UNIVERSIDADE D
COIMBRA

Miguel Ângelo Simão Dias

UNCOVERING THE REGULATORY T CELL TRANSCRIPTIONAL SIGNATURE IN THE HUMAN THYMUS

Dissertação no âmbito do Mestrado de Biologia Celular e Molecular, orientada pelo Doutor Alexandre Raposo e pela Doutora Emília Duarte, apresentada ao Departamento de Ciências da Vida, Faculdade de Ciências e Tecnologia da Universidade de Coimbra

Agosto de 2018

Mestrado em Biologia Celular e Molecular

UNCOVERING THE REGULATORY T CELL TRANSCRIPTIONAL SIGNATURE IN THE HUMAN THYMUS

Miguel Ângelo Simão Dias

Dissertação no âmbito do Mestrado de Biologia Celular e Molecular, orientada pelo Doutor Alexandre Raposo e pela Doutora Emília Duarte, apresentada ao Departamento de Ciências da Vida, Faculdade de Ciências e Tecnologia da Universidade de Coimbra

Agosto de 2018

"All you need is passion. If you have a passion for something,
you'll create the talent"
Yanni

AGRADECIMENTOS

Esta tese não seria possível sem a ajuda e presença de muitas pessoas que foram fundamentais ao longo desta fase, e das quais nunca me irei esquecer.

Em primeiro lugar, gostaria de agradecer à Doutora Ana Espada de Sousa por me ter recebido no seu laboratório, pela sua constante disponibilidade, pela confiança que depositou em mim, e por me ter dado a conhecer e a admirar a difícil, mas extraordinária, área que é a imunologia.

Ao meu orientador Doutor Alexandre Raposo, por me ter mostrado e ensinado todo um novo e fantástico mundo até então desconhecido. A sua exigência, incentivo e conhecimento foram fundamentais para a realização desta tese. Se hoje tenho as competências que tenho, a ele lho devo. Obrigado por toda a dedicação e esforço durante esta etapa.

Um obrigado à Doutora Emília Duarte por toda a sua disponibilidade e apoio mostrado ao longo deste ano.

Um agradecimento especial a todos os meus colegas do laboratório de imunologia clínica que me fizeram sentir como se fosse a minha segunda casa durante este ano. Em primeiro lugar à Ana Luísa por ter sido uma amiga, por me ter ajudado deste o princípio e por ter estado presente. À Helena por todo o conhecimento que me transmitiu. À Dra. Susana e à professora Conceição pela boa disposição e por estarem sempre dispostas a ajudar. Ao João, à Weronika, à Vitória, à Kikas, ao Pedro, à Chinita e à Catarina. Obrigada pelos momentos de galhofa e alegria que foram muito importantes nestes últimos meses, são os melhores colegas de laboratório!

A todo o pessoal do JBarata lab, em particular ao Bruno, à Ana Cachucho, à Mafalda Duque e à Mafalda Matos.

Ao grande GDIMM e aos seus craques por todas as jogatanas que tanta falta fizeram nos dias mais stressantes e frustrantes.

Gostaria também de agradecer em particular à Ana Sofia por toda a ajuda, por ter sido uma amiga, por termos partilhado a mesma etapa em conjunto, e por toda a companhia ao longo deste ano.

Um agradecimento muito especial à minha namorada, Marta, pela paciência, por me aturar há 8 anos, pelo amor, pelas parvoeiras, por nunca ter desistido quando a distância era demasiada, e por ter sido a melhor pessoa com quem eu podia ter vivido esta etapa.

Por último, aos meus pais, irmão e avós. Obrigado por terem acreditado e confiado em mim, por me apoiarem, por estarem sempre ao meu lado, pelo sacrifício que fizeram para que eu pudesse chegar aqui, por serem os melhores do mundo. Espero nunca vos desiludir. Esta tese é para vocês.

These thesis's results were presented by this thesis' author in national and international meetings, namely: Poster presentation at the XLIV Annual Meeting of the Portuguese Society for Immunology, 27th-29th of June 2018, Lisbon, Portugal; Oral presentation at the V European Congress of Immunology, 2nd-5th of September 2018, Amsterdam, Netherlands.

ABSTRACT

Regulatory T cells (Tregs) are key players in maintaining immune homeostasis, by preventing or limiting immune responses. They are particularly efficient in suppressing conventional T cells (Tconvs), and in this way control the immunopathology associated with immunity against pathogens and cancer as well as preventing allergy, autoimmune diseases, and chronic inflammation. An important Treg subset is generated during T cell development in the thymus, known as thymic-derived Treg. It is best defined by the expression of the forkhead box protein FOXP3, a transcription factor that plays a crucial role in Treg cell differentiation by repressing the expression of genes otherwise upregulated in Tconv cells, as well as by promoting to the activation of Treg specific genes, including *IL2Rα* (CD25) and *CTLA4*. However, recent studies have shown that Treg commitment may occur independently of FOXP3, indicating that other factors, presently unknown, are sufficient for the generation of Tregs.

Here I have investigated the differentiation and commitment of Treg cells in the human thymus, in order to identify novel factors potentially involved in this decision. To do this, we FACS sorted mature CD4 single-positive thymic Tregs (tTregs) and their conventional counterparts (tTconvs) based on the expression of CD27, CD25, and CD127 markers, from three human thymuses collected during pediatric corrective cardiac surgery, and generated their respective genome-wide expression profiles by RNA-seq. We ensure that these thymuses have an immunophenotype representative of all stages of T cell development and consistent with the one described in the literature.

Our comparative transcriptomic analysis identified 1047 genes significantly differentially expressed between tTreg and tTconv subsets, with 648 of these up-regulated in tTregs. Amongst these, I observed the prominent expression of Treg-associated genes, including *FOXP3*, the *IL2Rα* (CD25), *CTLA4*, *TNFRSF4* (OX40), *TNFRSF18* (GITR), *IKZF2* (HELIOS) and *IKZF4* (EOS). From these, I identified a set of 196 genes that are uniquely expressed in tTreg compared to tTconv, encoding proteins with relevance to Treg biology, such as TNFRSF8, LRRC32, and CCR8, as well as others with no previously reported activity in tTreg cells, as DNAH8 and TNFRSF11A. Whilst DNAH8 expression may be indicative of the formation of immunological synapses, the expression of TNFRSF11A may indicate an additional suppression mechanism by which Tregs prevents Tconv cell activation.

From the genes found to be up-regulated in tTreg cells compared to tTconvs, 46 were transcription factors. These include some known to be directly involved in Treg development, such as *FOXP3, IKZF2, IKZF4, FOXO1,* and *NR4A3*, as well as transcription factors involved in Tconv cell activation and differentiation, such as *TBX21*, *IRF4*, *STAT4*, *BATF* and *RORA*. In addition, several members of the NF-kB pathway (*REL, RELB, NFKB2*) are also up-regulated in tTregs, indicating the activated state of this pathway during tTreg differentiation. Importantly, a set of transcription factors with no previous reported role in human regulatory T cells, *IRF5,*

ZBTB38, KLF6, and *CREB3L2*, are overexpressed in tTregs, suggesting additional layers of transcriptional regulation of Treg cell differentiation and function.

Altogether, this thesis presents the first transcriptomic profile of the human thymic Treg and Tconv subsets, which analyses are absolutely necessary to the understanding of their development. So far, they allowed the identification of novel genes with unreported functions in these subsets, which might represent additional factors involved in the definition of thymic T cells; in the near future, these data will open several new lines of research aiming to clarify the pathways of Treg lineage commitment in the human thymus and will help in the definition of the expression signature of human tTreg subset.

Key words: Regulatory T cells, Human Thymus, RNA-seq, FOXP3, T-cell development.

RESUMO

Os linfócitos T reguladores (Tregs) desempenham um papel crucial na manutenção da homeostasia imunológica, impedindo ou limitando respostas imunes. As Tregs são particularmente eficazes a suprimir as células T convencionais (Tconvs), e desta forma limitam a imunopatologia associada à imunidade contra patogéneos e células cancerígenas, bem como os processos alérgicos, autoimunes e inflamatórios. Uma população significativa de Tregs é gerada durante o desenvolvimento das células T no timo conhecidas como Tregs naturais (ou derivadas do timo), sendo definidas pela presença da proteína FOXP3. Este fator de transcrição desempenha um papel crucial na diferenciação destas células, não só através da repressão de genes normalmente expressos em Tconvs, mas também promovendo a ativação de genes específicos de Tregs, incluindo *IL2RA* (CD25) e *CTLA4*. No entanto, vários estudos demostraram que a diferenciação e comprometimento destas células pode ser independente de FOXP3, indicando a existência de outros fatores, atualmente desconhecidos, os quais são suficientes para promover o desenvolvimento das Tregs.

Neste estudo, investiguei a diferenciação e o comprometimento das células Treg no timo humano, a fim de identificar novos fatores potencialmente envolvidos nesta decisão. Para isso, isolámos Tregs e Tconvs tímicas maduras CD4 positivas com base na expressão dos marcadores CD27, CD25 e CD127, a partir de tecido tímico humano removido durante cirurgias cardíacas pediátricas corretivas de três indivíduos, e geramos os seus respetivos perfis de expressão génica através da sequenciação do RNA (RNA-seq).

A nossa análise da comparação da transcrição identificou 1047 genes significativamente e diferencialmente expressos entre tTregs e tTconvs, dos quais 648 com sobre expressão em tTregs. Destes 648 genes, observei a expressão proeminente de alguns genes associados a este tipo celular, incluindo *FOXP3*, *IL2Rα* (CD25), *CTLA4*, *TNFRSF4* (OX40), *TNFRSF18* (GITR), *IKZF2* (HELIOS) e *IKZF4* (EOS). Para alem disso, identifiquei um conjunto de 196 genes unicamente expressos em tTreg em comparação com tTconv, alguns dos quais codificam proteínas com conhecida relevância na biologia destas células, incluindo *TNFRSF8* (CD30), *LRRC32* (GARP) e *CCR8*, bem como genes cuja função em tTregs é desconhecida, nomeadamente *DNAH8* e *TNFRSF11A*. Enquanto a expressão de DNAH8 pode ser indicativa da formação de sinapses imunológicas, a expressão de TNFRSF11A pode sugerir um mecanismo adicional de supressão através dos quais as tTregs previnem a ativação das tTconvs.

Dos genes encontrados sobre expressos nas tTregs em comparação com as tTconvs, 46 codificam fatores de transcrição. Estes incluem alguns já conhecidos como estando diretamente envolvidos na diferenciação destas células, como *FOXP3*, *IKZF2*, *IKZF4*, *FOXO1* e *NR4A3*, bem como fatores de transcrição envolvidos na ativação e diferenciação de Tconvs, como o *TBX21*, *IRF4*, *STAT4*, *BATF* e *RORA*. Além disso,

identifiquei também membros da via de sinalização NF-kB (*REL*, *RELB* e *NFKB2*), indicando o seu estado ativo durante a diferenciação destas células.

Por fim, encontrei um grupo de fatores de transcrição sobre expressos em Tregs e sem funções previamente descritas nestas células, nomeadamente *IRF5*, *ZBTB38*, *KLF6* e *CREB3L2*, o que sugere a presença de novas vias de regulação da transcrição envolvidas nos processos de diferenciação e função das células T reguladoras

Em conclusão, esta tese apresenta o primeiro perfil de transcrição de Tregs e Tconvs de timo humano, cujas análises serão fundamentais para a compreensão do seu desenvolvimento. Até à data, estes dados permitiram a identificação de novos genes com funções desconhecidas nestes grupos celulares, o que poderá representar fatores adicionais envolvidos na definição das células T no timo. No futuro, estes dados permitirão explorar novas linhas de investigação com o objetivo de esclarecer o desenvolvimento da linhagem Treg no timo humano, bem como ajudar na definição da assinatura de expressão destas células.

Palavras-chave: Células T reguladoras, Timo Humano, RNA-seq, FOXP3, Desenvolvimento de células T.

INDEX OF CONTENTS

INDEX OF FIGURES

INDEX OF TABLES

INDEX OF ANNEXES

LIST OF ABBREVIATIONS

AIRE – Autoimmune regulator

APC – Antigen presenting cells

APECED – Autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy

BAM – Binary Alignment Map

BLS – Bare lymphocyte syndrome

CD – Cluster of Differentiation

CD4ISP – CD4 Immature single positive

cjTREC – coding-joint TRECs

CM – Complete medium

CNS – Conserved non-coding sequence

CPM – Counts per million

cTEC – cortical thymic epitelial cells

CVID – Common Variable Immunodeficiency Disease

DCs – Dendritic cells

DGE – Differentially gene expression

EDP – Early double positive

FACS – Fluorescence-activated cell sorting

FBS – Fetal bovine serum

FC – Fold-change

FDR – False Discovery Rate

FTOC – Fetal thymic organ culture

FWD – Forward

GO – Gene ontology

Hh – Hedgehog

HPC – High-Performance Computing

HSC – Hematopoietic stem cell

IL– Interleukin

IPEX – Immunedysregulation, Polyendrocrinopathy, Enteropathy, X-linked syndrome

mAbs – Monoclonal antibodies

MDS – Multi-dimensional scalling

MHC – Major histocompatibility complex

mTEC – medullary thymic epithelial cells

mTOR – mammalian target of rapamycin

NB – Negative binomial

NGS – Next-Generation Sequencing

NK – Natural killer

OCT – Optimal Cutting Temperature compound

PCA – Principle-component analysis

PE – Paired-end

pTreg – Peripheral induced/adaptive Treg

pTα – Pre-TCRα

REV – Reverse

SAM – Sequence Alignment Map

SCF – Stem cell factor

SCID – Severe combined immune deficiencies

sjTREC – signaljoint TRECs

SP – Single positive

Tconv – Conventional T-cell

TCR – T-cell receptor

TEC – Thymic epithelial cells

TF – Transcription factor

Tfh – T follicular helper

Th – T-helper

TN – Triple negative

TREC – TCR excision circles

Treg – Regulatory T cell

TSAs – Tissue-specific antigens

tTconv – thymic-derived Tconv

tTreg – Thymic-derived naturally Treg

VST – Variance-stabilizing transformation

γc – Common γ chain

CHAPTER 1 - INTRODUCTION

1.1 The human thymus

The thymus is a primary lymphoid organ of the immune system and provides a specialised and architecturally organised microenvironment to support the differentiation of hematopoietic progenitor cells (HPC) into thymus-dependent (T) lymphocytes, or T cells[1]. This organ is located just above the heart and consists of several distinct anatomical compartments, including the sub-capsular area, the cortex, the cortical medullary junction, an inner medulla, and Hassall's bodies[2,3] (Figure 1.1).

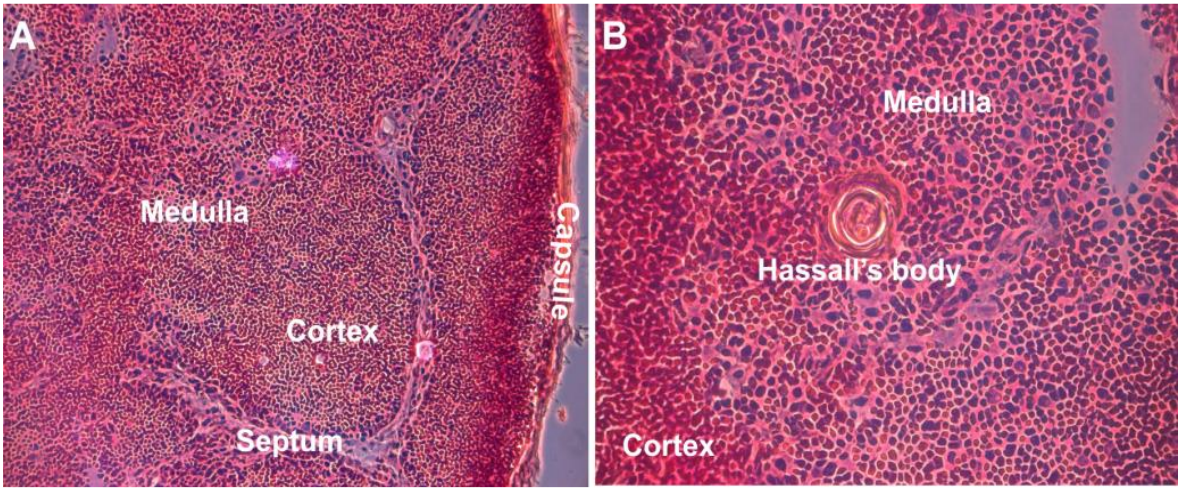


Figure 1.1 Structure of the human thymus. Hematoxylin/eosin staining of a section of a pediatric thymus (7 year old boy). The cortex is composed of densely packed thymocytes (A, B) while in the medulla thymocyte density is lower and contains Hassall's bodies (B). Magnification: 10x (A), 40x (B). Adapted from Nunes-Cabaço 2010[4].

The thymus is unique in its ability to host T cell differentiation and repertoire selection, functions that are mediated by thymic epithelial cells (TECs), the major constituents of the thymic stroma[5]. These cells can be divided into cortical (cTEC) and medullary (mTEC) epithelial cells, featuring different ultrastructural characteristics, antigenic expression, and the capacity to synthesize thymic factors that are important for intrathymic maturation and modulation of lymphocyte responsiveness[5]. The thymic stroma is also composed by several non-epithelial cell types, such as endothelial cells, which are important for the maintenance of the thymic vasculature (and therefore for thymus colonization by blood-borne T cell progenitors), and mesenchymal cells, which regulate the proliferation of TECs and T cell progenitors[6]. The thymic stroma also includes populations of haematopoietic origin, namely dendritic cells (DCs), B cells, and macrophages, all of which are involved in the shaping of the T cell repertoire and the elimination of apoptotic thymocytes[6,7].

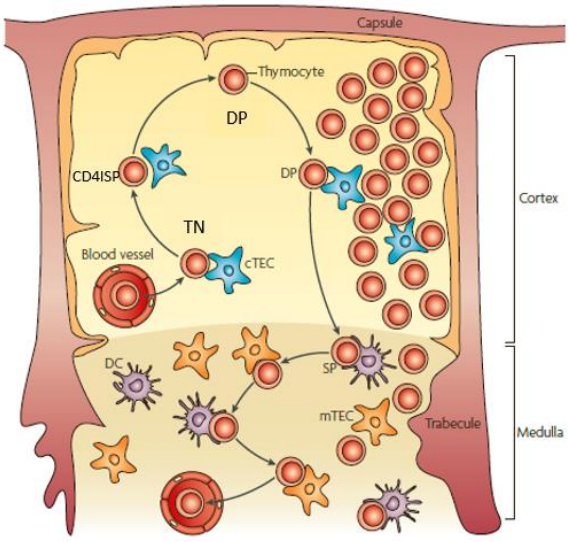


Figure 1.2 Thymus T cell development. T cell precursors enter the thymus through the blood vessels at the cortico–medullary junction (CMJ), and then begin a highly ordered differentiation programme, which is linked to migration through the thymic stroma. Development of CD4ISP cells is accompanied by an outward movement towards the subcapsular zone. $CD4^+CD8^+$ double positive (DP) cells move through the cortex and scan cTECs for positively selecting ligands. After positive selection and lineage commitment, CD4 or CD8 single positive (SP) cells move to the medulla, where they scan medullary antigen presenting cells, mainly DCs and mTECs, before the egress to the periphery. *Adapted from Klein et al. 2009*[8].

1.2 Stages of human T cell development

The thymus is first colonised by lymphoid progenitor cells, which are derived from hematopoeitic stem/progenitor cells in the bone marrow and then migrate via the blood stream to this organ, where they differentiate into T cells[1] (Figure 1.2). In humans, these precursors are part of a population of cells that express the membrane protein Cluster of Differentiation 34 (CD34)[9], a marker that is also expressed in pluripotent stem cells and in hematopoietic progenitor cells, but not in their differentiated progeny[10–12].

The sequential expression of CD cell-surface markers, particularly CD4, CD3, and CD8[1] (Figure 1.3) can be used to define the distinct stages of human T cell development. Developing thymocytes are first divided according to these three receptors as being triple negative (TN); $CD4^+$ immature single positive (CD4ISP); $CD3^-$ or $CD3^+$ double positive (DP); and $CD4^+$ or $CD8^+$ single positive (SP)[1,13]. The latter, CD4 and CD8, interact with the major histocompatibility complexes (MHC) class I and II, respectively, and are markers for the identification of the two major peripheral T cell subsets, $CD4^+$ or regulatory/helper T cells; and $CD8^+$, or cytotoxic T cells.

In humans, there are three distinct TN stages which can be identified by different combinations of markers: the $CD34^+CD38^-CD1a$ $stage^-$, the consecutive $CD34^+CD38^+CD1a^-$ stage, and $CD34^+CD38^+CD1a^+$ stage[14] (Figure 1.3). Cells expressing $CD34^+CD38^-CD1a^-$ represent the most immature thymic subset. Their identity is based on T cell receptor (TCR) rearrangement status[15] and on the preservation of the capacity to differentiate into other lineages in addition to T cells, eg, natural killer (NK) cells and dendritic cells[12,16]. Therefore, $CD34^+$ cord-blood precursors can differentiate into T cells, although they are not yet committed to the T cell lineage[17]. T cell commitment occurs within the thymus itself[13]. This is supported by the fact that TCR gene rearrangements, the most definite marker for T cell commitment, are not found in cord blood $CD34^+$ cells, and by the lack of expression of recombinase activating gene 1 (RAG1), CD1a, cytoplasmic CD3, CD2 and CD7, which

are all present in committed T cell-precursors in the thymus[17]. In addition, it has been established that the human thymus also hosts populations of multipotent precursors[13].

The transition from $CD34^+CD38^-CD1a^-$ to the $CD34^+CD38^+CD1a^+$ stage marks an important checkpoint in early T cell development. The expression of the CD1a molecule is strongly associated with the induction of T cell commitment, since the thymocytes at this stage largely lose the ability to develop into non-T cells, such as NK cells and DCs[12,18]. In addition, at this point these cells have already started their β, γ, and δ TCR loci rearrangements[19].

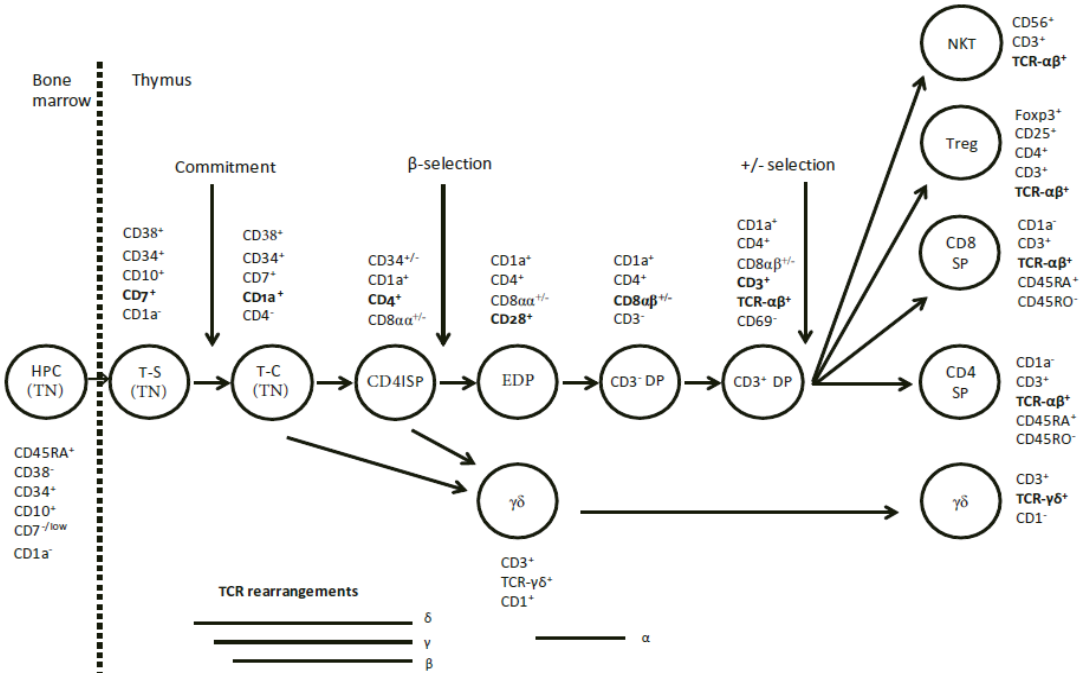


Figure 1.3 Schematic overview of the different developmental stages that characterize human T cell development. HPC, hematopoietic progenitor cell; T-S, T-lineage specified progenitor; T-C, T cell committed progenitor; TN, Triple positive; CD4ISP, CD4 immature single positive; EDP, early double positive; DP, double positive; NKT, natural killer T cell; Treg, regulatory T cell; CD8SP, CD8 single positive; CD4SP, CD4 single positive; γδ, T cell receptor-γδ positive cell. Adapted from Van de Walle *et al.* 2016[20].

Thymocyte proliferation, survival, and differentiation are controlled by a combination of several factors, most importantly of which are major signalling pathways. These include stromal cell-derived signals secreted by TECs (eg, interleukin-7, IL-7, and stem cell factor, SCF); Wnt signalling; Hedgehog (an essential positive regulator of T cell progenitor differentiation), and the Notch1 pathway[7,21–23].

Notch1 is a highly conserved transmembrane receptor involved in the regulation of cell-fate choices in many cell lineages[24] and, together with its ligands Delta-like 1 (DLL1) and DLL4, has been identified as a determinant factor for the choice between T- and B-cell fate[25,26]. For example, when cord blood and bone marrow $CD34^+$ cells are cultured in the presence of murine bone marrow stromal cells expressing DLL1 they

develop into full αβ T cells[27,28]. In addition, the use of an inhibitor of Notch signalling strongly impairs T cell development in different experimental systems[29,30].

The precursor immigrant cells enter the thymus through the cortex-medullary junction[31], where they are exposed to DLL1 expressed by the resident stromal cells[32,33]. This activates Notch signalling in the incoming precursors which promotes their differentiation along the T cell/NK cell lineage, but not the B-cell lineage[1] .

The development of $CD34^+$ human thymic T cell progenitors is also critically dependent on interleukine-7 (IL-7) cytokine signalling[34]. The IL-7 cytokine receptor consists of two chains, IL-7Rα (CD127) and a common γ chain (γc), shared with the receptors for IL-2, IL-4, IL-9, IL-15 and IL-21. Consistently, mutations in genes encoding for IL-7Rα[35,36], γc[37,38] or the Janus kinase Jak3, a component of the IL-7-induced signaling transduction pathway[39,40], result in profound T cell depletion and account for severe combined immune deficiency (SCID)[37]. Conversely, exposure of $CD34^+$ thymic T cell progenitors to an IL-7R signaling inhibitor in a fetal thymic organ culture (FTOC) efficiently blocks their development, abrogating the transition of TN cells into CD4ISP[34]. In addition, IL-7 is reported to be involved in the rearrangement of TCRα genes[41] and TCRβ genes in mice[42].

1.2.1 Early TCR arrangements and β-selection of thymocytes

During these early stages of T cell development, $CD34^+CD38^+CD1a^+$ thymocytes lose CD34 expression and start to express CD4 co-receptor, but not CD8 or surface CD3, thus becoming CD4ISP cells[43–45]. TCR loci undergo rearrangements to generate T cells that express a functional TCR, TCRαβ or TCRγδ[15,19,46]. In this rearrangement process, the variable domains of *TCRα, TCRβ, TCRγ,* and *TCRδ* (located within *TCRα*) genes are assembled following rearrangement of variable (V), diversity (D), and joining (J) gene segments by a process called V(D)J recombination[47]. V(D)J recombination uses the RAG1 and RAG2 enzymes that selectively target recombination signal sequences that flank V, D, and J segments[47]. The majority of mature T cells express a TCR composed of α and β chains (αβ T cells). Its variety is generated by rearrangement of the multiple germline-encoded segments (V-variable: 42Vα and 47Vβ; D-diversity: 2 Dβ segments; J-joining: 61 Jα and 13 Jβ gene segments; and non-germline-encoded N region insertions), as well as α and β chain pairing[48,49]. All these variations imply a possible repertoire of more than 10^{18} different human αβTCR[49]. The current model of TCR rearrangements suggests that recombination of TCR genes are sequential (*TCRδ* → *TCRγ* → *TCRβ* → *TCRα*)[15,19,46], although the cell phenotype at each rearrangement of a particular locus occurs still unclear.

Commitment to αβ and γδ lineages occurs after TCR expression is instructed by TCR signals[50]. The β, γ, and δ loci begin to undergo rearrangement almost simultaneously in developing thymocytes, and the two cell lineages diverge from a common precursor only after certain gene rearrangements have already occurred[51]. The decision of a precursor cell to commit to either γ:δ or α:β lineage depends on which type of receptor is expressed first during thymocyte development, i.e., if this is a functional γ:δ TCR or a α:β pre-TCR (functional β

chain paired with a surrogate un-rearranged α chain, called pre-T cell α chain, pTα) (Figure 1.4). The difference in fate results from different quality signals from the two types of receptors, with the γ:δ TCR delivering stronger signals than the α:β pre-TCR (reviewed in Hayes *et al.*, 2003[52]). Thus, if a complete γ:δ T cell receptor is formed before a successful β-chain gene rearrangement has led to the production of the pre-TCR, the thymocyte receives stronger signals through the γ:δ receptor, impairing further rearrangement of the β-chain gene and committing the cell to the γ:δ lineage. In contrast, if a functional β chain is formed before completion of a γ:δ receptor, it will pair with a pre-TCR α chain to generate the pre-TCR. In this case, the developing thymocyte receives a signal through the pre-TCR that shuts off rearrangements of the γ and δ loci, committing the cell to the α:β lineage (reviewed in Hayes *et al.*, 2003[52]).

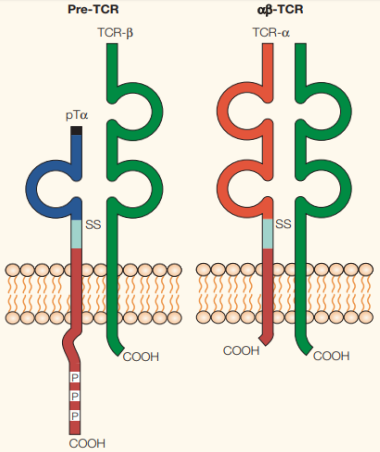


Figure 1.4 A representation of the structure of the pre-TCR (left) and the αβ -TCR (right). Adapted from Von Boehmer 2005[53].

In addition to pTα, the β chain also pairs with the lymphocyte co-receptor CD3. These molecules will be transported to the cell surface in a complex, called pre-TCR complex, providing the signaling components of T cell receptors[54–56]. Effective pre-TCR complex signalling is required for cell survival, rapid proliferation, arrest of further β-chain gene rearrangements (by promoting the degradation of RAG-2), and the initiation of *TCRα* gene rearrangements[55,57]. This process is known as β-selection and represents the first checkpoint of T cell development. Therefore, cells that fail to generate a productive TCRβ will not receive a survival and/or proliferation signal and do not proceed along the αβ lineage differentiation pathway[55]. Moreover, it also results in further differentiation of the CD4ISP subset into $CD4^+CD8^+$ thymocytes, which starts by only express the CD8*α* (early double-positive, EDP), and then both CD8*α* and CD8β molecules, becoming double-positive thymocytes expressing low levels of CD3 (DP $CD3^-$)[43,58].

In humans, several studies suggest that TCR β expression and β-selection are not directly associated with a specific stage and are not tightly coupled to the regulation of CD4 and CD8α/CD8β expression[13]. Thus, while a few cells already undergo β-selection before CD4 is expressed[15], a larger proportion is β-selected after upregulation of CD4 (CD4ISP stage)[19], and a third group of the pre-T cells upregulate CD4 and CD8α before initiating and completing TCR rearrangements[57,59]. In fact, CD28 expression, described as a marker of cells that

have passed the β-selection checkpoint, was observed in CD4ISP cells, indicating that β-selection can occur in this stage[60].

1.2.2 TCRα rearrangements

The *TCRα* locus initiates its rearrangements after the selection of TCRβ-expressing thymocytes exiting the cell cycle[1,49]. Each thymocyte rearranges its α-chain gene independently so that a single functional β-chain can be associated with many different α chains in the progeny cells[1,57]. However, as the *TCRδ* gene segments are embedded within the *TCRα* locus, the V-to-Jα rearrangements lead to deletion of the *δ* locus from the chromosome[1,61].Therefore, the δ locus excision generates DNA circles that persist as episomes, generating two types of TCR excision circles (TRECs)[62]. Rearrangement of δ-rec to Ψ-Jα locus generates a single TREC containing a signal-joint (sj) sequence (sjTREC). Subsequent Vα-Jα rearrangement deletes the remaining *Vδ* gene segment and produces coding-joint (cj) TRECs (cjTREC) (Figure 1.5).

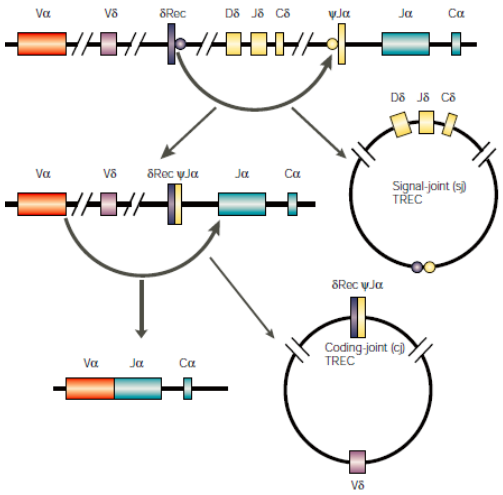


Figure 1.5 Generation of sjTRECs and cjTRECs. Simplified representation of the *TCRδ* locus flanked by portions of the *TCRα* locus (Vα, Jα and Cα). Adapted from Spits 2002[1].

TCRα rearrangements were initially proposed to occur mainly in the $CD3^-$ DP ($CD4^+CD8^+$) population[57,62]. However, more recent observations showed that δRec-ΨJα rearrangement can already be detected at the CD4ISP stage and even in the $CD34^+CD38^+CD1a^+$ TN stage[15].

1.2.3 Positive and negative selection of developing T cells

Before αβ receptor generation and translocation of the TCR complex to the cell surface, T cell development is independent of antigens. Thymocytes bearing αβTCR complexes than can engage self-peptides in MHC molecules expressed by TECs, DCs and other cells of the immune system present in the thymus, receive signals to survive, while those that are unable to recognize self-peptides (more than 90% of the thymocytes) undergo ''death by neglect''. This process is designated ''positive selection'', ensuring that only thymocytes with a functional TCR will differentiate into mature T cells[63,64]. Additionally, thymocytes with high-affinity binding of the TCR to self-peptide are also subjected to programmed cell death – "negative selection". This

process is also known as clonal deletion and results in the elimination of potentially self-reactive cells[1]. Although this ''affinity hypothesis'' (Figure 1.6) is supported by experimental affinity measurements[65,66], how can self-recognition be concomitantly essential for thymocyte survival, and able to induce cell death remains unclear[8].

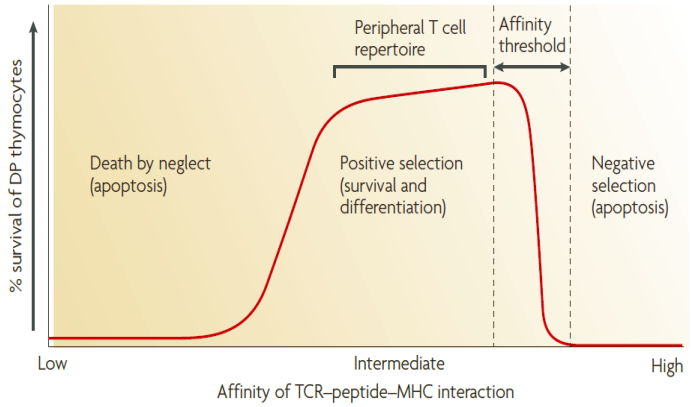


Figure 1.6 The affinity model of thymocyte selection. According to this model, only thymocytes with intermediate affinity receive survival signals and differentiate into mature T cells (positive selection). In contrast, those thymocytes that express TCRs with no or too low affinity die by neglect. High-affinity binding of the TCR to self-peptide–MHC complexes induces cell death by apoptosis – ''negative selection'' or ''clonal deletion''. Adapted from Klein *et al*. 2009[8].

Efficient positive selection of developing thymocytes requires interactions with self-peptide-MHC complexes displayed by cTECs[67–69], and some studies have also assigned a role for these cells in negative selection[6,70]. Namely, cTECs provide specialized accessory interactions that MHC$^+$ epithelial cells from other tissues do not[71]. These accessory molecules are poorly defined, although CD83 has been shown to be involved in CD4 T cell development[6,72]. Despite studies demonstrating that intrathymic expression of MHC molecules by non-TECs, including thymocytes themselves, can support positive selection, the relative efficiency of this process and the nature of the TCR repertoire that is selected still unclear[73–75].

The response of immature T cells to stimulation by self-antigen is the basis of negative selection. However, many tissue-specific antigens (TSAs), such as insulin, are not expected to be expressed in the thymus so that antigen presenting cells (APC) may present them in self-peptide:self-MHC. The expression of these ''tissue-specific'' proteins is promoted in mTECs in the thymic medulla by the autoimmune regulator *AIRE* and possibly other factors that induce the transcription of numerous TSAs[76]. Therefore, mutations in *AIRE* give rise to the human autoimmune disease known as autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED)[77,78].

1.2.4 CD4 and CD8 lineage decision of developing αβ T cells

Positively selected DP thymocytes expressing high levels of CD3 (DP CD3^{high}) differentiate into CD4$^+$ or CD8$^+$ SP mature effector T cells, a crucial decision known as CD4/CD8 lineage choice[79]. The specificity of the TCR for self-peptide:self-MHC molecule complexes is thought to determine which co-receptor a mature T cell will express. If the TCR is specific for an antigen presented by self-MHC class I molecules, the mature T cell will express the co-receptor CD8. Similarly, If the T cell receptor is specific for an antigen presented by self-MHC class II molecules, the mature T cell then express the co-receptor CD4[80]. The importance of MHC molecules is attested by the human immunodeficiency disease known as bare lymphocyte syndrome (BLS), which is caused by mutations in transcription factors that regulate MHC expression, resulting in a depletion of MHC in lymphocytes and thymic epithelial cells. Patients lacking MHC class II molecules have major CD4 T cell depletion with only few highly abnormal CD4 T cells, while those deficient for MHC class I molecules lack CD8 T cells (reviewed in Reith and Mach, 2001[81]).

The currently most accepted model for the CD4/CD8 lineage choice is the kinetic signaling model, which proposes that this choice is determined by TCR-signal duration, with cytokines, mainly IL-7, playing a determinant role dependent of TCR-signal duration (Figure 1.7)[82–85]. According to this model, positively selected mature DP thymocytes down-regulates CD8 expression, and if these CD4$^+$CD8^{low} intermediate thymocytes, received TCR signals they will be likely recognizing self-peptides-MHC class II molecules given the absence of CD8 and will be committed to the CD4 lineage [79,82]. Persistence of TCR signalling possibly by self-peptides:self-MHC Class II in CD4$^+$CD8^{low} intermediate thymocytes blocks IL-7 signalling and induces differentiation into mature CD4$^+$ T cells. Cessation or disruption of TCR signalling in the absence of CD8 in CD4$^+$CD8^{low} thymocytes allows IL-7 signalling, which promotes co-receptor reversal of CD4$^+$CD8^{low} intermediate thymocytes by enhancing *CD4* silencing and promoting re-initiation of *CD8* gene transcription, with cells ultimately differentiating into CD8$^+$ T cells[79,85].

