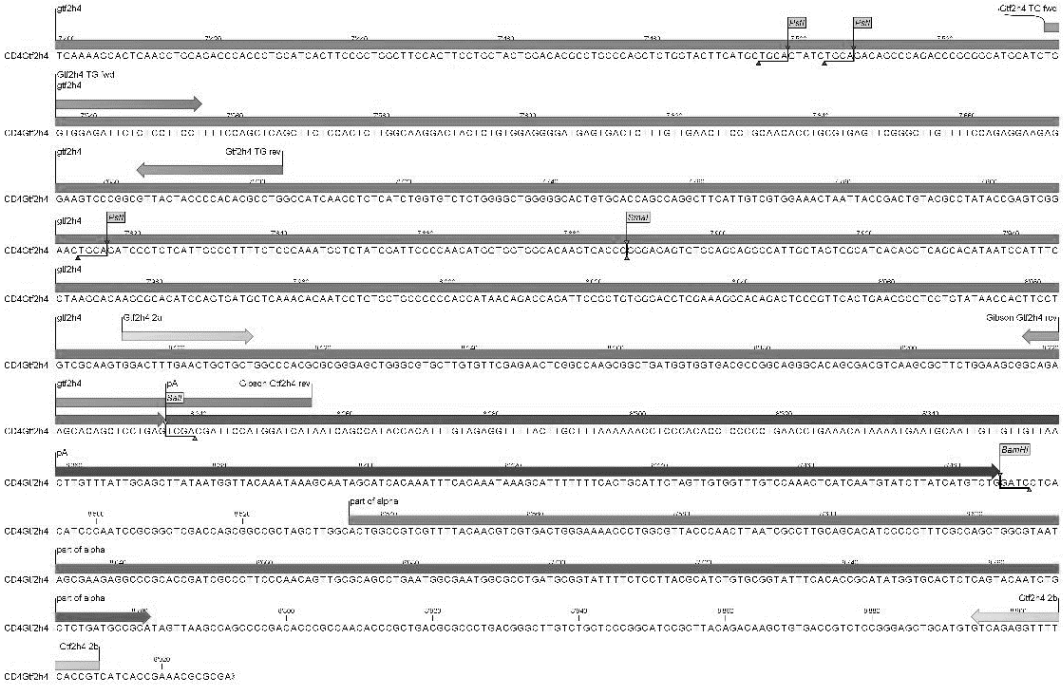

SpCas9
PX330 seqB
PstI
PX330 seqB
SpCas9
PX330 seqB
SpCas9
PX330 seqB
SpCas9
PX330 seqB
SpCas9
PX330 seqB
SpCas9
PX330 seqB
SpCas9
PX330 seqB
EcoRV
SpCas9
PX330 seqB
SpCas9
PstI
PX330 seqB
SpCas9
PstI
PstI
PX330 seqB
SpCas9
PX330 seqB
SpCas9
PX330 seqB
SpCas9
PX330 seqB
SpCas9
PX330 seqB
SpCas9
PX330 seqB
SpCas9
PstI
PX330 seqB
SpCas9
PX330 seqB
SpCas9
PX330 seqB
NLS
EcoRI
pGH polyA
SpCas9
PX330 seqB
pGH polyA
PX330 seqB
pGH polyA
R-ITR


SmaI
SmaI
PacI
f1 Origin
f1 Origin
f1 Origin
AmpR
AmpR
AmpR
AmpR
AmpR
pUC ori
pUC ori
pUC ori
pUC ori
pUC ori

6.4 Plasmid map of CD4Gtf2h4




"proximal enhancer"
XbaI
"CD4 Promoter"
"CD4 Exon 1"
Gt2h4 RT Primer fwd
CD4 Intron 1 del SII
XhoI
HindIII
BglII
BamHI
PstI
(Gibson Gt2h4 fwd
gt2h4
SalI
Gt2h4 RT Primer rev
Gt2h4 1b
SmaI


gtf2h4
Ctf2-4 TC fwd
CD4Gtf2h4
Gtf2h4 TG fwd
gtf2h4
Gtf2h4 TG rev
gtf2h4
gtf2h4
gtf2h4
Gibson Gtf2h4 rev
gtf2h4
pA
Gibson Ctf2h4 rev
pA
BamHI
part of alpha
part of alpha
Ctf2-4 2b
Ctf2-4 2b

7 Publications

Book Chapter

- J. Beil, I. M. Mansuy , and T. Buch; Conditional Gene Targeting: A Refined Method for Genetic Studies in Neurosciences; In Peter V. Nguyen (ed.), Multidisciplinary Tools for Investigating Synaptic Plasticity, Neuromethods, vol. 81, DOI 10.1007/978-1-62703-517-0_1, © Springer Science+Business Media, LLC 2013
- J. Beil and T. Buch; Generation of Bacterial Artificial Chromosome (BAC) Transgenic Mice; In Shree Ram Singh and Vincenzo Coppola (eds.), Mouse Genetics: Methods and Protocols, Methods in Molecular Biology, vol. 1194, DOI 10.1007/978-1-4939-1215-5_8, © Springer Science+Business Media New York 2014

Review article

- J. Beil, L. Fairbairn, P. Pelczar, and T. Buch; Is BAC Transgenesis Obsolete? State of the Art in the Era of Designer Nucleases; Journal of Biomedicine and Biotechnology Volume 2012 (2012)

Paper

- J. Beil-Wagner, G. Dössinger, K. Schober, J. vom Berg, A. Tresch, M. Grandl, P. Palle, F. Mair, M. Gerhard, B. Becher, D. H. Busch, T. Buch; T cell-specific inactivation of mouse CD2 by CRISPR/Cas9; Scientific Reports 6, Article number: 21377 (2016); doi:10.1038/srep21377

www.nature.com/scientificreports

OPEN

T cell-specific inactivation of mouse CD2 by CRISPR/Cas9

Jane Beil-Wagner[1,2], Georg Dössinger[1], Kilian Schober[1], Johannes vom Berg[2], Achim Tresch[3,4], Martina Grandl[1], Pushpalatha Palle[2], Florian Mair[5], Markus Gerhard[1], Burkhard Becher[5], Dirk H. Busch[1] & Thorsten Buch[1,2]

Received: 16 February 2015
Accepted: 22 January 2016
Published: 23 February 2016

The CRISPR/Cas9 system can be used to mutate target sequences by introduction of double-strand breaks followed by imprecise repair. To test its use for conditional gene editing we generated mice transgenic for CD4 promoter-driven Cas9 combined with guide RNA targeting CD2. We found that within $CD4^+$ and $CD8^+$ lymphocytes from lymph nodes and spleen 1% and 0.6% were not expressing CD2, respectively. T cells lacking CD2 carryied mutations, which confirmed that Cas9 driven by cell-type specific promoters can edit genes in the mouse and may thus allow targeted studies of gene function *in vivo*.