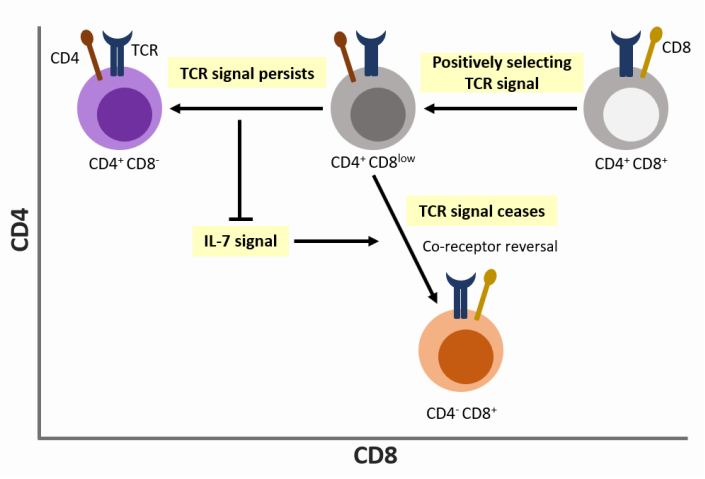


Figure 1.7 Kinetic signaling model of CD4/CD8 lineage choice in T cell development. Positively-selecting TCR signals induce DP thymocytes to terminate *CD8* transcription and thus become CD4$^+$CD8^{low}. TCR-signal duration and IL-7 signaling then determine whether a given thymocyte commits to the CD4 or the CD8 lineage. Adapted from Singer *et al*. 2008[79].

Several transcription factors involved in the regulation of *CD4* and *CD8* gene transcription have been identified. For instance, T-helper inducing POZ/Kruppel-like factor (Th-POK) seems to be both necessary and sufficient for CD4-lineage specification, probably acting as a master regulator of CD4 T cell differentiation[86,87], while runt-related transcription factor (RUNX) proteins, in particular RUNX3, seem to promote CD8 T cell differentiation[88,89]. These two factors inhibit the expression of each other, thus reinforcing lineage choices[90–92]. TOX (thymus-high mobility group (HMG) box protein)[93] and GATA3 (GATA-binding protein)[94] also appear to play a role in lineage commitment, particularly in CD4 lineage choice[79].

Finally, these CD4SP and CD8SP thymocytes with a stringently selected TCR repertoire, mature and acquire the naïve-associated marker CD45RA, exit the thymus and migrate to the periphery, incorporating the naïve T cell pool[1,7].

1.3 Regulatory T cells

The thymus is responsible for the generation of the T cell lineage committed to the regulatory fate (Tregs). This population is specialized in the maintenance of immune homeostasis by preventing or limiting the effects of the excessive immune responses harmful to the host that are mediated by conventional non-regulatory T cells (Tconv) and other cells of the immune system[95]. Depletion or disruption of its development or function results in autoimmune and inflammatory diseases[96,97]. In addition, Tregs are also important in suppressing allergy; establishing tolerance to organ grafts; preventing graft-versus-host disease after bone marrow transplantation; and promoting feto-maternal tolerance[96]. Thus, Treg functions have been intensively studied and can be become the basis of promising therapeutic approaches.

Tregs exert their action by several mechanisms, and these may be classified into four main groups (Figure 1.8): competition for IL-2, a crucial interleukin required for effector T cell survival that Tregs are unable to produce; production of inhibitory cytokines, such as IL-10, IL-35, and transforming growth factor β (TGF-β); modulation of DC maturation and/or function, which are required for the activation of effector T cells; and cytolysis of effector T cells mediated through the secretion of granzymes and perforin (reviewed in Vignali *et al*., 2008[98]).

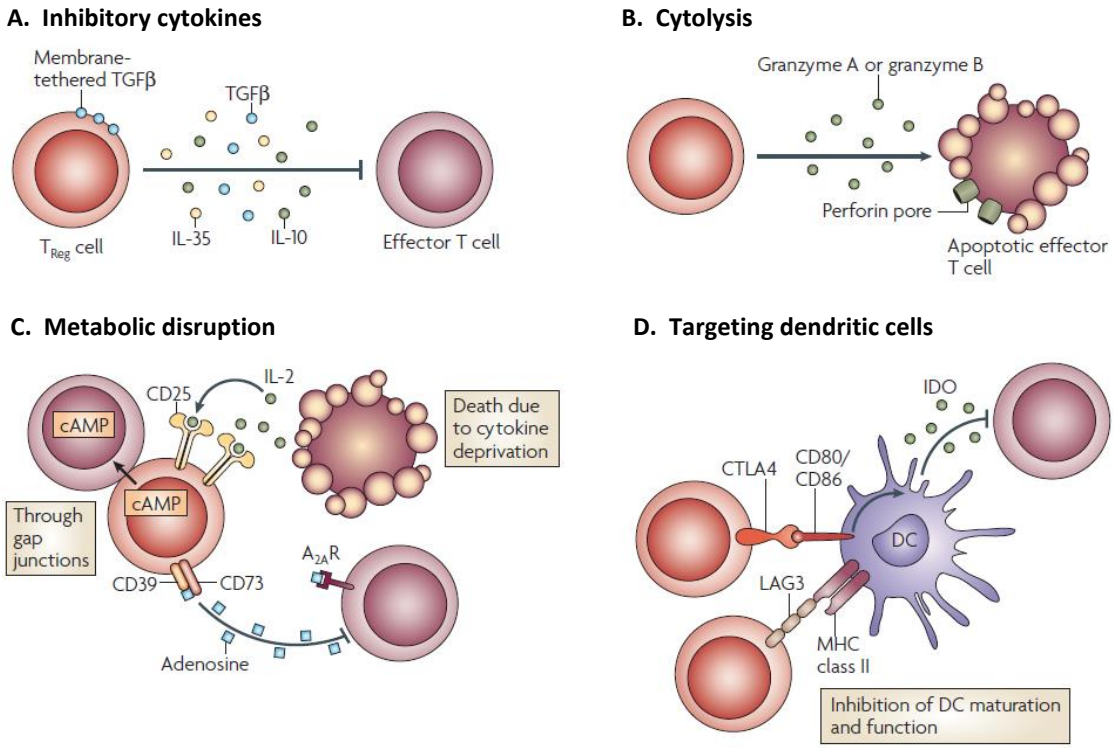


Figure 1.8 Basic mechanisms of action mediated by Treg-cells. Suppression mechanism used by Treg-cells can be grouped into four basic mechanisms of action. **A)** Suppression by inhibitory cytokines. **B)** Suppression by cytolysis. **C)** Suppression by metabolic disruption. **D)** Suppression by modulation of DC maturation or function. Adapted from Vignali *et al*. 2008[98].

There are several regulatory populations in the human body, including $CD4^+$ or $CD8^+$ regulatory T cells, regulatory B cells, and myeloid-derived supressor cells[99,100]. Here I'm focusing on $CD4^+$ Treg population, as they are the best characterized regulatory cell population and have the main role in the maintenance of immunological self-tolerance and immune homeostasis[101,102]. $CD4^+$ Tregs are divided into two main groups, according to the place of formation: Thymic-derived naturally occurring Treg (tTregs) and Peripheral induced/adaptive Tregs (pTreg) (reviewed in Nedoszytko *et al*. 2017[103]).

The first evidence of thymic Tregs came from neonatal thymectomy experiments studies performed by Sakaguchi, Nishizuka and others[104–107] in mice. In these experiments, thymectomy performed 3 days after birth resulted in T cell-mediated tissue inflammation and mice predisposition to systemic autoimmune diseases, which could be prevented by the transfer of thymocytes or splenocytes from adult euthymic mice[104–107]. Later, these were described as $CD4^+$ T population expressing the IL-2 receptor α-chain (CD25) with suppressive functions that in mice differentiate in the thymus after birth, and were termed regulatory T cells[108].

Importantly, a homologous population of $CD4^+CD25^+$ Tregs are also found in humans[109–111]. However, as CD25 expression in human peripheral blood also represents a state of activation in conventional CD4 Tconv-

cells with non-regulatory properties, only 2-4% of this $CD4^+CD25^+$ cells in the peripheral blood can be classified as Tregs, which express the highest levels of CD25 ($CD4^+CD25^{high}$)[112].

In addition to these, several other proteins have also been proposed as tTreg-cell markers, namely: CTLA-4 (cytotoxic T-lymphocyte antigen-4, CD152); GITR (glucocorticoid-induced TNF receptor family related protein); HLA-DR, and the constitutive expression of the specific transcription factor Forkhead box P3 (FOXP3)[113–117]. Low expression of CD127 (IL-7 receptor α chain) is often used to identify human Tregs. Indeed, there is a 90% correlation between $CD4^+CD25^+CD127^{low}$ T cells and FOXP3 expression[118,119], although $CD4^+$ Tconv cells tend to downregulate CD127 expression after activation[120]. A list of markers associated with Tregs is presented in Table 1.1. However, none can be used as unambiguous identifiers of human Tregs, as they can be expressed in other cell types or under certain physiologic conditions[121].

Table 1.1 Regulatory T cell markers

Transcription Factor	Activation and memory	Homing and origin	Suppressive and effector function	Apoptosis, survival or other
FOXP3	CD45RA	CD62L	CTLA4	CD27
IKZF2	CD45RO	CCR4	ICOS	OX40
IKZF4	CD25	CCR6	CD39-CD73	CD95
	HLA-DR	CCR9	LAP	PD1
	Lack of CD127	CD103	Granzyme B	GITR
	CD69	CD304	Galectin 1	Galectin 3
		CD31	Galectin 10	GARP
		Lack of CD49d	RANKL	MS4A4B
			CD80 and CD86	IL-1R
			IL-10	CD6
			IL-17	
			CD2	
			Lack of IL-2	

CCR, CC-chemokine receptor; GARP, glycoprotein A repetitions predominant; ICOS, inducible T cell co-stimulator; IL, interleukin; LAP, latency-associated peptide; MS4A4B, membrane-spanning 4-domains, subfamily A, member 4B; PD1, programmed cell death 1; R, receptor; TRANCE, TNF-related activation-induced cytokine. Adapted from Sakaguchi *et al*. 2010[122].

1.3.1 Role of FOXP3 in regulatory T cells

The transcription factor FOXP3 belongs to the forkhead-winged-helix family of transcription factors (TFs) and it is crucial for the differentiation, maintenance, and function of tTregs, as well as for restraining the expansion and function of the conventional T cells[123,124]. Members of the FOX family can be either transcriptional activators or transcriptional repressors - or both - and many have been implicated in regulating immune system development and function[125].

The importance of FOXP3 in this regulatory subset was first observed in mice mutant for *scurfy*, an X-linked recessive mutation in the forkhead domain that is lethal within a month after birth. These mice fail to develop FOXP3$^+$ Tregs thus exhibiting hyperactivation of autoreactive CD4$^+$ T cells and overproduction of proinflammatory cytokines, resulting in a fatal lymphoproliferative syndrome with multi-organ inflammation[126]. In humans, mutations in *FOXP3* gene, located at Xq11.23-Xq13-3 of the X chromosome, result in the fatal Immunedysregulation, Polyendocrinopathy, Enteropathy, X-linked syndrome (IPEX) autoimmune syndrome[127], the human equivalent of mouse *scurfy* phenotype[128,129]. Importantly, and since random X-chromosome inactivation in females ensures that a number of T cells will express one copy of wild-type *FOXP3*, *FOXP3* mutations affect only males. [130–132]. Patients with this genetic disorder manifest severe symptoms within the first few months of life, eventually leading to death[127,133].

FOXP3 can activate or repress transcription of hundreds of target genes upon direct binding to their regulatory regions (approximately 6-10%), including those encoding signal transduction molecules, transcription factors, cytokines (such as IL-2), cell-surface molecules (such as IL-2Rα and CTLA-4), enzymes for cell metabolism and intergenic microRNAs (such as miR-155)[134–136]. In addition, FOXP3 can also regulate gene transcription, by physically interacting with a number of DNA-binding co-factors[137,138], including the nuclear factor of activated T cells (NFAT), RUNX1 (also known as AML1), GATA3, c-REL, and RORγt[137,139–142], among several others (reviewed in Lu *et al.* 2017[143]) (Figure 1.9).

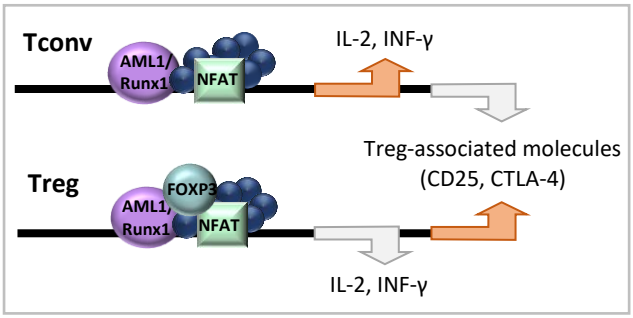


Figure 1.9 Indirect control of gene transcription by FOXP3. The transcriptional complexes involving NFAT and AML1/Runx1 activate or repress the genes encoding cytokines and several cell-surface molecules in Treg and non-Treg (Tconv), depending on the presence of FOXP3. Adapted from Sakaguchi *et al*. 2008[99].

Epigenetic changes play a pivotal role in the regulation of Treg function and differentiation. In fact, Foxp3 regulates transcription of its target genes by remodelling the chromatin structure and interacting with histone acetyl transferase (HAT)/histone deacetylase complex (HDAC)[144]. FOXP3, when attached to *IL-2* gene, recruits the acetyltransferase TIP60, which leads to the acetylation of Foxp3 promoter[145], enhancing its synthesis and its biding to the *IL-2* and *INF-γ* promoter, augmenting their repression. In contrast, when FOXP3 attach to the gene encoding CD25 and CTLA-4 (Treg-associated proteins), histone acetylation of their promoters is stimulated, and expression of these genes is up-regulated as well[146].

1.3.2 Transcriptional regulation of FOXP3 gene expression

The human *FOXP3* gene is located at the p-arm of the X chromosome, includes 11 exons and is highly conserved between humans and mice[126,127]. Regulation of FOXP3 expression is tightly controlled, with several transcription factors downstream of TCR signaling binding to the promoter to activate expression (see detailed discussion below). These include NFAT, AP-1, FOXO1 and FOXO3 proteins, and cAMP-responsive element-binding protein (CREB)-activating transcription factor 1 (ATF1) complexes[147–149]. The control of FOXP3 expression, including its initiation and maintenance, is also highly dependent on the binding of other TFs to four conserved intronic non-coding sequences (CNSs) (Figure 1.11). The NF-kB family transcription factor member Proto-Oncogene c-REL binds to CNS3 to control the initiation of FOXP3 transcription during Treg cell development in the thymus. The CNS2 region is important to maintain the expression of FOXP3 in thymic-derived Tregs after their egress from the thymus. This region becomes extensively hypomethylated during tTreg cell development[150], enabling the binding of several transcription factors that will, ultimately, enhance FOXP3 transcription, namely c-REL, CREB-ATF1, ETS1, RUNX1 and signal transducer and activator of transcription 5 (STAT5)[147,151].

Treg cells can be generated in the periphery (pTregs) from FOXP3$^-$ CD4$^+$ Tconv cells (Figure 1.10), under TCR stimulation and a combination of TGF-β and IL-2 cytokines[152], and the CNS1 region is particularly important for the induction of FOXP3 expression in extrathymic Tregs[153]. Specifically, TGF-β induces SMAD signalling, resulting in the activation of SMAD3 and its binding to CNS1 region. pTreg cells accumulate mostly at barrier sites (such as the gut) where they maintain immune homeostasis. In addition, other peripherally-induced FOXP3$^-$ populations of cells with regulatory properties have been described, such as type 1 regulatory T cells (Tr1) and T helper-3 cells (Th3), in which suppression is mainly mediated by the production of IL-10, and these also may play a role in the control of the development of autoimmunity and in promotion of transplantation tolerance (Figure 1.10)[154].

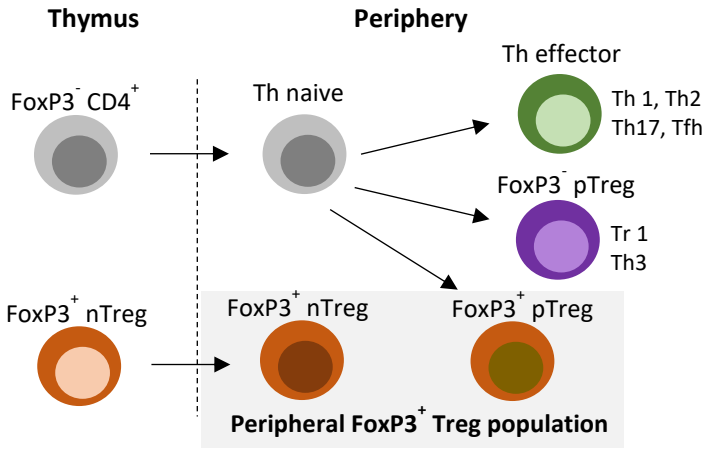


Figure 1.10 Thymic and Peripheral Generation of FOXP3⁺ Treg Cells. The peripheral Treg population comprises both thymic-derived Tregs (tTregs), which differentiate in the thymus and migrate to the peripheral tissues, and peripheral Tregs (pTreg), which differentiate in secondary lymphoid organs and tissues. Adapted from Lafaille *et al*. 2009[155].

Very recently, Kitagawa et al., discovered another regulatoy element in the FOXP3 locus named CNS0, located on an intron of the neighboring gene 5' of its locus (Figure 1.11)[156]. In adition, they found that Satb1 bind to this region, modifing the epigenetic status of the foxp3 locus to a poised state, allowing other transcription factors to bind to the other regulatory elements[156]. Therefore, at the early stages of thymic Treg development, SATB1 act as a pioneer factor, by activating Treg cell-specific super enhancers of the Foxp3 gene, as well as other Tregs-associated genes, including IL2RA and CTLA4[156]. However, Satb1 also acts on general thymic T-cell development[157]. Thus, its induction and binding to Treg cell-specific super enhancers in a Treg cell-specific manner is unclear.

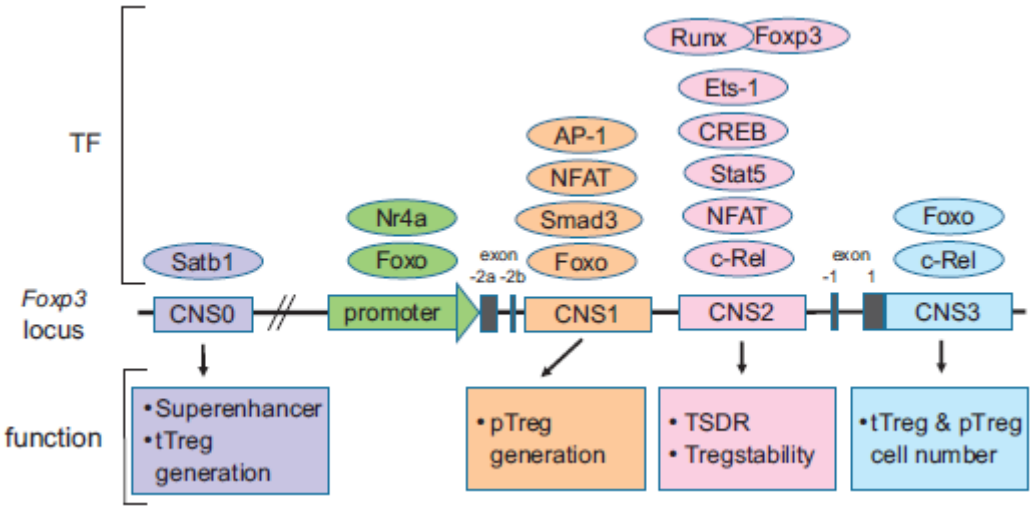


Figure 1.11 Schematic diagram of transcriptional regulation of the FOXP3 locus. Regulatory regions of the FOXP3 locus including the promoter CNS1, CNS2, CNS3, and recently discovered CNS0 are shown. Transcription factors (TFs) binding to each regulatory region and the function of each regulatory region are shown. TSDR: Treg-specific demethylated region. Adapted from Lee and Lee 2018[158].

Therefore, although FOXP3 is currently considered the most reliable tTreg marker available for humans[115] and mice[124,159], the existence of this non-regulatory FOXP3^{low} T cell population in normal individuals has raised doubts regarding the reliability of FOXP3 as a marker for all Tregs[160]. In fact, several recent studies of FOXP3 function in Treg cells during the last decade support the existence of several additional factors likely to be involved in Treg-cell development. Namely, Lin *et al*. 2007 and shown in mice that the commitment to Treg cell lineage didn't require the expression of functional FOXP3 protein, although indispensable for effector functions[161]. More recently, Ohkura *et al.* 2012 and Toker *et al.* 2013 shown that FOXP3 is insufficient to establish the Treg cell lineage, as it requires a specific epigenetic pattern that occurs independently of Foxp3 expression, and that thymic Treg precursors were already primed before FOXP3 expression[150,162].

1.3.3 TCR signalling in the control of Treg cell differentiation

As mentioned previously, thymocyte development is a tightly controlled process involving: positive selection of thymocytes recognizing self-MHC, essential for thymocyte survival; commitment to either $CD4^+$ or $CD8^+$ T cell lineage; and negative selection of T cells with TCRs of high avidity for class I and class II MHC molecules presenting self-antigens[63]. During these processes, thymocyte fate is determined by the TCR signaling strength, which results from distinctive characteristics such as functional avidity and duration of self-peptide-MHC complexes interactions (Figure 1.12).

Human $FOXP3^+$ thymocytes usually express markers of positive selection and maturation, such as CD69 and CD27, respectively[163–165]. Indeed, it is thought that thymocytes with TCR-MHC:self-peptide interactions stronger than those provided by Tconv cells, but not strong enough to induce negative selection, escape deletion and differentiate into Tregs (reviewed in: Singer *et al*., 2008[75]; Josefowicz et al., 2012[151]). Thus, the Treg TCR repertoire feature a higher degree of self-reactivity as compared to Tconv cells[166].

In line with this role for agonist-driven TCR stimulation in tTreg cell generation, studies in mice have shown that mutations in several TCR signalling molecules (such as the tyrosine protein kinase ZAP70, the transmembrane protein LAT, and the protein RASGRP) result in impaired Treg cell differentiation with much reduced impact in the development of Tconv cells[167–169].

TCR signaling activates other downstream signals and transcription factors, including AKT, mammalian target of rapamycin (mTOR), NFAT, and nuclear factor-kB (NF-kB). The NF-kB pathway appears to promote tTreg cell differentiation through the activation of c-REL, which directly regulates *FOXP3* expression[170]. In support of this, *REL* knockout mice show severe reduction of $CD25^+$CD4SP thymocytes[171]. Also, NF-kB activation induces the upregulation of CD25 expression, thus favouring the differentiation of Treg cells by facilitating the generations of $CD25^+$ Treg cell precursors. Calcium (Ca^{2+}) signaling is also important during thymic Treg cell development, as it is an upstream activator of NFAT. Stromal interaction molecules (STIM) are key players in this signalling, and mice with deletions of STIM1 and STIM2, show a substantial decrease in number and function of Treg cells[172,173]. On the other hand, the activation of the AKT-mTOR pathway inhibits Treg cell differentiation[174,175], through the phosphorylation and consequently inhibition of FOXO1 and FOXO3 transcription factors[148,176].

Importantly, cell fate is determined by both intensity and duration of TCR stimulation: continuous TCR stimulation induces both stimulatory and inhibitory signalling pathways and results in clonal deletion (Figure 1.12 A). In fact, the impact of TCR pathway is clearly demonstrated in mice with mutations in genes encoding components of this signalling cascade, which display dramatic decreases in the number of thymic Treg cells (reviewed in Feuerer *et al.* 2009[177]).

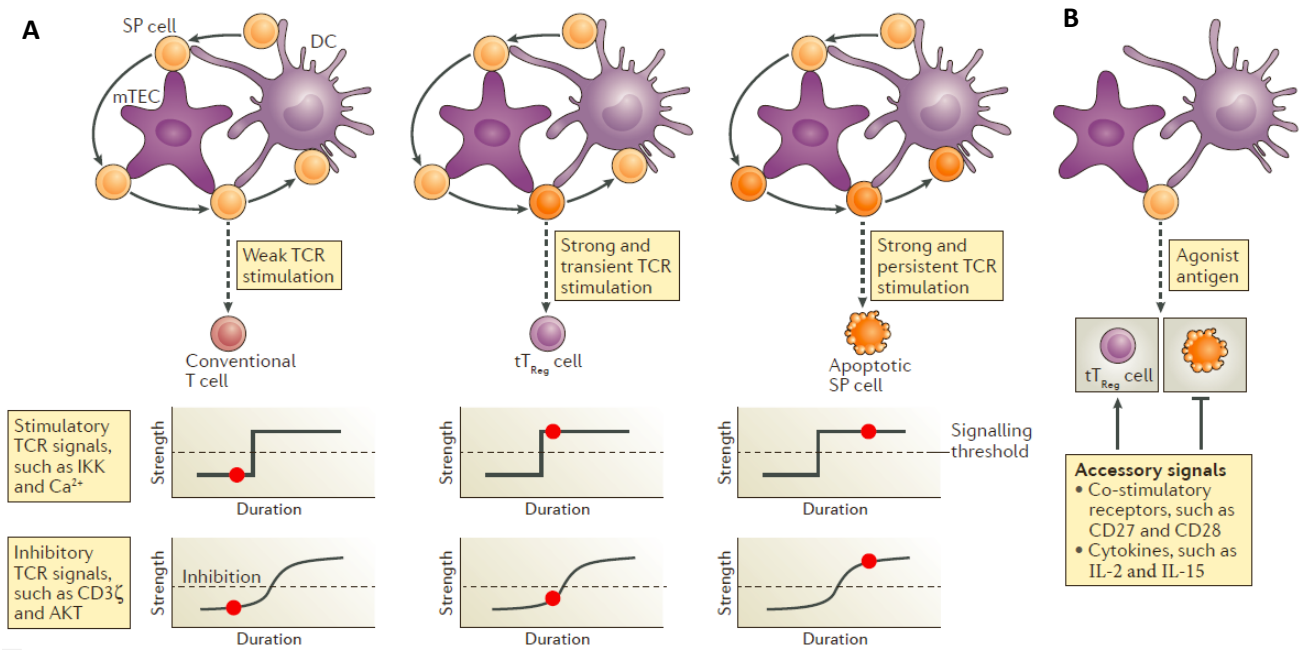


Figure 1.12 TCR and accessory signals in tTreg cell fate specification. A) Antigen-triggered T cell receptor (TCR) signalling plays a crucial role in dictating the fate of SP cells. Weak TCR stimulation promotes the continuous maturation of SP cells into conventional T cells (Tconvs). Transient stimulation of SP cells by high-affinity antigens is probably sufficient to activate the regulatory T (Treg) cell-stimulatory signalling pathways, including IkB kinase (IKK) and Ca^{2+}, whereas the activities of Treg cell-inhibitory signalling modules, including CD3ζ and AKT, may not reach an optimal level (dashed line). However, persistent stimulation SP cells by high-affinity antigens activates both Treg cell stimulatory and inhibitory signalling pathways, which may result in T cell clonal deletion. **B)** Accessory signals provided by co-stimulatory receptors such as CD27 and CD28, as well as cytokines, including TGF-β and IL-2 promote tTreg cell differentiation by suppressing T cell clonal deletion. Adapted from Li and Rudensky 2016[178].

1.3.4 The role of co-stimulation and cytokines in tTreg development

Taken together, the interactions between TCR and self-peptide/MHC are crucial for normal tTreg development. The importance of TCR signalling is consistent with the finding that many highly expressed Treg-associated molecules are induced upon TCR stimulation, including CD25, CTLA-4, and FOXP3. However, TCR stimulation alone is not sufficient initiate the tTreg differentiation programme, and other signals and factors are also required (Figure 1.12 B). The co-stimulatory signals provided by CD28-CD80/CD86 interaction have a well-established role in tTreg differentiation. CD28-deficient mice have a strong reduction in the frequency of thymic Tregs[179,180], whereas humanized mice treated with a super agonist anti-CD28 resulted in increased numbers of $CD25^{+}FOXP3^{+}$ CD4SP thymocytes[181].

CD28 co-stimulation promotes the activation of NF-kB signalling pathway, which enhances the efficiency of Treg cell development, or the survival of thymocytes undergoing Treg differentiation, by sparing self-reactive tTreg progenitors from apoptosis, and promotes FOXP3 expression[180,182,183]. Furthermore, TCR and CD28 signalling induce the expression of tumour necrosis factor receptor (TNFR) superfamily proteins, including

GITR, OX40 and TNFR2, which collectively promote tTreg generation[184]. The co-stimulatory receptor CD27 also prevents apoptosis of developing tTregs and is likely to act subsequently of CD28 engagement[185].

Cytokines have also crucial functions in the control of Treg differentiation, namely the IL-2R-γC cytokine family, which includes IL-2, IL-7 and IL-15. IL-2 activates STAT5, through γC chain-associated Janus Kinase 3 (JAK3), which subsequently binds to the promoter region of *FOXP3* to positively regulate its expression[177,186–188]. Consistent with this role of IL-2 in *FOXP3* expression, $Jak3^{-/-}$ and $Stat5^{-/-}$ mice having few or no circulating $Foxp3^{+}$ cells[188,189]. Moreover, developing Treg precursors in the thymus express CD25 and are highly attuned to IL-2, a competitive advantage in the IL-2 poor environment of the thymus[190]. In humans, a recent study showed that IL-2 and IL-15, but not IL-4, IL-7 or IL-21, can equally drive human tTreg precursor differentiation into $FOXP3^{+}$ cells, and promote tTreg proliferation and survival[191].

Based on these TCR-dependent and -independent processes, Lio and Hsieh proposed a two-step model of tTreg differentiation[192] (Figure 1.13). According to this model, tTreg precursors are selected from CD4SP thymocytes [193] whose TCRs engage with high affinity self-peptides-MHC class II complexes[99] presented by APCs, including mTECs, cTECs and DCs[194,195], in the presence of CD28-CD80/CD86 co-stimulation[180]. These signals then lead to the activation of downstream pathways, namely the NF-kB activation, which induces the expression of CD25 and remodelling of the FoxP3 locus. From here, Treg precursors do not require further TCR stimulation, as cytokine signals mediated by IL-2 and IL-15, facilitate the induction of *FOXP3* expression in a Stat5-regulated manner, resulting in the differentiation into the Treg lineage[192,196].

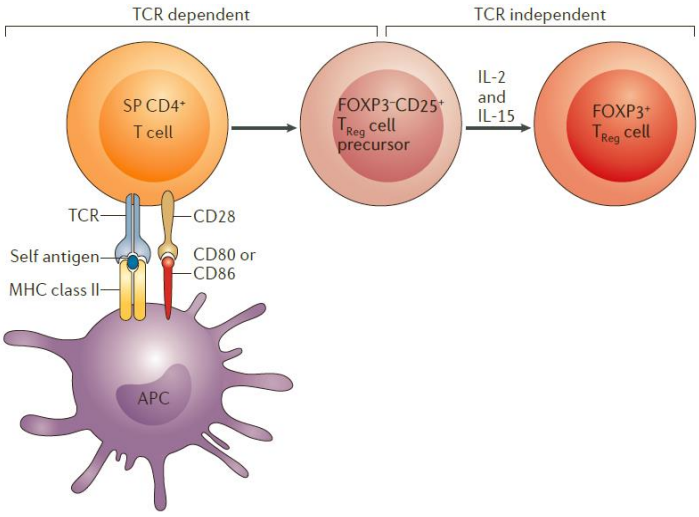


Figure 1.13 A two-step model for Treg cell development. Thymocytes that recognize self-peptide–MHC class II complexes in the presence of co-stimulatory signals (CD28-CD80/CD86) and with enough high per cell avidity, are selected to become $FoxP3^{-}CD25^{+}$ Treg cell precursors. TCR signal leads to the activation of several downstream pathways, including NF-κB activation, and results in the remodelling of the *FOXP3* locus, rendering it permissive to the induction of FOXP3 expression by IL-2 signalling (not shown). TCR-independent step includes signals mediated by IL-2 and IL-15, which facilitate the induction of *FOXP3* expression. Adapted from Hsieh *et al*. 2012[197].

1.3.5 tTreg commitment in the human thymus

The exact moment when a thymocyte becomes committed to the Treg lineage is still unknown[198]. The human thymic primordium is colonised by T cell progenitors during 8th week of gestation[199–201], and CD4^{+} tTregs are already present in the thymus at the 12th to 13th gestational weeks[163,202,203], as identified by their elevated expression of CD25[163]. In addition, fetal human thymic CD4^{+}CD25^{+} Tregs express *FOXP3* mRNA, as well as other markers related to their suppressive phenotype (CTLA-4, CD45RO, CD62L, and GITR), and have the ability to suppress T cell proliferation[163,202].

Studies in human fetal and post-natal thymuses have shown that, besides CD4SP thymocytes, CD8SP (CD4^{-}CD8^{+}), DP, and thymocytes in early pre-DP stages, namely at the triple-negative and CD4ISP stages, also can express CD25 and FOXP3 Treg markers[163–165,202,204–208]. Thus, it seems that the commitment to the tTreg lineage can occur at various stages of human T cell development (Figure 1.14).

DP thymocytes expressing FOXP3 and/or CD25 are clearly identified in the human thymus, which already express CTLA-4, CD39, and GIRT[163,165,202], all Treg function-associated markers, and exhibit immunosuppressive functions[165,206]. Although this population has been reported to express *rag-2* mRNA, a feature of immaturity, the majority of these cells express high levels of CD3 and CD27, which are markers associated with positive selection and maturity, respectively[165,202].

Altogether, DP tTreg cells appear to be the main contributors to the CD4SP tTreg pool in humans, which represents the major population of FOXP3^{+} human thymocytes[163,165,202,209,210]. This may be partly due to Treg accumulation during cell maturation in the thymic medulla[165]. Moreover, the contribution of recirculating peripheral Tregs to this compartment cannot be excluded[165].

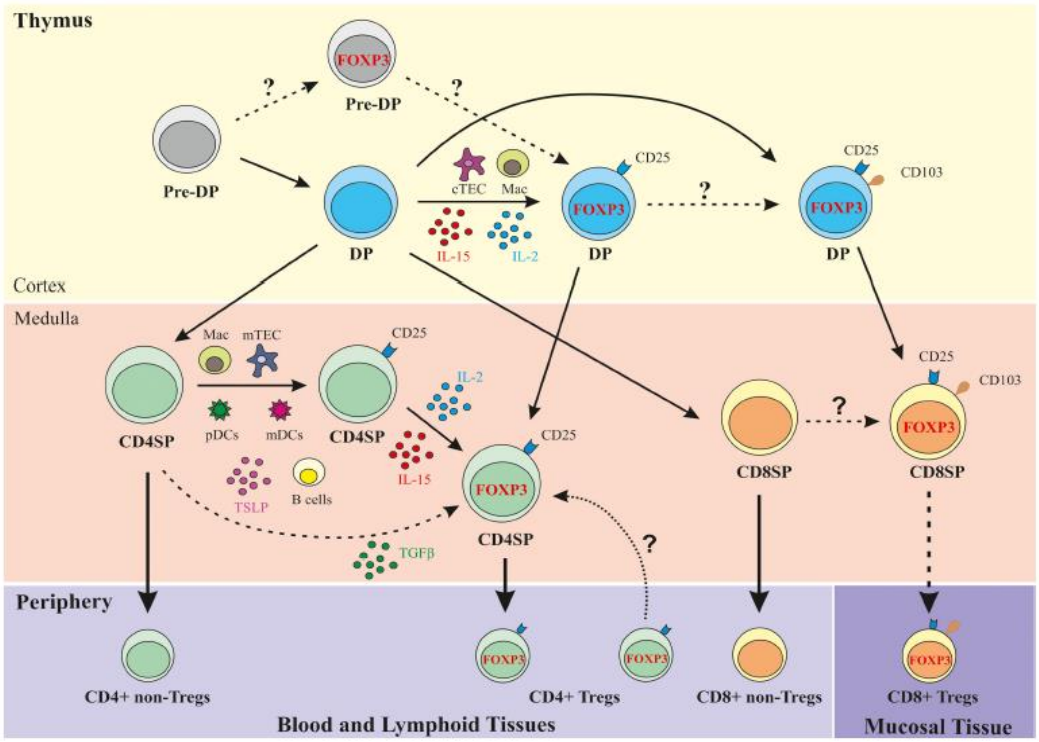


Figure 1.14 Schematic representation of human Treg development in the human thymus. DP, double-positive ($CD4^+CD8^+$); CD4SP, CD4 single-positive ($CD4^+CD8^{neg}$); CD8SP, CD8 single-positive ($CD8^+CD4^{neg}$); cTEC, cortical thymic epithelial cell; mTEC, medullary TEC; Mac, macrophage; FOXP3, Forkhead box P3; TSLP, thymic stromal lymphopoietin. Adapted from Caramalho *et al.* 2015[198].

Treg development in the human thymus remains largely undefined. To the best of our knowledge, a direct comparison regarding the transcriptional signature of CD4SP $FOXP3^+$ and $FOXP3^-$ subsets in the human thymus has not been reported. This assessment is of utmost importance to better characterise the commitment of tTreg populations and to provide us with important clues regarding the differentiation and development of this organ in humans. Additional studies are also required to provide valuable knowledge on the generation of this fundamental lineage in immune homeostasis.

1.4 Aims

To understand how Tregs regulate the immunological homeostasis it is necessary to characterise how thymic-derived CD4SP FOXP3$^+$ Treg cells are generated during their development. Although T cell commitment is known to occur earlier during development, CD4SP tTregs represent the most abundant fully-committed population expressing FOXP3 in the thymus that can be easily isolated in suitable numbers using surface markers[198,208]. My main purpose is to expand the current knowledge regarding the factors controlling the differentiation and commitment of human T cells, by generating the transcriptional profile of CD4SP tTregs and tTconvs.

To do this, I made use of the population discerning potential of cell sorting at a protein level; the massive resolution of Next Generation Sequencing (NGS) technologies at a transcript level; and the quantification and statistical powers of bioinformatics.

NGS-based RNA-seq is currently the preferred methodology for the study of gene expression[211,212]. It allows the analysis at an unknown depth, high sensitivity and resolution of all the transcripts present in a sample, including characterizing their sequences, and quantifying their abundances at the same time. Briefly, millions of short strings, called 'reads', are sequenced from random positions of the input mRNAs. These reads are then computationally mapped to a reference genome to reveal a 'transcriptional map', where the number of reads aligned to each gene gives a measure of its level of expression. This technology has been particularly important and valuable to compare transcriptional regulatory programs that define different cell subsets, developmental stages, or physiological conditions. With this approach, the study of gene expression and differential gene expression, can reveal several important aspects about the cell states, even though mRNAs are not the final products of the transcription-translation process. Nevertheless, the powerful features of RNA-seq have boosted a huge progress of transcriptomics research and the production of an impressive amount of data.

In human immunology research, RNA-seq has been widely used as an approach to analyse the transcriptome profile of several CD4$^+$ T subsets in the peripheral blood , including Treg and Tconvs, in both naïve and memory compartments[213–219]. In the human thymus, however, the currently available NGS-based RNA-seq data are mainly focused on progenitor T cells and mature CD4/CD8 SP populations[220,221].Therefore, and to the best of our knowledge, there are no whole-genome expression data regarding the CD4SP tTconv and tTreg lineage specification in the human thymus, as well as the regulators involved in this commitment.

With this thesis, I expect to have contributed to:

1. Characterise and compare the genome-wide expression profile of highly purified CD4SP tTconv and tTreg thymocytes population by RNA-seq.

2. Identify potential novel factors involved in the definition of tTreg identity.

3. Clarify the pathways and determinants of human tTreg lineage commitment and signature.

CHAPTER 2 – METHODS

2.1 – Tissue and cell preparation

2.1.1 Human thymic samples

Thymic specimens were obtained from routine thymectomy performed during pediatric corrective cardiac surgery at the Hospital de Santa Cruz (Carnaxide, Portugal), after written informed consent by parents or legal representatives. Thymic specimens are collected by clinical indication and would otherwise be discarded. Individuals with known immunodeficiency or with syndromic features associated with diseases potentially involving the immune system were excluded from this study. Three thymuses were used for RNA-seq (T274: 7 weeks, female; T276: 5 months, female; T277: 2 years, male), while the remaining 16 becoming the first samples in their own new biobank. This study was approved by the Ethical Board of the Faculty of Medicine of the University of Lisbon.

2.1.2 Isolation of thymocytes

Thymocytes were recovered through tissue dispersion followed by a Ficoll-Paque Plus (GE Healthcare) density gradient, where 10ml of Ficoll-Paque solution were added to the bottom of a 50ml falcon tube containing 30ml of thymic cells. After 15 minutes of centrifugation at 2100 rpm, the resulting ring containing the thymocytes were placed in a universal tube. Thymocytes were then washed twice with RPMI 1640 (Rosewell Park Memorial Institute, Gibco) supplemented with 2% of fetal bovine serum (FBS) by centrifuging at 1600 rpm for 5 minutes. Thymocytes were then kept at 4ºC in 10 ml of complete medium (CM), containing 88,5% of RPMI 1640, 10% of FBS, 1% of 2mM L-glutamine (L-Glu), and 0,5% of 200U/ml Penicillin and 200µ/ml Streptomycin (Pen/Strep). Next, thymocytes were diluted in Trypan blue to exclude dead cells and counted using a Neubauer chamber.

2.1.3 Flow cytometry phenotype assessment

Immunostaining of membrane proteins was performed for 30 minutes at room temperature with conjugated monoclonal antibodies (mAbs) listed below (Table 2.1). Thymocytes were then washed with PBS/BSA/Azide at 1800 rpm 5 minutes followed by a fixation and permeabilization step for 30 minutes at 4°C using FOXP3 intracellular staining kit (eBioscience). Intracellular staining was performed for 30 minutes at 4°C. After washed with PBS/BSA/Azide at 1800 rpm for 5 minutes, cells were acquired on a BD LSRFortessa (BD

Biosciences) and analysed using FlowJo software (TreeStar, Ashland, Oregon, USA). Double exclusion was confirmed using Side Scatter Width.

Table 2.1 List of antibodies used in this study

Molecule	Clone	Fluorochrome	Brand
CD127	FAB306P	PE	R&D Systems
CD25	2H3	PE-Cy7	BD Biosciences
CD27	O323	FITC	eBiosciences
CD3	OKT3	BV605	BioLegend
CD4	RPA-T4	PerCP-Cy 5.5	Invitrogen
CD45RA	HI100	BV510	BioLegend
CD8α	RPA-T8	APC-Cy7	eBiosciences
Foxp3	PCH101	eFluor 450	BioLegend
Ki-67	B56	APC	BD Biosciences

2.1.4 CD4SP regulatory and conventional T cell sorting

After isolation, thymocytes were washed with PBS 2% FBS and centrifuged for 7 minutes at 1600 rpm. Immunostaining was performed for 30 minutes at room temperature, using CD4 CD8, CD27, CD25, and CD127 anti-human mAbs (Table 2.1). Cells were then washed with PBS 2% FBS for 7 minutes at 1600 rpm, resuspended in PBS 2% FBS and filtered through a 70µM strainer. Selection for human CD4SP tTregs ($CD4^+CD8^-CD27^+CD25^{+/high}CD127^-$) and CD4SP tTconvs ($CD4^+CD8^-CD27^+CD25^-CD127^{-/+}$) subsets was performed by a BD FACS Aria III high-speed Cell Sorter (BD Biosciences), which were then recovered in 1,5 ml Eppendorf's containing PBS 2% FBS. Cell pellets of 300,000 cells from sorted subsets were obtained centrifugation at 13200 rpm at 4°C for 5 minutes and stored -80ºC for a better preservation of their RNA. Specifics for the sorting strategy are shown in Figure 2.1.

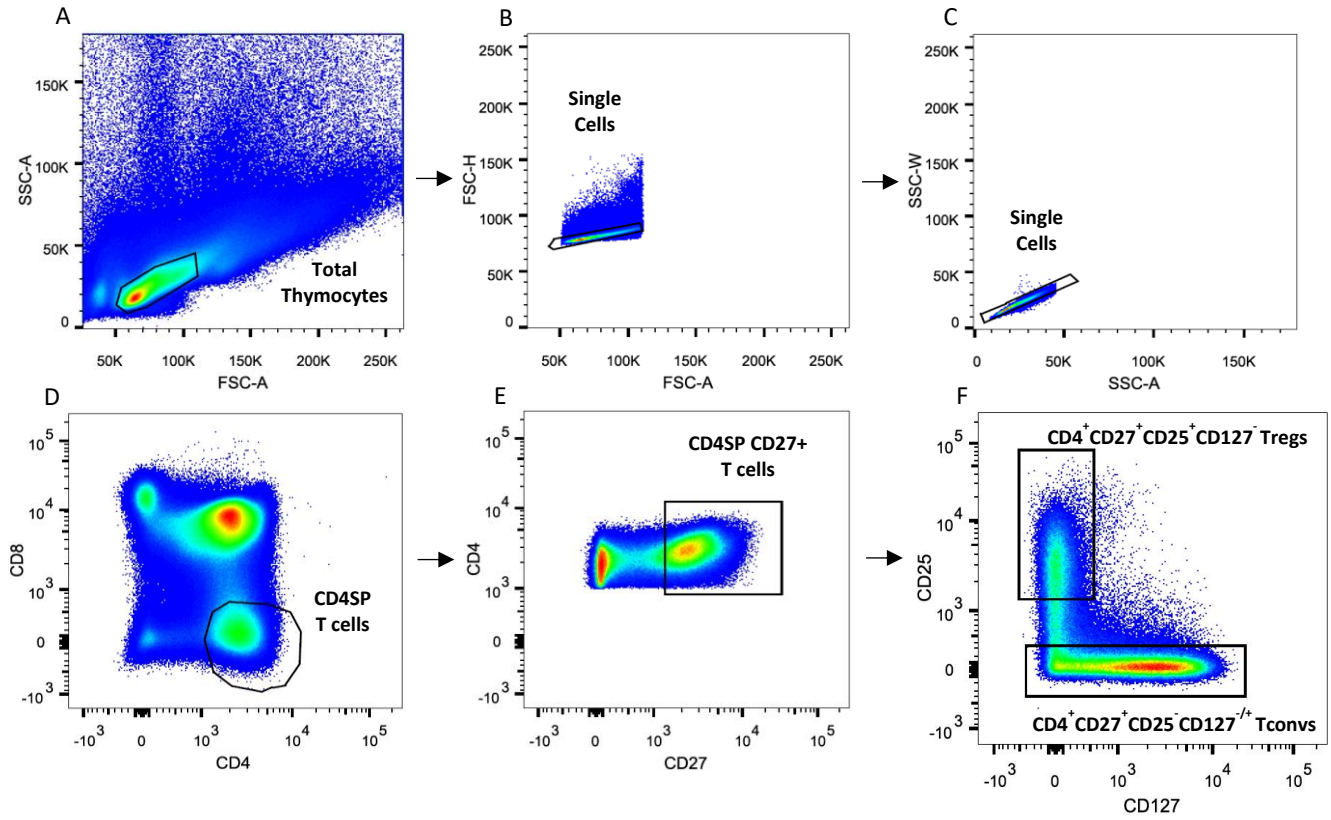


Figure 2.1 Gating strategies for cell sorting. A-F: Sequential gating strategies for CD4SP Tregs and Tconvs cell sorting.

2.1.5 RNA extraction

RNA was extracted from the cell pellets of 600,000 CD4SP Tregs and Tconvs, using the All prep® DNA/RNA micro Kit (Qiagen), and following the manufacturer's instructions. RNA quantity and quality were assessed by AATI Fragment Analyzer before RNA-seq library preparation. RNA quality number (RQN) and RNA concentration obtained is depicted in table X.

Table 2.2 RNA quality and concentration

	RQN	Concentration (ng/µl)
tTreg274	9,4	3,7
tTconv274	9,2	3,1
tTreg276	8,5	3,0
tTconv276	7,3	3,0
tTreg277	8,9	2,7
tTconv277	9,0	2,3

2.2 Next Generation Sequencing and Data Analysis

2.2.1 RNA-seq library preparation and sequencing

High-throughput sequencing was carried out by the Illumina HiSeq™4000 at the Beijing Genomics Institute (BGI) and according to the following protocol: after the total RNA extraction and DNase I treatment, magnetic beads with Oligo (dT) were used to isolate mRNA from the total RNA. Mixed with the fragmentation buffer, the mRNA was fragmented into short fragments. Then, cDNA was synthesised using the mRNA fragments as templates. Short fragments were purified and resolved with EB buffer for end reparation and single nucleotide A (adenine) addition. After that, the short fragments were connected with adapters, and the suitable fragments were selected for the PCR amplification as templates. During the Quality Control (QC) steps, Agilent 2100 Bioanaylzer and ABI StepOnePlus Real-Time PCR System were used in quantification and qualification of the sample library. Libraries were then sequenced on an Illumina HiSeq™4000, with paired-end (PE) 100bp long reads, estimating a minimum output of 200,000,000 reads per library. Finally, adaptor sequences, contamination and low-quality reads from raw reads were removed.

2.2.2 RNA-seq data analysis

Sequencing data was received in FASTQ format and analysed on the iMM High-Performance Computing (HPC) Cluster LOBO (https://insidehpc.com/hpc-basic-training/what-is-hpc). Quality of the data was assessed with FastQC[222]. Then, reads were mapped to reference genome hg38 with TopHat2[223] using default parameters (Annexe 1 A). Mapped data was further processed with with Samtools[224] and BigWig files for visualisation in the genome browsers UCSC Genome Browser[225] (http://genome.ucsc.edu/.) and Integrative Genomics Viewer (IGV)[226,227] were created with deepTools2 with the feature BamCoverage[228] (Annexe 1 C). HTseq-count was used to quantify total expression for each sample[229] (Annexe 1 B).

2.2.3 Differential expression analysis of RNA-seq data

All the analyses and graphics were produced using the R/Bioconductor environment[230].
Differential gene expression (DEG), Principal Component Analysis (PCA) and Multi-Dimensional Scaling (MDS) was determined and generated with edgeR[231] and DESeq2[232]. Each step of the differential expression analysis is described in detail in chapter 3. *P*-values were adjusted using the Benjamini–Hochberg procedure[233]. A gene was considered to be significantly differentially expressed between subsets when fold-change (FC) ≥ 2 and FDR < 0.05. Gene ontology was performed using WebGestalt web tool[234], using Ensembl IDs as identifiers. Most of the graphics displayed in this work were generated using the *ggplot* function from the *ggplot2* R package[235].

2.2.4 Identification of transcription factors

Genes encoding transcription factors (TFs) were identified by comparison with the most recently published peer-reviewed catalogue of human transcription factors[236]. This catalogue is a result of a manually examination of 2,765 proteins compiled by combining putative TFs lists from several sources. Each protein was classified based on the likelihood of each protein to be a TF, its DNA binding mode, and known motifs for each protein along with available DNA-protein structures. The final tally encompassed 1,639 known or likely human TFs, which was used to identify DEG encoding TFs.

CHAPTER 3 – RESULTS AND DISCUSSION

To understand human thymic Treg and Tconv differentiation, my strategy was to characterise their expression profiles at an RNA level and then compare them in order to find the common and different genes and pathways involved in their commitment to their respective identities and specific functions. Finally, I describe a set of genes up-regulated in tTreg cells that might represent potential novel factors involved in the definition of this subset.

3.1 Thymocyte subsets are representative of normal human T cell development

We selected three human thymic specimens (T274: 7 weeks, Female; T276: 5 months, Female; T277: 2 years, Male) from routine thymectomy performed during pediatric corrective cardiac surgery and assessed the phenotype of the whole thymocytes in the three thymuses by flow cytometry using the following cell-surface and intracellular antigens (Fig. 3.1): CD3, CD4 and CD8 to distinguish the major stages of T cell development ($CD3^-CD4^-CD8^-$ - Triple Negative, TN; $CD3^-CD4^+CD8^-$ - CD4 Immature Single Positive, CD4ISP; $CD3^{-/+}CD4^+CD8^+$ - Double Positive, DP; $CD3^+CD4^+CD8^-$ - CD4 Single Positive, CD4SP; and $CD3^+CD4^-CD8^+$ - CD8 Single Positive, CD8SP); CD25, CD127, and FoxP3 markers to distinguish and characterise Tconv and Treg subsets[118,119]; CD27 as a maturation marker[1]; CD45RA as both T cell precursor and a naïve cell marker[1]; and Ki-67 as a marker of proliferating cells[237].

Our analysis of the three thymuses is displayed in (Figure 3.2)

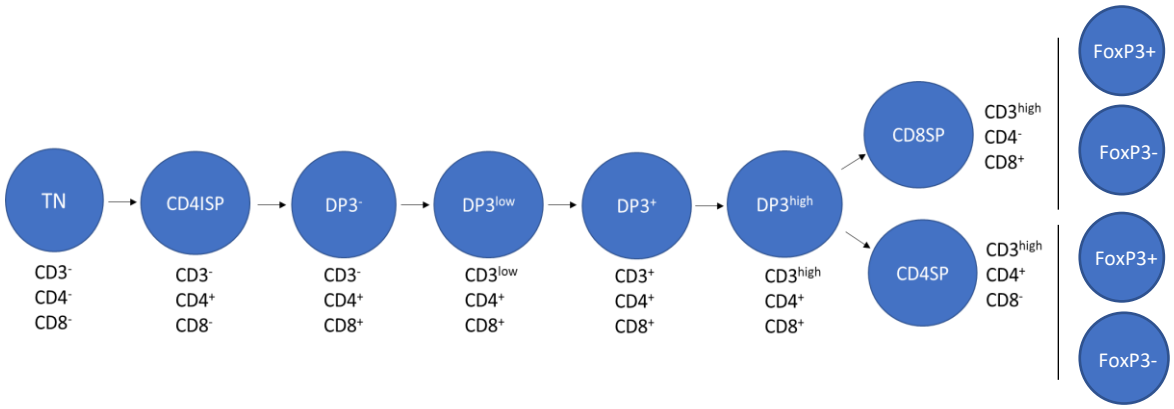


Figure 3.1 The major stages of T cell development defined by the expression of CD3, CD4 and CD8 markers. T-cell progenitors contained within the early $CD3^-CD4^-CD8^-$ triple negative population (TN) initially acquire CD4 (CD4 immature single positive cells, CD4ISP) and subsequently CD8 to become DP cells. A progressive increase in surface expression of CD3 occurs in parallel with surface TCRαβ in DP cells, followed by differentiation into CD4 or CD8 SP cells. Within the CD4/CD8 SP stage, two populations can be distinguished according to the expression of FoxP3.

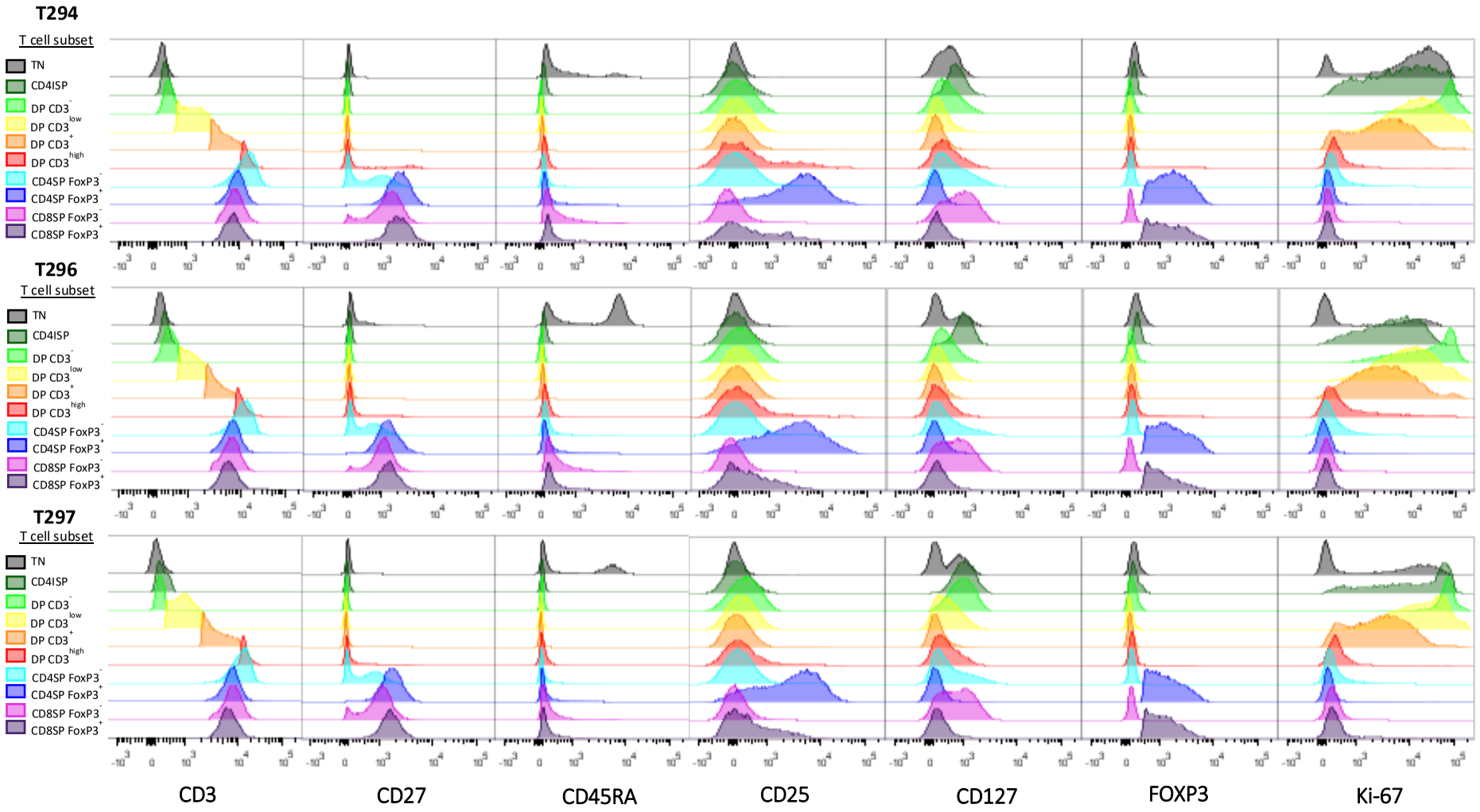


Figure 3.2 Phenotypical analysis of thymocytes subsets. Histogram plots showing the expression levels of each marker in the defined T cell subsets for the three thymuses. Each marker fluorescence intensity in a Log_{10} scale for the frequency of events detected at that intensity in each subset is represented in x- and y-axis, respectively.

Of note, the $CD3^-CD4^-CD8^-$ T cells, or TN population also expressed CD45RA. In contrast, the expression of this marker was not observed in the following stage, the CD4 immature single positive or CD4ISP population. This is in agreement with the pattern of high expression of CD45RA in T cell precursors, but not in the early CD4ISP stage of T cell development[1]. Moreover, in this stage, the level of expression IL-7 receptor (CD127) is consistent with its essential role in human T-cell development, by promoting thymocytes survival, proliferation, and growth induced by IL-7 stimulation[238,239]. Additionally, both TN and CD4ISP populations express elevated levels of Ki-67 and not of CD27 marker, consistent with an immature and proliferating phenotype, as previously observed in these cells[208].

The initiation of CD8 expression in CD4ISP thymocytes marks the transition to the double positive (DP) stage. This population was, as expected, the most abundant (52,4% mean of total thymocytes) for all thymuses, and featured higher rate of proliferation, as indicated by Ki-67 expression (Figure 3.2, in blue). Furthermore, DP cells showed different levels of CD3 expression, which correlates with the decrease of Ki-67 expression. These observations are in accordance with the progressive increase in the expression of the marker CD3 along with surface TCRαβ, and a decrease in the proliferation activity in double positive thymocytes during human T cell development[13].

Previous studies have shown that DP thymocytes represent the most immature population where significant amounts of FOXP3 can be detected[164,208]. Consistently, I found amongst the mature CD3-expressing DP population (DP $CD27^+$ $CD3^{low/+/high}$) a direct correlation between the expression of CD27 and CD25 and of FoxP3, with the cells expressing higher levels of CD27 also expressing higher levels of both CD25 and FoxP3 markers Figure 3.3).

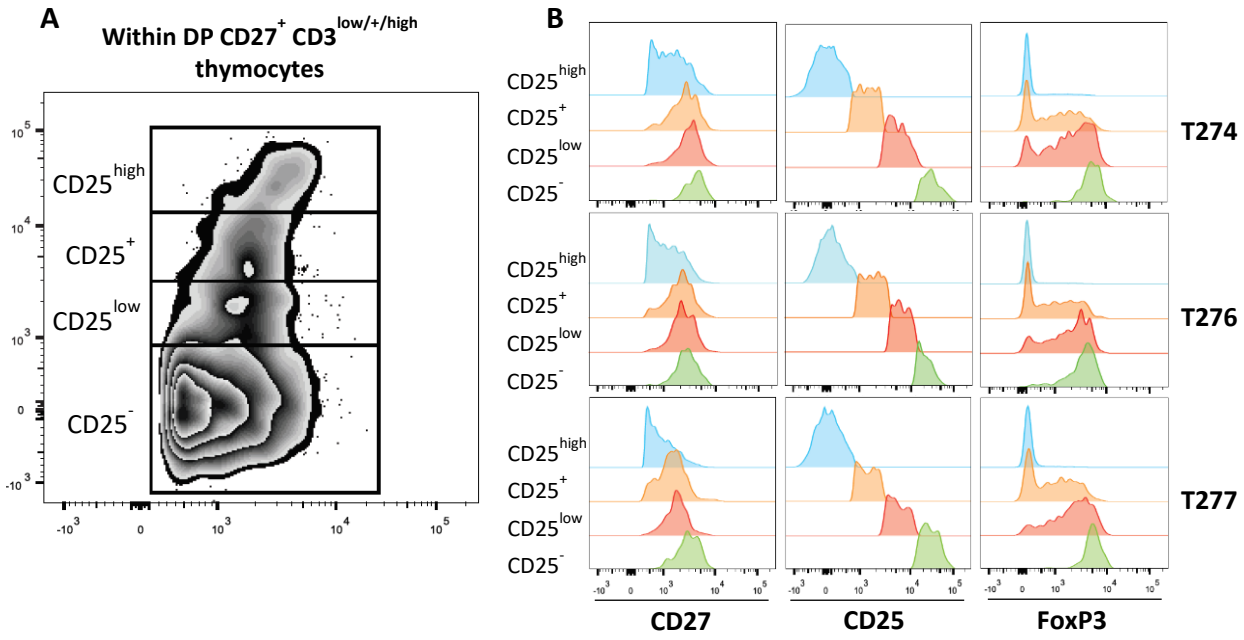


Figure 3.3 Expression of CD27, CD25 and FoxP3 in DP CD3$^{low/+/high}$ mature thymocytes. A, Contour plot of a representative thymus sample showing the gating of DP CD27^{+} CD3$^{low/+/high}$ thymocytes with different levels of CD25 expression. Intensity of CD25 and CD27 is represented along the respective axis, in a Log_{10} scale. B, Histogram plots showing the association of CD27, CD25 and FoxP3 markers expression levels within DP CD3$^{low/+/high}$ expressing CD27 and different levels of CD25 for the three thymuses. Each marker fluorescence intensity in a Log_{10} scale for the frequency of events detected at that intensity in each subset is represented in x- and y-axis, respectively.

The majority of FoxP3^{+} DP CD3^{high} thymocytes in the three thymuses express high levels of CD27, corresponding to the mature phenotype associated with this population (Figure 3.4)[1,165]. In addition, DP CD3^{high} thymocytes express the highest level of FoxP3 (Figure 3.2) within the DP population, an observation that is consistent with the dependency on TCR signalling for its induction in DP thymocytes[131].

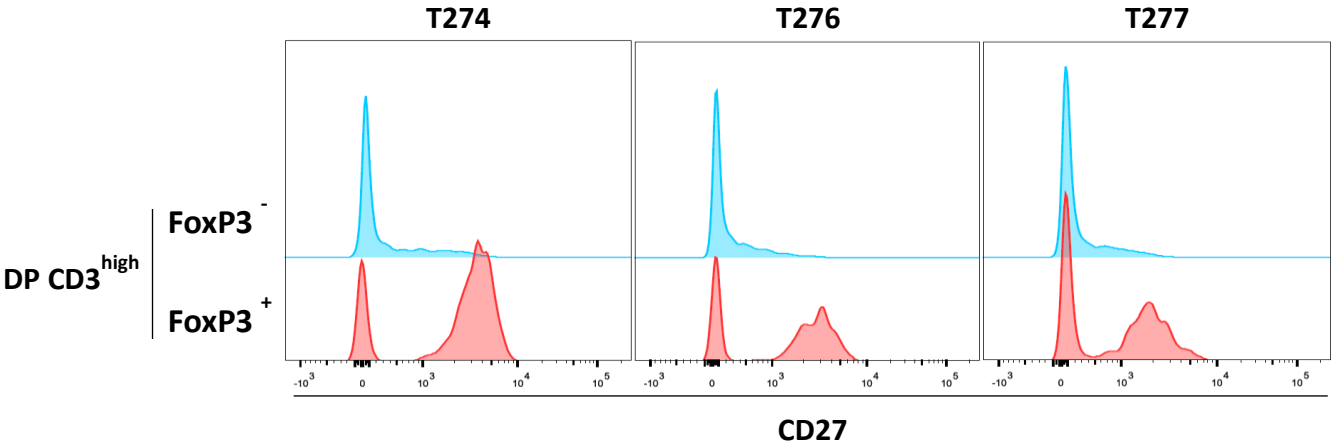


Figure 3.4 Expression levels of CD27 in FoxP3^{-} and FoxP3^{+} DP CD3^{high} thymocytes. Histograms plots showing the CD27 fluorescence intensity in a Log_{10} scale for the frequency of events detected at that intensity in both subsets of each thymus.

In these DP thymocytes, the expression of FoxP3 correlates inversely with the expression of IL-7Ra (CD127), with cells that express high FoxP3 also expressing low levels of IL-7Ra (Figure 3.2). Although it has been reported that DP CD3^{high} and FoxP3^{+} cells may sometimes express high levels of CD127[165], the correlation found is in agreement with previous studies that associate low levels of CD127 to the regulatory phenotype[9,10].

I also observed that CD3^{high} DP cells express significantly less CD8 than CD3$^{-/dim}$ DP thymocytes (Figure 3.5), a reliable reproduction of the down-regulation of CD8 expression in mature DP thymocytes, an event that precedes the CD4/CD8 lineage choice during T cell development[79].

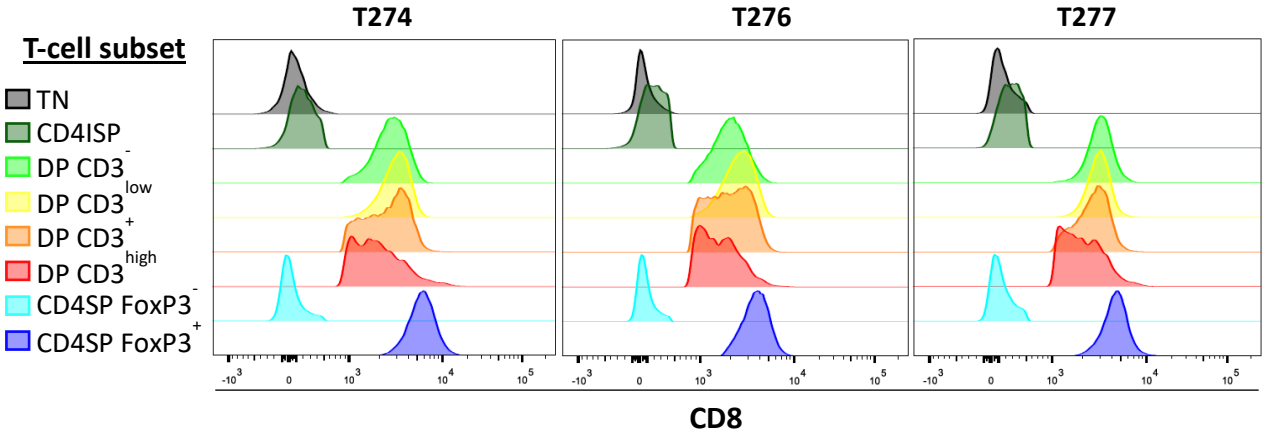


Figure 3.5 Expression of CD8 during the stages of T cell development. Histograms plots showing the CD8 fluorescence intensity in a Log_{10} scale for the frequency of events detected at that intensity in each subset. The down-regulation can be observed in the DP CD3^{+} (orange) and, more accentuated, in the DP CD3^{high} stage (red). Thymocytes that differentiate into CD8SP re-express higher levels of CD8, while CD4SP completely lose the expression of this co-receptor.

Regarding the single positive stages, I observed that CD4SP and CD8SP FoxP3^{+} thymocytes have a CD127$^{-/low}$CD25^{+} phenotype, in contrast to CD4SP and CD8SP FoxP3^{-} thymocytes (Figure 3.6). This is in accordance to the levels of CD127, CD25, and FoxP3 markers that are commonly used to identify Tconv (FoxP3^{-} CD25^{-}CD127^{+}) and Treg (FoxP3^{+}CD25^{+}CD127$^{-/low}$) subsets[118,119] within both CD4SP and CD8SP populations. Furthermore, CD4SP FoxP3^{+} cells express higher levels of CD25 than CD8SP FoxP3^{+} (Figure 3.6), in accordance to what is described regarding the levels of CD25 in CD4SP FoxP3^{+} and CD8SP FoxP3^{+} thymocytes[165,191,205].

The decrease of proliferation activity and Ki-67 expression was also observed in CD4SP and CD8SP subsets, in agreement with DP CD3$^{+/high}$ thymocytes maturing and differentiating into CD4SP or CD8SP thymocytes (Figure 3.2).

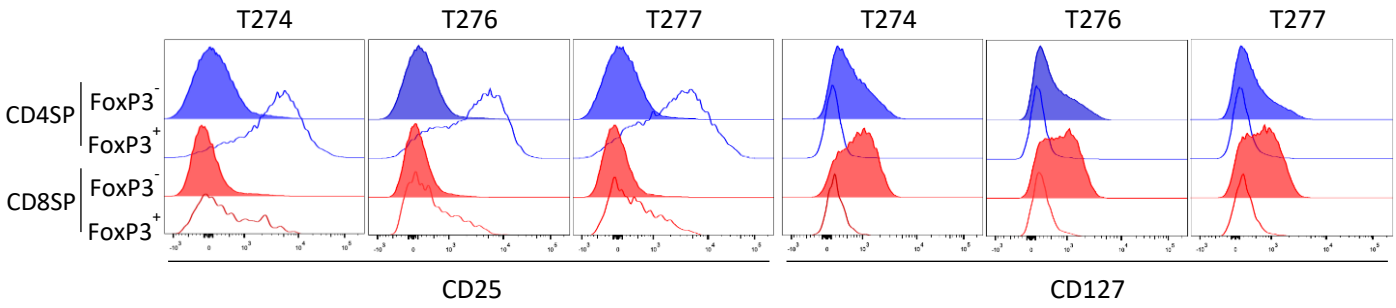


Figure 3.6 Expression of CD127 and CD25 in FoxP3$^+$ and FoxP3$^-$ thymocytes within CD4SP and CD8SP stage. Histograms plots showing the CD25 and CD127 fluorescence intensity in a Log_{10} scale for the frequency of events detected at that intensity in each subset.

In addition, the expression of the naïve-associated marker CD45RA was observed in a subset of mature CD4SP and CD8SP thymocytes. This is in agreement with the gradually expression of this marker in positive-selected and functional mature single positive cells, which exit the thymus and migrate to the periphery to incorporate the naïve T-cell pool[1,7] (Figure 3.2).

I have also analysed the levels of FoxP3 within tTregs and tTconvs subsets. As expected, the majority of CD4SP CD25$^+$CD127$^-$ tTreg cells expressed FoxP3 protein (84% mean), comparing to the respective tTconv counterparts, which only 0.52% (mean) of them presented FoxP3 expression (Figure 3.7).

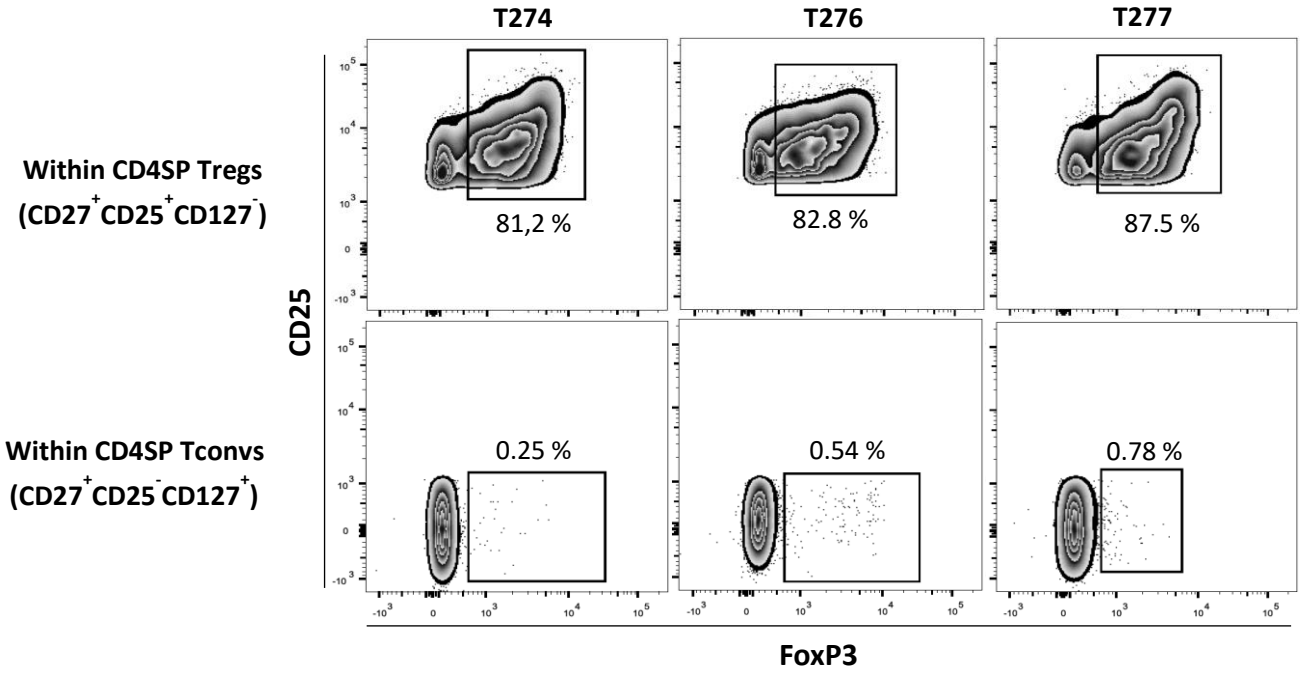


Figure 3.7 Percentage of FoxP3 expressing cells in tTreg and tTconv subsets. Contour plot showing the percentage of FoxP3$^+$ cells in both CD4SP tTreg (CD27$^+$CD25$^+$CD127$^-$) and tTconv (CD27$^+$ CD25$^-$ CD127$^+$) subsets in each thymus. Intensity of CD25 and FoxP3 is represented along the respective axis, in a Log_{10} scale.

As previously described, these markers were used to sort tTreg and tTconv subsets. To ensure that the populations obtained are not contaminated by cells with a different phenotype it is important to assess the purity of sorted cells. This was done through the phenotypical analyses of a representative number of cells from each sorted subset, in order to verify whether each subset had the proper phenotype. Sorted tTreg was composed by more than 97,2%; and tTconv sorted at 99,5% with the associated phenotype, respectively.

These observations confirm our sorting strategy (figure 2.1) that was used to isolate CD4SP Tconv and Treg subsets, without significant levels of contamination (Table 3.1).

Table 3.1 Sorting purity

Subset name	Purity (%) within CD4SP CD27+ population		
	T274	T276	T277
CD25- CD127+ Tconvs	100.0	98.6	100.0
CD25+ CD127- Tregs	99.0	95.2	97.6

The percentage of CD4SP, CD8SP, and DP cells, as well as the CD4SP/CD8SP ratio is also a good method to assess if our samples correspond to individuals with a normal T cell development[240]. With exception of CD4SP (higher) and DP (lower) cells of the T276 thymus, all the populations extracted from the three individuals are representative of age group and subset, according to a study on the distribution of thymocyte subsets with age of human healthy pediatric thymus[240]. Although the associated CD4SP/CD8SP ratio appears to be higher (T274: 4,1; T276: 4,3; T277: 2,8) than those considered standard for each thymus age (T274: 2,3; T276: 2,1; T277: 1,3), it is possible that the difference is due to the authors calculating this ratio using the median value of each subset in every age group, and not taking into account the range values.

Overall, our results shown that the expression pattern of each marker was very similar in all the three thymic samples and in agreement to what has been described in the literature. Therefore, we are confident that these thymic samples can be used as a representation of normal T cell development.

Generation of a Human Thymic Biobank

In this project I have also created a new collection of thymic samples to be used by the lab in the future validation studies. At the moment of writing this thesis, this is composed of unsorted cells collected from the thymuses of 16 individuals (thymus identification – millions of total thymocytes: T290 - 46M; T292 - 252M; T296 - 340M; T298 - 700M; T299 - 2x538M; T300 - 3x650M; T301 - 387M; T302 - 157M; T303 - 2x340M; T304 - 3x400M; T305 - 4x475M; T306 - 375M; T307 - 2x350M; T310 - 3x449M; T311 - 350M; T312 - 3x428M). These

thymuses were phenotypically validated following the methodology described above. I also stored in the biobank also stores thymic tissue samples in paraffin blocks and cryopreserved in OCT (Optimal Cutting Temperature compound). As they are, these samples can also be used to subsequent validation of findings described in this thesis, or other projects.

3.2 Whole-Genome expression of CD4SP Treg and Tconv thymocytes

The characterization of the genome-wide expression profile of tTreg and tTconv CD4SP thymocytes involves several processing steps of the raw RNA sequenced data, ultimately resulting in a matrix containing the expression values for each gene in all the samples. Finally, these genes will be compared and analysed for differential expression analysis between each subset. Figure 3.8 illustrates the pipeline used for the processing and differential expression analysis of the RNA-seq data.