Genome modification in the mouse has become an essential research tool for addressing gene function *in vivo*. Since 1987, targeted mutagenesis in the mouse has been performed by manipulating embryonic stem (ES) cells *in vitro*[1]. Chimeric mice that were partially derived from these ES cells were used to find new, manipulated mouse lines. As an extension to this technology, recombinases such as Cre are used to avoid the use of resistance genes, to allow embryonic-lethal mutations and to make inducible, cell-type specific knockouts and knock-ins[2–4]. While, recombinase-mediated conditional gene targeting has drastically widened the scope of the technology, it still relies on ES cells for introduction of the recombinase targets such as loxP sites. Further, mouse lines carrying a loxP-containing allele need to be intercrossed with a Cre transgene to address target gene function. Altogether, the minimal time for generation of such conditional mutants is more than a year. Recently, a paradigm shift in germline mutagenesis has taken place through DNA editing by designer nucleases. Zinc finger nucleases, Tal effector nucleases, and the CRISPR/Cas9 system have facilitated the generation of gene-modified animals without the use of ES cells[5–9]. These designer nucleases can be tailored to introduce single- or double-strand breaks into target genes, which are repaired through cell-intrinsic repair mechanisms. Non-homologous end joining-mediated repair leads to insertions or deletions[10] and homologous recombination in the presence of a donor fragment facilitates introduction of defined mutations[11]. In the mouse, designer nuclease-mediated mutagenesis is achieved through injection of nuclease-encoding plasmids, RNA or protein into zygotes followed by screening of the founder animals[5,6]. Recently, the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) in combination with CRISPR associated gene 9 (Cas9) system has surpassed the other designer nucleases in terms of efficiency and flexibility. It has proven to allow both insertions/deletions and specific targeted mutagenesis in the germline of the mouse[8,12–14] and in human cells[14,15]. The original CRISPR/Cas9 system from *Streptococcus pyogenes*[16] was modified for application in molecular biology and now relies on a single 102 bp long guide RNA (gRNA), of which 20 bp determine the target sequence[17]. Since, the exact choice of gRNA is subject to further restrictions, such as presence of the *Protospacer Adjacent Motif* (PAM) sequence (NGG) and uniqueness within the target genome, online tools have been developed for designing functionally competent gRNAs[18,19]. The CRSIPR/Cas9 system has been used to modify genes in the germline by transient expression in the zygote and was found to facilitate somatic mutations in the mouse when expression of the system was under control of the Cre/loxP system and gRNA was delivered through viruses or transfection[20]. Only recently, it was shown in mice that doxycycline-induced Cas9 allows conditional *in vivo* gene editing[21]. In *Drosophila melanogaster*, Cas9 was shown to enable efficient cell-specific gene editing after its placement under a tissue-specific promoter[22]. However, direct gRNA/Cas9-mediated conditional gene editing using cell-type specific promoters similar to conditional mutagenesis by Cre has been so far not reported for the mouse (Fig. 1a). Such an approach would allow

[1]Institute for Medical Microbiology, Immunology and Hygiene, Technische Universität München, Germany. [2]Institute of Laboratory Animal Science, University of Zurich, Schlieren, Switzerland. [3]Max-Planck-Institute for Plant Breeding Research, Cologne, Germany. [4]Department of Biology, Albertus-Magnus University, Cologne, Germany. [5]Institute of Experimental Immunology, University of Zurich, Zurich, Switzerland. Correspondence and requests for materials should be addressed to T.B. (email: thorsten.buch@uzh.ch)

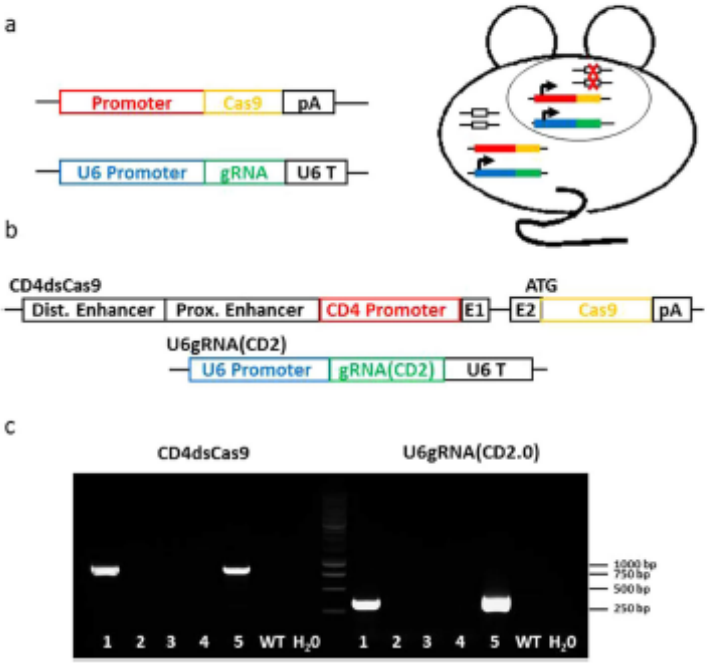


Figure 1. Conditional gene editing. (**a**) Scheme of the concept of conditional gene editing. In the Cas9 driver strain, the nuclease is placed under control of a cell type or lineage specific promoter. The gRNA construct is driven by the ubiquitous U6 promoter. Both transgenes are co-injected into oocytes. In double-transgenic animals, cell-type specific gene deletions are induced. (**b**) Scheme of constructs used for the CD4dsCas9/U6gRNA(CD2) mouse strain. The two used linearized plasmids are shown. First, distal and proximal enhancer, CD4 promoter followed by exon 1, part of exon 2 and Cas9 with a PolyA at the end. Second, the U6 promoter driven gRNA specific for CD2 followed by the U6 terminator (U6 T). (**c**) PCR analysis of tail biopsies for presence of CD4dsCas9 (835 bp amplicon) (lanes 1 and 5) and U6gRNA(CD2.0) (407 bp amplicon) (lanes 1 and 5) by PCR. DNA from a wildtype (WT) mouse as well as H_20 were run as a negative control. 2, 3 and 4 were non-transgenic litter mates.

single-step analysis of gene function within a cell lineage circumventing tedious gene targeting and crossing necessary for Cre/loxP-mediated conditional gene ablation.