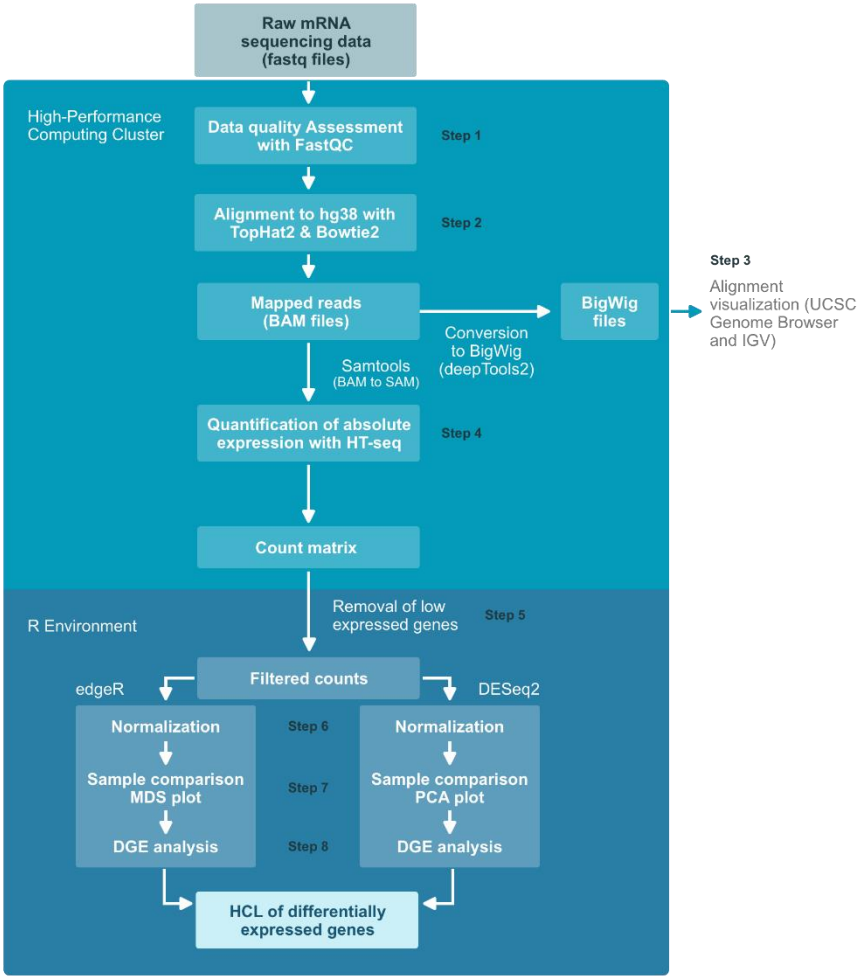


Figure 3.8 Processing and differential expression analysis pipeline of RNA-seq data. Quality assessment of raw RNA-seq data was performed with FastQC followed by the alignment to hg38 reference genome with TopHat2 and Bowtie2. Alignment results were then visualized with UCSC Genome Browser and IGV, and absolute expression quantified with HTseq. After removal of low expressed genes, read counts were normalized, compared, and analysed for differential gene expression (DGE) using edgeR and DESeq2 tool, to finally obtain a high confidence list containing the significantly DEG resulted from both tools.

3.2.1 Next Generation Sequencing (NGS) library strategy

To identify which genes are expressed in each cell type, and at what levels, I generated the genome-wide expression profile of human Treg and Tconv CD4SP thymocytes. To select the coding RNA content, libraries were built using poly-A enrichment before being sequenced by mass parallel sequencing, mRNA-seq (TruSeq RNA library). mRNA-seq allows the precise measurement of gene expression levels between several conditions, as well as the discovery of novel and rare transcripts, and the detection of alternative splicing events[241]. The sequencing coverage allowing all these analyses (equation 3.1) is difficult to calculate for RNA-seq experiments, since transcripts can be expressed at vastly different levels[242]. Instead, and in humans, others have estimated that a sequencing depth of approximately 200 million per sample is required[243,244], together with sequencing both ends of mRNA fragments to allow the detection and quantification of different isoforms and alternative splicing events. Therefore, an excess of 250 million PE reads, at 100bp of length each per sample was planned.

$$C = \frac{L\ x\ N}{G}$$

Equation 3.1. Average genome coverage, *C*; *L,* read length in number of bases; *N*, total number of reads; *G*, the haploid genome length in number of bases.

Sequencing was performed in an Illumina HiSeq™4000 (Methods – Section 2.2.1) and we obtained the sequencing depths indicated within the magnitude required.

Table 3.2 RNA-sequencing depth

Sample Name	Clean Reads
Treg274	186.140.484
Tconv274	210.108.472
Treg276	277.964.152
Tconv276	246.253.892
Treg277	238.926.142
Tconv277	249.283.098

3.2.2 Quality Control of RNA-seq data

The currently high throughput sequencers have the capacity to generate millions of reads in a single run. Therefore, it is highly recommended to perform quality control checks to ensure that there is no problems or biases in the raw data (Figure 3.8 – step 1). I assessed the quality of the sequenced data with FastQC software[222], which reports a wide range of information related to the quality profile of the total reads, including: the overall sequences quality per base; the proportion of each DNA nucleotides across all bases; the percentage of GC pairs along the sequences; and the level of sequence duplication. As representative examples, the forward (FWD) fastq file from T276 Treg and T276 Tconv sequenced subsets are shown. All the above mentioned FastQC plots for every file are in Annexes (Annexe 2 A-D).

Per base sequence quality gives a range of quality values across all bases at each position for a given fastq file (Figure 3.9). The quality values are given by Phred score (Q) which is an integer quantifying the probability that the corresponding base call is incorrect (p), where[245]:

$$Q = -10\, Log_{10}\, p$$

For example, a Q of 20 corresponds to one error in every 100 base calls (p = 0.01), or 99% accuracy. The maximum Phred score is 41. It is very common to observe a decrease in the quality on most platforms, as a function of the read length. When this occurs, a quality trimming process may be required, in which the read bases with poor quality are removed[222]. However, as represented in Figure 3.9, all the base calls throughout the length of the reads were within the higher quality score (Q > 30; green area), which doesn't require additional processing steps at this stage.

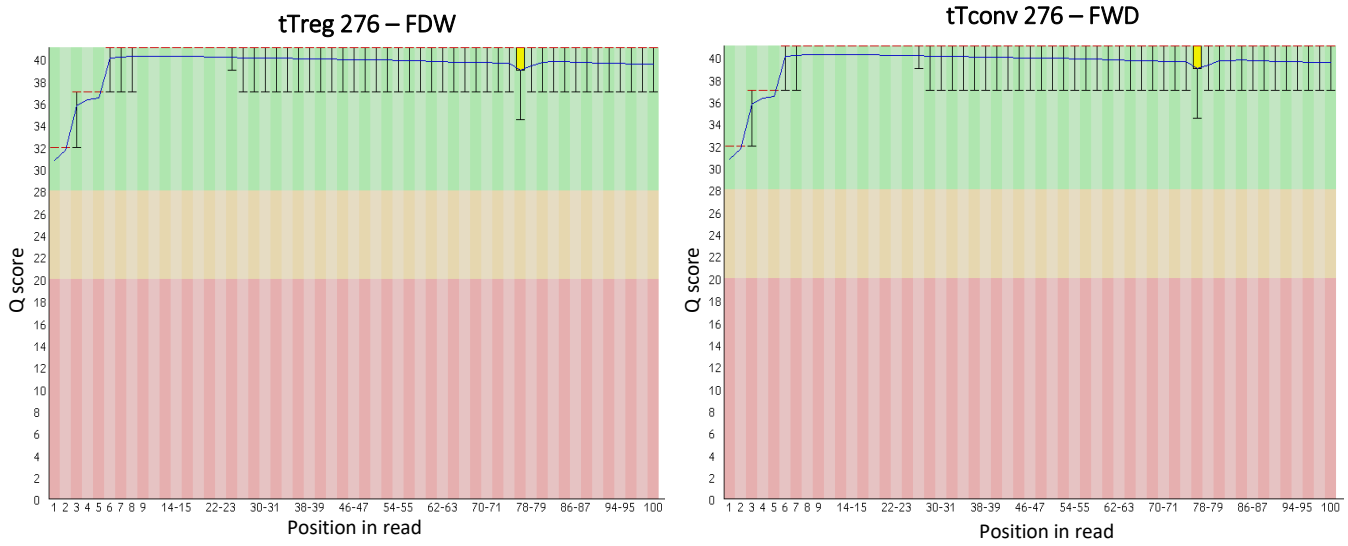


Figure 3.9 Per base sequence quality FastQC plot. Box plots of Phred quality (Q) scores generated by FastQC. The yellow box plots (box: interquartile ranges 25–75%, whiskers: 10–90%; blue and red lines: mean and median quality value, respectively) show the nucleotide-calling Q score for each base across all reads. Q value of each base call ranges from 0 to 41 (y-axis) and is divided by background colours into very good quality (green), reasonable quality (yellow), and poor-quality (red).

Per base sequence content gives the proportion of each individual DNA bases (A, C, G and T) across the length of the reads, which reflects their overall amount in the genome (Figure 3.10). In a good library[222], I would expect little to no differences between the percentages of the different bases in the total sequences along the run, i.e., parallel lines. Indeed, with an exception for the initial bases, no visible differences were observed in the percentages of each base, constantly at 25%. The imbalance observed at the first 15 bases is a technical bias produced in the positions at which the reads start[222]. The pattern is frequently produced when sequencing RNA in Illumina machines due to a biased selection of random primers, although it doesn't represent any individually biased sequences neither adversely affect the downstream analysis[222].

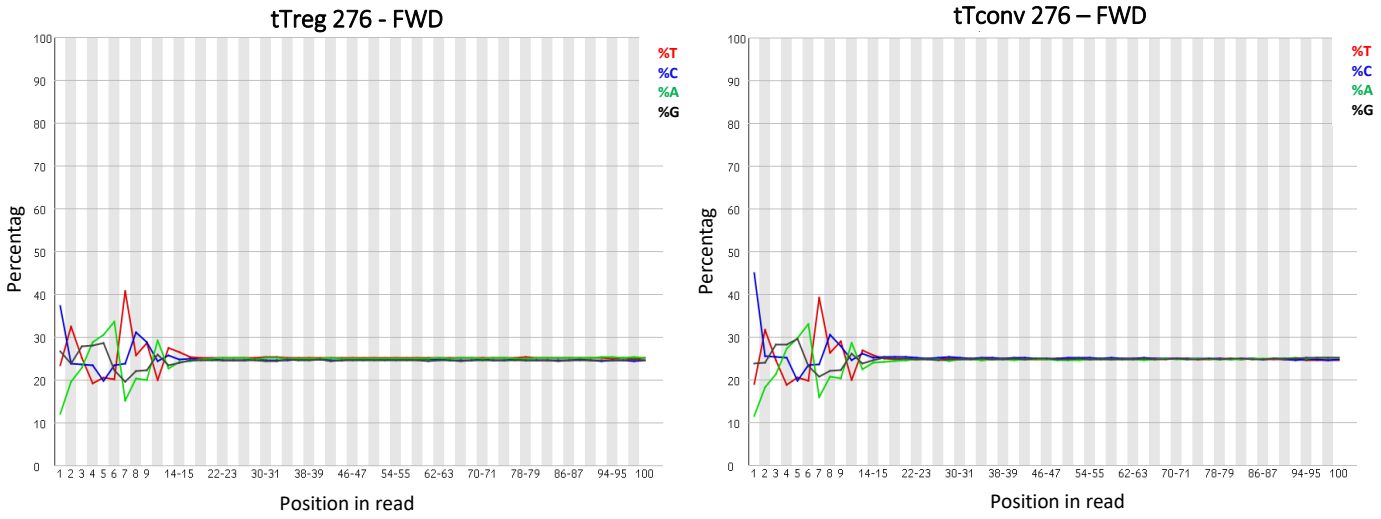


Figure 3.10. Per base sequence content FastQC plot. Percentage of each DNA base across the read length.

Another important measure is the GC content per read of our data (Figure 3.11 - red curve), which is expected to approximately follow the theoretical distribution[222] (Figure 3.11 - blue curve). The theoretical distribution is a computed Gaussian distribution parameterized based on the average and variance of the GC proportion of the input reads. An unusually shaped could represent a contamination with rRNA or organisms with a higher GC content (bacteria, fungi), while a deviated distribution may indicate a biased subset with a GC/AT enrichment. As observed in Figure 3.11, the GC distribution of the data is similar (in both shape and position) to that indicated by the theoretical distribution. Furthermore, the mean percentage of GC content in all samples were 48,18 %.

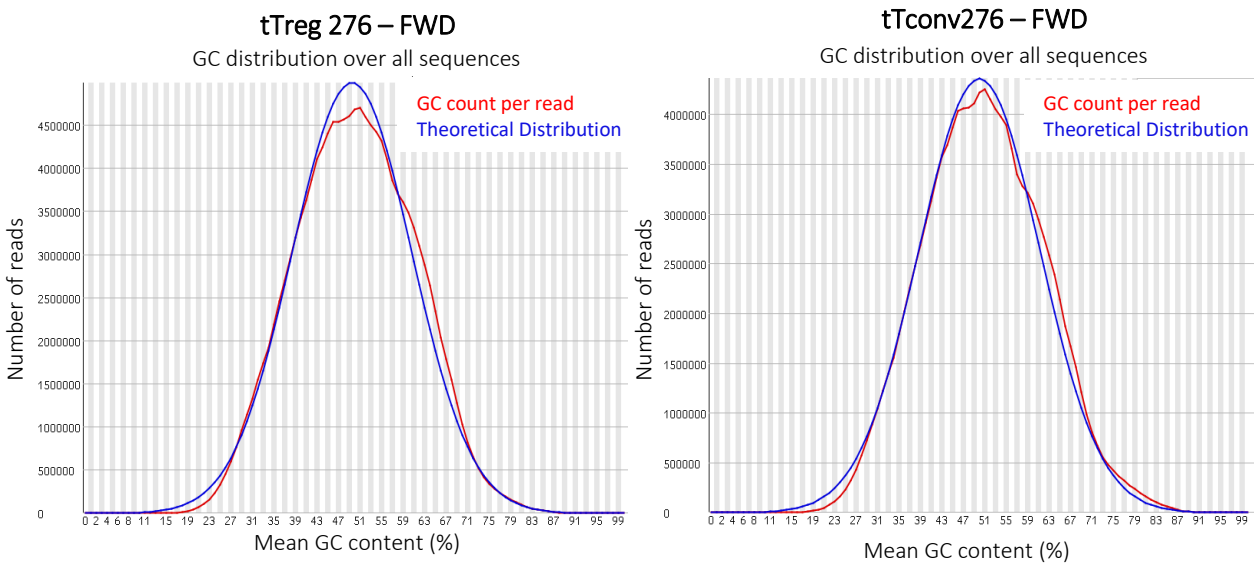


Figure 3.11 Per sequence GC content FastQC plot. A comparison between the observed and the theoretical GC distribution represented by the red and blue curves, respectively.

The duplication levels plot represents the percentages of sequences in the library with different degrees of duplication (Figure 3.12). Due to memory requirements, FastQC only analyse the first 100,000 sequences and this is generally considered enough to get a good estimation for the duplication levels in the whole file[222]. While a low level of duplication may indicate high levels of coverage, high levels of duplication should indicate enrichment bias related with PCR over amplification. However, this assumption is not valid for RNA-Seq libraries, as these are expected to be largely dynamic, resulting in large differences, between lowly and highly expressed genes. The result, as it can be observed in our data, is the presence of sequences with high duplication levels from highly expressed genes (Figure 3.12 – blue line).

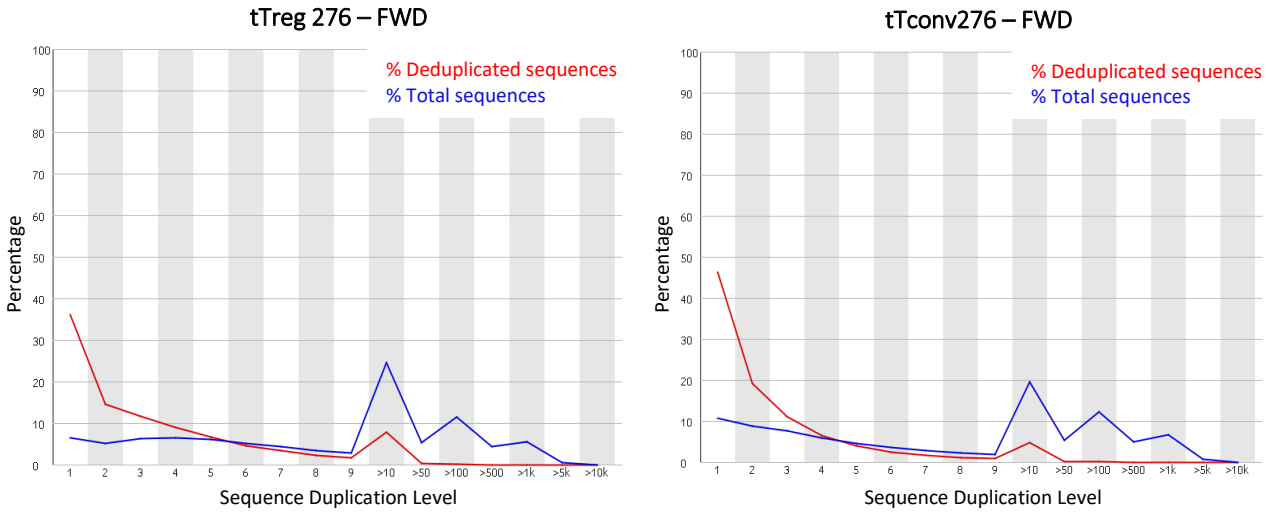


Figure 3.12 Duplicated sequences FastQC plot. The blue line shows the duplication levels distribution of full sequence set, while the red line represents the proportions of the deduplicated set which come from different duplication levels in the original data.

Overall, all of the sequencing data obtained from our samples are with quality in accordance to what is expected for a good library[222], with no evidences for contaminations and/or bias. Given this assessment, I decided there was no need for quality trimming.

3.2.3 Mapping of sequences to the human genome

As a starting point to generate the full transcriptome of each of our samples, the sequencing data of each sample must be mapped to the coordinates provided by a reference genome (Figure 3.8 - step 2), in this case, the latest assembly of the human genome, hg38 (https://www.ncbi.nlm.nih.gov/assembly/GCF_000001405.38). To do this, I chose the tool TopHat2, which is a highly accurate splice junction mapper optimised to work with paired-end reads in a wide range of RNA-seq experimental designs[223]. TopHat2 is indicated to align RNA-seq reads to mammalian-sized genomes, including the human genome, using the ultrafast and memory-efficient high-throughput short read aligner Bowtie2[246]. TopHat2 uses Bowtie2 in each sample to pair the forward and reverse sequence .fastq files and extract their coordinates. Then it will annotate to the respective gene using a reference annotation file (.gtf).

TopHat2 outputs a summary text file containing all the information regarding de alignment, including: the number of inputted reads; the number of pairs that aligned to the reference genome; the number of pairs with multiple alignments; the number of discordant alignments; and the percentage of concordant aligned pairs (Table 3.3). I obtained more than 80% (mean of 85,2%) of concordant pair alignment. In addition, amongst all the samples, the difference in this percentage was not bigger than 7,3%. This range was highly influenced by

the elevated number of discordant alignments identified in Treg274 sample, which was also the one with the lowest number of inputted reads. Overall, the rate of concordant pair alignment obtained was in accordance to the minimal percentage of read pairs expected to map to the reference genome in experiments using human samples[242] .

Table 3.3 Alignment results

	Input reads	Aligned pairs	Pairs w/multiple alignments	Discordant alignments	Concordant pair alignment rate (%)
Tconv274	210,108,472	89,295,957	2.424.239 (2.7%)	2.022.898 (2.3%)	**83,1**
Tconv276	246,253,892	110,130,003	3.291.931 (3.0%)	3.914.759 (3.6%)	**86,3**
Tconv277	249,283,098	112,186,956	2.945.840 (2.6%)	3.127.077 (2.8%)	**87,5**
Treg274	186,140,484	83,070,760	2.464.479 (3.0%)	8.394.300 (10.1%)	**80,2**
Treg276	277,964,152	124,920,212	3.338.530 (2.7%)	3.904.009 (3.1%)	**87,1**
Treg277	238,926,142	106,059,785	2.669.549 (2.5%)	1.940.314 (1.8%)	**87,2**

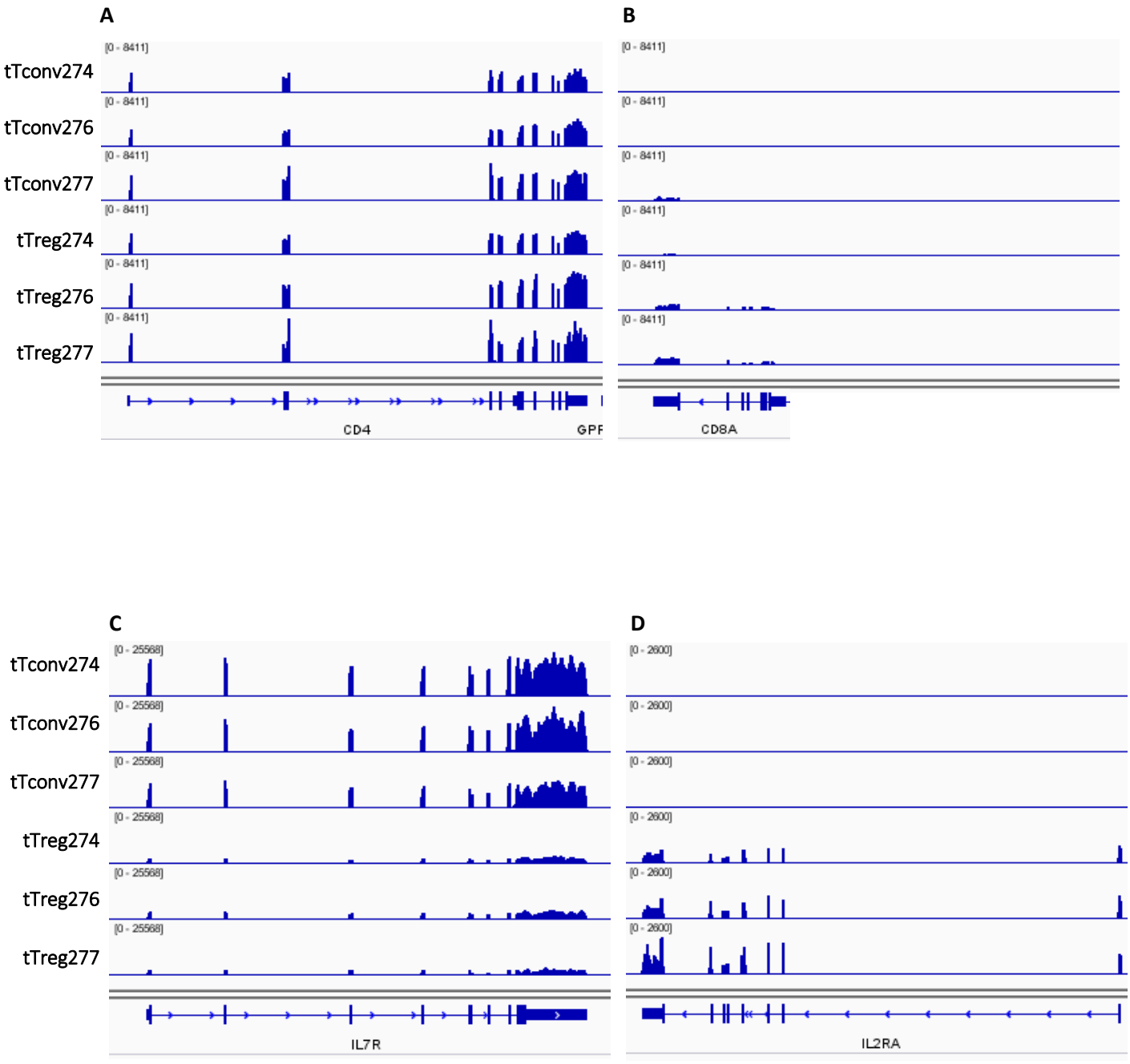
Additionally, TopHat2 also outputs two Binary Alignment Map (BAM) files. One of them contains all the unmapped sequences, while the other one contains all the sequences that aligned to the reference genome. The BAM file containing the mapped sequences can be further converted into a more compressed and indexed binary file, known as BigWig file, useful for alignment visualization (Figure 3.8 – step 3). The high-performance visualization Integrative Genomic Viewer (IGV)[226,227] tool uses the BigWig files and display for each sample, the regions of the genome in which there was read alignment (Figure 3.13)

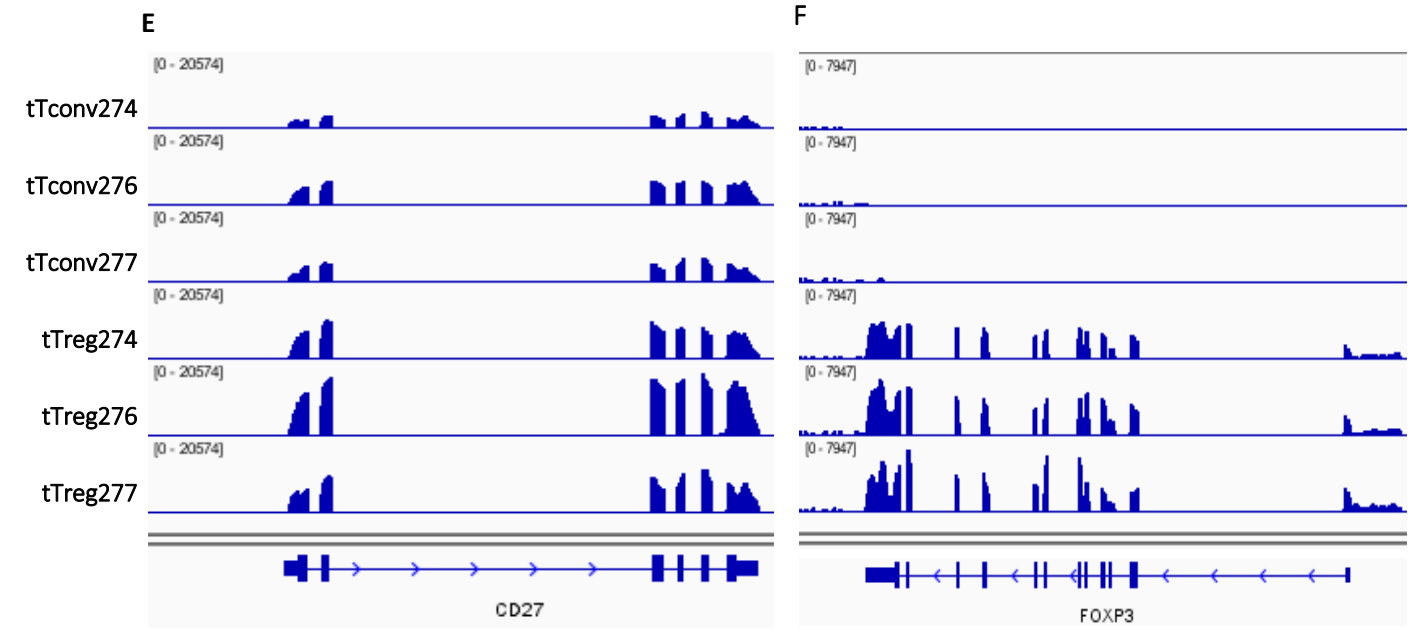
With this, and in accordance to the protein levels that allowed us the isolation of both subsets, I confirmed the expression, in mRNA, of the genes encoding these proteins. Namely, I observed the high expression of *CD4* gene and a low expression in *CD8*α gene in all the samples (Figure 3.13 A and B). In addition, the expression of *IL2Rα* is only observed in tTreg samples, whereas *IL7R* gene expression is more accentuated in tTconv comparing to tTreg samples (Figure 3.13 C and D). Furthermore, all the samples express higher levels of the gene-encoding the maturation marker CD27 (Figure 3.13 E).

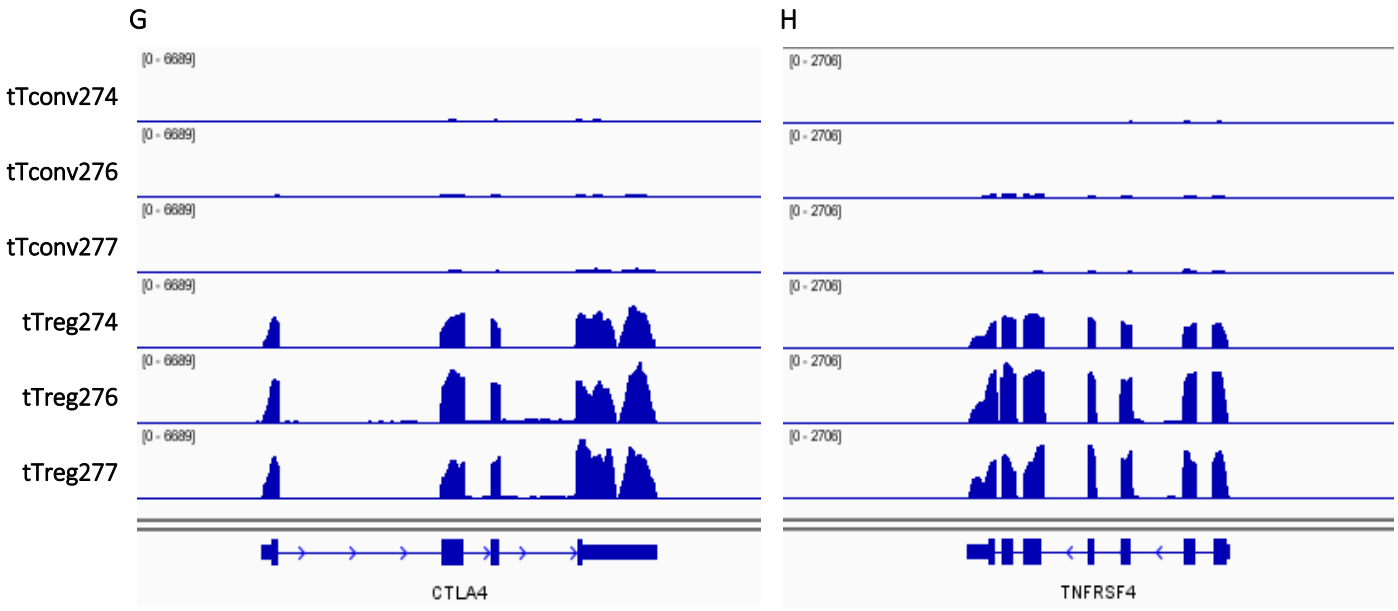
These observations strongly support both the sorting and the alignment steps, since each sample follows a mRNA gene expression in accordance to the phenotype, in protein, that distinguishes tTreg ($CD4^+CD8^-CD27^+CD25^+CD127^{low}$) from tTconv ($CD4^+CD8^-CD27^+CD25^-CD127^+$) subset.

Additionally, I observed that tTregs and tTconvs cells samples expresses genes associated with each subset. These include: FOXP3, the main marker of regulatory T cells; *CTLA4*, which mediates the suppressive function of Treg cells[113]; *TNFRSF4* (OX40), *TNFRSF18* (GITR), *IKZF4* (Eos) and *LRRC32* (GARP), which are all genes

with an expression associated to a Treg phenotype[247–250] (Figure 3.13 F-K); and *CD40LG,* a gene expressed by Tconv cells[251] (Figure 3.13 L).




E
F
tTconv274
tTconv276
tTconv277
tTreg274
tTreg276
tTreg277
[0 - 20574]
[0 - 7947]
CD27
FOXP3


G
H
tTconv274
tTconv276
tTconv277
tTreg274
tTreg276
tTreg277
[0 - 6889]
[0 - 2706]
CTLA4
TNFRSF4

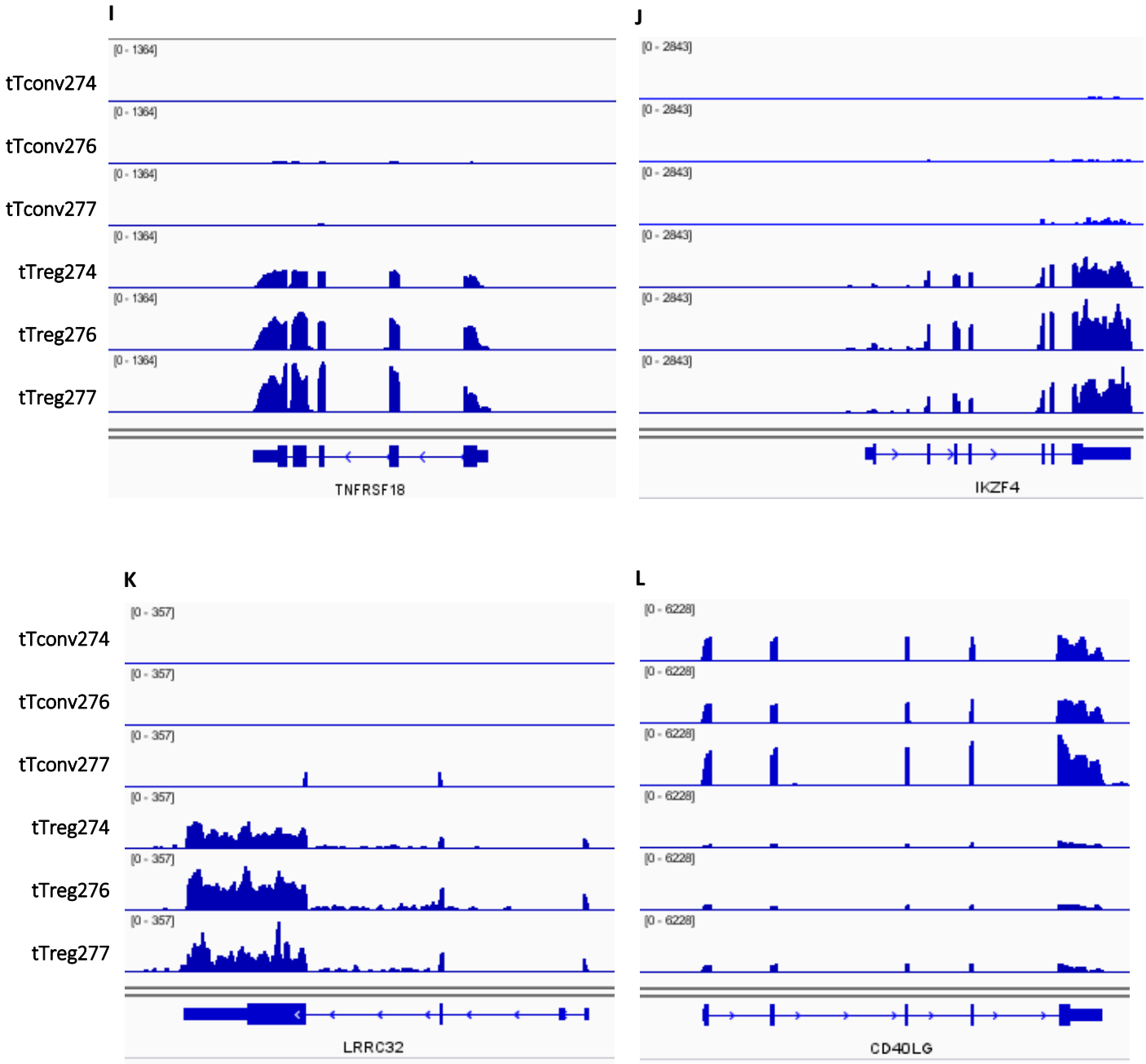


Figure 3.13 Visualization of alignment results. A-L: Screenshots of Integrative Genomics Viewer (IGV) tool showing the distribution of aligned reads across each gene for all samples (horizontal rows).

Overall, mRNA expression found for the above-mentioned genes are in accordance with the presence of these proteins that characterize each subset.

3.3 Expression Analysis between tTreg and tTconv

3.3.1 Quantification of absolute expression

Once the sequenced data are correctly mapped and annotated, it is necessary to quantify it so to determine how much of each gene is being expressed in these different thymocytes (Figure 3.8 – step 4). To do this, I used HTseq tool, a fast and highly-performance tool with the capacity to process paired-end data, with the feature *count*, which counts how many reads map to each gene[229]. In addition, the HTseq *count* method is designed specifically for expression analysis, since it only counts those reads that mapped unambiguously to a single gene, whereas the reads aligned to multiple positions or overlapping with more than one gene are discarded[229]. To do this, HT-seq tool uses the BAM file produced by TopHat2 that contains all the sequences that mapped to the reference genome during the alignment. However, as HT-seq *count* cannot count sequences in a binary format, the BAM file was first converted into its uncompressed text-based format, known as Sequence Alignment Map (SAM) file (Figure 3.8).

As a result, HT-seq provides a table with the number of uniquely mapped RNA-seq reads (counts) for each gene in all the samples. In addition, HT-seq also reports the number of reads with no genomic feature, and the reads that mapped ambiguously (Table 3.4), which were not considered in this analysis.

Table 3.4 Number of aligned reads per sample

	tTconv274	tTconv276	tTconv277	tTreg274	tTreg276	tTreg277
Uniquely mapped	81,235,084	99,326,328	103,319,398	72,039,890	111,011,513	97,794,943
No feature	5,106,409	6,831,335	5,433,583	5,064,574	10,925,970	5,442,710
Ambiguous	8,536,981	13,062,939	11,280,912	13,610,991	13,449,505	9,327,560

Using the mean expression distribution of each gene in the two subsets - taken from raw counts for all three replicates of each subset - (Figure 3.14), it is possible to detect that, overall, there are several genes with higher absolute expression in one subset than the other, as well as genes with the similar levels in both subsets. In addition, the large majority of the genes have very low levels of in both subsets (Figure 3.14 A - bottom left corner).

It is assumed that, within the limitations of detectability and reproducibility of RNA-seq signal, a very low count across all libraries is of little significance, as well as it decreases the statistical power in future calculations of differential expression[252]. In addition, a gene must be expressed at some minimal level before it is likely to be translated into a protein or to be biologically important. Therefore, it is important to distinguish

real signal from noise in my analysis and be confident that samples are considered only where there is true expression (Figure 3.8 – step 5).

To do this, I used counts-per-million (CPM) rather than filtering on the counts directly (Equation 3.2), as the latter does not account for the differences obtained in library sizes between samples (Table 3.4).

$$CPMi = \frac{Xi}{N} \times 10^6$$

Equation 3.2 – Counts-per-million of mapped reads, CPM; X, read counts for each gene in each sample; N, library size.

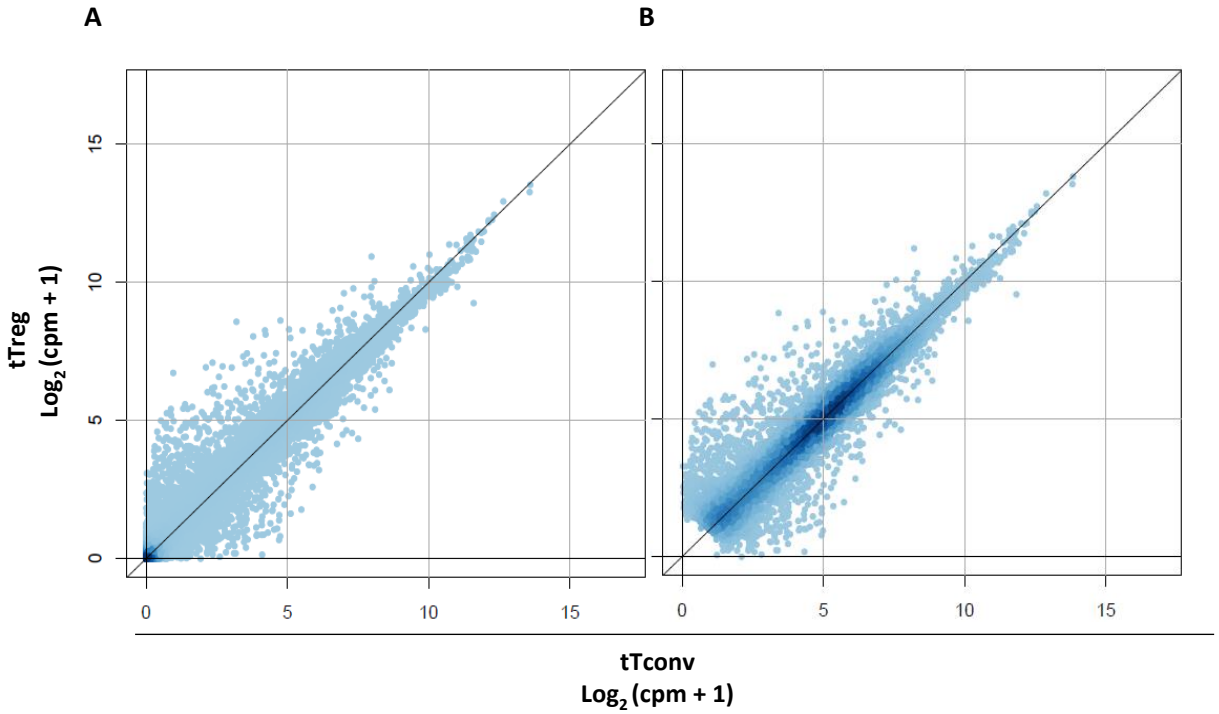


Figure 3.14 Counts distribution in tTreg and tTconv subsets. Raw counts distribution before (A) and after (B) filtering of low expressed genes in tTconv and tTreg subset. Each dot represents the mean expression for each gene in tTconv and tTreg subsets. Overlapping genes are indicated by dark blue colours. Cpm, counts per million.

Therefore, all genes without at least 1 cpm in the sum of the 3 replicates of each subset were removed. With this, I obtained a total of 12,945 genes with an expression value higher enough to be used for the subsequent differential expression analysis (Figure 3.14 B), with a higher proportion (52,6%) of them with more absolute expression in tTconv samples than in tTreg. In addition, I observed several genes with similar expression, as indicated by the bisector line.

As previously noted, although sequencing depth was very similar for all samples, the data sets are obtained from the sequencer with slightly different library sizes (Table 3.2). Therefore, these need to be

normalised, so that the samples may be comparable in their expression profiles across subsets (Figure 3.8 – Step 6). This normalisation is also an essential step in the analysis of differential gene expression[253].

To do this I used both edgeR[231] and DESeq2[232] R packages. These tools are the most widely used and the recommended ones for differential gene expression analysis of RNA-seq data with 3 replicates per subset. Comparing to other tools, both edgeR and DESeq2 have a superior identification rate of true positives and a well-controlled significance values at lower differences of expression (fold changes)[254]. Furthermore, by using both tools I have higher confidence over the results.

Each tool has its own normalization method. edgeR uses the trimmed mean of the log expression ratios (trimmed mean of M values, TMM) and then estimates an additional normalisation factor to account for sample-specific effects, such as diversity. Analogously, DESeq2 defines a virtual reference sample by taking the median of each gene's values across samples and then computes size factors as the median of rations of each sample to the reference sample. In practice, the normalisation factors are similar. Dividing each column of the count table by the corresponding size factors yields normalized count values, which can be scaled to give a counts per million interpretation (Figure 3.15 B and C). Before normalisation, the median values of raw counts for all the genes is not the same for all samples (Figure 3.15 – A). After the normalisation within edgeR or DESeq2 tools, all the samples have an equally distributed centred median values of the normalised counts for all the genes in each sample (Figure 3.15 B and C).

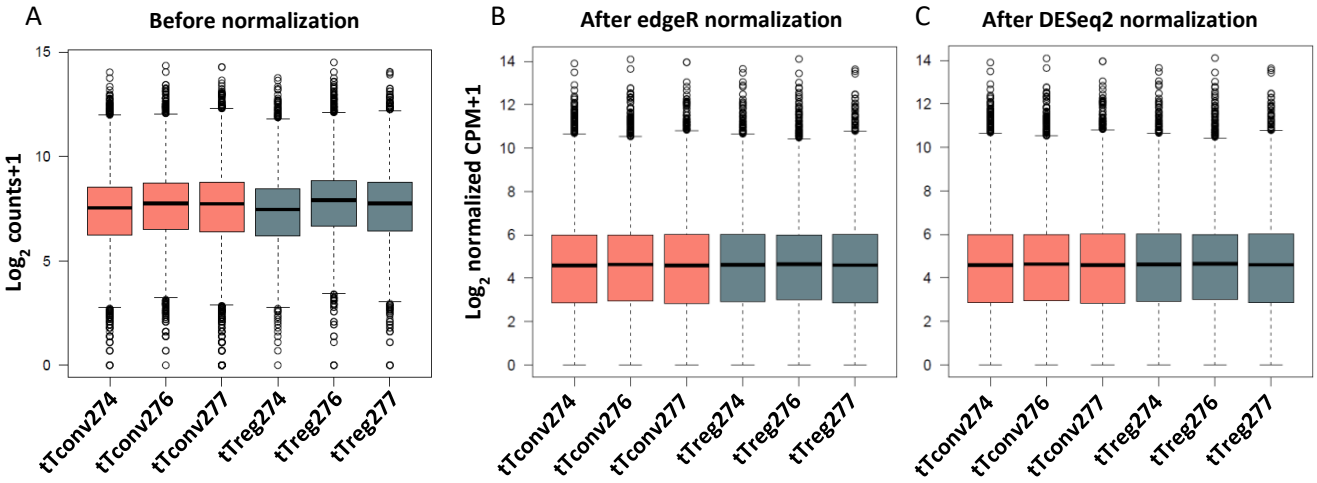
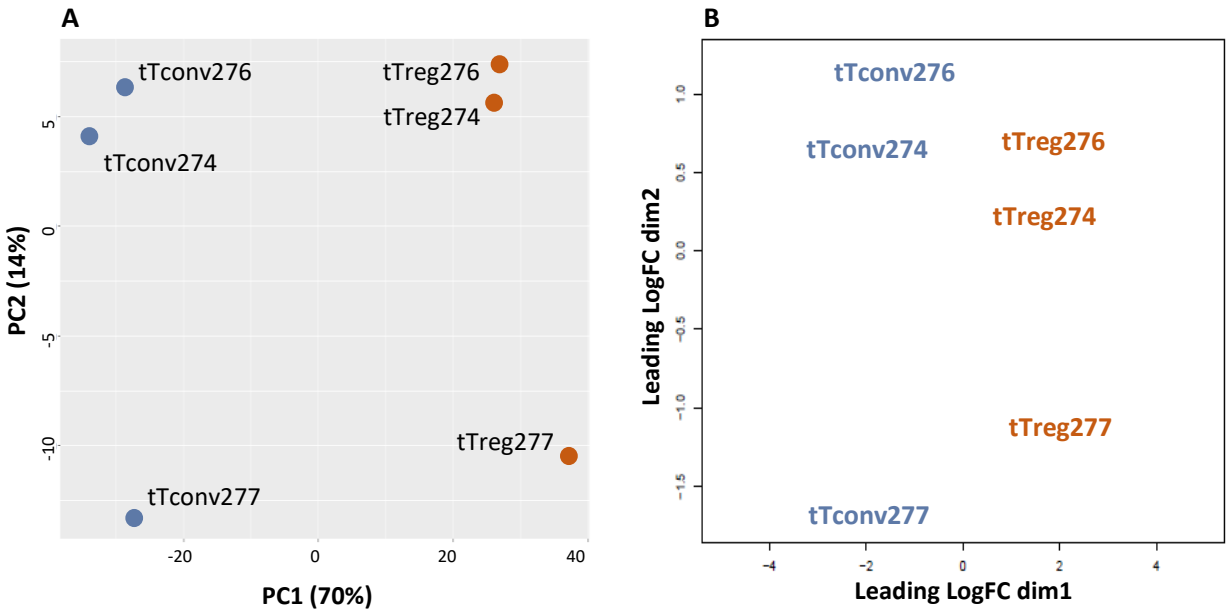


Figure 3.15 Libraries size normalizations. Box plots representing the counts distribution of all genes for each sample before (A) and after normalization with edgeR (B) and DESeq2 (C) tools. Box: median and interquartile range; whiskers: minimum and maximum; extremes: outliers CPM, counts per million.

To compare all the samples (Figure 3.8 – Step 7) and have an overview regarding the similarities and differences between them, as well as within the replicates of each subset, I confronted the overall expression

profiles with both edgeR and DESeq2 tools (Figure 3.16). edgeR draws a multi-dimensional scaling (MDS) plot of the samples in which the distances correspond to the leading log-fold-changes between each pair of samples (Figure 3.16 A). The leading log-foldchange is the average of the largest absolute log-fold-changes between each pair of samples[231,255]. Similarly, DESeq2 draws principal component analysis (PCA) plot (Figure 3.16 B) representing the variance-stabilizing transformation (VST) of transformed count data[256]. In addition, the VST can also be used to produce an unsupervised hierarchical clustering heatmap showing the Euclidean distances between the samples (Figure 3.16 C).

These plots allow the analysis of the degree of variability between the biological replicates and identify possible outlying samples. As represented in the PCA plot, tTreg replicates segregate from tTconv replicates (Figure 3.16 A) to a VST of 70% on the 1st principal component, PC1, suggesting that the expression profiles are very distinct between the two cell types. In contrast, replicates appear to cluster for each subset for PC2, suggesting similarity of expression profiles between individuals. In addition, there seems to exist a tendency for either older and/or male individuals (T277) to show differences comparing to both female and younger individuals (T274 and T276). In fact, in an unsupervised hierarchical clustering heatmap showing the sample-to-sample distances based on the VST data, this tendency is represented by clustering both T274 and T276 individuals apart from the T277 in both subsets (Figure 3.16 C). A similar result is obtained with MDS plot (Figure 3.16 B), in which tTreg subset segregate from tTconv subset to a LogFC around 4 on the 1st dimension (dim1), while the replicates cluster for each subset (dim2). Taken together, all replicates appear to be equally distinct across cell types and similar within them.



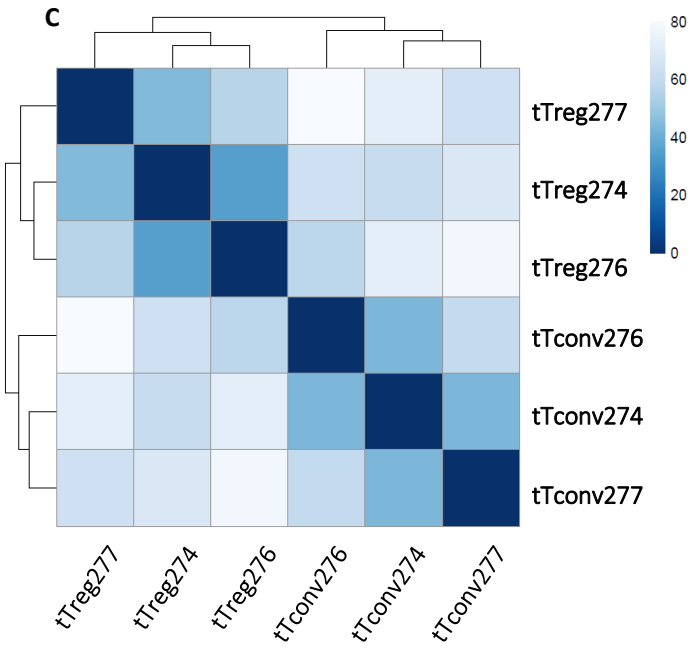


Figure 3.16 Sample-to-sample comparison. A. Principle Component Analysis (PCA) plot of tTconv and tTreg samples. The distance between and within subsets is indicated by PC1 and PC2, respectively. B. Multidimensional Scaling (MDS) showing the distance between each pair of samples based on the leading fold change between each pair of samples. C. Hierarchical heatmap with the Euclidean distances between the samples as calculated from the variance-stabilizing transformation of the count data. Brighter blue colours reflect bigger differences across sample pairs.

3.3.2 Defining the pattern of expression that distinguish tTreg from tTconv

The identification of genes that are differentially expressed between tTreg and tTconv subsets may provide valuable clues regarding the pathways involved in the control of the differentiation and homeostasis of thymic T cells (Figure 3.8 – Step 8). Both edgeR and DESeq2 calculate the differential expression (referred as fold change, FC) and a statistical significance associated with this differential expression for each gene. The fold change is based on ration of the mean count across replicates in each condition and is expressed as a logarithm 2 of the expression ratio (Log_2FC). The statistical significance calculated by edgeR and DESeq2 is calculated using same model, known as negative binomial (NB) model, which assumes that the number of reads for each gene in a sample has a mean and variance, in which the dispersion represents the overdispersion relative to the Poisson distribution[257]. This NB model has been shown to be a good fit to analyse RNA-seq data, and it is flexible enough to account for biological variability[255]. Both tools measure the statistical significance by returning a Benjamini-Hochberg corrected *P*-values of FDR (False Discovery Rate), which controls for the expected proportion of false positives[233]. Figure 3.17 represents a volcano plot with the distribution of the FDR values (expressed as a minus logarithm 10 of the FDR, $- Log_{10}$ FDR) associated with the FC for each gene, calculated by edgeR (Figure 3.17 A) and DESeq2 (Figure 3.17 B) tools.

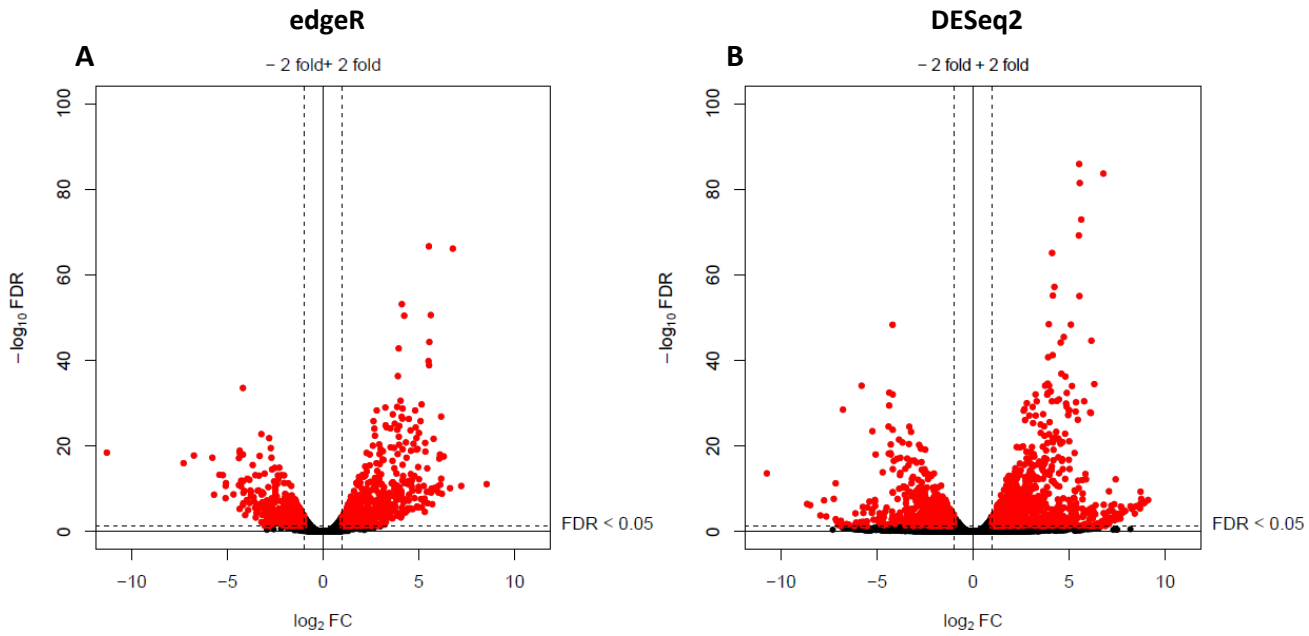


Figure 3.17 Fold Change and FDR values distribution of tTreg and tTconv DE genes. Volcano plots obtained from edgeR (A) and DESeq2 (B) tools. Vertical and horizontal dashed lines represent fold change and FDR thresholds, respectively. According to these thresholds, significantly differentially expressed genes are highlighted (red).

For a gene to be differentially expressed at a statistically significant level, it needs to be quantitatively different between the 2 biological groups and relatively consistent within each group of replicates. With this, to identify the significantly differentially expressed genes between tTreg and tTconv subsets, I applied two cut-offs. First, I selected only the statistically significant genes by choosing an FDR value lower than 0.05 (5%). Second, to select the genes that are differentially expressed between the two subsets, only those with a fold change equal or above 2-fold were considered (FC ≥ 2). While a higher FC cut-off value could result in the exclusion of genes with a considerable difference of expression between tTregs and tTconvs, applying a lower FC cut-off value could result in the inclusion of genes whose difference of expression is not sufficiently pronounced between the subsets.

Following these criteria, I obtained a total of 1588 and 1077 DE genes from DESeq2 and edgeR tools, respectively (Figure 3.17 – red highlighted genes). Although both tools implement general differential analysis on the basis of the NB model, they differ in how the dispersion is determined. Specifically, edgeR moderates feature-level dispersion estimates toward a trended mean according to the dispersion-mean relationship. In contrast, DESeq2 takes the maximum of the individual dispersion estimates and the dispersion-mean trend[256]. This results in DESeq2 being less sensitive to the presence of outliers, whereas edgeR is more selective[256]. In fact, DESeq2 returns gene FCs with a significance close to the cut-off (338 compared to 173 genes between 0.05 and 0.01 FDR values obtained from DESeq2 and edgeR tool, respectively). Conversely, edgeR selection includes genes with lower statistical significance. Overall, the differential expression did not differ significantly

between tools, as there were no contradictions of genes being up-regulated in one tool and down-regulated in the other.

To obtain a high confidence list (HCL) containing the significantly DEG resulted from both tools, I intersected the 1588 genes obtained with DESeq2 with the 1077 genes obtained with edgeR (Figure 3.8). With this, I obtained 1047 genes significantly differentially expressed between tTreg and tTconv subsets (Figure 3.18), with 648 of these (62%) being up-regulated in tTregs.

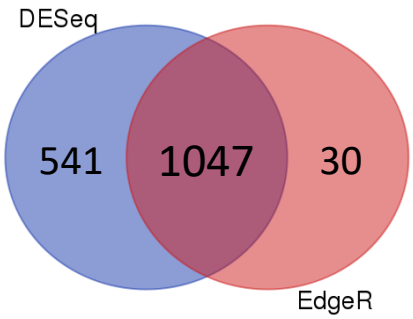


Figure 3.18 Venn Diagram of differentially expressed genes identified in both edgeR and DESeq2 tools.

The relative expression for each identified gene in each sample to the mean expression of that gene in all the replicates is represented in Figure 3.19. In both subsets there is a clear distinction between the genes with higher expression from those with lower expression comparing to the mean. Consistently with our previous observations, T277 individual is clustered separately from the T274 and T276 individuals in both subsets. However, this is not marked by strongly differences in expression, as the relative expression profile of each replicate is similar across the respective subset.

Confirming the previous observations from the alignment visualization, I observed the prominent expression of already known Treg-associated genes (Figure 3.19 B), including those encoding for the lineage-specifying transcription factor *FOXP3*, the IL-2 receptor α chain (CD25) and *CTLA4*, which is constitutively expressed on Treg and mediates suppressive function[258]. Additionally, *TNFRSF4* (OX40)[248], *TNFRSF18* (GITR)[249] and *IKZF4* (Eos)[250] up-regulation were also observed, all genes associated with a Treg phenotype. In addition, the relative expression observed for these genes is consistent along the replicates. *IKZF2* (Helios) expression was also observed in tTregs which has been proposed to be exclusive to thymic-derived Treg, as opposed to peripheral-derived Treg[259]. However, the use of this gene as a thymic-derived Treg marker has been discussed, since it was found to be expressed in peripheral induced FoxP3$^+$ Treg cells[260–262].

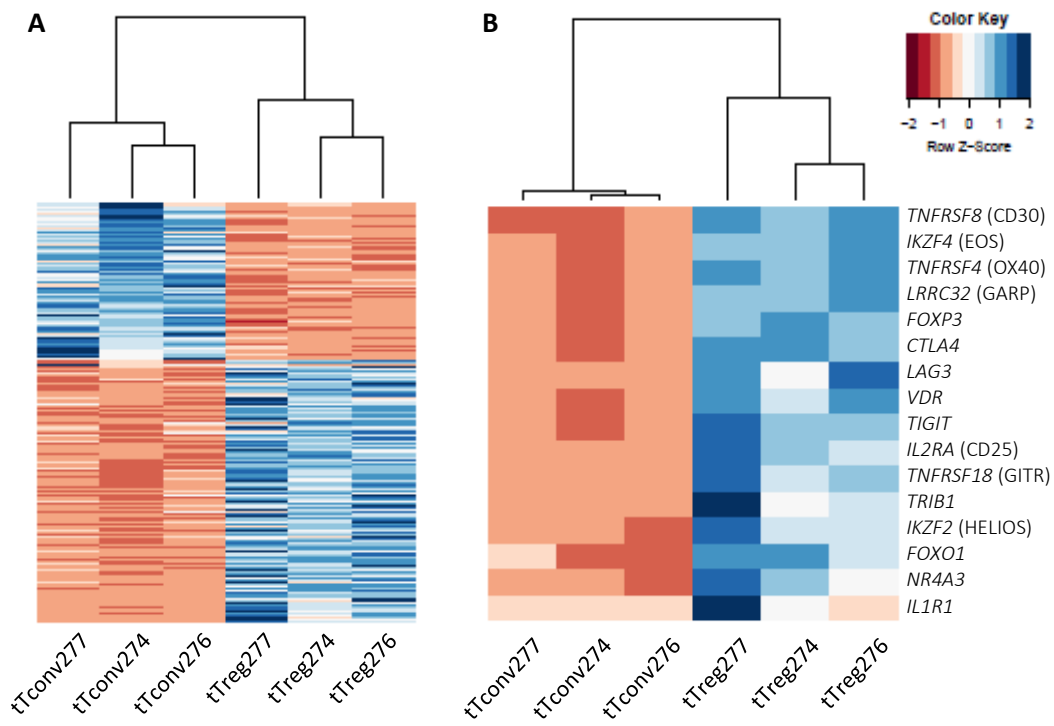


Figure 3.19 Relative expression of differential expressed genes between tTreg and tTconv. A, Hierarchical clustering heatmap of the 1047 differentially expressed genes between tTreg and tTconv subsets. B, Hierarchical clustering heatmap of identified Treg-associated genes. Colour scale represents mean centred fold expression level of normalized cpm across samples.

Taken together, these observations gave confidence that our RNA-seq data can be used to detect changes in gene expression relevant to Treg biology. Having this set differentially expressed genes between tTreg and tTconv (the full identity of the 1047 DE genes is detailed in annexes (Annexe 3)), it is important to have a first overview of the biological processes and molecular pathways associated to this this differential expression.

To do this, I used Gene Ontology (GO), which provides controlled vocabularies of defined terms representing gene product properties. In addition, GO allows one to know if the association between a GO term and a group of genes is significant, by comparing this set of genes to a random list of genes. The most significant enriched biological processes and pathways terms associated with each set of genes are represented in Figure 3.20 and 3.21, respectively.

For the genes overexpressed in tTregs, GO revealed several biological process terms (Figure 3.20) associated with the negative regulation of cell activation for a group of genes, including: *CTLA4*, *TIGIT*, *LRRC32*, and *HMOX1*. The identification of genes involved with this process is in accordance to the main function of regulatory T cells, which is the suppression of T cell activation[98]. Not surprising, GO also revealed genes wit associated processes related with NF-kB signalling, including *NFKBIA*, *NFKB2*, *MAP3K14*, and *BIRC3*. Indeed, this pathway is deeply involved in the differentiation of regulatory T cells and activated downstream of TCR

signalling[263,264]. Interestingly, for the set of genes overexpressed in tTconvs, GO showed several genes associated with the negative regulation of phosphorylation process, such as the phosphatases *DUSP14* and *CTDSPL,* and the kinases *PRKCA* and *AKAP6* genes. Additionally, biological processes involving the regulation of cell-cell adhesion, adaptive immune response, and response to pathogens were also associated with over-expressed tTreg and tTconv genes. These are all processes that feature both T cell subset and consistent with the importance of them in its biology.

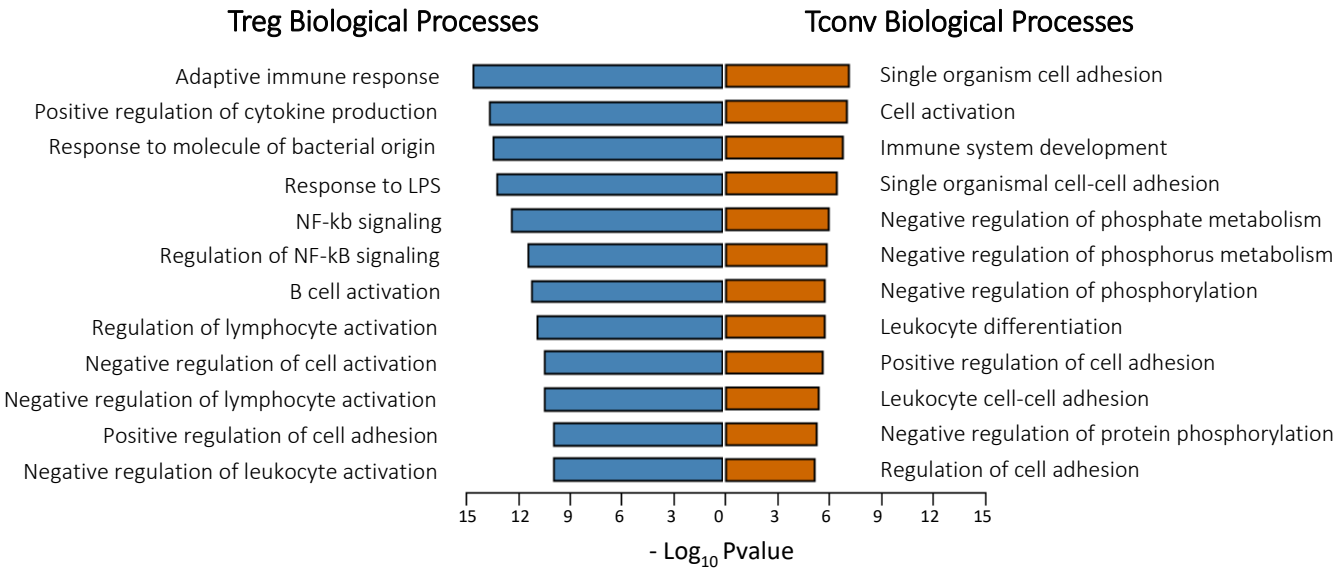


Figure 3.20 Biological processes associated with overexpressed genes in tTreg and tTconv subsets. Gene Ontology results on the significant biological processes associated with tTreg (left blue bars) and tTconv (right orange bars) overexpressed genes.

Regarding the pathways for the set of overexpressed genes in tTreg subset (3.21), the most significant associated term was NF-kB signaling pathway, as previously observed in GO of biological processes. In addition, several tTreg up-regulated genes were also associated with in immune-related pathways, namely cytokine-cytokine receptor interaction (e.g. *CCR8, CCL22* and *TNFRSF4*), TNF signaling (e.g. *TNFAIP3*, *TRAF1* and *TRAF2*, and cell adhesion (e.g. *CDH1*, *ALCAM* and *ICAM1*) pathways. Surprisingly, a group of genes over-expressed in tTregs were significantly associated with effector Tconv cell differentiation pathways, including *TBX21*, *STAT4*, *IL4R*, and *RORA.* These observations will be further discussed in section 3.4. Interestingly, for the set of genes over-expressed in tTconvs, a group of genes (*e.g. WNT5A*, *WNT10B* and *FZD1)* were significantly associated with some cancer-related pathway terms, in agreement with studies reporting the involvement of each in cancer[265–267].

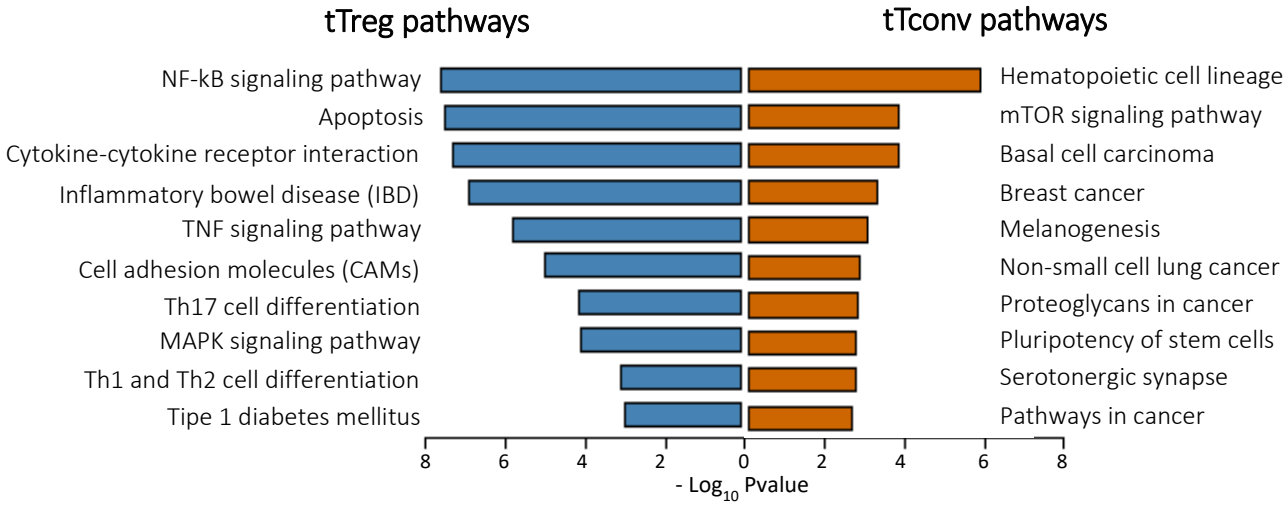


Figure 3.21. KEGG pathways associated with overexpressed genes in tTreg and tTconv subsets. Gene Ontology results on the significant pathways associated with tTreg (left blue bars) and tTconv (right orange bars) overexpressed genes.

3.3.3 A Treg signature for the human thymus

Assessing the first expression signature of human thymic Treg cells is of utmost importance to give new perspectives in the identification and definition of these cells in humans, as well as to provide tools to better understand their development and function. Genes that are up-regulated in tTregs and not in tTconvs are likely involved in the determination of their respective cell fate. From these, the expression that is specific to one subset and not the other is central in the definition of a Treg signature in the human thymus. In addition, its assessment could reveal novel factors specifically involved in the biology of these cells.

I therefore selected a set of genes that are expressed in tTregs and not in tTconvs. To this purpose, I used the mean absolute expression levels of *IL2RA* gene in tTconvs as the threshold to determine this set of genes, given the fact that this gene has very low levels of expression in thymic Tconvs, in contrast to its constitutive expression in thymic Tregs[108]. In addition, this strategy is consistent with the sorting strategy used to isolate tTregs from tTconvs, based on the presence of CD25 marker, encoded by *IL2RA* gene. Therefore, for the set of genes up-regulated in tTregs, those with an expression mean value, in tTconvs, below this cut-off were considered to be uniquely expressed in tTreg subset.

With this strategy, 196 genes (Annexe 4) were found to be uniquely expressed amongst the 648 up-regulated genes in tTregs compared to tTconv. Using Gene Ontology in this set of genes, I observed some of them with a strong statistical association for biological process terms related with the regulation of cell

migration *(e.g. FN1, CCL22, LMNA, LAMA2)*, cytokine pathways (*e.g.*, IRF5, *IL12RB2, IL1RL1, EBI3*), and ion homeostasis (*e.g. RYR1, ACTN2, HMOX1, CHRNA6, TMPRSS6*) (Figure 3.22).

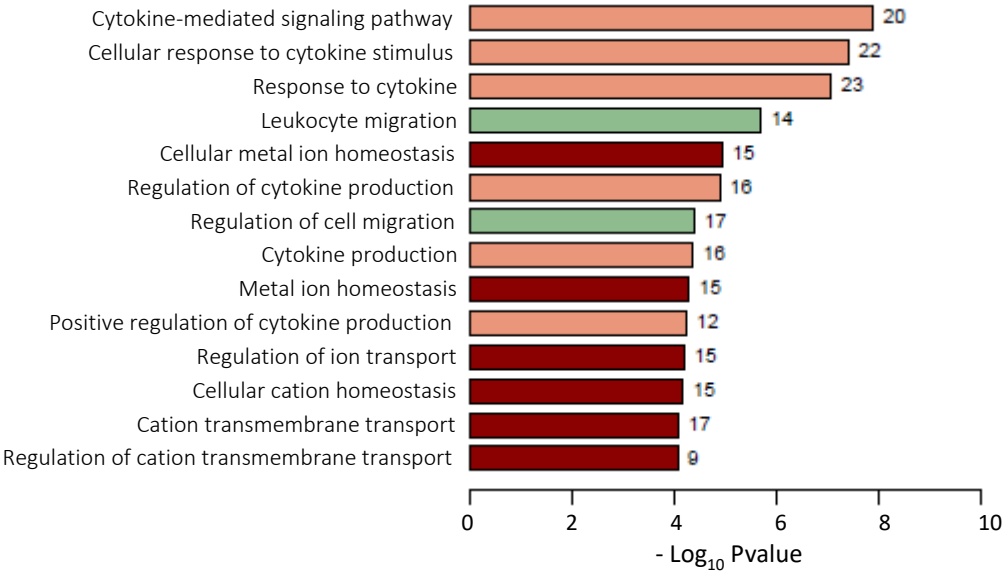


Figure 3.22 Biological processes associated with the uniquely expressed genes in tTreg subsets. Gene Ontology results on the significant biological processes associated with tTreg uniquely expressed genes. Number of genes associated with each biological process term is indicated after each bar. Regulation of cell migration, cytokine pathways and ion homeostasis biological processes are grouped by colour.

Consistent with these observations, several studies using mouse models have shown that thymic T-cell development is coupled with a highly ordered migratory pattern, called intrathymic cell migration[268,269]. In fact, *FN1* (fibronectin 1), which is the most expressed gene in this list of 196 genes, is involved in cell migration processes, and has been shown to be capable of regulating the development of T regulatory cells in mouse[270].

In addition and, ion channels and ion transporters are important regulators of the intracellular concentration of different ions, including calcium (Ca^{2+}), which plays crucial roles in lymphocyte function and immunity[271]. Calcium release from endoplasmic reticulum (ER) is induced by TCR signaling and is, in part, mediated by Ryanodine receptor (RYR) channels. Accordingly, *RYR1* was uniquely expressed in our tTreg subsets and has already been found in T and B lymphocytes[272].

From this list of uniquely expressed genes, I have also identified some genes with relevance to Treg biology including *TNFRSF8* (also known as CD30), *LRRC32* (also known as GARP), and *CCR8*. CD30 is preferentially expressed on activated compared to resting lymphocytes and is relevant to Treg function in the context of autoimmunity through limitation of autoreactive CD8 T cell proliferation[273], while GARP is highly expressed on activated Tregs and contributes to their suppressive activity[247,274]. CCR8 is mainly expressed on Treg cells and is known to be critical for Treg function[275].

Among this group of uniquely differentially expressed genes, I identified two genes with no reported functions in tTreg cells, namely *DNAH8* and *TNFRSF11A*.

Dyneins are microtubule-associated motor protein complexes, composed of several light, intermediate, and heavy chains[276]. *DNAH8* encodes a type 8 heavy chain of an axonemal dynein, involved in sperm and respiratory cilia motility[277,278]. According to our observation, previous studies in humans have also reported the up-regulation of *DNAH8* in both resting and CD3/CD28-stimulated human peripheral $CD4^+$ Tregs compared to $CD4^+$ Tconvs[14]. Furthermore, the over-expression of this gene was also identified in human breast tumour-resident $CD4^+$ Treg cells compared to the respective Tconv cells[280]. In addition, *DNAH8* was found to be down-regulated in *IKZF2* (Helios) knockdown compared to wild-type Treg cells in mice[281], which is known to regulate Treg functional stability[282].

Following the role of axonemal dyneins in cilia structures, there have been some controversy regarding the existence of a cilia in T lymphocytes, with presumably roles in T cell activation and in the formation of immunological synapses[283]. However, cilia have not been previously observed in lymphocytes and the evidence for their presence in these cells is indirect and functionally based. In fact, with the basis that Hedgehog (Hh) signalling in mammalian cells requires primary cilia, De la Roche *et al*. has shown that Hh signalling play a role in cytotoxic T lymphocyte function, proposing that the immunological synapse may represent a modified cilium[284]. Therefore, DNAH8 might be an important player in these structures.

Supporting the putative involvement of DNAH8 in immune cells, a recent paper has reported a predicted damaging missense variant in the *DNAH8* gene of a consanguineous family with CVID (Common Variable Immunodeficiency Disease)[285]. Moreover, they observed that one of the main symptoms displayed by one of the sisters of this family, was chronic sinusitis and respiratory failure, which could be related to the reported mutation in *DNAH8*, considering its role in respiratory cilia motility. In fact, primary cilia dyskinesia, a genetically disorder that leads to chronic pulmonary diseases[286], have been related with mutations in other genes belonging to this family, including *DNAH11*[287], *DNAH5*[288], and *DNAH6*[289], this last showing an up-regulation in our tTconvs cells. However, no additional genes belonging to this family were found to be dysregulated in our tTregs or in tTconvs cells.

Therefore, although the function of DNAH8 in regulatory T cells remains to be confirmed and further explored, there are strong evidences that this protein might be playing important roles in regulatory T cells biology.

Another gene uniquely expressed and with unreported functions in Tregs was the Tumour Necrosis Factor Receptor Superfamily member 11A (*TNFRSF11A*). This receptor, also known as RANK (receptor activator of NF-kB), together with its ligand (RANKL), play a critical role in the development and functions of diverse tissues. Namely, RANK-RANKL signalling is best known for their essential role in the regulation of bone

homeostasis through osteoclast differentiation[290,291], though it was originally identified as regulator of dendritic cells function[292,293].

Regarding the thymus, the RANKL-RANK signaling is crucial for the development and maturation of medullary Thymic Epithelial Cells (mTECs)[294,295]. Since mTECs are necessary for the elimination of self-antigen reactive thymocytes, RANKL-RANK signaling is therefore important in the context of autoimmunity. Moreover, interactions between RANKL (expressed in conventional $CD4^+$ T cells) and RANK (expressed by dendritic cells, DCs), promotes the survival of DCs[296], while ensuring T cell priming and activation, thereby enhancing the acquired immune responses[293]. Interestingly, there is a study showing that Tregs are the main source of RANKL during primary tumour growth[297].