To test whether gRNA/Cas9 could drive cell-type specific mutagenesis in the mouse, we placed Cas9 under control of a CD4 promoter (CD4dsCas9), thus initiating the expression in T lymphocytes during their maturation in the thymus (Fig. 1b)[23]. In a second construct, we placed a gRNA expression cassette under control of a U6 promoter (U6gRNA(CD2)). The gRNA was directed against a target sequence in exon 2 of CD2 (Fig. 1b), an easily detectable surface marker found on all T cells. In addition, we chose this target because its deficiency does not influence cell viability[24] and does not lead to changes in the cell's phenotype, that could impair this proof of concept study. The two constructs were co-injected into the pronuclei of oocytes of FVB/N mice and offspring were screened for presence of the transgenes. Two transgenic founders carrying both constructs were identified (Fig. 1c). We also obtained two founders carrying solely the Cas9 and two carrying solely the gRNA construct (data not shown). Further analysis of peripheral blood by flow cytometry revealed that in one of the double-transgenic founders a fraction of $CD4^+$ (0.4%) and $CD8^+$ (0.5%) cells lacked expression of CD2 (Fig. 2). No such populations were found in wildtype controls (data not shown), the other double-transgenic (Suppl. Fig. 1) or Cas9 and gRNA single-transgenic founders (Fig. 2). Furthermore, we analyzed pooled lymph nodes and spleen cells of two double transgenic F1 mice as well as of two wildtype littermates to verify the initial blood analysis. Here we observed that 1% of $CD4^+$ and 0.6% of $CD8^+$ T cells were not expressing CD2 anymore (Fig. 3). A transgene copy number analysis by real-time PCR was performed for this transgenic line and indicated numbers of integrated Cas9 being 26 and 41 and of the gRNA of 12 and 6/7 (Fig. 4a). In addition, CD4-Cas9 mRNA expression was found solely in T cells and thymocytes but not B cells of double transgenic (dTG) mice or wildtype thymocytes (Fig. 4b). U6-driven gRNA(CD2.0) expression was detected in total spleen and thymus of dTGs but not in WT controls (Fig. 4c). To confirm that the CD2 locus was edited by gRNA/Cas9, we single cell-sorted T cells lacking CD2 and sequenced a fragment around the target site after single cell PCR amplification and cloning (Fig. 5a). We observed that eight out of the nine sequenced $CD2^-CD4^+$ T cells carried a mutation in the amplicon (Fig. 5b). In $CD2^-CD8^+$ T cells two of the three amplicons were wildtype and one mutated (Fig. 5b). In all amplicons, we only found a single sequence, either mutated or wildtype. This may be either result of a technical bias leading to amplification of only one allele or alternatively an outcome of repair of the second allele by homologous

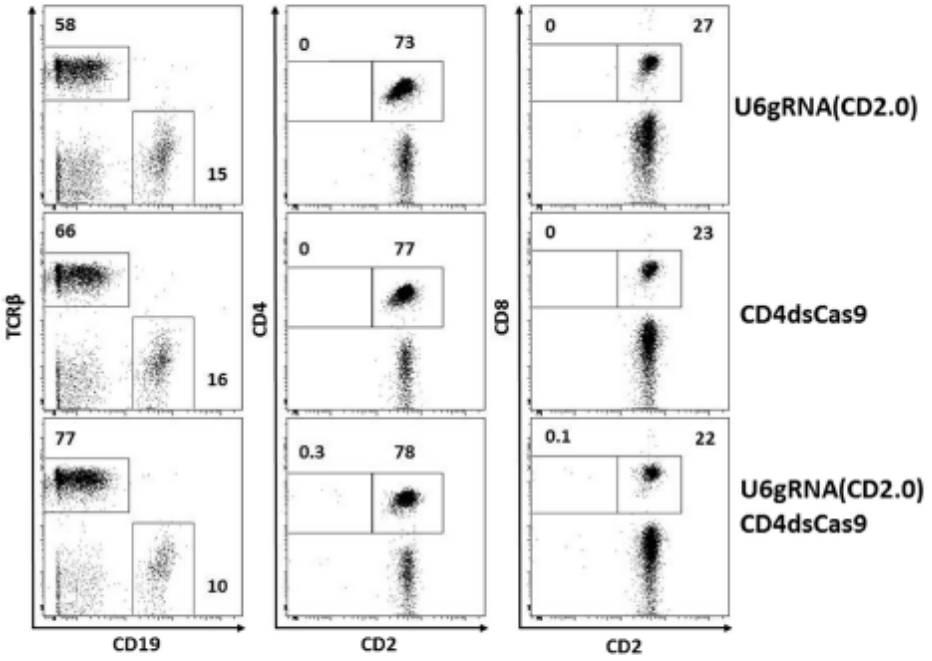


Figure 2. Abrogation of CD2 expression on a small population of peripheral blood T cells. Analysis of blood lymphocytes of the indicated transgenic mice by flow cytometry. Shown are live cells within a lymphocyte gate. CD4 and CD8 cells are additionally gated on TCRβ. The percentages of cells found within the marked gates of the dot plot analysis are shown.

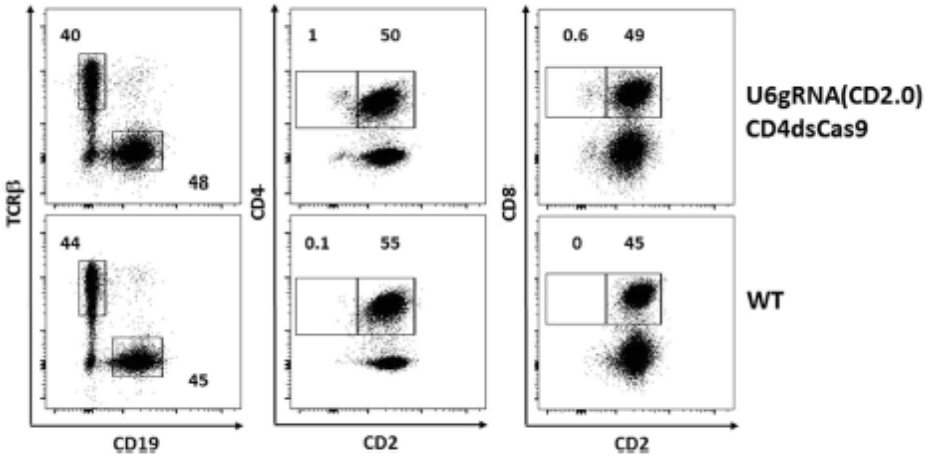
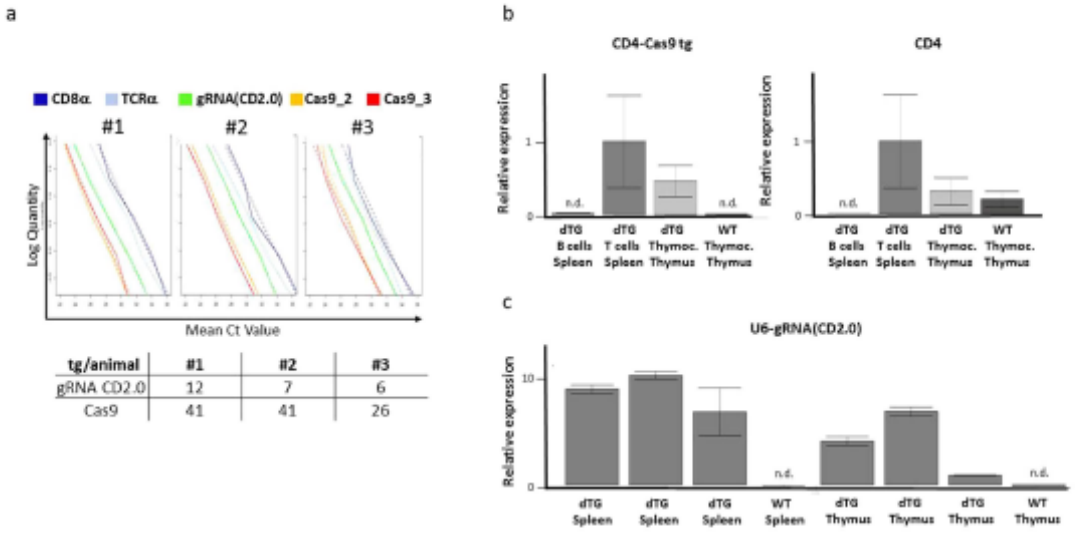


Figure 3. Abrogation of CD2 expression on a small population of lymph node/splenic T cells. Flow cytometric analysis of pooled lymphocytes from lymph nodes and spleens of two double transgenic and wildtype mice. Shown are live cells within a lymphocyte gate. The four plots on the right are additionally gated on TCRβ$^+$ cells. The percentages of cells found within the marked gates of the dot plot analysis are indicated.

recombination using an already mutated donor allele. In all 17 amplicons from CD2$^+$ T cells as well as nine amplicons from CD19$^+$ B cells, we obtained solely unmutated sequences (Fig. 5b). As an additional control, we performed single cell PCR on FVB/N wildtype CD4$^+$ and CD8$^+$ T cells and sequenced the PCR products, again yielding solely wildtype sequences (Fig. 5b). Sequence analysis of the mutations found within the amplicons from



tg/animal	#1	#2	#3
gRNA CD2.0	12	7	6
Cas9	41	41	26

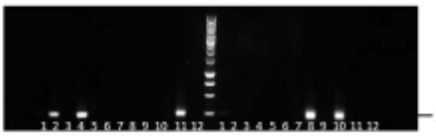
Figure 4. Analysis of transgene copy numbers and mRNA expression levels. (**a**) Serial dilutions of genomic DNA from three F1 offspring of founder #5 (Fig. 1c) were analyzed by real-time PCR specific for sequences within the indicated (trans)genes. A regression analysis was performed to calculate the copy number of the transgenes in each individual mouse, as indicated in the table. (**b**) Relative expression of CD4-Cas9 transgenic and endogenous CD4 mRNA. Transgenic splenic B cells, T cells and thymocytes as well as wildtype thymocytes were magnetically enriched and analyzed by RT-PCR. The CD4-Cas9 product is spanning the intron between CD4 Exon 1 and Cas9 ORF. Cxxc1 and Ywhaz served as housekeeping gene. Data was normalized to "dTG T cells Spleen". (**c**) Relative expression of gRNA(CD2.0). Splenocytes and thymocytes of three different dTG offspring as well as of a wildtype control were analyzed by RT-PCR. Cxxc1 and Ywhaz served as housekeeping gene. Data was normalized to "dTG Thymus" of the third offspring.

$CD2^-$ T cells revealed deletions spanning up to 19 bp and insertions of one to four base pairs, five were found as out of frame and one in frame (Fig. 5c). Two mutations were repeatedly found (Fig. 5c).

Even though we observed $CD2^-$ T cells in a gRNA/Cas9 double-transgenic situation, we were surprised that the number of these cells was extremely low. This may have been a specific result of this particular founder due to expression variegation or may constitute a general problem of the approach. We therefore analyzed the T cell compartments in our different transgenic mice in greater detail. The ratio of $CD4^+$ $TCR\beta^+$ and $CD8^+$ $TCR\beta^+$ T to $CD19^+$ B cells was found to be slightly increased in double-transgenic (8 ($CD4^+$ $TCR\beta^+$) and 2.2 ($CD8^+$ $TCR\beta^+$)) versus Cas9 single (4.8 ($CD4^+$ $TCR\beta^+$) and 1.4 ($CD8^+$ $TCR\beta^+$)) and gRNA transgenic animals (4.9 ($CD4^+$ $TCR\beta^+$) and 1.8 ($CD8^+$ $TCR\beta^+$)), thus ruling out a general counter-selection of $CD4^+$ and $CD8^+$ T cells. Since expression of nucleases such as Cre recombinase were reported to lead to apoptosis in situations of high proliferative activity[25] and also for Cas9 toxic *in vivo* effects were described[22], we analyzed $CD4^+$ and $CD8^+$ T cells for cell death (Suppl. Fig. 2a,b). We obtained no evidence of an increased number of dead cells in lymphocytes of the gRNA/Cas9 double-transgenic situation compared to Cas9 or gRNA single transgenic mice as well as wildtype mice. Thus, in our model we cannot find evidence for genotoxicity of gRNA and Cas9. Another important point regarding the use of CRISPR/Cas9 for *in vivo* and *ex vivo* use are off-target effects. Whereas some groups report high specificity of CRISPR/Cas9[26,27] others show high frequency of off-target effects[28,29]. Several groups found that off-target effects can be reduced by the use of optimized gRNAs[28,30,31]. In our case off-target effects were not analyzed because our on-target frequency of modification was already low[32].

Taken together, our data indicate that gRNA/Cas9 can be used to mutate target genes in specific cell lineages in the mouse. The frequency of such mutations in our system was low. We excluded that this was result of the actual gene deficiency by using a target whose absence was shown to have no effect over T cell survival[24]. Also, we did not find evidence that nuclease activity would increase cell death and thereby remove the edited cells from the population, which could explain the low frequency of mutations. One reason for inefficient abrogation of CD2 expression may, of course, be a feature of this particular transgenic founder. Defined and characterized Cas9-driver lines that are crossed to gRNA lines may overcome such problems in the future. To address this point we analyzed gRNA expression and Cas9 expression on mRNA and protein level. gRNA was expressed in both spleen and thymus. However, while we clearly detected Cas9 mRNA, we failed to see expression of the protein (0.25 ng Cas9 protein detection limit, data not shown). Thus, low protein expression levels in our particular founder line may contribute to the low efficiency of the system. To further address whether or not the efficiency of the used gRNA contributed to the observed low mutation frequencies we compared by *in vitro* digestion of a PCR product the used gRNA (gRNA(CD2.0) to three other gRNAs (gRNA(CD2.1,CD2.2,CD2.3)) targeting the same region (Fig. 6). We found that gRNA(CD2.1) showed almost no cutting and gRNA(CD2.2) complete digest. The gRNA(CD2.0) used for generating the transgenic mice presented with good but not complete cutting

a



b

Genotype	Population	# PCR Products	Sequence
CD4dsCas9/ U6gRNA(CD2.0)	$CD4^+CD2^+$	12	all WT
CD4dsCas9/ U6gRNA(CD2.0)	$CD8^+CD2^+$	5	all WT
CD4dsCas9/ U6gRNA(CD2.0)	$CD19^+$	9	all WT
CD4dsCas9/ U6gRNA(CD2.0)	$CD4^+CD2^-$	9	8 KO/1 WT
CD4dsCas9/ U6gRNA(CD2.0)	$CD8^+CD2^-$	5	1 KO/2 WT
Wildtype	$CD4^+$	7	all WT
Wildtype	$CD8^+$	7	all WT

c

		Origin	Frame
WT	5'-GGGATGACTAGGCTGGAGAAGG----ACCTGGACGTGAGGATTCTGGGTAAGTCTCC-3'	WT	
3x	5'-GGGATGACTAGGCTGGAGAAGG--GGACCTGGACGTGAGGATTCTGGGTAAGTCTCC-3'	CD4	OF
1x	5'-GGGATGACTAGGCTGGAGAAGG----------GGGTAAGTCTCC-3'	CD4	OF
2x	5'-GGGATGACTAGGCTGGAGAAGG---AACCTGGACGTGAGGATTCTGGGTAAGTCTCC-3'	CD4	OF
1x	5'-GGGATGACTAGGCTGGAGAAGGGTCGACCTGGACGTGAGGATTCTGGGTAAGTCTCC-3'	CD4	OF
1x	5'-GGGATGACTAGGCTGGAGA----------GGATTCTGGGTAAGTCTCC-3'	CD4	IF
1x	5'-GGGATGACTAGGCTGGAGAAGG---GACCTGGACGTGAGGATTCTGGGTAAGTCTCC-3'	CD8	OF