Supporting our observations, recent studies have also reported an over expression of *TNFRSF11A* in naturally-derived, comparing to peripheral-derived Tregs in humans[298], as well as in human breast tumour-resident $CD4^+$ Treg cells compared to the respective Tconv cells[280]. However, the existence of RANK protein in Tregs remain speculative and require further confirmation. Nevertheless, it is not surprising if the presence of RANK in Treg cells could represent an additional mechanism to prevent the activation of $CD4^+$ Tconv cells, by competing with the binding sites of this receptor in dendritic cells.

3.4 Identification of transcriptional programs activated in thymic Tregs

Transcription factors are proteins that orchestrate the expression of other genes – and themselves – upon binding directly or indirectly to regulatory regions of the genome. Transcription factors can either be upstream transcriptional regulation or part of the downstream cascade of events triggered by so-called master regulators of differentiation[299,300]. Given the long-term aim of identifying genes and pathways responsible for the differentiation of tTregs in the human thymus, I characterised the transcription factors up-regulated in tTregs, as these are most likely controlling the transcriptional programs that define tTregs instead of tTconvs[301,302].

Genes encoding transcription factors were identified by comparison with the most recently published peer-reviewed catalogue of human transcription factors[236]. This catalogue is a result of a manually examination of 2,765 proteins compiled by combining putative transcription factors from several sources. Each protein was then classified based on the likelihood of each of them to be a transcription factor, its DNA binding mode, and known motifs for each protein along with available DNA-protein structures. The final tally encompassed 1,639 known or likely human transcription factors, which was used to identify the 46 transcription factors up-regulated in tTreg (Annexe 5), relative to the tTconv subset (Figure 3.23)[236].

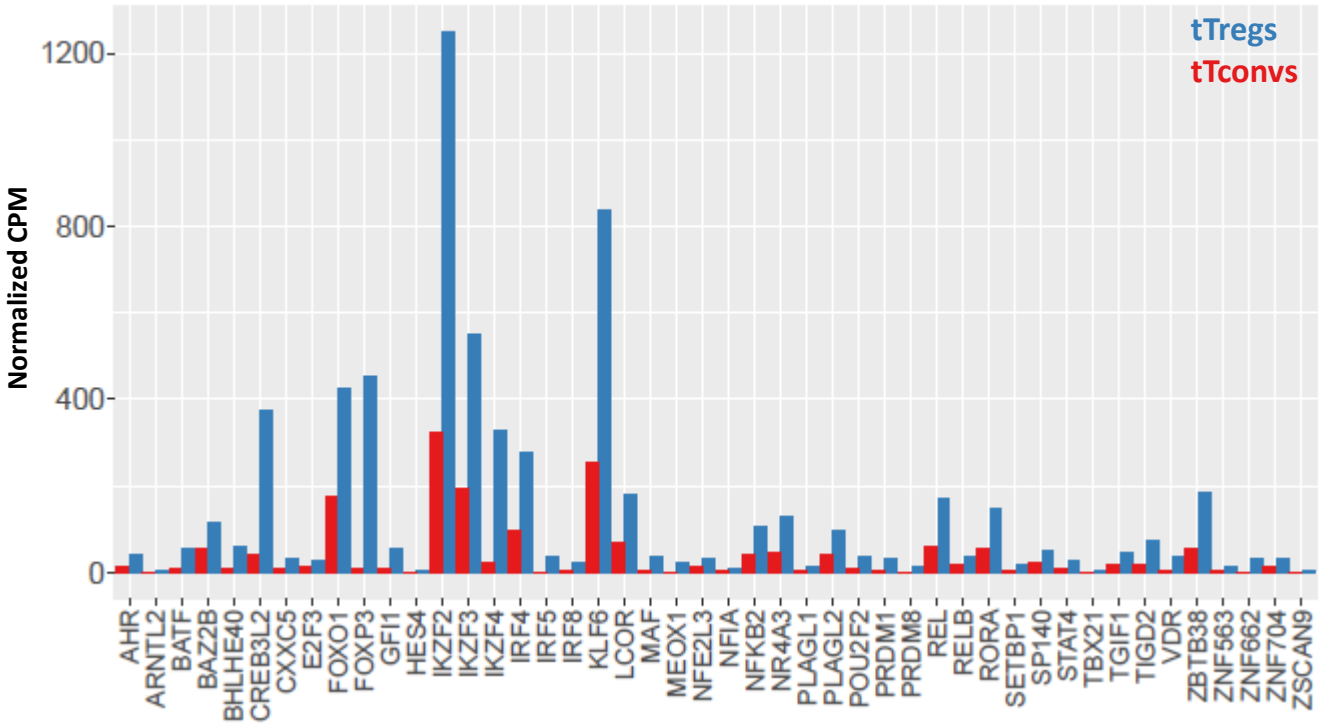


Figure 3.23 Absolute expression levels of Treg up-regulated genes encoding transcription factors. Blue and red bars indicate the absolute expression value, in normalized counts per million, of each gene in tTreg and tTconv subsets, respectively.

As previously observed, *FoxP3*, *IKZF2* and *IKZF4* are TFs up-regulated in tTreg subset. Additional genes encoding for transcription factors that are known to be involved in Treg development were: *FOXO1*; *NR4A3*; and *REL*. FOXO1 binds to *FOXP3* promoter, activating its expression[148,176]; NR4A3 (nuclear receptor subfamily 4, group A, member 3) has been shown to support thymic Treg cell development by 'translating' the strength of TCR signalling[303]; c-REL is the most important transcription factor of the NF-kB family for the induction of FOXP3, by binding to both CNS2 and CNS3 region of this gene, downstream of TCR signalling[151,170,171,264].

Interestingly, and in addition to this canonical NF-kB pathway member, I have also identified TFs that are members of the alternative NF-kB pathway, namely *RELB* and *NFKB2* (NF-kB subunit 2)[263,304]. Accordingly, Grinberg-Bleyer *et al*., has recently demonstrated that this pathway is crucial for the maintenance of Treg homeostasis and function, by preventing excessive RELB activation, since deletion of *NFKB2* can drive uncontrolled Treg activation[305]. Supporting this observation, I have also identified the up-regulation in our tTregs of TFs activators of this alternative pathway compared to tTconvs, namely *TNFRSF4* (OX40)[306], *TNFRSF18* (GITR)[307], and *MAP3K14* (also known as NF-Kappa-Beta-Inducing Kinase, NIK), which is an essential upstream kinase in alternative NF-kB signalling and whose overexpression can lead to increased numbers of Treg cells[306].

Therefore, these analyses suggest that the alternative NF-kB pathway might be active and playing an important role in the differentiation of regulatory T cells.

The *IKZF3* gene-encoding the transcription factor AIOLOS, was also highly expressed in our tTreg cells. Indeed, this gene is known to be involved in the differentiation of FOXP3^{+}Helios^{-} peripheral-induced Tregs[308], as well as in Th17 differentiation by suppressing *IL-2* gene expression[309]. Therefore, AIOLOS might be inhibiting IL-2 expression during thymic-derived Treg cells differentiation, a cytokine that is known to be expressed by and necessary for the differentiation of effector Tconv cells[310].

Another transcription factor identified was the vitamin D_3 ligand-dependent transcription factor vitamin D receptor (VDR). This observation is in accordance with the roles of this transcription factor in inducing *FOXP3* expression and enhance the suppressive activity of Treg cells, by binding to the CNS regions of *FOXP3* gene in response to its ligand[311]. In addition, VDR is important for Treg function, by in promoting the production of TGF-β1 (a Treg-associated anti-inflammatory cytokine)[312] while suppressing Th17 cell proliferation and IL-17 production[313]. There are, however, studies in mice lacking VDR presenting normal numbers of functional splenic and thymic Tregs, without the development and autoimmune disorders, suggesting an indispensable role of VDR for Treg generation[314,315].

I have further identified the over-expression of the ligand-activated transcription factor aryl hydrocarbon receptor (AHR) in our tTregs. This transcription factor has been shown to regulate the differentiation of both regulatory T and T-helper 17 cells in mice[316]. However, the role of AHR in Tregs remains controversial, with conflicting data showing AHR expression in Tregs and either positive or negative regulation of Treg differentiation by this transcription factor[317–319]. Nevertheless, a recent study has shown that AHR may be a marker and a promoter of peripheral induced regulatory T cells in the gut[320].

Amongst the 46 genes identified in Treg population, I identify some of them encoding for transcription factors known to be involved in the differentiation of naïve Tconv cells upon activation into different effector populations, as previously observed in GO of Pathways (Figure 3.23). These effector Tconv cell lineages are functionally heterogeneous and divided according to the transcription factors[321]. I have identified several of these genes over-expressed in Tregs, namely: *TBX21* (T-bet), a T helper type 1 (Th1)-associated transcription factor[322]; the Interferon Regulatory Factor 4 (IRF4*)*, a Th2-associated transcription factor that has also been shown to be important in Th1[323], Th17[324] and in Th9[325] effector function and differentiation; *IRF8*, which is implicated in Th17[326] and in Th9 differentiation[327]; *RORA*, a regulator of Th17 differentiation[328]; *STAT4*, which has an essential role in Th1[329], and appear to be involved in both Th2[330] and Th17 cells[331]; and *BATF*, which is essential for T follicular helper (Tfh) cell generation[332], and seems also to be involved in controlling Th17[333], Th9[334] and Th2[332] cell differentiation.

These observations confirm several studies that have shown association between activated Treg cells and the expression or activation of specific Tconv-associated transcription factors for a more targeted regulation of effector T conv responses. Namely, *TBX21* expression in FOXP3-expressing Treg occurs in response

to Th1 cytokines IFN-γ and IL-27. This expression then leads to the expression of the Th1-associated chemokine receptor CXCR3, a gene that was also found to be upregulated in our tTregs, with accumulation of T-bet positive Treg being seen during type 1 inflammation[335,336]. Furthermore, IRF8 was also recently reported to control Th1-like regulatory T cell function in a T-bet independent manner[337].

Analogously, FOXP3 induces the transcription factor essential for Th2 effector cell differentiation, IRF4, to facilitate efficient suppression the responses mediated by this subset[338]. In addition, the gene-encoding transcription factor *PRDM1* (Blimp1) was also identified in our tTreg cells and seems to act downstream of IRF4 in the differentiation of Th2-like Treg cells by promoting IL-10 expression[339].

Therefore, these observations support our identification of tTreg up-regulated genes-encoding transcription factors associated with effector Tconv cells. However, it's not possible to decipher whether these transcription factors are already being expressed in developing thymocytes, or if it is due to the recirculation of T cells back into the thymus. In fact, it has been shown that mature peripheral T cells can recirculate back to this organ, including activated Tregs, which can further supress the development of their thymic precursors[340,341].

Another identified gene-encoding a member of the IRF family of transcription factors and also uniquely expressed in our tTreg cells was the *IRF5* gene. IRF5 is known in promoting the activation of proinflammatory cytokine genes, such as IL-6, IL-12 and IL-23, as well as effector Th1 and Th17 responses [342,343]. The evidences that IRF5 might have an involvement in the differentiation of thymic-derived Treg cells are very scarce and requires further studies. However, a study has observed an upregulation of IRF5 in tumour-infiltrating Tregs upon blockade of TIM3/PD-1 suppressive signaling pathways, suggesting a role of this transcription factor in regulatory T cells function[344].

I have also observed the over-expression of *BHLHE40* gene in tTregs, which encodes the transcription factor basic helix-loop-helix (bHLH) family member e40 (Figure 3.23). Supporting the role of this transcription factor in Treg cells, Miyazaki *et al.* 2010 observed reduced numbers of Foxp3$^+$ Treg cells in *BHLHE40*-deficient mice[345]. In addition, they shown that *BHLHE40* could induce *IL2Rα* (CD25) gene expression during mice thymocyte development[345]. However, BHLHE40 was recently demonstrated to act as a repressor of IL-10 production during an infection[346,347], an anti-inflammatory cytokine and one of the mechanisms by which Tregs can exert their function[98]. This repression seems to be through the repression of *MAF* (c-MAF), a major regulator of *IL10* gene expression, and a transcription factor that was also up-regulated in our tTreg cells (Figure 3.23)[348].

The *CREB3L2* gene (cyclic AMP-Responsive Element-Binding Protein 3-Like Protein 2) encodes a member of the oasis bZIP transcription factor family[349], and is mainly involved in the differentiation of chondrocytes[350,351]. We identify the over-expression of this gene in our tTreg subset, with an expression value

very similar to the expression of *FOXP3* gene (Figure 3.23). However, as far as I known, this transcription factor has never been found to be over-expressed in thymic-derived Treg cells.

Nevertheless, there are a few studies reporting an involvement of this gene in Tconv cells. Namely, it was observed an up-regulation of this gene in $CD161^+$ Tconvs comparing to $CD161^-$ T conv cells[352], a marker for human IL-17 producing T-cell subsets[353], while other study observed that *CREB3L2* was positively regulated by STAT6[354], a transcription factor that influences naïve Tconvs cell fate decision towards the effector Th2 cell subset[355]. Interestingly it is known that CREB-ATF1 complexes bind to the promoter and CNS2 region in the *FOXP3* gene, activating its expression[149]. Therefore, and considering the similar levels of *CREB3L2* and *FOXP3* expression observed in our tTreg cells, *CREB3L2* might as well have important functions in promoting *FOXP3* expressing during thymic regulatory T cell differentiation.

Another identified transcription factor was ZBTB38 (zinc finger and BTB domain containing 38). As far as I know, the functions of this gene in regulatory T cells are completely unknown, although previous studies have reported the expression of this gene in Treg cells[356,357].

The gene encoding the Kruppel-like transcription factor-6 (KLF6) was the second most expressed in the identified genes-encoding transcription factors (Figure 3.23). The identification of a member of this zing-finger family is not surprising, since this family are well known in regulating critical cellular processes of leukocytes, including development, differentiation, proliferation and function[358,359]. Particularly, KLF10 constitutes an important player in both regulatory T cell differentiation, by targeting *TGF-β1* and *FOXP3* promoters, and suppressive function by increasing the expression of *TGF-β1*[360,361].

However, the involvement of the identified KLF6 transcription factor in regulatory T cells is unknown, although it has been found to be up-regulated in mice $CD4^+CD25^+$ naïve Tregs compared to tumour-isolated Tregs[362]. Furthermore, and similarly to KLF10, *TGF-β1* gene is a well-established target of KLF6[363],one of the inhibitory cytokines used by Tregs in the suppression of effector Tconv cells[98]. However, and to the best of my knowledge, *KLF6* expression in human thymic-derived Treg cells has never been reported and might represent an additional player in the differentiation and function of this subset.

Overall, I have identified the over-expression of a set of transcription factors with previously unreported functions related with human regulatory T cells, namely IRF5, ZBTB38, KLF6 and CREB3L2 transcription factors. These findings suggest the involvement of novel players in the regulation of Treg cell differentiation and function.

CHAPTER 4 - CONCLUSIONS AND FUTURE PERSPECTIVES

In this thesis, I present the first genome-wide expression profile of human regulatory and conventional $CD4^+$ thymocytes, in an effort to reveal previously unknown aspects of thymic regulatory and conventional T cell development. This was only possible because of the proximity of the lab to the clinic, giving me access to fresh human thymic samples which allowed this work.

The aim of this thesis was to characterise and explore the gene expression profile of human thymic Treg cells and, with it, to identify potential novel factors involved in the pathways regulating the definition and commitment of this lineage. Indeed, I have uncovered novel candidate factors involved in the development of regulatory T cells, which expression pattern in these cells now should be experimentally validated and their relevance further investigated to better characterise their involvement of in human Treg biology. Future analysis should also extend the scope of this thesis and make use of the quality of our raw data and sequencing depth to identify of differential isoform expression and the investigation of possible role of alternative splicing in the regulation of Treg identity.

This thesis describes genes that are expressed in Tregs and not in Tconvs and whose functions on Tregs are so far unknown. One of these is the type 8 heavy chain of an axonemal dynein (DNAH8), involved in cilia structures of sperm and respiratory tract. The expression of this gene was already reported in other studies in both mice and human Treg cells. In fact, some authors suggest that this protein, and dyneins in general, might have roles in T cell activation and in the formation of immunological synapses through a modified cilium. In addition, a mutation in this gene was found in a few cases of CVID, together with respiratory failure symptoms. Therefore, these evidences strongly support the possible role of this protein in regulatory T cells biology. Another gene uniquely expressed and with unreported functions in Tregs described in this thesis was the Tumour Necrosis Factor Receptor Superfamily member 11A (*TNFRSF11A*), or RANK. Together with other studies, RANK expression in tTregs might indicate an additional mechanism to prevent the activation of $CD4^+$ Tconv cells, through the competition with the binding sites of this receptor present in dendritic cells.

Although validation is still required, (e.g., by qRT-PCR), this is a promising set of gene expression that could be used as a signature of expression for Tregs in the human thymus. This signature may and should be complemented in the future with similar studies for Tregs in the periphery to obtain a complete signature of the Treg lineage. Currently, we are producing mRNA-seq data from Treg and Tconv subsets in both naïve and memory compartments of the peripheral blood. The comparison of these three compartments, will allow us to understand the dynamics of expression that characterize the changes of the two subsets in each compartment throughout T cell differentiation and the factors that determine their quiescence and life-long maintenance.

An important group of genes up-regulated in Tregs versus Tconvs are Transcription Factors, given they regulatory role in orchestrating the expression of genes. Here I identified the over-expression of transcription factors known to be important in tTreg development, including: FOXP3, FOXO1, IKZF2, IKZF4. In addition, I identify members of the canonical and alternative NF-kB pathway (REL, RELB and NFKB2) suggesting the active stage of this pathway during tTreg development. Interestingly, I have also identified some genes encode for transcription factors known to be involved in the differentiation of different populations of naïve Tconv cells upon activation, namely *TBX21*, *IRF4*, *RORA*, *STAT4*, *BATF* and *IRF8*. These observations might represent recirculating activated Treg cells, which can express Tconv-associated transcription factors for a more targeted regulation its response. Finally, I describe the over-expression of a set of transcription factors with previously unreported functions related with human regulatory T cells, namely IRF5, ZBTB38, KLF6 and CREB3L2, which might represent novel players in the regulation of Treg cell differentiation and function.

One relevant line of research stemming from this thesis could be the identification of the downstream targets of these transcription factors, e.g., those directly regulated by FOXP3. To do this, we can use another next-generation sequencing technology, a Chromatin Immunoprecipitation assay followed by massive parallel sequencing (ChIP-seq) and compare it with our tTconv data set, as a surrogate "FOXP3 knockdown" to better identify the direct targets of FOXP3 in tTregs. In addition, studies on the regulation of transcription should further be complemented with a full assessment of the accessible chromatin landscape in both tTregs and tTconvs by ATAC-seq (Assay for Transposase Accessible Chromatin), for a better characterisation of transcriptional programs in thymic Treg cells.

To conclude, this work presents the first transcriptomic profile of human thymic Treg and Tconv subsets, and contributes to a better understanding of their development, namely by identifying novel genes with unreported functions in these subsets. As described above, the data here presented open several new lines of research aiming to characterise the regulation of Treg lineage in the human thymus.

REFERENCES

1. Spits, H. Development Of αβ T Cells In The Human Thymus. *Nat. Rev. Immunol.* **2,** 760–772 (2002).
2. Muller-Hermelink, H. K., Wilisch, A., Schultz, A. & Marx, A. Characterization of the Human Thymic Microenvironment: Lymphoepithelial Interaction in Normal Thymus and Thymoma. *Arch. Histol. Cytol.* **60,** 9–28 (1997).
3. Anderson, G. & Jenkinson, E. J. Lymphostromal interactions in thymic development and function. *Nat. Rev. Immunol* **1,** 31–40 (2001).
4. Nunes-Cabaço, H. (2010). Development of regulatory T cells in the human thymus: one step beyond (PhD dissertation). Faculdade de medicina da universidade de Lisboa. Retrieved from http://repositorio.ul.pt/handle/10451/2277.
5. Pearse, G. Normal Structure, Function and Histology of the Thymus. *Toxicol. Pathology* **34,** 504–514 (2006).
6. Anderson, G., Lane, P. J. L. & Jenkinson, E. J. Generating intrathymic microenvironments to establish T - cell tolerance. *Nat. Rev. Immunol.* **7,** 954–963 (2007).
7. Takahama, Y. Journey through the thymus: stromal guides for T-cell development and selection. *Nat. Rev. Immunol.* **6,** 127–135 (2006).
8. Klein, L., Hinterberger, M., Wirnsberger, G. & Kyewski, B. Antigen presentation in the thymus for positive selection and central tolerance induction. *Nat. Rev. Immunol.* **9,** 833–844 (2009).
9. Payne, K. J. & Crooks, G. M. Human hematopoietic lineage commitment. *Immunol. Rev.* **187,** 48–64 (2002).
10. Terstappen, L. W. M. M., Huang, S., Safford, M., Lansdorp, P. M. & Loken, M. R. Sequential Generations of Hematopoietic Colonies Derived From Single Nonlineage-Committed CD34+CD38- Progenitor Cells. *Blood* **77,** 1218–1227 (1991).
11. Baum, C. M., Weissman, I. L., Tsukamoto, A. N. N. S., Buckle, A. & Peault, B. Isolation of a candidate human hematopoietic stem-cell population. *Proc. Natl, Acad. Sci. USA* **89,** 2804–2808 (1992).
12. Spits, H. *et al.* Early stages in the development of human T, natural killer and thymic dendritic cells. *Immunol. Rev.* **165,** 75–86 (1998).
13. Blom, B. & Spits, H. Development of Human Lymphoid Cells. *Annu. Rev. Immunol.* **24,** 287–320 (2006).
14. Staal, F. J. T., Weerkamp, F., Langerak, A. W., Hendriks, R. W. & Clevers, H. C. Transcriptional Control of T Lymphocyte Differentiation. *Stem Cells* **19,** 165–179 (2001).
15. Dik, W. A. *et al.* New insights on human T cell development by quantitative T cell receptor gene rearrangement studies and gene expression profiling. *J. Exp. Med.* **201,** 1715–1723 (2005).
16. Res, B. P. *et al.* CD34+CD38dim Cells in the Human Thymus Can Differentiate Into T, Natural Killer, and Dendritic Cells But Are Distinct From Pluripotent Stem Cells. *Blood* **87,** 5196–5206 (1996).
17. Blom, B., Res, P., Weijer, K. & Spits, H. Prethymic CD34+ Progenitors Capable of Developing into T Cells Are Not Committed to the T Cell Lineage. *J. Immunol.* **38,** 3571–3577 (1997).
18. Spits, H., Couwenberg, F., Bakker, A. Q., Weijer, K. & Uittenbogaart, C. H. Id2 and Id3 Inhibit Development of CD34+ Stem Cells into Predendritic Cell (Pre-DC)2 but Not into Pre-DC1: Evidence for a Lymphoid Origin of Pre-DC2. *J. Exp. Med.* **192,** 1775–1783 (2000).
19. Blom, B. B. *et al.* TCR Gene Rearrangements and Expression of the Pre-T Cell Receptor Complex During Human T-Cell Differentiation. *Blood* **93,** 3033–3044 (1999).
20. Van De Walle, I., Davids, K. & Taghon, T. Characterization and Isolation of Human T Cell Progenitors. *Methods Mol*

Biol **1323,** 221–237 (2016).

21. Andaloussi, A. El *et al.* Hedgehog signaling controls thymocyte progenitor homeostasis and differentiation in the thymus. *Nat Immunol* **7,** 418–426 (2006).

22. Petrie, H. T. & Z, J. C. Zoned Out: Functional Mapping of Stromal Signaling Microenvironments in the Thymus. *Annu. Rev. Immunol.* **25,** 649–679 (2007).

23. Artavanis-Tsakonas, S., Rand, M. D. & Lake, R. J. Notch Signaling: Cell Fate Control and Signal Integration in Development. *Science (80-.).* **284,** 770–776 (1999).

24. Allman, D., Punt, J. A., Izon, D. J., Aster, J. C. & Pear, W. S. An Invitation to T and More : Notch Signaling in Lymphopoiesis. *Cell* **109,** 1–11 (2002).

25. Radtke, F. *et al.* Deficient T Cell Fate Specification in Mice with an Induced Inactivation of Notch1. *Immunity* **10,** 547–558 (1999).

26. Hozumi, K. *et al.* Delta-like 4 is indispensable in thymic environment specific for T cell development. *J. Exp. Med.* **205,** 2507–2513 (2008).

27. Motte-mohs, R. N. La, Herer, E. & Zu, J. C. Induction of T-cell development from human cord blood hematopoietic stem cells by Delta-like 1 in vitro. *Blood* **105,** 1431–1440 (2005).

28. Smedt, M. De, Hoebeke, I. & Plum, J. Human bone marrow CD34+ progenitor cells mature to T cells on OP9-DL1 stromal cell line without thymus microenvironment. *Blood Cells, Mol. Dis.* **33,** 227–232 (2004).

29. Smedt, M. D. microenvironment, Hoebeke, I., Reynvoet, K., Leclercq, G. & Plum, J. Different thresholds of Notch signaling bias human precursor cells toward B-, NK-, monocytic/dendritic-, or T-cell lineage in thymus microenvironment. *Blood* **106,** 3498–3507 (2017).

30. Dontje, W. *et al.* Delta-like1 – induced Notch1 signaling regulates the human plasmacytoid dendritic cell versus T-cell lineage decision through control of GATA-3 and Spi-B. *Blood* **107,** 2446–2453 (2005).

31. Lind, E. F., Prockop, S. E., Porritt, H. E. & Petrie, H. T. Mapping Precursor Movement through the Postnatal Thymus Reveals Specific Microenvironments Supporting Defined Stages of Early Lymphoid Development. *J. Exp. Med.* **194,** 127–134 (2001).

32. Jaleco, A. C. *et al.* Differential Effects of Notch Ligands Delta-1 and Jagged-1 in Human Lymphoid Differentiation. *J. Exp. Med.* **194,** 991–1001 (2001).

33. Han, W., Ye, Q. & Moore, M. A. S. A soluble form of human Delta-like-1 inhibits differentiation of hematopoietic progenitor cells. *Blood* **95,** 1616–1625 (2000).

34. Plum, B. J., Smedt, M. De, Leclercq, G., Verhasselt, B. & Vandekerckhove, B. Interleukin-7 Is a Critical Growth Factor in Early Human T-cell Development. *Blood* **88,** 4239–4245 (1996).

35. Puel, A., Ziegler, S. F., Buckley, R. H. & Leonard, W. J. Defective IL7R expression in T- B+ NK+ severe combined immunodeficiency. *Nat. Genet.* **20,** 394–397 (1998).

36. Giliani, S. *et al.* Interleukin-7 receptor a (IL-7R a) deficiency : cellular and molecular bases . Analysis of clinical , immunological , and molecular features in 16 novel patients. *Immunol. Rev.* **203,** 110–126 (2005).

37. Noguchi, M. *et al.* Interleukin-2 Receptor gamma Chain Mutation Results in X-Linked Severe Combined Immunodeficiency in Humans. *Cell* **73,** 147–157 (1993).

38. Russell, S. M. *et al.* Interaction of IL-2Rbeta and Yc Chains with Jak1 and Jak3 : Implications for XSCID and XCID. *Science (80-.).* **266,** 1042–1045 (1994).

39. Macchi, P. *et al.* Mutations of Jak-3 gene in patients with autosomal severe combined immune deficiency (SCID). *Nature* **377,** 65–68 (1995).

40. Russell, S. M. *et al.* Mutation of Jak3 in a Patient with SCID: Essential Role of Jak3 in Lymphoid Development. *Science (80-.).* **270,** 797–799 (1995).

41. Boudil, A. *et al.* IL-7 coordinates proliferation, differentiation and Tcra recombination during thymocyte beta-selection. *Nat Immunol* **16,** 397–405 (2015).

42. Tsuda, S., S Rieke, Hashimoto, Y., Nakauchi, H. & Takahama, Y. Il-7 supports D-J but not V-DJ rearrangement of TCR-beta gene in fetal liver progenitor cells. *J. Immunol.* **156,** 3233–3242 (1996).

43. Galy, B. A. *et al.* Precursors of CD3+ CD4+ CD8+ Cells in the Human Thymus Are Defined by Expression of CD34 . Delineation of Early Events in Human Thymic Development. *J. Exp. Med.* **178,** 391–401 (1993).

44. Kraft, B. D. L., Weissman, I. L. & Waller, E. K. Differentiation of CD3- 4- 8- Human Fetal Thymocytes In Vivo: Characterization of a CD3- 4+ 8- Intermediate. *J. Exp. Med.* **178,** 265–277 (1993).

45. Alvarez-Vallina, L., González, A., Gambón, F., Kreisler, M. & Díaz-Espada, F. Delimitation of the proliferative stages in the human thymus indicates that cell expansion occurs before the expression of CD3 (T cell receptor). *J. Immunol.* **150,** 8–16 (2017).

46. Verschuren, M. C. M., Blom, B., Bogers, A. J. J. C., Spits, H. & Dongen, J. J. M. Van. PJA-BP expression and TCR δ deletion during human T cell differentiation. *Int. Immunol.* **10,** 1873–1880 (1998).

47. van Dongen, J. J. M., Staal, F. J. & Langerak, A. W. Developmental and functional biology of T lymphocytes. Non-Hodgkin's Lymphoma's. in *2004* (ed. P.M. Mauch, J.O. Armitage, B. Coiffier, R. Dalla-Favera, and N.L. Harris, editors, Lippincott, W. & W.) 787–808

48. Wilson, R. K., Lai, E., Concannon, P., Barth, R. K. & Hood, L. E. Structure, Organization and Polymorphism of Murine and Human T-Cell Receptor alfa and beta Chain Gene Families. *Immunol. Rev.* **101,** 149–172 (1988).

49. Arstila, T. P. *et al.* A Direct Estimate of the Human alfabeta T Cell Receptor Diversity. *Science (80-.).* **286,** 958–961 (1999).

50. Kreslavsky, T., Garbe, A. I., Krueger, A. & Boehmer, H. Von. T cell receptor - instructed alfabeta versus gammadelta lineage commitment revealed by single-cell analysis. *J. Exp. Med.* **205,** 1173–1186 (2008).

51. Petrie, H. T., Scollay, R. & Shortman, K. Commitment to the T cell receptor-alfabeta or -gammadelta lineages can occur just prior to the onset of CD4 and CD8 expression among immature thymocytes. *Eur. J. Immunol.* **22,** 2185–2188 (1992).

52. Hayes, S. M., Shores, E. W. & Love, P. E. An architectural perspective on signaling by the pre- , alfabeta and gammadelta T cell receptors. *Immunol. Rev.* **191,** 28–37 (2003).

53. Von Boehmer, H. Unique features of the pre-T-cell receptor alpha-chain: not just a surrogate. *Nat. Rev. Immunol.* **5,** 571–577 (2005).

54. Groettrup, M. *et al.* A Novel Disulfide-Linked Heterodimer on Pre-T Cells Consists of the T Cell Receptor beta Chain and a 33 kd Glycoprotein. *Cell* **75,** 283–294 (1993).

55. Boehmer, H. Von & Hans, J. Structure And Function of the Pre-T Cell Receptor. *Annu. Rev. Immunol.* **15,** 433–452 (1997).

56. Ramiro, A. R., Trigueros, C., Márquez, C., Millán, J. L. S. & Toribio, M. L. Regulation of Pre-T Cell Receptor (pTalfa-TCRbeta) Gene Expression during Human Thymic Development. *J. Exp. Med.* **184,** 519–530 (1996).

57. Trigueros, C. *et al.* Identification of a Late Stage of Small Noncycling pTalfa- Pre-T Cells as Immediate Precursors of T Cell Receptor alfa/beta+ Thymocytes. *J. Exp. Med.* **188,** 1401–1412 (1998).

58. Hori, T. *et al.* Identification of a novel human thymocyte subset with a phenotype of CD3-CD4+CD8 alpha+ beta-1. Possible progeny of the CD3- CD4- CD8- subset. *J. Immunol.* **146,** 4078–4084 (1991).

59. Carrasco, Y. R., Trigueros, C., Ramiro, A. R., Yébenes, V. G. de & Toribio, M. L. beta-Selection Is Associated With the Onset of CD8beta Chain Expression on CD4+CD8alfaalfa+ Pre-T Cells During Human Intrathymic Development. *Immunobiology* **94,** 3491–3499 (1999).

60. Taghon, T. *et al.* Notch signaling is required for proliferation but not for differentiation at a well-defined beta-selection checkpoint during human T-cell development. *Blood* **113,** 3254–3264 (2009).

61. Paul, W. E. *Fundamental Immunology, 6th Edition*. (Lippincott Williams and Wilkins, 2008).

62. Verschuren, M. C. *et al.* Preferential rearrangements of the T cell receptor-delta-deleting elements in human T cells. *J. Immunol.* **158,** 1208–1216 (1997).

63. Starr, T. K., Jameson, S. C. & Hogquist, K. A. Positive and Negative Selection of T cells. *Annu. Rev. Immunol.* **21,** 139–176 (2003).

64. Vrisekoop, N., Monteiro, P., Mandl, J. N. & Germain, R. N. Revisiting Thymic Positive Selection and the Mature T Cell Repertoire for Antigen. *Immunity* **41,** 181–190 (2014).

65. Ashton-Rickardt, P. G. *et al.* Evidence for a Differential Avidity Model of T Cell Selection in the Thymus. *Cell* **76,** 651–663 (1994).

66. Alam, S. M. *et al.* T cell receptor affinity and thymocyte positive selection. *Nature* **381,** 616–620 (1996).

67. Bill, J. & Palmer, E. Positive selection of CD4+ Tcells mediated by MHC class II-bearing stromal cell in the thymic cortex. *Nature* **341,** 649–651 (1989).

68. Anderson, B. G., Owen, J. J. T., Moore, N. C. & Jenkinson, E. J. Thymic Epithelial Cells Provide Unique Signals for Positive Selection of CD4 + CD8 + Thymocytes In Vitro. *J. Exp. Med.* **179,** 2027–2031 (1994).

69. Benoist, C. & Mathis, D. Positive Selection of the T Cell Repertoire: Where and When Does It Occur ? *Cell* **58,** 1027–1033 (1999).

70. Goldman, K. P., Park, C., Kim, M., Matzinger, P. & Anderson, C. C. Thymic cortical epithelium induces self tolerance. *Eur. J. Immunol.* **35,** 709–717 (2005).

71. Hare, K. J., Jenkinson, E. J. & Anderson, G. Specialisation of thymic epithelial cells for positive selection of CD4+8+ thymocytes. *Cell Mol Biol* **47,** 119–127 (2001).

72. Fujimoto, Y. *et al.* CD83 Expression Influences CD4+ T Cell Development in the Thymus. *Cell* **108,** 755–767 (2002).

73. Martinic, M. M. *et al.* Efficient T cell repertoire selection in tetraparental chimeric mice independent of thymic epithelial MHC. *Proc. Natl, Acad. Sci. USA* **100,** 1861–1866 (2002).

74. Choi, E. Y. *et al.* Thymocyte-Thymocyte Interaction for Efficient Positive Selection and Maturation of CD4 T Cells. *Immunity* **23,** 387–396 (2005).

75. Li, W. *et al.* An Alternate Pathway for CD4 T Cell Development: Thymocyte-Expressed MHC Class II Selects a Distinct T Cell Population. *Immunity* **23,** 375–386 (2005).

76. Mathis, D. & Benoist, C. Aire. *Annu. Rev. Immunol.* **27,** 287–312 (2009).

77. Nagamine, K. *et al.* Positional cloning of the APECED gene. *Nat. Genet.* **17,** 393–398 (1997).

78. Björses, P., Aaltonen, J., Horelli-kuitunen, N., Yaspo, M. & Peltonen, L. Gene defect behind APECED : a new clue to autoimmunity. *Hum. Mol. Genet.* **7,** 1547–1553 (1998).

79. Singer, A., Adoro, S. & Park, J. Lineage fate and intense debate: myths, models and mechanisms of CD4/CD8 lineage choice. *Nat. Rev. Immunol.* **8,** 788–801 (2008).

80. Boehmer, H. von *et al.* The Expression of CD4 and CD8 Accessory Molecules on Mature T Cells is not Random but Correlates with the Specificity of the alfabeta Receptor for Antigen. *Immunol. Rev.* 143–151 (1989).

81. Reith, W. & Mach, B. The Bare Lymphocyte Syndrome and the Regulation of MHC Expression. *Annu. Rev. Immunol.* **19,** 331–373 (2001).

82. Brugnera, E. *et al.* Coreceptor Reversal in the Thymus: Signaled CD4+8+ Thymocytes Initially Terminate CD8 Transcription Even When Differentiating into CD8+ T Cells. *Immunity* **13,** 59–71 (2000).

83. Singer, A. New perspectives on a developmental dilemma: the kinetic signaling model and the importance of signal duration for the CD4/CD8 lineage decision. *Curr. Opin. Immunol.* **14,** 207–215 (2002).

84. Singer, A. & Bosselut, R. CD4/CD8 Coreceptors in Thymocyte Development, Selection, and Lineage Commitment:

Analysis of the CD4/CD8 Lineage Decision. *Adv. Immunol.* **83,** 91–131 (2004).

85. Yu, Q., Erman, B., Bhandoola, A., Sharrow, S. O. & Singer, A. In Vitro Evidence That Cytokine Receptor Signals AreRequired for Differentiation of Double Positive Thymocytes into Functionally Mature CD8+ T Cells. *J. Exp. Med.* **197,** 476–487 (2003).

86. He, X. *et al.* The zinc finger transcription factor Th-POK regulates CD4 versus CD8 T-cell lineage commitment. *Nature* **433,** 826–833 (2005).

87. Kappes, D. J., He, X. & He, X. Role of the transcription factor Th-POK in CD4:CD8 lineage commitment. *Immunol. Rev.* **209,** 237–252 (2006).

88. Woolf, E. *et al.* Runx3 and Runx1 are required for CD8 T cell development during thymopoiesis. *Proc. Natl. Acad. Sci. USA* **100,** 7731–7736 (2003).

89. Sato, T. *et al.* Dual Functions of Runx Proteins for Reactivating CD8 and Silencing CD4 at the Commitment Process into CD8 Thymocytes. *Immunity* **22,** 317–328 (2005).

90. Setoguchi, R. *et al.* Repression of the Transcription Factor Th-POK by Runx Complexes in Cytotoxic T Cell Development. *Science (80-.).* **319,** 822–825 (2008).

91. Egawa, T. & Littman, D. R. ThPOK acts late in specification of the helper T cell lineage and suppresses Runx-mediated commitment to the cytotoxic T cell lineage. *Nat. Immunol.* **9,** 1131–1139 (2008).

92. Egawa, T. & Taniuchi, I. Blood Cells , Molecules , and Diseases Antagonistic interplay between ThPOK and Runx in lineage choice of thymocytes. *Blood Cells, Mol. Dis.* **43,** 27–29 (2009).

93. Aliahmad, P. & Kaye, J. Development of all CD4 T lineages requires nuclear factor TOX. *J. Exp. Med.* **205,** 245–256 (2008).

94. Hernández-hoyos, G., Anderson, M. K., Wang, C., Rothenberg, E. V & Alberola-ila, J. GATA-3 Expression Is Controlled by TCR Signals and Regulates CD4/CD8 Differentiation. *Immunity* **19,** 83–94 (2003).

95. Sakaguchi, S., Wing, K. & Miyara, M. Regulatory T cells – a brief history and perspective. *Eur. J. Immunol.* **37,** 116–123 (2007).

96. Sakaguchi, S. & Sakaguchi, N. Regulatory T Cells in Immunologic Self-Tolerance and Autoimmune Disease. *Int. Rev. Immunol.* **24,** 211–226 (2005).