Figure 5. Single cell PCR analysis of the target region within the CD2 locus. Peripheral blood lymphocytes were surface stained for CD2, CD4, CD8, TCRβ and CD19 and the following populations within live cell and lymphocyte gates single cell-sorted by flow cytometry: $TCR\beta^+$ $CD4^+$ $CD2^+$, $TCR\beta^+$ $CD8^+$ $CD2^+$, $TCR\beta^+$ $CD4^+$ $CD2^-$, $TCR\beta^+$ $CD8^+$ $CD2^-$, $CD19^+$. A 253 bp long region including the gRNA target was amplified by two rounds of nested PCR. Products were cloned in pGEM-T and pGEM-Teasy and sequenced. For wildtype controls the single cell PCR products were column-purified and sequenced directly. (**a**) Agarose gel showing PCR products of single cell amplicons of the target region within the CD2 gene locus from sorted peripheral blood single cells. Lanes 6 and 12 on both sides of the marker are H_20 negative controls. (**b**) Table showing the number of obtained mutations in amplicons of double-transgenic and wildtype cells. (**c**) Alignment of the obtained sequences from the CD2 gene amplicon. Indicated is the number (left side) and the cell type as well as the outcome of the mutation (OF: out of frame, IF: in frame) (right side) of the respective sequence. – indicates a deletion and red bases insertions.

efficiency, along with gRNA(CD2.3) (Fig. 6). It thus appears unlikely that the low *in vivo* efficiency of CD4-Cas9/gRNA(CD2.0) is sole result of suboptimal gRNA performance. Methylation of target loci as factor contributing to efficiency of Cas9-mediated genome editing has been also discussed and assessed, though with contradictory results[18,32], but chromatin structure in general appears to influence Cas9 binding[32,33]. Interestingly, Cre-induced Cas9 expression from within the *gt(ROSA)26Sor* locus combined with viral gRNA expression also resulted in only infrequent gene editing in 0.1% to 1.8% of lung cells[20] while a recent report of doxycycline-regulated Cas9-mediated gene editing showed significantly higher frequencies of mutated cells[21]. gRNA library transduction experiments in EL4 cells, a mouse thymic cell line, was found to also lead to only 0.94–3.60% mutant cells when analyzing different surface markers[34]. Furthermore, we were not able to efficiently mutate CD2 in T cell transfection experiments (data not shown). While certainly the evidence is still anecdotal, T cells may be difficult to modify by CRISPR/Cas9. One way to increase the frequency of gene editing events leading to inactivation of gene functionality may be the use of multiple gRNAs towards the same target. Yet, even in such an approach only 2/3 of alleles will carry an out-of-frame mutation. CRISPR/Cas9-mediated *in vivo* targeted mutagenesis may thus work best when a critical region, such as an enzymatic site, is functionally compromised even through in-frame in/del mutations. Nevertheless, also in the here-described simple version conditional DNA editing is extremely useful. It allows rapid assessment of cell-type specific phenotypic changes in a time frame incomparable to classical conditional mutagenesis, especially when mutant cells can be easily identified for phenotypic analysis. Additionally, for the investigation of onco- and tumor suppressor genes the observed low editing frequency may prove to be advantageous. Taken together, this proof of concept study shows that direct conditional gene editing by gRNA/Cas9 transgenesis may extend the toolbox for genetic studies in the mouse.

Methods

Methods were carried out in accordance with the approved guidelines.

Animal Models. hCas9 (a gift from George Church (Addgene plasmid # 41815))[15] was PCR-amplified and placed into the second exon of CD4 in a construct (CD4-DEPE, gift from Marc Schmidt-Supprian) consisting of the CD4 promoter, the distal and the proximal enhancer, exon 1 and parts of exon 2 but lacking the intronic silencer, thus supporting expression in all αβ T cells. Subsequently, the plasmid was digested with NotI. The Crispr design tool at crispr.mit.edu was used to identify the protospacer specific for CD2 Exon 2 (5′-GACTAGGCTGGAGAAGGACC-3′), which was cloned via BbsI into a modified px330 vector (px330ccdBChR), thus replacing a ccdB/chloramphenicol cassette. The transgene was cut out by BciVI and XbaI. Both

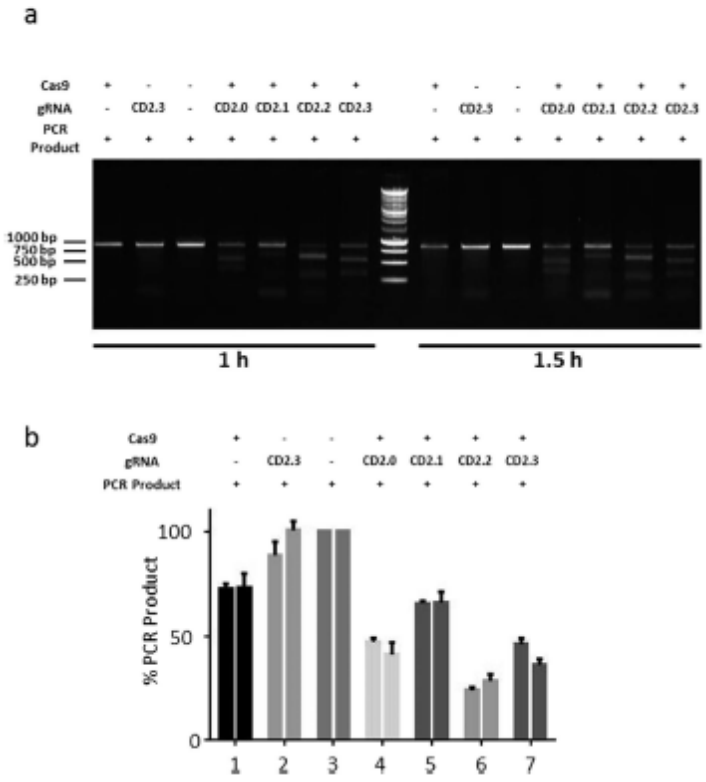


Figure 6. Functional analysis (**a**) *In vitro* test of gRNA efficiency. gRNA(CD2.0) as well as three controls (CD2.1,CD2.2, CD2.3) were incubated with a PCR product for the indicated period of time. Digests were separated on an agarose gel. (**b**) Analysis of the gRNA efficiency. The intensities of the bands resulting from gRNA/Cas9-digested PCR product of three different experiments were analyzed using ImageJ. Shown is the mean and standard error of the mean.

constructs were injected into FVB/N oocytes. The founders were screened by PCR using the following primers: CD4 Cas9 typ fwd: 5′ - tgc tca caa ccc ttt agt tt-3′, CD4 Cas9 typ rev: 5′ -ctt ttt atc ctc ctc cac c-3′ (product length: 835bp); U6 fwd: 5′ -gag ggc cta ttt ccc atg att cc-3′, T7 gRNA rev: 5′ -gca cgc gct aaa aac gga-3′ (product length: 407bp). Animals were kept in barrier-SPF level animal facilities at Technische Universität München according to German Animal Protection Law. Experiments were conducted under the license number 55.2-1-54-2532-2-12 and approved by the Regierung von Oberbayern.

Transgene copy number determination. We estimated genomic transgene abundance by performing a RT-PCR using SsoFast EvaGreen (BioRad) with following primers: Cas9F2: 5′ -aag aga acc cga tca acg a-3′, Cas9R2: 5′ -aga tta cca aac agg ccg t-3′; Cas9F3: 5′ -gcg cta ggc tgt cca aat-3′, Cas9R3: 5′ -att taa agt tgg ggg tca gcc-3′; gRNA(CD2.0)F2: 5′ -tgg aaa gga cga aac acc ga-3′, gRNA(CD2.0)R2: 5′ -cac gcg cta aaa acg gac-3′; CD8aFb: 5′ -caa gga agc aag tgg tat gaa-3′, CD8aRb: 5′ -ttt cca gat tta ccg tac cg-3′; TCRaF: 5′ -tcc ggc caa acc atc tgt-3′ and TCRaR: 5′ -cgc tgg ggg aga tga cta t-3′. A linear regression was performed, in which the dependent variable, the (decadic) logarithm of the amount of DNA used in each PCR experiment was explained by the mean Ct-value (number of PCR cycles until detection level) in replicate experiments and gene-specific primer pairs were used.

$$1.\ \log_{10}(\text{Quantity of genomic DNA}) \sim \text{Ct}_{\text{value}} + \text{Gene}$$

Here, "Ct value" is a continuous variable, and "Gene" is a nominal variable with 5 distinct values coding for the 5 primer pairs used ("CD8a","TCRa","gRNA CD2.0","Cas9 2","Cas9 3"). As a result, we obtain one coefficient for "Ct_{value}", β_{Cvalue}, and 5 coefficients for "Gene", β_{CD8a}, β_{TCRa}, $\beta_{\text{gRNA(CD2.0)}}$, $\beta_{\text{Cas9\_2}}$, $\beta_{\text{Cas9\_3}}$ (one of the latter 5 is redundant and set to zero. For our purpose, it is irrelevant which one is chosen). As a result, we obtain a regression line for each primer pair (see plots Fig. 1d). The difference in the coefficients for each factor level of "Gene" is the log10 of the fold difference in genomic abundance of the corresponding amplicons (genes). We chose, the mean $2.\ \gamma = (\beta_{\text{CD8a}} + \beta_{\text{TCRa}})/2$ of the coefficients β_{CD8a} and β_{TCRa} as reference value, corresponding to two genomic copies of an amplicon. The genomic abundance of gRNA CD2.0 was then estimated as $3.\ 2^{*}10^{\wedge}(\gamma - \beta_{\text{TCRa}})$ (the factor 2 accounts for the two copies of the reference genes), and the genomic abundance of Cas9 was estimated

as 4. $2^{*}10^{\wedge}(\gamma - (\beta_{Cas9\_2} + \beta_{Cas9\_3})/2)$. In the latter case, we averaged the two abundance measurements for Cas9. Calculations were carried out in R (Version 3.2.0)[35].

RT-PCR. The determination of CD4-Cas9 expression was performed on MACS-isolated B and T cells using mouse CD19 microbeads and mouse Pan T Cell Isolation Kit II according to manufacturer's instruction (Miltenyi Biotec). RNA was isolated using RNeasy Mini Kit (Qiagen) and transcribed with iScript cDNA Synthesis Kit (BioRad). RT-PCR was performed with SsoFast EvaGreen Supermix (BioRad) and template-free negative controls were included. For the quantification of U6-gRNA(CD2.0) the total RNA of thymus and spleen was isolated using RNeasy Mini Kit (Qiagen) and reverse-transcribed with Moloney Murine Leukemia Virus Reverse Transcriptase RNase H– Point Mutant (Promega) according to the manufacturer's instructions. RT-PCR was performed with KAPA SYBR FAST qPCR Universal Master Mix (Peqlab) and template-free and reverse transcriptase-free negative controls were included. Both transcript abundances were assessed using the Bio-Rad CFX384 systems and following primers were used: CD4-Cas9 F: 5′-agc cct cat ata cac aca cct-3′, CD4-Cas9 R: 5′-acg ctg ttt gtg ccg ata-3′; CD4 F: 5′-aga act ggt tcg gca tga ca-3′, CD4 R: 5′-aag gag aac tcc gct gac-3′; CD2 F: 5′-ccc atg att cct tca tat ttg ca-3′ and CD2 R: 5′-ggt cct tct cca gcc tag t-3′; Cxxc1 F: 5′-ggt tgt tgc acg ggg tcc ag-3′, Cxxc1 R: 5′-ccc cat tct cag act tgc tgt cg-3′; Ywhaz F: 5′-gtt act tgg ccg agg ttg ctg ct-3′, Ywhaz R: 5′-ggt gtg tcg gct gca tct cct t-3′. Cxxc1 and Ywhaz were used as reference genes. The $\Delta\Delta$Cq Analysis was performed with BioRad CFX Manager 3.1.

Flow cytometry. Blood samples and splenocytes were treated with red blood cell lysis buffer and along with lymph node cells and thymocytes were washed with phosphate buffered saline (PBS). They were stained in FACS buffer (PBS, 10% FCS, 0.01% NaN_3) with anti-CD2 FITC, anti-TCRβ PE, anti-CD8 PerCP, anti-CD19 APC and anti-CD4 PB (all Biolegend). Live/Dead discrimination was performed by propidium iodide or Live Dead Fixable Aqua Dead Cell Stain Kit (Life Technologies). Single cells were sorted (BD, MoFlow) onto AmpliGrid slides and processed immediately. For FACS analysis, cells were acquired at BD Canto II and Aria 5. Analysis was performed with FlowJo 9.4 and 10.

Molecular Biology. Single cell PCR was performed as a nested PCR. The outer PCR was performed with Advantage2 (Clontech) directly on the AmpliGrid slide. For the inner PCR, Herculase II Fusion Enzyme with dNTP combos (Agilent) was used. PCRs were performed with the following primers: out fwd: 5′-atc acc ctg aac atc ccc aac-3′, out rev: 5′-act gga gtc ttc tt gtg ggc-3′ (product length: 382bp); in fwd: 5′-ctg gtc gca gag ttt aaa agg-3′, in rev: 5′-gct gct ccc caa ctt tct ac-3′ (product length 253bp). One AmpliGrid slide consists of 48 wells, of which every sixth sample was not loaded with a cell and served as negative control.