97. Singh, B. *et al.* Control of intestinal inflammation by regulatory T cells. *Immunol. Rev.* **182,** 190–200 (2001).

98. Vignali, D. A. A., Collison, L. W. & Workman, C. J. How regulatory T cells work. *Nat. Rev. Immunol.* **8,** 523–532 (2008).

99. Sakaguchi, S., Yamaguchi, T., Nomura, T. & Ono, M. Regulatory T Cells and Immune Tolerance. *Cell* **133,** 775–787 (2008).

100. Jonuleit, H. & Schmitt, E. The Regulatory T Cell Family: Distinct Subsets and their Interrelations. *J. Immunol.* **171,** 6323–6327 (2003).

101. Sakaguchi, S., Sakaguchi, N., Asano, M., Itoh, M. & Toda, M. Immunologic Self-Tolerance Maintained by Activated T Cells Expressing IL-2 Receptor alpha-Chains (CD25). *J. Immunol.* **155,** 1151–1164 (1995).

102. Sakaguchi, S. Naturally Arising CD4+ Regulatory T Cells for Immunologic Self-Tolerance and Negative Control of Immune Responses. *Annu. Rev. Immunol.* **22,** 531–562 (2004).

103. Nedoszytko, B. *et al.* The role of regulatory T cells and genes involved in their differentiation in pathogenesis of selected inflammatory and neoplastic skin diseases. Part I: Treg properties and functions. *Adv. Dermatol. Allergol* **34,** 285–294 (2017).

104. Bonomo, A. *et al.* Pathogenesis of post-thymectomy autoimmunity. Role of syngeneic MLR-reactive T cells. *J. Immunol.* **154,** 6602–6611 (1995).

105. Nishizuka, Y. & Sakakura, T. Thymus and Reproduction; Sex-linked Dysgenesia of the Gonad after Neonatal

Thymectomy in Mice. *Science (80-.).* **166,** 753–755 (1969).

106. Sakaguchi, S., Takahashi, T. & Nishizuka, Y. Study on Celullar Events in Postthymectomy Autoimmune Oophoritis in Mice. *J. Exp. Med.* **156,** 1565–1576 (1982).

107. Asano, M., Toda, M., Sakaguchi, N. & Sakaguchi, S. Autoimmune Disease as a Consequence of Developmental Abnormality of a T Cell Subpopulation. *J. Exp. Med.* **184,** 387–396 (1996).

108. Sakaguchi, S., Sakaguchi, N., Asano, M., Itoh, M. & Toda, M. Immunologic self-tolerance maintained by activated T cells expressing IL-2 receptor alpha-chains (CD25). Breakdown of a single mechanism of self-tolerance causes various autoimmune diseases. *J. Immunol.* **155,** 1151–1164 (1995).

109. Jonuleit, B. H. *et al.* Identification and Functional Characterization of Human CD4+ CD25+ T Cells with Regulatory Properties Isolated from Peripheral Blood. *J. Exp. Med.* **193,** 1285–1294 (2001).

110. Levings, B. M. K., Sangregorio, R. & Roncarolo, M. Human CD25+ CD4+ T Regulatory Cells Suppress Naive and Memory T Cell Proliferation and Can Be Expanded In Vitro without Loss of Function. *J. Exp. Med.* **193,** (2001).

111. Baecher-allan, C. *et al.* CD4 + CD25 high Regulatory Cells in Human Peripheral Blood. *J. Immunol.* **167,** 1245–1253 (2001).

112. Baecher-allan, C., Viglietta, V. & Hafler, D. A. Human CD4 + CD25 + regulatory T cells. *Semin. Immunol.* **16,** 89–97 (2004).

113. Wing, K. *et al.* CTLA-4 Control over Foxp3+ Regulatory T Cell Function. *Science (80-.).* **322,** 271–275 (2008).

114. Sakaguchi, S., Wing, K., Onishi, Y., Prieto-martin, P. & Yamaguchi, T. Regulatory T cells: how do they suppress immune responses ? *Int. Immunol.* **21,** 1105–1111 (2009).

115. Yagi, H. *et al.* Crucial role of FOXP3 in the development and function of human CD25+ CD4+ regulatory T cells. *Int. Immunol.* **16,** 1643–1656 (2004).

116. Shimizu, J., Yamazaki, S., Takahashi, T., Ishida, Y. & Sakaguchi, S. Stimulation of CD25+ CD4+ regulatory T cells through GITR breaks immunological self-tolerance. *Nat. Immunol.* **3,** 135–142 (2002).

117. Baecher-allan, C., Wolf, E. & Hafler, D. A. MHC Class II Expression Identifies Functionally Distinct Human Regulatory T Cells. *J. Immunol.* **176,** 4622–4631 (2006).

118. Seddiki, N. *et al.* Expression of interleukin (IL)-2 and IL-7 receptors discriminates between human regulatory and activated T cells. *J. Exp. Med.* **203,** 1693–1700 (2010).

119. Liu, W. *et al.* CD127 expression inversely correlates with FoxP3 and suppressive function of human CD4+ T reg cells. *J. Exp. Med.* **203,** 1701–1711 (2006).

120. Mazzucchelli, R. & Durum, S. K. Interleukin-7 receptor expression: intelligent design. *Nat. Rev. Immunol* **7,** 144–154 (2007).

121. Antonio, G. *et al.* Thymic versus induced regulatoryT cells – who regulates the regulators ? *Front. Immunol.* **4,** 1–22 (2013).

122. Sakaguchi, S., Miyara, M., Costantino, C. M. & Hafler, D. A. FOXP3+ regulatory T cells in the human immune system. *Nat. Rev. Immunol.* **10,** 490–500 (2010).

123. Zheng, Y. & Rudensky, A. Y. Foxp3 in control of the regulatory T cell lineage. *Nat. Rev. Immunol.* **8,** 1–6 (2007).

124. Fontenot, J. D., Gavin, M. A. & Rudensky, A. Y. Foxp3 programs the development and function of CD4+ CD25+ regulatory T cells. *Nat. Immunol.* **4,** 330–336 (2003).

125. Coffer, P. J. & Burgering, B. M. T. FORKHEAD-BOX TRANSCRIPTION FACTORS AND THEIR ROLE IN THE IMMUNE SYSTEM. *Nat. Rev. Immunol.* **4,** 889–899 (2004).

126. Brunkow, M. E. *et al.* Disruption of a new forkhead/winged-helix protein, scurfin, results in the fatal lymphoproliferative disorder of the scurfy mouse. *Nat. Genet.* **27,** 68–73 (2001).

127. Bennett, C. L. *et al.* The immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome (IPEX) is caused by mutations of FOXP3. *Nat. Genet.* **27,** 20–21 (2001).

128. Wildin, R. S. *et al.* X-linked neonatal diabetes mellitus, enteropathy and endocrinopathy syndrome is the human equivalent of mouse scurfy. *Nat. Genet.* **27,** 18–20 (2001).

129. Ochs, H. D., Ziegler, S. F. & Torgerson, T. R. FOXP3 acts as a rheostat of the immune response. *Immunol. Rev.* **203,** 156–164 (2005).

130. Godfrey, V. L., Rouse, B. T. & Wilkinsont, J. E. Transplantation of T Cell-Mediated, Lymphoreticular Disease from the Scurfy (sf) Mouse. *Am. J. Pathol.* **145,** 281–286 (2009).

131. Fontenot, J. D. *et al.* Regulatory T Cell Lineage Specification by the Forkhead Transcription Factor Foxp3. *Immunity* **22,** 329–341 (2005).

132. Gavin, M. A. *et al.* Foxp3-dependent programme of regulatory T-cell differentiation. *Nature* **445,** 771–775 (2007).

133. Ramsdell, F. & Ziegler, S. F. FOXP3 and scurfy: how it all began. *Nat. Rev. Immunol.* **14,** 343–349 (2014).

134. Zheng, Y. *et al.* Genome-wide analysis of Foxp3 target genes in developing and mature regulatory T cells. *Nature* **445,** 1–5 (2007).

135. Samstein, R. M. *et al.* Foxp3 Exploits a Pre-Existent Enhancer Landscape for Regulatory T Cell Lineage Specification. *Cell* **151,** 153–166 (2012).

136. Marson, A. *et al.* Foxp3 occupancy and regulation of key target genes during T-cell stimulation. *Nature* **445,** 931–935 (2007).

137. Rudra, D. *et al.* Transcription factor Foxp3 and its protein partners form a complex regulatory network. *Nat Immunol* **13,** 1010–1019 (2012).

138. Hori, S. The Foxp3 interactome: a network perspective of Treg cells. *Nat. Immunol.* **13,** 943–945 (2012).

139. Wu, Y. *et al.* FOXP3 Controls Regulatory T Cell Function through Cooperation with NFAT. *Cell* **126,** 375–387 (2006).

140. Ono, M. *et al.* Foxp3 controls regulatory T-cell function by interacting with AML1/Runx1. *Nature* **446,** 685–689 (2007).

141. Zhang, F., Meng, G. & Strober, W. Interactions among the transcription factors Runx1, RORct and Foxp3 regulate the differentiation of interleukin 17-producing T cells. *Nat. Immunol.* **9,** 1297–1306 (2008).

142. Loizou, L., Andersen, K. G. & Betz, A. G. Foxp3 Interacts with c-Rel to Mediate NF-kB Repression. *PLoS One* **6,** 1–7 (2011).

143. Lu, L., Barbi, J. & Pan, F. The regulation of immune tolerance by FOXP3. *Nat. Rev. Immunol.* **17,** 703–717 (2017).

144. Li, B. & Greene, M. I. FOXP3 actively represses transcription by recruiting the HAT/HDAC complex. *Cell Cycle* **6,** 1432–1436 (2007).

145. Dhuban, K. Bin *et al.* Suppression by human FOXP3+ regulatory T cells requires FOXP3-TIP60 interactions. *Autoimmunity* **2,** 1–13 (2017).

146. Chen, C., Rowell, E. A., Thomas, R. M., Hancock, W. W. & Wells, A. D. Transcriptional Regulation by Foxp3 Is Associated with Direct Promoter Occupancy and Modulation of Histone Acetylation. *J. Biol. Chem.* **281,** 36828–36834 (2006).

147. Josefowicz, S. Z., Lu, L. & Rudensky, A. Y. Regulatory T Cells : Mechanisms of Differentiation and Function. *Annu. Rev. Immunol.* **30,** 531–564 (2012).

148. Ouyang, W. *et al.* Foxo proteins cooperatively control the differentiation of Foxp3+ regulatory T cells. *Nat. Immunol.* **11,** 618–627 (2010).

149. Kim, H.-P. & Leonard, W. J. CREB/ATF-dependent T cell receptor–induced *FoxP3* gene expression: a role for DNA methylation. *J. Exp. Med.* **204,** 1543–1551 (2007).

150. Toker, A. *et al.* Active Demethylation of the Foxp3 Locus Leads to the Generation of Stable Regulatory T Cells within the Thymus. *J. Immunol.* **190,** 3180–3188 (2013).

151. Zheng, Y. *et al.* Role of conserved non-coding DNA elements in the Foxp3 gene in regulatory T-cell fate. *Nature* **463,** 808–812 (2010).

152. Pillai, V. & Karandikar, N. J. Human Regulatory T Cells : A Unique , Stable Thymic Subset or a Reversible Peripheral State of Differentiation? *Immunol. Lett.* **114,** 9–15 (2007).

153. Schlenner, S. M., Weigmann, B., Ruan, Q., Chen, Y. & von Boehmer, H. Smad3 binding to the foxp3 enhancer is dispensable for the development of regulatory T cells with the exception of the gut. *J. Exp. Med.* **209,** 1529–1535 (2012).

154. Bettini, M. L. & Vignali, D. A. A. Development of Thymically-Derived Natural Regulatory T Cells. *Ann N Y Acad Sci* **3,** 1–12 (2010).

155. Lafaille, M. A. C. De & Lafaille, J. J. Natural and Adaptive Foxp3+ Regulatory T Cells: More of the Same or a Division of Labor ? *Immunity* **30,** 626–635 (2009).

156. Kitagawa, Y. *et al.* Guidance of regulatory T cell development by Satb1-dependent super-enhancer establishment. *Nat. Immunol.* **18,** 173–183 (2017).

157. Alvarez, J. D. *et al.* The MAR-binding protein SATB1 orchestrates temporal and spatial expression of multiple genes during T-cell development. *Genes Dev.* **14,** 521–535 (2000).

158. Lee, W. & Lee, G. R. Transcriptional regulation and development of regulatory T cells. *Exp. Mol. Med.* **50,** 1–10 (2018).

159. Hori, S. Control of Regulatory T Cell Development by the Transcription Factor Foxp3. *Science (80-.).* **299,** 1057–1061 (2003).

160. Schmidt, A., Oberle, N. & Krammer, P. H. Molecular mechanisms ofTreg-mediatedT cell suppression. *Front. Immunol.* **3,** 1–20 (2012).

161. Lin, W. *et al.* Regulatory T cell development in the absence of functional Foxp3. *Nat. Immunol.* **8,** 359–368 (2007).

162. Ohkura, N. *et al.* T Cell Receptor Stimulation-Induced Epigenetic Changes and Foxp3 Expression Are Independent and Complementary Events Required for Treg Cell Development. *Immunity* **37,** 785–799 (2012).

163. Darrasse-je, G., Marodon, G., Salomon, L., Catala, M. & Klatzmann, D. Ontogeny of CD4+CD25+ regulatory/suppressor T cells in human fetuses. *Blood* **105,** 4715–4722 (2005).

164. Tuovinen, H., Pekkarinen, P. T., Rossi, L. H., Mattila, I. & Arstila, T. P. The FOXP3+ subset of human CD4+ CD8+ thymocytes is immature and subject to intrathymic selection. *Immunol. Cell Biol.* **86,** 523–529 (2008).

165. Nunes-Cabaço, H., Caramalho, Í., Sepúlveda, N. & Sousa, A. E. Differentiation of human thymic regulatory T cells at the double positive stage. *Eur. J. Immunol.* **41,** 3604–3614 (2011).

166. Romagnoli, P., Hudrisier, D. & Van, J. P. M. Preferential Recognition of Self Antigens Despite Normal Thymic Deletion of CD4+ CD25+ Regulatory T Cells. *J. Immunol.* **168,** 1644–1648 (2002).

167. Tanaka, S. *et al.* Graded Attenuation of TCR Signaling Elicits Distinct Autoimmune Diseases by Altering Thymic T Cell Selection and Regulatory T Cell Function. *J. Immunol.* **185,** 2295–2305 (2010).

168. Koonpaew, S., Shen, S., Flowers, L. & Zhang, W. LAT-mediated signaling in CD4+CD25+ regulatory T cell development. *J. Exp. Med.* **203,** 119–129 (2006).

169. Chen, X. *et al.* Preferential Development of CD4 and CD8 T Regulatory Cells in RasGRP1-Deficient Mice. *J. Immunol.* **180,** 5973–5982 (2008).

170. Long, M., Park, S., Strickland, I., Hayden, M. S. & Ghosh, S. Nuclear Factor-kB Modulates Regulatory T Cell Development by Directly Regulating Expression of Foxp3 Transcription Factor. *Immunity* **31,** 921–931 (2009).

171. Ruan, Q. *et al.* Development of Foxp3+ Regulatory T Cells Is Driven by the c-Rel Enhanceosome. *Immunity* **31,** 932–

940 (2009).

172. Bopp, T. *et al.* NFATc2 and NFATc3 transcription factors play a crucial role in suppression of CD4+ T lymphocytes by CD4+ CD25+ regulatory T cells. *J. Exp. Med.* **201,** 181–187 (2005).

173. Oh-hora, M. *et al.* Dual functions for the endoplasmic reticulum calcium sensors STIM1 and STIM2 in T cell activation and tolerance. *Nat Immunol* **9,** 432–443 (2008).

174. Sauer, S. *et al.* T cell receptor signaling controls Foxp3 expression via PI3K, Akt, and mTOR. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 7797–7802 (2008).

175. Haxhinasto, S., Mathis, D. & Benoist, C. The AKT–mTOR axis regulates de novo differentiation of CD4 $^{+}$ Foxp3 $^{+}$ cells. *J. Exp. Med.* **205,** 565–574 (2008).

176. Kerdiles, Y. M. *et al.* Foxo Transcription Factors Control Regulatory T Cell Development and Function. *Immunity* **33,** 890–904 (2010).

177. Feuerer, M., Hill, J. A., Mathis, D. & Benoist, C. Foxp3 + regulatory T cells : differentiation , specification , subphenotypes. *Nat. Immunol.* **10,** 689–695 (2009).

178. Li, M. O. & Rudensky, A. Y. T cell receptor signalling in the control of regulatory T cell differentiation and function. *Nat. Rev. Immunol.* **16,** 220–233 (2016).

179. Lohr, J., Knoechel, B., Kahn, E. C. & Abbas, A. K. Role of B7 in T Cell Tolerance. *J. Immunol.* **173,** 5028–5035 (2004).

180. Tai, X., Cowan, M., Feigenbaum, L. & Singer, A. CD28 costimulation of developing thymocytes induces Foxp3 expression and regulatory T cell differentiation independently of interleukin 2. *Nat. Immunol.* **6,** 152–162 (2005).

181. Legrand, N. *et al.* Transient accumulation of human mature thymocytes and regulatory T cells with CD28 superagonist in 'human immune system' Rag2 -/- ?? -/- mice. *Blood* **108,** 238–245 (2006).

182. Hinterberger, M., Wirnsberger, G. & Klein, L. B7/CD28 in central tolerance: Costimulation promotes maturation of regulatory T cell precursors and prevents their clonal deletion. *Front. Immunol.* **2,** 1–12 (2011).

183. Chan-Wang, L., Dodson, L. F., Deppong, C. M., Hsieh, C.-S. & Green, J. M. CD28 facilitates the generation of Foxp3-cytokine responsive regulatory T cell precursors. *J. Immunol.* **184,** 6007–6013 (2011).

184. Mahmud, S. A. *et al.* Tumor necrosis factor receptor superfamily costimulation couples T cell receptor signal strength to thymic regulatory T cell differentiation. *Nat. Immunol.* **15,** 473–481 (2014).

185. Coquet, J. M. *et al.* Epithelial and dendritic cells in the thymic medulla promote CD4+ Foxp3+ regulatory T cell development via the CD27–CD70 pathway. *J. Exp. Med.* **210,** 715–728 (2013).

186. Zorn, E. *et al.* IL-2 regulates FOXP3 expression in human CD4+ CD25+ regulatory T cells through a STAT-dependent mechanism and induces the expansion of these cells in vivo. *Blood* **108,** 1571–1580 (2006).

187. Burchill, M. A. *et al.* IL-2 Receptor β-Dependent STAT5 Activation Is Required for the Development of Foxp3+ Regulatory T Cells. *J. Immunol.* **178,** 280–290 (2017).

188. Yao, Z. *et al.* Nonredundant roles for Stat5a/b in directly regulating Foxp3. *Blood* **109,** 4368–4376 (2007).

189. Mayack, S. R. & Berg, L. J. Cutting Edge: An Alternative Pathway of CD4+ T Cell Differentiation Is Induced Following Activation in the Absence of gamma-Chain-Dependent Cytokine Signals. *J. Immunol.* **176,** 2059–2063 (2006).

190. Cheng, G., Yu, A., Dee, M. J. & Malek, T. R. IL-2R signaling is essential for functional maturation of T regulatory cells during thymic development. *J. Immunol.* **190,** 1567–1575 (2013).

191. Caramalho, I. *et al.* Human regulatory T-cell development is dictated by Interleukin-2 and -15 expressed in a non-overlapping pattern in the thymus. *J. Autoimmun.* **56,** 98–110 (2015).

192. Lio, C. J. & Hsieh, C. A two-step process for thymic regulatory T cell development. *Immunity* **28,** 100–111 (2008).

193. Fontenot, J. D., Dooley, J. L., Farr, A. G. & Rudensky, A. Y. Developmental regulation of Foxp3 expression during ontogeny. *J. Exp. Med.* **202,** 901–906 (2005).

194. Aschenbrenner, K. *et al.* Selection of Foxp3 + regulatory T cells specific for self antigen expressed and presented by Aire + medullary thymic epithelial cells. *Nat. Immunol.* **8,** 351–358 (2007).

195. Liston, A. *et al.* Differentiation of regulatory Foxp3+ T cells in the thymic cortex. *Proc. Natl, Acad. Sci. USA* **105,** 11903–11908 (2008).

196. Burchill, M. A. *et al.* Linked T Cell Receptor and Cytokine Signaling Govern the Development of the Regulatory T cell Repertoire. *Immunity* **28,** 112–121 (2008).

197. Hsieh, C. S., Lee, H. M. & Lio, C. W. J. Selection of regulatory T cells in the thymus. *Nat. Rev. Immunol.* **12,** 157–167 (2012).

198. Caramalho, Í., Nunes-cabaço, H., Foxall, R. B. & Sousa, A. E. Regulatory T-cell development in the human thymus. *Front. Immunol.* **6,** 1–7 (2015).

199. Haynes, B. B. F. & Heinly, C. S. Early Human T Cell Development: Analysis of the Human Thymus at the Time of Initial Entry of Hematopoietic Stem Cells into the Fetal Thymic Microenvironment. *J. Exp. Med.* **181,** 1445–1458 (1995).

200. Farley, A. M. *et al.* Dynamics of thymus organogenesis and colonization in early human development. *Development* **140,** 2015–2026 (2015).

201. Haddad, R. *et al.* Dynamics of Thymus-Colonizing Cells during Human Development. *Immunity* **24,** 217–230 (2006).

202. Cupedo, T., Nagasawa, M., Weijer, K., Blom, B. & Spits, H. Development and activation of regulatory T cells in the human fetus. *Eur. J. Immunol.* **35,** 383–390 (2005).

203. Michaëlsson, J., Mold, J. E., Mccune, J. M. & Nixon, D. F. Regulation of T Cell Responses in the Developing Human Fetus. *J. Immunol.* **176,** (2006).

204. Annunziato, F. *et al.* Phenotype, Localization, and Mechanism of Suppression of CD4+ CD25+ Human Thymocytes. *J. Exp. Med.* **196,** 379–387 (2002).

205. Cosmi, L. *et al.* Human CD8+CD25+ thymocytes share phenotypic and functional features with CD4+CD25+ regulatory thymocytes. *Blood* **102,** 4107–4115 (2003).

206. Liotta, F. *et al.* Functional features of human CD25 + regulatory thymocytes. *Microbes Infect.* **7,** 1017–1022 (2005).

207. Tuovinen, H., Kekäläinen, E., Rossi, L. H. & Arstila, T. P. Cutting Edge: Human CD4- CD8- Thymocytes Express FOXP3 in the Absence of a TCR. *J. Immunol.* **180,** 3651–3654 (2008).

208. Nunes-Cabaço, H. *et al.* Foxp3 induction in human and murine thymus precedes the CD4+ CD8+ stage but requires early T-cell receptor expression. *Immunol. Cell Biol.* **88,** 523–528 (2010).

209. Martín-gayo, E., Sierra-filardi, E., Corbí, A. L. & Toribio, M. L. Plasmacytoid dendritic cells resident in human thymus drive natural Treg cell development. *Blood* **115,** 5366–5376 (2010).

210. Ito, T. *et al.* Two functional subsets of Foxp3+ regulatory T cells in human thymus and periphery. *Immunity* **28,** 870–880 (2008).

211. Wang, Z., Gerstein, M. & Snyder, M. RNA-Seq : a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* **10,** 57–63 (2009).

212. Shendure, J. The beginning of the end for microarrays? *Nat. Methods* **5,** 585–587 (2008).

213. Tuomela, S. *et al.* Comparative analysis of human and mouse transcriptomes of Th17 cell priming. *Oncotarget* **7,** 13416–13428 (2016).

214. Kanduri, K. *et al.* Identification of global regulators of T-helper cell lineage specification. *Genome Med.* **7,** 1–11 (2015).

215. Van Den Broek, T. *et al.* Neonatal thymectomy reveals differentiation and plasticity within human naive T cells. *J. Clin. Invest.* **126,** 1126–1136 (2016).

216. Barski, A. *et al.* Rapid Recall Ability of Memory T cells is Encoded in their Epigenome. *Sci. Rep.* **7,** 1–10 (2017).

217. Hertweck, A. *et al.* T-bet Activates Th1 Genes through Mediator and the Super Elongation Complex. *Cell Rep.* **15,** 2756–2770 (2016).

218. Schmidt, A. *et al.* Time-resolved transcriptome and proteome landscape of human regulatory T cell (Treg) differentiation reveals novel regulators of FOXP3. *BMC Biol.* **16,** 1–35 (2018).

219. Bhairavabhotla, R. *et al.* Transcriptome profiling of human FoxP3+ regulatory T cells. *Hum. Immunol.* **77,** 201–213 (2016).

220. Casero, D. *et al.* Long non-coding RNA profiling of human lymphoid progenitor cells reveals transcriptional divergence of B cell and T cell lineages. *Nat. Immunol.* **16,** 1282–1291 (2015).

221. Ha, V. L. *et al.* The T-ALL related gene BCL11B regulates the initial stages of human T-cell differentiation. *Leukemia* **31,** 2503–2514 (2017).

222. Andrews S. 2015. FastQC: A Quality Control Tool for High Throughput Sequence Data. Babraham Institute. (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/).

223. Kim, D. *et al.* TopHat2 : accurate alignment of transcriptomes in the presence of insertions , deletions and gene fusions. *Genome Biol.* **14,** 0–9 (2013).

224. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25,** 2078–2079 (2009).

225. Kent, W. J. *et al.* The Human Genome Browser at UCSC. *Genome Res.* **12,** 996–1006 (2002).

226. Thorvaldsdóttir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration. *Brief. Bioinform.* **14,** 178–192 (2012).

227. Robinson, J. T. *et al.* Integrative Genome Viewer. *Nat. Biotechnol.* **29,** 24–26 (2011).

228. Ramírez, F. *et al.* deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44,** 1–6 (2016).

229. Anders, S., Pyl, P. T. & Huber, W. HTSeq - A Python framework to work with high-throughput sequencing data. *Bioinformatics* **31,** 166–169 (2015).

230. R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

231. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26,** 139–140 (2010).

232. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15,** 1–21 (2014).

233. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc.* **57,** 289–300 (1995).

234. Wang, J., Vasaikar, S., Shi, Z., Greer, M. & Zhang, B. WebGestalt 2017: A more comprehensive, powerful, flexible and interactive gene set enrichment analysis toolkit. *Nucleic Acids Res.* **45,** W130–W137 (2017).

235. Wickham, H. (2009). ggplot2: elegant graphics for data analysis (Springer-Verlag).

236. Lambert, S. A. *et al.* The Human Transcription Factors. *Cell* **172,** 650–665 (2018).

237. Scholzen, T. & Gerdes, J. The Ki-67 protein: From the known and the unknown. *J. Cell. Physiol.* **182,** 311–322 (2000).

238. Hong, C., Luckey, M. A. & Park, J. Intrathymic IL-7: The where, when, and why of IL-7 signaling during T cell development. *Semin. Immunol.* **24,** 151–158 (2012).

239. Vicente, R. *et al.* Molecular and cellular basis of T cell lineage commitment. *Semin Immunol.* **22,** 270–275 (2010).

240. Weerkamp, F. *et al.* Age-related changes in the cellular composition of the thymus in children. *J. Allergy Clin.*

Immunol. **115,** 834–840 (2005).

241. Marguerat, S. & Bähler, J. RNA-seq: From technology to biology. *Cell. Mol. Life Sci.* **67,** 569–579 (2010).

242. Sims, D., Sudbery, I., Ilott, N. E., Heger, A. & Ponting, C. P. Sequencing depth and coverage: Key considerations in genomic analyses. *Nat. Rev. Genet.* **15,** 121–132 (2014).

243. Liu, Y. *et al.* Evaluating the Impact of Sequencing Depth on Transcriptome Profiling in Human Adipose. *PLoS One* **8,** 1–10 (2013).

244. Sonia Tarazona, Fernando García-Alcalde, Dopazo, J., Ferrer, A. & Conesa, A. Differential expression in RNA-seq: a matter of depth. *Genome Res.* **21,** 2213–2223 (2011).

245. Ewing, B. & Green, P. Base-Calling of Automated Sequencer Traces Using Phred. II. Error Probabilities. *Genome Res.* **8,** 186–194 (2005).

246. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9,** 357–359 (2012).

247. Sun, L., Jin, H. & Li, H. GARP: a surface molecule of regulatory T cells that is involved in the regulatory function and TGF-β releasing. *Oncotarget* **7,** 42826–42836 (2016).

248. So, T., Seung-Woo, L. & Croft, M. Immune Regulation and Control of Regulatory T cells by OX40 and 41BB. *Cytokine Growth Factor Rev* **19,** 253–262 (2008).

249. Watts, T. H. TNF/TNFR Family Members in Costimulation of T Cell Responses. *Annu. Rev. Immunol.* **23,** 23–68 (2005).

250. Pan, F. *et al.* Eos mediates Foxp3-dependent gene silencing in CD4+ regulatory T cells. *Science (80-.).* **325,** 1142–1146 (2009).

251. Grewal, I. S. & Flavell, R. A. A central role of CD40 ligand in the regulation of CD4+ T-cell responses. *Immunol. Today* **17,** 410–414 (1996).

252. Bourgon, R., Gentleman, R. & Huber, W. Independent filtering increases detection power for high-throughput experiments. *Pnas* **107,** 9546–9551 (2010).

253. Robinson, M. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* **11,** 1–9 (2010).

254. Schurch, N. J. *et al.* How many biological replicates are needed in an RNA-seq experiment and which differential expression tool should you use? *RNA* **22,** 839–851 (2016).

255. McCarthy, D. J., Chen, Y. & Smyth, G. K. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* **40,** 4288–4297 (2012).

256. Anders, S. *et al.* Count-based differential expression analysis of RNA sequencing data using R and Bioconductor. *Nat. Protoc.* **8,** 1765–1786 (2013).

257. Robinson, M. D. & Smyth, G. K. Moderated statistical tests for assessing differences in tag abundance. *Bioinformatics* **23,** 2881–2887 (2007).

258. Read, S., Malmström, V. & Powrie, F. Cytotoxic T Lymphocyte-associated Antigen 4 Plays an Essential Role in the Function of CD25+ CD4+ Regulatory Cells that Control Intestinal Inflammation. *J. Exp. Med* **192,** 295–302 (2000).

259. Thornton, A. M. *et al.* Expression of Helios, an Ikaros Transcription Factor Family Member, Differentiates Thymic-Derived from Peripherally Induced Foxp3+ T Regulatory Cells. *J. Immunol.* **184,** 3433–3441 (2010).

260. Akimova, T., Beier, U. H., Wang, L., Levine, M. H. & Hancock, W. W. Helios expression is a marker of T cell activation and proliferation. *PLoS One* **6,** 1–13 (2011).

261. Gottschalk, R. A., Corse, E. & Allison, J. P. Expression of Helios in Peripherally Induced Foxp3+ Regulatory T Cells. *J. Immunol.* **188,** 976–980 (2012).

262. Himmel, M. E., MacDonald, K. G., Garcia, R. V., Steiner, T. S. & Levings, M. K. Helios+ and Helios- Cells Coexist within

the Natural FOXP3+ T Regulatory Cell Subset in Humans. *J. Immunol.* **190,** 2001–2008 (2013).

263. Sun, S.-C. The noncanonical NF-kB pathway. *Immunol. Rev.* **246,** 125–140 (2012).

264. Isomura, I. *et al.* c-Rel is required for the development of thymic Foxp3 + CD4 regulatory T cells. *J. Exp. Med.* **206,** 3001–3014 (2009).

265. Kirikoshi, H., Inoue, S., Sekihara, H. & Katoh, M. Expression of WNT10A in human cancer. *Int. J. Oncol.* **19,** 997–1001 (2001).

266. Asem, M. S., Buechler, S., Wates, R. B., Miller, D. L. & Stack, M. S. Wnt5a signaling in cancer. *Cancers (Basel).* **8,** 1–18 (2016).

267. Corda, G. *et al.* Functional and prognostic significance of the genomic amplification of frizzled 6 (FZD6) in breast cancer. *J. Pathol.* **241,** 350–361 (2017).

268. Dzhagalov, I. & Phee, H. How to find your way through the thymus: a pratical guide for aspiring T cells. *Cell Mol. Life Sci.* **69,** 663–682 (2012).

269. Mißlitz, A., Bernhardt, G. & Förster, R. Trafficking on serpentines: Molecular insight on how maturating T cells find their winding paths in the thymus. *Immunol. Rev.* **209,** 115–128 (2006).

270. Assi, K. *et al.* Role of epithelial integrin-linked kinase in promoting intestinal inflammation: Effects on CCL2, fibronectin and the T cell repertoire. *BMC Immunol.* **12,** 12–42 (2011).

271. Feske, S., Skolnik, E. Y. & Prakriya, M. Ion channels and transporters in lymphocyte function and immunity. *Nat. Rev. Immunol.* **12,** 532–547 (2012).

272. Hosoi, E. *et al.* Expression of the Ryanodine Receptor Isoforms in Immune Cells. *J. Immunol.* **167,** 4887–4894 (2001).

273. Dai, Z. *et al.* CD4+ CD25+ regulatory T cells suppress allograft rejection mediated by memory CD8+ T cells via a CD30-dependent mechanism. *J. Clin. Invest.* **113,** 310–317 (2004).

274. Wang, R., Wan, Q., Kozhaya, L., Fujii, H. & Unutmaz, D. Identification of a regulatory T cell specific cell surface molecule that mediates suppressive signals and induces Foxp3 expression. *PLoS One* **3,** 1–9 (2008).

275. Barsheshet, Y. *et al.* CCR8+ FOXp3+ T reg cells as master drivers of immune regulation. *Proc. Natl. Acad. Sci.* **114,** 6086–6091 (2017).

276. Holzbaur, E. Dyneins: Molecular Structure and Cellular Function. *Annu. Rev. Cell Dev. Biol.* **10,** 339–372 (1994).

277. Chapelin, C. *et al.* Isolation of several human axonemal dynein heavy chain genes: Genomic structure of the catalytic site, phylogenetic analysis and chromosomal assignment. *FEBS Lett.* **412,** 325–330 (1997).

278. Neesen, J. *et al.* Identification of dynein heavy chain genes expressed in human and mouse testis: Chromosomal localization of an axonemal dynein gene. *Gene* **200,** 193–202 (1997).

279. Albert, M. H. *et al.* MiRNome and transcriptome aided pathway analysis in human regulatory T cells. *Genes Immun.* **15,** 303–312 (2014).

280. Plitas, G. *et al.* Regulatory T cells exhibit distinct features in human breast cancer. *Immunity* **45,** 1122–1134 (2016).

281. Katarzyna Polak. (2015). Role of the Eos and Helios transcription factors in regulatory T cell biology (PhD dissertation). Université de Strasbourg. Retrieved from https://tel.archives-ouvertes.fr/tel-01647132/document.

282. Kim, H.-J. *et al.* Stable inhibitory activity of regulatory T cells requires the transcription factor Helios. *Science (80-.).* **350,** 334–339 (2015).

283. Le Borgne, M. & Shaw, A. S. Do T cells have a cilium? *Science (80-.).* **342,** 1177–1178 (2013).

284. Roche, M. De *et al.* Hedgehog Signaling Controls T Cell Killing at the Immunological Synapse. *Science (80-.).* **342,** 1247–1250 (2013).

285. Di Pierro, V. *et al.* Consanguinity and polygenic diseases: A model for antibody deficiencies. *Hum. Hered.* **77,** 144–149 (2014).

286. Knowles, M. R., Daniels, L. A., Davis, S. D., Zariwala, M. A. & Leigh, M. W. Primary ciliary dyskinesia: Recent advances in diagnostics, genetics, and characterization of clinical disease. *Am. J. Respir. Crit. Care Med.* **188,** 913–922 (2013).

287. Bartoloni, L. *et al.* Mutations in the DNAH11 (axonemal heavy chain dynein type 11) gene cause one form of situs inversus totalis and most likely primary ciliary dyskinesia. *PNAS* **99,** 10282–10286 (2002).

288. Olbrich, H. *et al.* Mutations in DNAH5 cause primary ciliary dyskinesia and randomization of left-right asymmetry. *Nat. Genet.* **30,** 143–144 (2002).

289. Li, Y. *et al.* DNAH6 and Its Interactions with PCD Genes in Heterotaxy and Primary Ciliary Dyskinesia. *PLoS Genet.* **12,** 1–20 (2016).

290. Nakagawa, N. *et al.* RANK is the essential signaling receptor for osteoclast differentiation factor in osteoclastogenesis. *Biochem. Biophys. Res. Commun.* **253,** 395–400 (1998).

291. Hsu, H. *et al.* Tumor necrosis factor receptor family member RANK mediates osteoclast differentiation and activation induced by osteoprotegerin ligand. *Proc. Natl. Acad. Sci. USA* **96,** 3540–3545 (1999).

292. Wong, B. R. *et al.* TRANCE Is a Novel Ligand of the Tumor Necrosis Factor Receptor Family That Activates c-Jun N-terminal Kinase in T Cells. *J. Biol. Chem.* **272,** 25190–25194 (1997).

293. Anderson, D. M. *et al.* A homologue of the TNF receptor and its ligand enhance T-cell growth and dendritic-cell function. *Nature* **390,** 175–179 (1997).

294. Rossi, S. W. *et al.* RANK signals from CD4+CD3− inducer cells regulate development of Aire-expressing epithelial cells in the thymic medulla. *J. Exp. Med.* **204,** 1267–1272 (2007).

295. Akiyama, T. *et al.* The Tumor Necrosis Factor Family Receptors RANK and CD40 Cooperatively Establish the Thymic Medullary Microenvironment and Self-Tolerance. *Immunity* **29,** 423–437 (2008).

296. Wong, B. R. *et al.* TRANCE (Tumor Necrosis Factor [TNF]-related Activation-induced Cytokine), a New TNF Family Member Predominantly Expressed in T cells, Is a Dendritic Cell–specific Survival Factor. *J. Exp. Med.* **186,** 2075–2080 (1997).

297. Tan, W. *et al.* Fibroblast-recruited, tumour-infiltrating CD4+ T cells stimulate mammary cancer metastasis through RANKL-RANK signaling. *Nature* **470,** 548–553 (2012).

298. Johdi, N. A., Ait-Tahar, K., Sagap, I. & Jamal, R. Molecular signatures of human regulatory T cells in colorectal cancer and polyps. *Front. Immunol.* **8,** 1–9 (2017).

299. Oestreich, K. J. & Weinmann, A. S. Master regulators or lineage-specifying? Changing views on CD4+ T cell transcription factors. *Nat. Rev. Immunol.* **12,** 799–804 (2012).

300. Chan, S. S.-K. & Kyba, M. What is a Master Regulator? *J. Stem Cell Res. Ther.* **3,** 10–13 (2013).

301. Robinette, M. L. *et al.* Transcriptional programs define molecular characteristics of innate lymphoid cell classes and subsets. *Nat. Immunol.* **16,** 306–317 (2015).

302. Reid, J. E., Ott, S. & Wernisch, L. Transcriptional programs: Modelling higher order structure in transcriptional control. *BMC Bioinformatics* **10,** 1–12 (2009).

303. Sekiya, T. *et al.* Nr4a receptors are essential for thymic regulatory T cell development and immune homeostasis. *Nat. Immunol.* **14,** 230–237 (2013).