The Cas9-gRNA efficiency test of four different gRNAs targeting CD2 was performed *in vitro* on column-purified 839 bp CD2 PCR products (Wizard SV Gel and PCR Clean-UP System, Promega) which were amplified using Taq PCR Master Mix Kit (Qiagen) and following primers: CD2 cut F: 5′-cct gac aga aag aac tc-3′, CD2 cut R: 5′-ctc act gct cct agg ca-3′. Either PCR product or annealed oligos (PCR primer: T7 CD2.0 F: 5′-taa tac gac tca cta tag ggg act agg ctg gag aag gac c-3′, T7 CD2.0 R: 5′-gca gcg gct aaa aac gga-3′; oligos: T7 promoter F: 5′-taa tac gac tca cta tag gg-3′, CD2.1 R: 5′-aaa agc acc gac tcg gtg cca ctt ttt caa gtt gat aac gga cta gcc tta ttt taa ctt gct att tct agc tct aaa aca aga cac ccc aga tgg tct ccc cta tag tga gtc gta tta-3′, CD2.2 R: 5′-aaa agc acc gac tcg gtg cca ctt ttt caa gtt gat aac gga cta gcc tta ttt taa ctt gct att tct agc tct aaa acg aac atc ccc aac ttt caa acc cta tag tga gtc gta tta-3′, CD2.3 R: 5′-aaa agc acc gac tcg gtg cca ctt ttt caa gtt gat aac gga cta gcc tta ttt taa ctt gct att tct agc tct aaa act cgc acc tca tca ata tca tcc cta tag tga gtc gta tta-3′) were used to transcribe RNA with MEGAshortscript T7 Transcription Kit (ThermoFisher Scientific) and purified with MEGAclear Transcription Clean-Up Kit (ThermoFisher Scientific). gRNA and Cas9 Nuclease, S pyogenes (NEB) were preincubated for 10 min at room temperature. After adding the CD2 PCR product the mix was incubated for either 1 h or 1.5 h at 37 °C and finally analyzed on an agarose gel. The following fragment sizes are expected: gRNA(CD2.0): 499 bp and 340 bp, gRNA(CD2.1): 251 bp and 588 bp, gRNA(CD2.2): 269 bp and 570 bp, gRNA(CD2.3): 312 bp and 527 bp. The signal quantification was performed using ImageJ 1.50a.

References

1. Hooper, M., Hardy, K., Handyside, A., Hunter, S. & Monk, M. HPRT-deficient (Lesch-Nyhan) mouse embryos derived from germline colonization by cultured cells. *Nature* **326**, 292 (1987).
2. Orban, P. C., Chui, D. & Marth, J. D. Tissue- and site-specific DNA recombination in transgenic mice. *Proc Natl Acad Sci USA* **89**, 6861–6865 (1992).
3. Gu, H., Marth, J. D., Orban, P. C., Mossmann, H. & Rajewsky, K. Deletion of a DNA polymerase beta gene segment in T cells using cell type-specific gene targeting [see comments]. *Science* **265**, 103–106 (1994).
4. Kühn, R., Schwenk, F., Aguet, M. & Rajewsky, K. Inducible gene targeting in mice. *Science* **269**, 1427–1429 (1995).
5. Meyer, M., de Angelis, M. H., Wurst, W. & Kuhn, R. Gene targeting by homologous recombination in mouse zygotes mediated by zinc-finger nucleases. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 15022–15026, doi: 10.1073/pnas.1009424107 (2010).
6. Hermann, M. *et al.* Evaluation of OPEN zinc finger nucleases for direct gene targeting of the ROSA26 locus in mouse embryos. *PLoS One* **7**, e41796, doi: 10.1371/journal.pone.0041796 (2012).
7. Sung, Y. H. *et al.* Knockout mice created by TALEN-mediated gene targeting. *Nat Biotechnol* **31**, 23–24, doi: 10.1038/nbt.2477 (2013).
8. Wang, H. *et al.* One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* **153**, 910–918, doi: 10.1016/j.cell.2013.04.025 (2013).
9. Peng, Y. *et al.* Making designer mutants in model organisms. *Development* **141**, 4042–4054, doi: 10.1242/dev.102186 (2014).
10. Bibikova, M., Golic, M., Golic, K. G. & Carroll, D. Targeted chromosomal cleavage and mutagenesis in Drosophila using zinc-finger nucleases. *Genetics* **161**, 1169–1175 (2002).
11. Bibikova, M., Beumer, K., Trautman, J. K. & Carroll, D. Enhancing gene targeting with designed zinc finger nucleases. *Science* **300**, 764, doi: 10.1126/science.1079512 (2003).
12. Ran, F. A. *et al.* Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* **154**, 1380–1389, doi: 10.1016/j.cell.2013.08.021 (2013).

13. Mashiko, D. *et al.* Generation of mutant mice by pronuclear injection of circular plasmid expressing Cas9 and single guided RNA. *Sci Rep* **3**, 3355, doi: 10.1038/srep03355 (2013).
14. Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823, doi: 10.1126/science.1231143 (2013).
15. Mali, P. *et al.* RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826, doi: 10.1126/science.1232033 (2013).
16. Wiedenheft, B., Sternberg, S. H. & Doudna, J. A. RNA-guided genetic silencing systems in bacteria and archaea. *Nature* **482**, 331–338, doi: 10.1038/nature10886 (2012).
17. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821, doi: 10.1126/science.1225829 (2012).
18. Hsu, P. D. *et al.* DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol* **31**, 827–832, doi: 10.1038/nbt.2647 (2013).
19. Heigwer, F., Kerr, G. & Boutros, M. E-CRISP: fast CRISPR target site identification. *Nat Methods* **11**, 122–123, doi: 10.1038/nmeth.2812 (2014).
20. Platt, R. J. *et al.* CRISPR-Cas9 knockin mice for genome editing and cancer modeling. *Cell* **159**, 440–455, doi: 10.1016/j.cell.2014.09.014 (2014).
21. Dow, L. E. *et al.* Inducible *in vivo* genome editing with CRISPR-Cas9. *Nat Biotechnol* **33**, 390–394, doi: 10.1038/nbt.3155 (2015).
22. Port, F., Chen, H. M., Lee, T. & Bullock, S. L. Optimized CRISPR/Cas tools for efficient germline and somatic genome engineering in Drosophila. *Proc Natl Acad Sci USA* **111**, E2967–2976, doi: 10.1073/pnas.1405500111 (2014).
23. Sawada, S., Scarborough, J. D., Killeen, N. & Littman, D. R. A lineage-specific transcriptional silencer regulates CD4 gene expression during T lymphocyte development. *Cell* **77**, 917–929 (1994).
24. Killeen, N., Stuart, S. G. & Littman, D. R. Development and function of T cells in mice with a disrupted CD2 gene. *EMBO J* **11**, 4329–4336 (1992).
25. Schmidt-Supprian, M. & Rajewsky, K. Vagaries of conditional gene targeting. *Nat Immunol* **8**, 665–668, doi: 10.1038/ni0707-665 (2007).
26. Veres, A. *et al.* Low incidence of off-target mutations in individual CRISPR-Cas9 and TALEN targeted human stem cell clones detected by whole-genome sequencing. *Cell Stem Cell* **15**, 27–30, doi: 10.1016/j.stem.2014.04.020 (2014).
27. Gilbert, L. A. *et al.* CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. *Cell* **154**, 442–451, doi: 10.1016/j.cell.2013.06.044 (2013).
28. Lin, Y. *et al.* CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences. *Nucleic acids research*, doi: 10.1093/nar/gku402 (2014).
29. Tsai, S. Q. *et al.* GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nature biotechnology* **33**, 187–197, doi: 10.1038/nbt.3117 (2015).
30. Cho, S. W. *et al.* Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome research* **24**, 132–141, doi: 10.1101/gr.162339.113 (2014).
31. Kim, D. *et al.* Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. *Nat Meth* **12**, 237–243, doi: 10.1038/nmeth.3284 (2015).
32. Wu, X. *et al.* Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nat Biotech* **32**, 670–676, doi: 10.1038/nbt.2889 (2014).
33. Kuscu, C., Arslan, S., Singh, R., Thorpe, J. & Adli, M. Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nat Biotech* **32**, 677–683, doi: 10.1038/nbt.2916 (2014).
34. Doench, J. G. *et al.* Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nat Biotech* **32**, 1262–1267, doi: 10.1038/nbt.3026 (2014).
35. R. Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/ (2013).

Acknowledgements

We thank Karin Mink and Julia Ritter for oocyte injections. We are grateful for Martin Skerhut's help with mouse breeding. We thank the animal caretaker team for their support of this study. Furthermore, we thank the Flow Cytometry Facilities of the Institute for Med. Microbiology, Immunology and Hygiene at the Technische Universität München and of the University of Zurich, especially Lynette Henkel and Claudia Dumrese. We are grateful to access to machinery of the Foundation for People with Rare Diseases, of the Neurobiochemical Laboratory of *the Kinder- und Jugendpsychiatrischer Dienst* of the Canton of Zurich, of the Center for Molecular Cardiology of the University of Zurich, the Research Unit Internal Medicine of the University Hospital of Zurich, and of the Institute of Virology of the Technische Universität München. We thank Sabine Specht and Daniela Schildknecht for assistance with western blot establishment as well as Stefan Haak for logistical support.

Author Contributions

J.B. and T.B. designed the experiments. J.B., G.D., P.P., J.B., F.M., M.G. and K.S. performed the experiments under supervision of T.B., D.B., M.G. and B.B. J.B., T.B. and A.T. performed data analysis. J.B. and T.B. co-wrote and all authors revised the manuscript.

Additional Information

Supplementary information accompanies this paper at http://www.nature.com/srep

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Beil-Wagner, J. *et al.* T cell-specific inactivation of mouse CD2 by CRISPR/Cas9. *Sci. Rep.* **6**, 21377; doi: 10.1038/srep21377 (2016).

8 Acknowledgements

Nach zehn Jahren Studium kann ich nun mit Stolz meine fertige Doktorarbeit in der Hand halten. All das wäre allerdings nicht ohne die Hilfe vieler Menschen möglich gewesen, die mich sowohl wissenschaftlich als auch privat unterstützt haben. Bei einigen möchte ich mich an dieser Stelle bedanken.

Anfangen möchte ich mit meinem Doktorvater Prof. Dr. Thorsten Buch. Auch wenn du nun offiziell nach unserem Laborumzug in die Schweiz nicht mehr mein Betreuer bist, hast du doch einen großen Teil zum Entstehen dieser Arbeit beigetragen und standst mir wissenschaftlich jederzeit mit Rat und Tat zur Seite. Danke auch für die Freiheiten und Möglichkeiten, die du mir eröffnet hast.

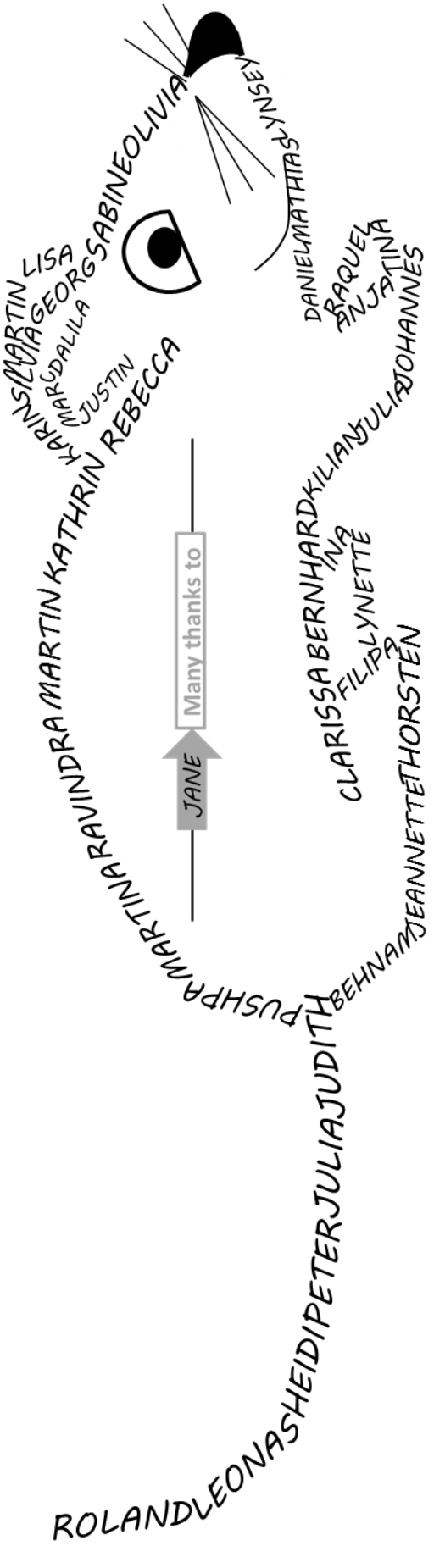
Vielen Dank auch an meine Komiteemitglieder apl. Prof. Dr. Clarissa Prazeres da Costa und Prof. Dr. Bernhard Küster für die Betreuung während der Arbeit und das wertvolle Feedback.

Einen großen Dank gilt der alten AG Buch und meinen neuen Institutskollegen der LTK aber auch dem gesamten Institut für med. Mikrobiologie, Immunologie und Hygiene. Eure Hilfestellung und Motivation sowie die daraus entstehende freundschaftliche und lustige Atmosphäre haben auch den trübsten Labortag erhellt und jedes unmögliche Ziel erreichen lassen.

Nicht zuletzt möchte ich meiner Familie und meinen Freunden danken. Mein Studium war nicht nur wissenschaftlich eine Reise, auf der ich viele Persönlichkeiten traf. So wurden aus manchen Kommilitonen auch enge Freunde, die mein Leben immer noch begleiten und mich durch reges Interesse jeden Tag motiviert haben.

Ein besonderer Dank gilt auch meinen Eltern für die jahrelange Unterstützung, die Motivation, Förderung und Reisebereitschaft mich auf jedem meiner Schritte zu begleiten. Ihr habt mich zu dem gemacht, was ich heute bin.

Zum Schluss möchte ich meinem Mann danken. Deine positive Lebenseinstellung und deine Liebe haben mich auch aus dem größten Dissertationstief wieder rausgeholt. Du hast mich jederzeit mit Motivation und Verständnis unterstützt und mir das schönste Geschenk der Welt gemacht.


Many thanks to
JANE
ROLAND LEONA SHEIDI PETER JULIA JUDITH
PUSHPA MARTINA RAVINDRA MARTIN KATHRIN REBECCA
CLARISSA BERNHARD KILIAN JULIA JOHANNES
BEHNAM JEANNETTE THORSTEN
FILIPA LYNETTE
adapted from Dr. Judith Habicher