304. Hayden, M. S. & Ghosh, S. Shared Principles in NF-κB Signaling. *Cell* **132,** 344–362 (2008).

305. Grinberg-Bleyer, Y. *et al.* The Alternative NF-κB Pathway in Regulatory T Cell Homeostasis and Suppressive Function. *J. Immunol.* **200,** 2362–2371 (2018).

306. Murray, S. E. *et al.* NF-κB–inducing kinase plays an essential T cell–intrinsic role in graft-versus-host disease and lethal autoimmunity in mice. *J. Clin. Invest.* **121,** 4775–4786 (2011).

307. Lu, L., Gondek, D. C., Scott, Z. A. & Noelle, R. J. NF-kB-Inducing Kinase Deficiency Results in the Development of a Subset of Regulatory T Cells, which Shows a Hyperproliferative Activity upon Glucocorticoid-Induced TNF Receptor

Family-Related Gene Stimulation. *J. Immunol.* **175,** 1651–1657 (2005).

308. Gandhi, R. *et al.* Activation of the aryl hydrocarbon receptor induces human type 1 regulatory T cell-like and Foxp3+ Regulatory T cells. *Nat. Immunol.* **11,** 846–853 (2010).

309. Quintana, F. J. *et al.* Aiolos promotes Th17 differentiation by directly silencing Il2 expression. *Nat. Immunol.* **13,** 770–777 (2013).

310. Smith, K. A. Interleukin-2: Inception, Impact, and Implications. *Science (80-.).* **240,** 1169–1176 (1988).

311. Kang, S. W. *et al.* 1,25(OH)2 vitamin D3 promotes FOXP3 expression via binding to vitamin D response elements in its conserved non-coding sequence region. *J. Immunol.* **6,** 5276–5282 (2013).

312. Zhou, Q. *et al.* 1,25(OH)2D3 induces regulatory T cell differentiation by influencing the VDR/PLC-γ1/TGF-β1/pathway. *Mol. Immunol.* **91,** 156–164 (2017).

313. Ghoreschi, K. *et al.* Generation of Pathogenic Th17 Cells in the Absence of TGF-b Signaling. *Nature* **467,** 967–971 (2010).

314. Yu, S., Bruce, D., Froicu, M., Weaver, V. & Cantorna, M. T. Failure of T cell homing, reduced CD4/CD8aa intraepithelial lymphocytes, and inflammation in the gut of vitamin D receptor KO mice. *PNAS* **105,** 20834–20839 (2008).

315. Froicu, M. *et al.* A Crucial Role for the Vitamin D Receptor in Experimental Inflammatory Bowel Diseases. *Mol. Endocrinol.* **17,** 2386–2392 (2003).

316. Quintana, F. J. *et al.* Control of Treg and Th17 cell differentiation by the aryl hydrocarbon receptor. *Nature* **453,** 65–71 (2008).

317. Zhou, L. AHR Function in Lymphocytes: Emerging Concepts. *Trends Immunol.* **37,** 17–31 (2016).

318. Nguyen, N. T., Hanieh, H., Nakahama, T. & Kishimoto, T. The roles of aryl hydrocarbon receptor in immune responses. *Int. Immunol.* **25,** 335–343 (2013).

319. Stockinger, B., Meglio, P. Di, Gialitakis, M. & Duarte, J. H. The Aryl Hydrocarbon Receptor: Multitasking in the Immune System. *Annu. Rev. Immunol.* **32,** 403–432 (2014).

320. Ye, J. *et al.* The Aryl Hydrocarbon Receptor Preferentially Marks and Promotes Gut Regulatory T Cells. *Cell Rep.* **21,** 2277–2290 (2017).

321. Zhu, J., Yamane, H. & Paul, W. Differentiation of effector CD4 T cell populations. *Annu Rev Immunol.* **28,** 445–489 (2010).

322. Szabo, S. J. *et al.* A novel transcription factor, T-bet, directs Th1 lineage commitment. *Cell* **100,** 655–669 (2000).

323. Mahnke, J. *et al.* Interferon Regulatory Factor 4 controls Th1 cell effector function and metabolism. *Sci. Rep.* **6,** 1–12 (2016).

324. Sha, Y. & Markovic-Plese, S. Activated IL-1RI signaling pathway induces Th17 cell differentiation via Interferon Regulatory Factor 4 signaling in patients with relapsing-remitting multiple sclerosis. *Front. Immunol.* **7,** 1–9 (2016).

325. Staudt, V. *et al.* Interferon-Regulatory Factor 4 Is Essential for the Developmental Program of T Helper 9 Cells. *Immunity* **33,** 192–202 (2010).

326. Ouyang, X. *et al.* Transcription factor IRF8 directs a silencing programme for TH17 cell differentiation. *Nat. Commun.* **2,** 1–12 (2011).

327. Humblin, E. *et al.* IRF8-dependent molecular complexes control the Th9 transcriptional program. *Nat. Commun.* **8,** 1–13 (2017).

328. Castro, G. *et al.* RORγt and RORα signature genes in human Th17 cells. *PLoS One* **12,** 1–22 (2017).

329. Nishikomori, R. *et al.* Activated STAT4 Has an Essential Role in Th1 Differentiation and Proliferation That Is Independent of Its Role in the Maintenance of IL-12R b2 Chain Expression and Signaling. *J. Immunol.* **169,** 4388–

4398 (2002).

330. Moriggl, R. *et al.* Activation of STAT Proteins and cytokine genes in human Th1 and Th2 cells generated in the absence of IL-12 and IL-4. *J. Immunol.* **160,** 3385–3392 (1998).

331. Xu, J. *et al.* STAT4 Is Critical for the Balance between Th17 Cells and Regulatory T Cells in Colitis. *J. Immunol.* **6,** 6597–6606 (2011).

332. Ise, W. *et al.* BATF controls the global regulators of CSR in both B and T cells. *Nat. Immunol.* **12,** 536–543 (2011).

333. Schraml, B. U. *et al.* The AP-1 transcription factor Batf controls Th17 differentiation. *Nature* **460,** 405–409 (2009).

334. Jabeen, R. *et al.* Th9 cell development requires a BATF-regulated transcriptional network. *J. Clin. Invest.* **123,** 4641–4653 (2013).

335. Koch, M. a *et al.* T-bet controls regulatory T cell homeostasis and function during type-1 inflammation. *Nat. Immunol.* **10,** 595–602 (2009).

336. Hall, A. O. H. *et al.* The cytokines Interleukin 27 and Interferon-γ promote distinct Treg cell populations required to limit infection-induced pathology. *Immunity* **37,** 511–523 (2013).

337. Lee, W., Kim, H. S., Baek, S. Y. & Lee, G. R. Transcription factor IRF8 controls Th1-like regulatory T-cell function. *Cell. Mol. Immunol.* **13,** 785–794 (2016).

338. Zheng, Y. *et al.* Regulatory T-cell suppressor program co-opts transcription factor IRF4 to control Th2 responses. *Nature* **458,** 351–356 (2009).

339. Cretney, E. *et al.* The transcription factors Blimp-1 and IRF4 jointly control the differentiation and function of effector regulatory T cells. *Nat. Immunol.* **12,** 304–312 (2011).

340. Hale, J. S. & Fink, P. J. Back to the Thymus: Peripheral T Cells Come Home. *Immunol Cell Biol* **87,** 58–64 (2009).

341. Thiault, N. *et al.* Peripheral regulatory T lymphocytes recirculating to the thymus suppress the development of their precursors. *Nat. Immunol.* **16,** 628–634 (2015).

342. Krausgruber, T. *et al.* IRF5 promotes inflammatory macrophage polarization and TH1-TH17 responses. *Nat. Immunol.* **12,** 231–238 (2011).

343. Ouyang, X. *et al.* Cooperation between MyD88 and TRIF pathways in TLR synergy via IRF5 activation. *Biochem. Biophys. Res. Commun.* **354,** 1045–1051 (2007).

344. Sakuishi, K. *et al.* TIM3+FOXP3+ regulatory T cells are tissue-specific promoters of T-cell dysfunction in cancer. *Oncoimmunology* **2,** 1–9 (2013).

345. Miyazaki, K. *et al.* The Role of the Basic Helix-Loop-Helix Transcription Factor Dec1 in the Regulatory T Cells. *J. Immunol.* **185,** 7330–7339 (2010).

346. Lin, C. C. *et al.* Bhlhe40 controls cytokine production by T cells and is essential for pathogenicity in autoimmune neuroinflammation. *Nat. Commun.* **5,** 1–13 (2014).

347. Yu, F. *et al.* The transcription factor Bhlhe40 is a switch of inflammatory versus antiinflammatory Th1 cell fate determination. *J. Exp. Med.* **215,** 1–9 (2018).

348. Gabryšová, L. *et al.* c-Maf controls immune responses by regulating disease-specific gene networks and repressing IL-2 in CD4+ T cells. *Nat. Immunol.* **19,** 497–507 (2018).

349. Kondo, S. *et al.* BBF2H7, a Novel Transmembrane bZIP Transcription Factor, Is a New Type of Endoplasmic Reticulum Stress Transducer. *Mol. Cell. Biol.* **27,** 1716–1729 (2007).

350. Imamura, K. *et al.* Human immunodeficiency virus type 1 enhancer-binding protein 3 is essential for the expression of asparagine-linked glycosylation 2 in the regulation of osteoblast and chondrocyte differentiation. *J. Biol. Chem.* **289,** 9865–9879 (2014).

351. Saito, A. *et al.* Chondrocyte Proliferation Regulated by Secreted Luminal Domain of ER Stress Transducer

BBF2H7/CREB3L2. *Mol. Cell* **53,** 127–139 (2014).

352. Duurland, C. L., Brown, C. C., O'Shaughnessy, R. F. L. & Wedderburn, L. R. CD161+ Tconv and CD161+ Treg share a transcriptional and functional phenotype despite limited overlap in TCRβ repertoire. *Front. Immunol.* **8,** 1–18 (2017).

353. Maggi, L. *et al.* CD161 is a marker of all human IL-17-producing T-cell subsets and is induced by RORC. *Eur. J. Immunol.* **40,** 2174–2181 (2010).

354. Jargosch, M. *et al.* Data integration for identification of important transcription factors of STAT6-mediated cell fate decisions. *Genet. Mol. Res.* **15,** 1–17 (2016).

355. Yamane, H. & Paul, W. E. Early signaling events that underlie fate decisions of naive CD4+ T cells towards distinct T-helper cell subsets. *Immunol. Rev.* **252,** 12–23 (2014).

356. Regateiro, F. S. *et al.* Foxp3 Expression Is Required for the Induction of Therapeutic Tissue Tolerance. *J. Immunol.* **189,** 3947–3956 (2012).

357. De Simone, M. *et al.* Transcriptional Landscape of Human Tissue Lymphocytes Unveils Uniqueness of Tumor-Infiltrating T Regulatory Cells. *Immunity* **45,** 1135–1147 (2016).

358. Cao, Z., Sun, X., Icli, B., Wara, A. K. & Feinberg, M. W. Role of Kruppel-like factors in leukocyte development, function, and disease. *Blood* **116,** 4404–4415 (2010).

359. Hart, G. T., Hogquist, K. A. & Jameson, S. C. Kruppel-like Factors in Lymphocyte Biology. *J. Immunol.* **188,** 521–526 (2012).

360. Xiong, Y. *et al.* Polycomb antagonizes p300/CREB-binding protein-associated factor to silence FOXP3 in a Kruppel-like factor-dependent manner. *J. Biol. Chem.* **287,** 34372–34385 (2012).

361. Cao, Z. *et al.* Kruppel-like Factor KLF10 Targets Transforming Growth Factor-β1 to Regulate CD4+ CD25− T Cells and T Regulatory Cells. *J. Biol. Chem.* **284,** 24914–24924 (2009).

362. Needham, D. J., Lee, J. X. & Beilharz, M. W. Intra-tumoural regulatory T cells: A potential new target in cancer immunotherapy. *Biochem. Biophys. Res. Commun.* **343,** 684–691 (2006).

363. Kojima, S. *et al.* Transcriptional activation of urokinase by the Krüppel-like factor Zf9/COPEB activates latent TGF-β1 in vascular endothelial cells. *Blood* **95,** 1309–1316 (2000).

ANNEXES

Annexe 1 Batch scripts used for RNA-seq data processing. Batch scripts used in TopHat2/Bowtie2 alignment (A), HTseq quantification (B) and BigWig conversion (C) steps of RNA-seq data processing.

A

```
#!/bin/bash
#SBATCH --job-name=TopHat
#SBATCH --time=10:00:00
#SBATCH --mem-per-cpu=3G
#SBATCH --nodes=2
#SBATCH --ntasks=6
#SBATCH --cpus-per-task=12
#SBATCH --workdir=/mnt/beegfs/scratch/PRECISE/miguel.dias/

export GENOME=/mnt/nfs/lobo/IMM-NFS/genomes/hg38/Sequence/Bowtie2Index/genome
export ANNOTATION=/mnt/nfs/lobo/IMM-NFS/genomes/hg38/Annotation/gencode.v26.annotation.gtf
export Thymus_RNAseq_data=/mnt/nfs/lobo/PRECISE-NFS/miguel.dias/Human_Genome/Thymus_RNAseq_data/
export srun="srun --nodes=1 --ntasks=1 --cpus-per-task=$SLURM_CPUS_PER_TASK"
parallel="parallel --delay 0.2 -j $SLURM_NTASKS --joblog /mnt/beegfs/scratch/PRECISE/miguel.dias/
$SLURM_JOB_ID.log"
export workdir=/mnt/beegfs/scratch/PRECISE/miguel.dias/

ls -d $Thymus_RNAseq_data/*/ | $parallel --dry-run '$srun shifter --image=docker:araposo/tophat:1
.0 tophat2 -g 1 -p $SLURM_CPUS_PER_TASK -G $ANNOTATION -o $workdir/$(basename {}) $GENOME {}
*_conc1.fq.gz {}*_conc2.fq.gz'

echo "Statistics for job $SLURM_JOB_ID:"
sacct --format="JOBID,Start,End,Elapsed,CPUTime,AveDiskRead,AveDiskWrite,MaxRSS,MaxVMSize,exitcode
,derivedexitcode" -j $SLURM_JOB_I
```

B

```
#!/bin/bash
#SBATCH --job-name=HTseq
#SBATCH --time=20:00:00
#SBATCH --mem-per-cpu=2G
#SBATCH --nodes=1
#SBATCH --ntasks=4
#SBATCH --cpus-per-task=10
#SBATCH --workdir=/mnt/beegfs/scratch/PRECISE/miguel.dias/Thymus_HTseq_results/

export ANNOTATION=/mnt/nfs/lobo/IMM-NFS/genomes/hg38/Annotation/gencode.v26.annotation.gtf
export TopHat_results=/mnt/nfs/lobo/PRECISE-NFS/miguel.dias/Human_Genome/Thymus_TopHat_results
export srun="srun --nodes=1 --ntasks=1 --cpus-per-task=$SLURM_CPUS_PER_TASK"
parallel="parallel --delay 0.2 -j $SLURM_NTASKS --joblog /mnt/beegfs/scratch/PRECISE/miguel.dias
/Thymus_HTseq_results/$SLURM_JOB_ID.log"
export workdir=/mnt/beegfs/scratch/PRECISE/miguel.dias/Thymus_HTseq_results

ls -d $Thymus_TopHat_results/*/ | $parallel '$srun shifter --image=docker:argrosso/htstools
 samtools sort -@ $SLURM_CPUS_PER_TASK -n -o $workdir/sortedbyname_$(basename {}).bam {
}accepted_hits.bam; $srun shifter --image=docker:argrosso/htstools samtools view -@
$SLURM_CPUS_PER_TASK -o $workdir/sortedbyname_$(basename {}).sam  $workdir/sortedbyname_$
(basename {}).bam; $srun shifter --image=araposo/tophat:1.0 htseq-count -s no -a 10 $workdir
/sortedbyname_$(basename {}).sam $ANNOTATION > $workdir/$(basename {}).count'

echo "Statistics for job $SLURM_JOB_ID:"
sacct --format="JOBID,Start,End,Elapsed,CPUTime,AveDiskRead,AveDiskWrite,MaxRSS,MaxVMSize
```

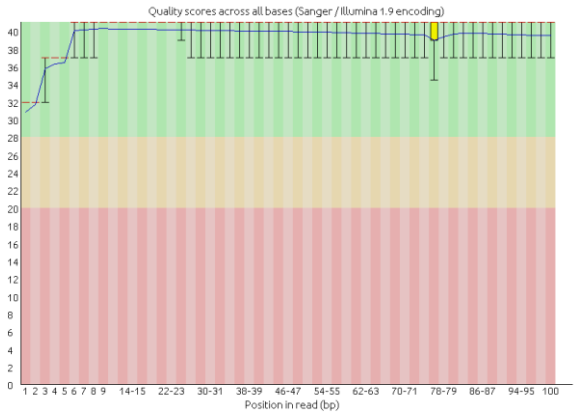
C

```
#!/bin/bash
#SBATCH --job-name=Bam2BW
#SBATCH --time=3:00:00
#SBATCH --mem-per-cpu=6G
#SBATCH --nodes=5
#SBATCH --ntasks=18
#SBATCH --cpus-per-task=10
#SBATCH --image=argrosso/htstools:0.1.1
#SBATCH --workdir=/mnt/beegfs/scratch/PRECISE/miguel.dias/Bam_2_BigWig/all_test

export Thymus_TopHat_results=/mnt/nfs/lobo/PRECISE-NFS/miguel.dias/Human_Genome/Thymus_TopHat_results
export PBMC_TopHat_results=/mnt/nfs/lobo/PRECISE-NFS/miguel.dias/Human_Genome/PBMC_TopHat_results
export mTconv_TopHat_Results=/mnt/nfs/lobo/PRECISE-NFS/miguel.dias/Human_Genome/mTconv_TopHat_Results
export srun="srun --nodes=1 --ntasks=1 --cpus-per-task=$SLURM_CPUS_PER_TASK"
parallel="parallel --delay 0.2 -j $SLURM_NTASKS --joblog /mnt/beegfs/scratch/PRECISE/miguel.dias
/Bam_2_BigWig/all_test/$SLURM_JOB_ID.log"
export workdir=/mnt/beegfs/scratch/PRECISE/miguel.dias/Bam_2_BigWig/all_test

## -- Sort the BAM file, creates an index and convert to bigwig for USCS Genome Browser/IGV --##

ls -d $Thymus_TopHat_results/*/ | $parallel '$srun shifter samtools sort -@ $SLURM_CPUS_PER_TASK -o
$workdir/sorted_$(basename {}).bam {}accepted_hits.bam; $srun shifter samtools index $workdir/sorted_$
(basename {}).bam $workdir/sorted_$(basename {}).bam.bai; $srun shifter bamCoverage -b $workdir
/sorted_$(basename {}).bam -o $workdir/sorted_$(basename {}).bw -of bigwig --binSize 1 -
-numberOfProcessors max/2'

echo "Statistics for job $SLURM_JOB_ID:"
sacct --format="JOBID,Start,End,Elapsed,CPUTime,AveDiskRead,AveDiskWrite,MaxRSS,MaxVMSize,exitcode
,derivedexitcode" -j $SLURM JOB ID
```

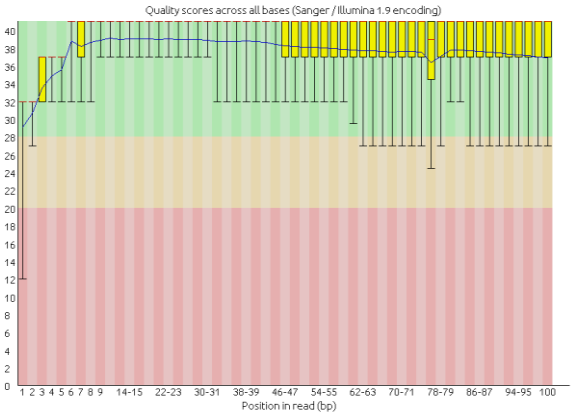
Annexe 2 FastQC report plots. FastQC plots of per base sequence quality (A), per base sequence content (B), per sequence GC content (C), duplication levels (D) from all files of each sample.

A

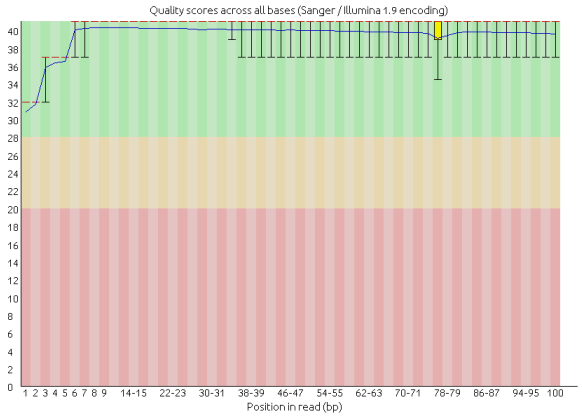
tTreg274 FWD



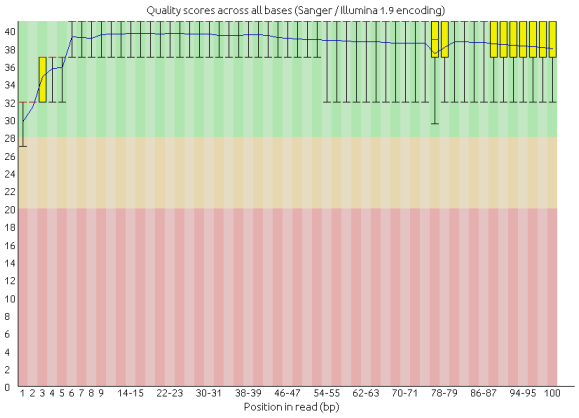
tTreg274 REV



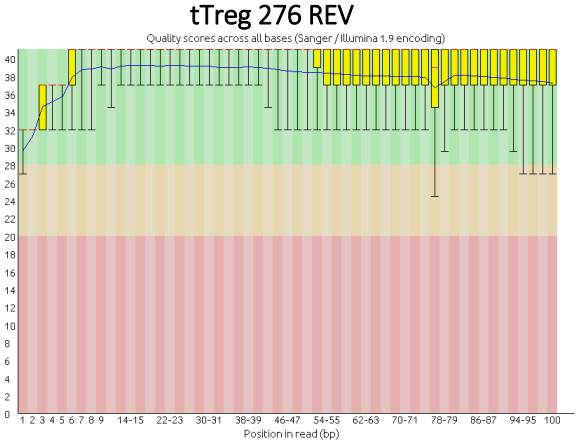
tTconv274 FWD



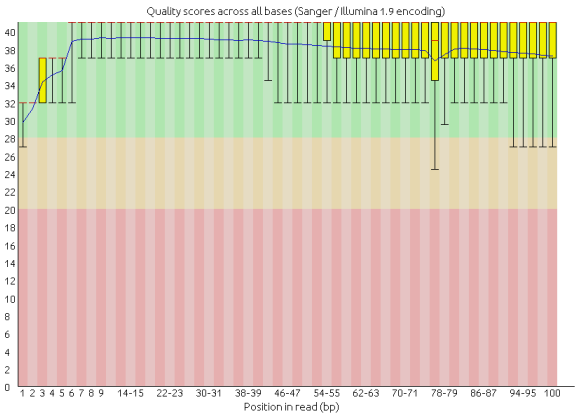
tTconv274 REV



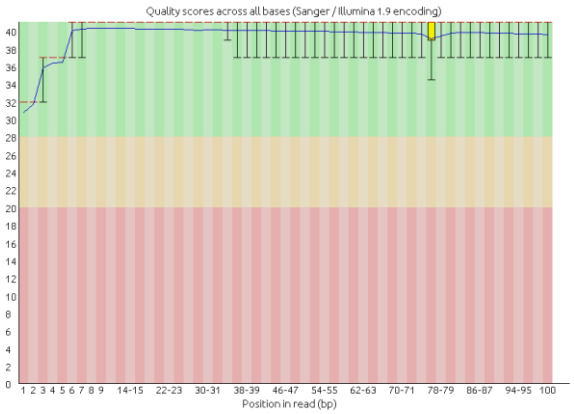
tTreg 276 REV



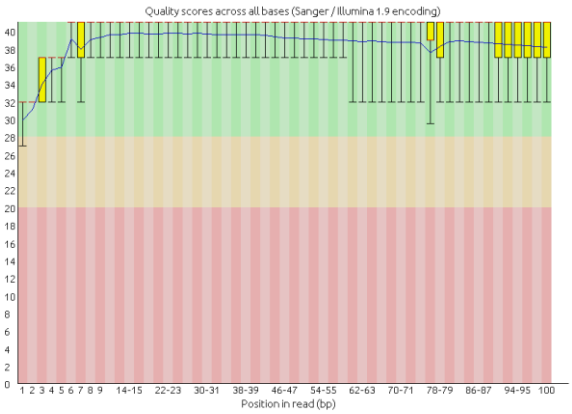
tTconv 276 REV



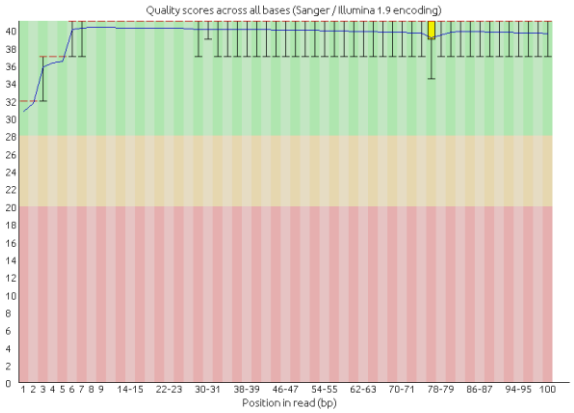
tTreg277 FWD



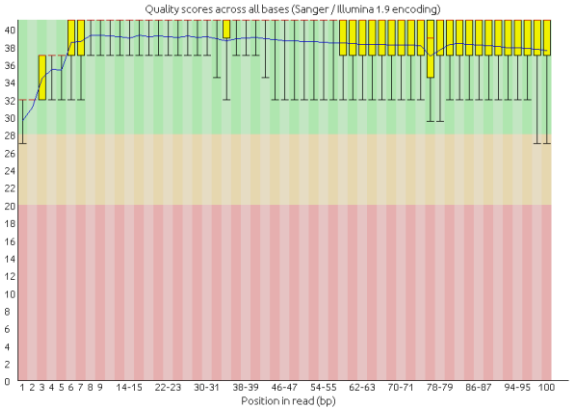
tTreg277 REV



tTconv277 FWD

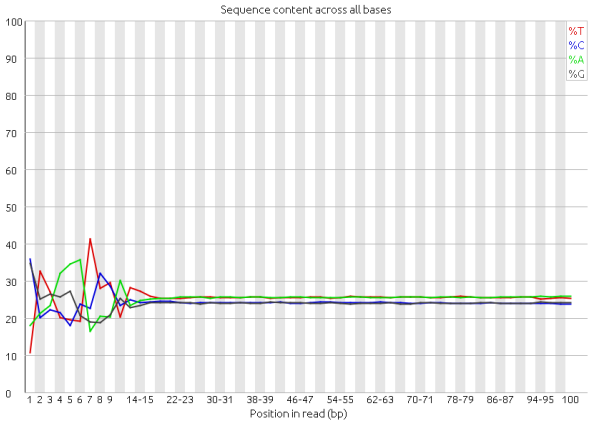


tTconv277 REV

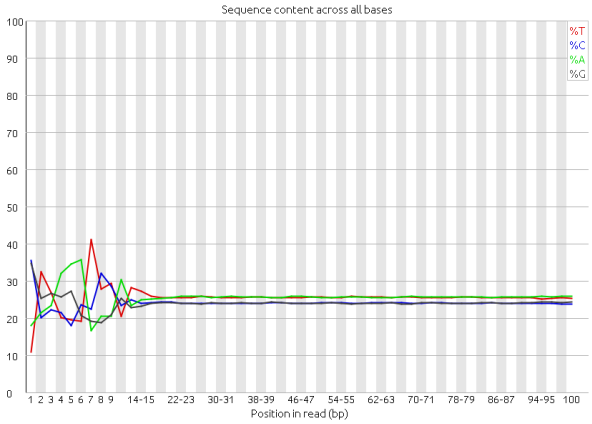


B

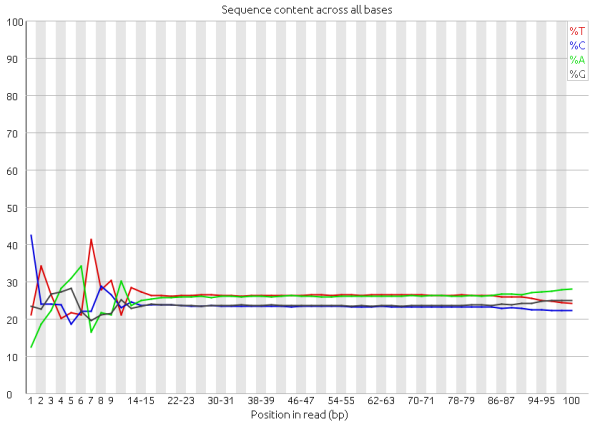
tTreg274 FWD



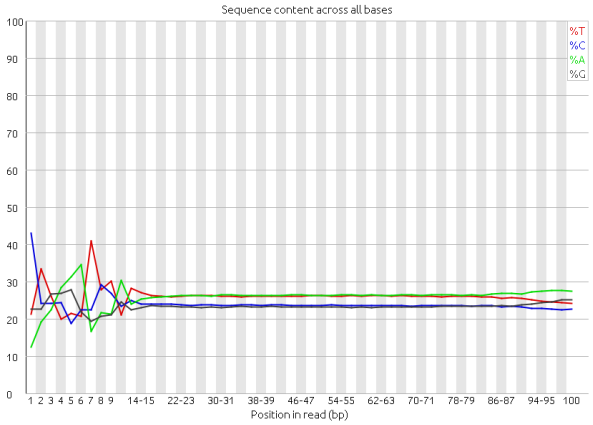
tTreg274 REV



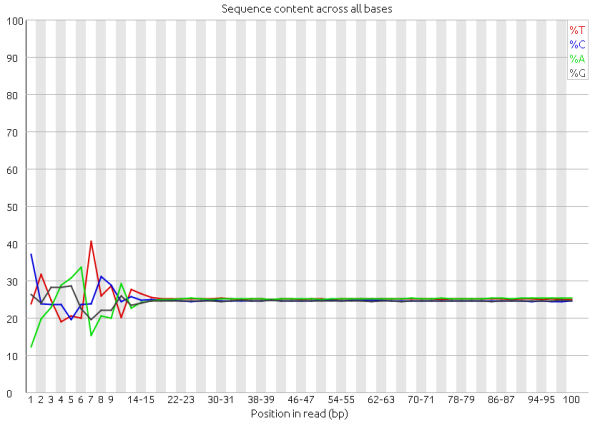
tTconv274 FWD



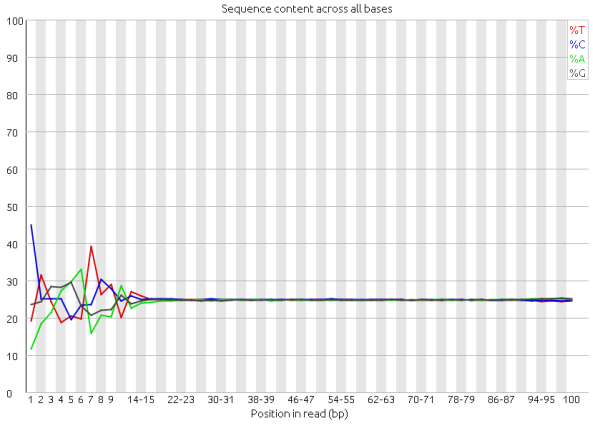
tTconv274 REV



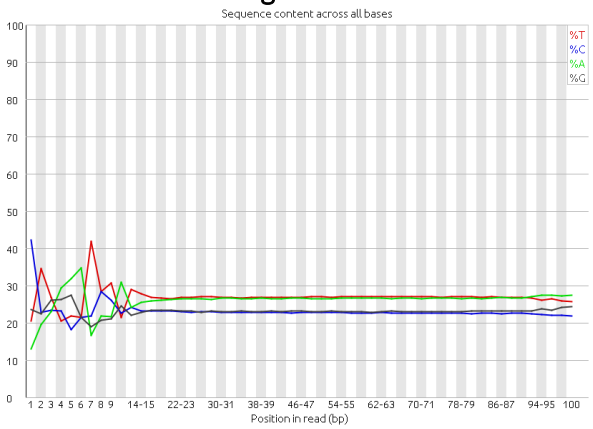
tTreg 276 REV



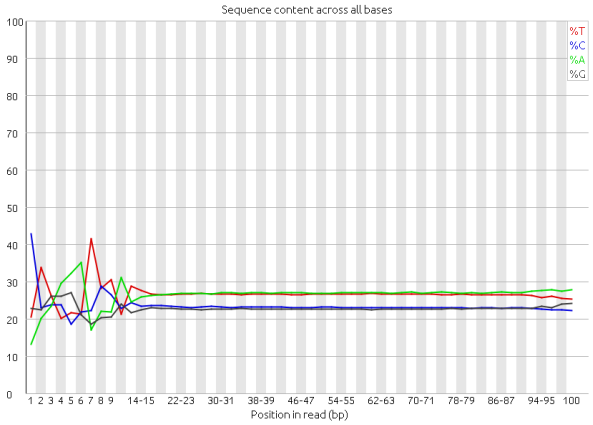
tTconv 276 REV



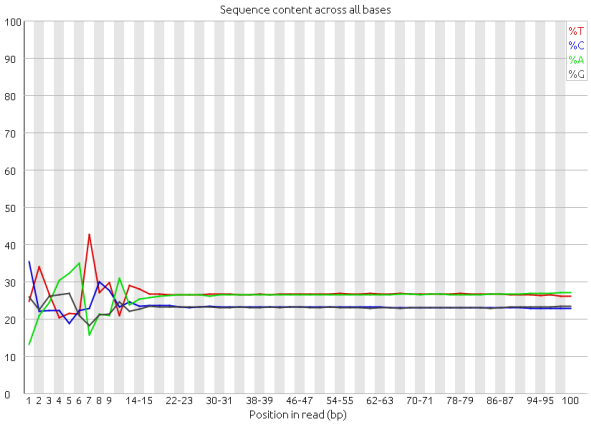
tTreg277 FWD



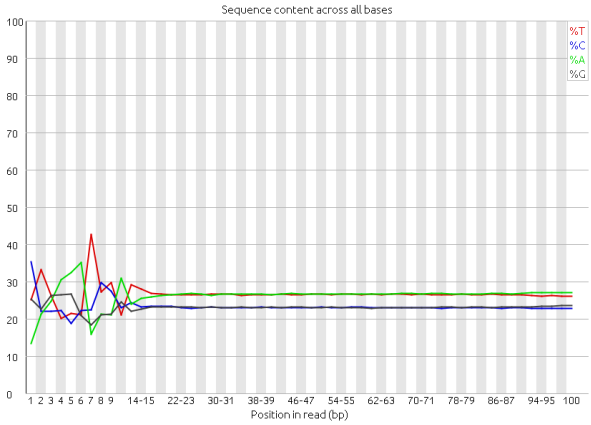
tTreg277 REV



tTconv277 FWD

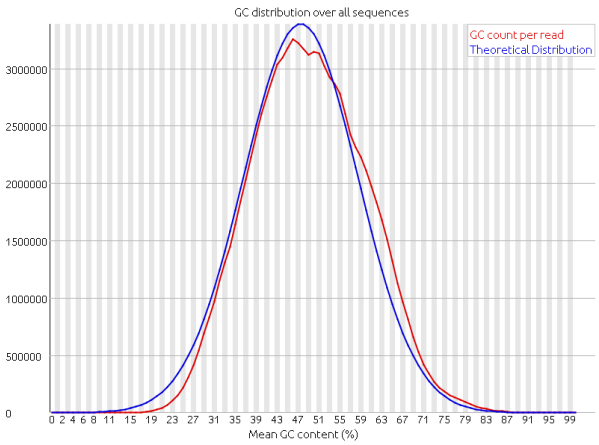


tTconv277 REV

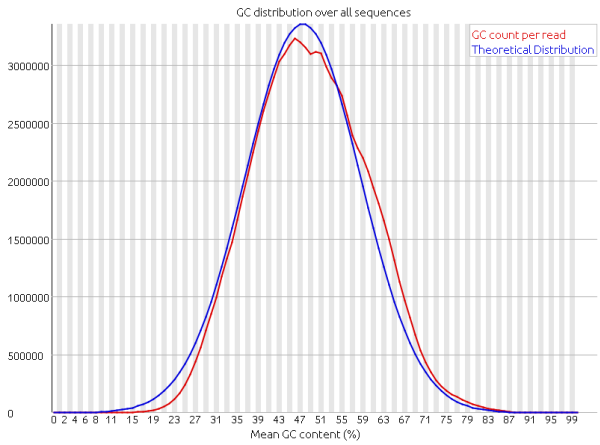


C

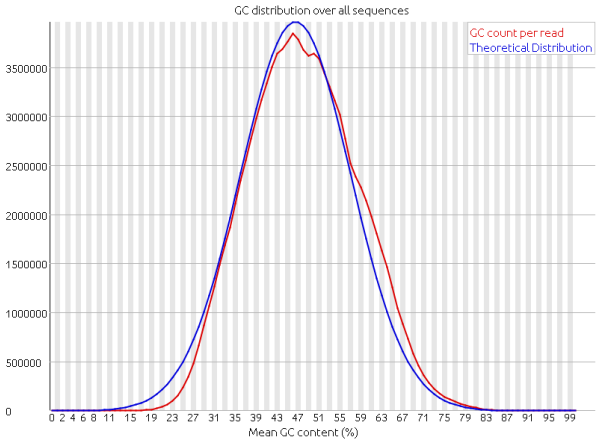
tTreg274 FWD



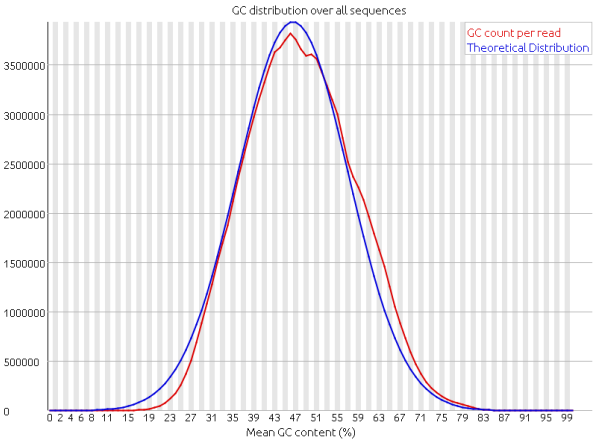
tTreg274 REV



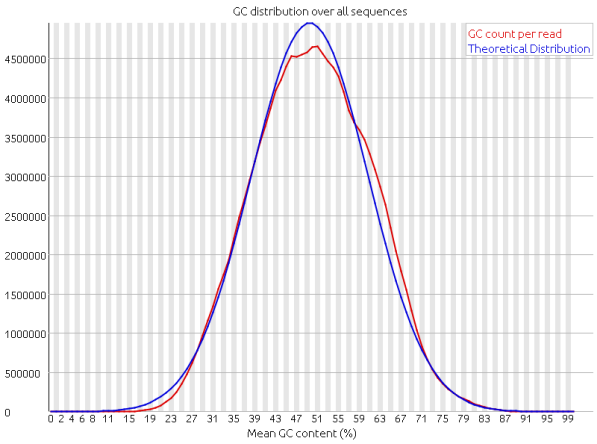
tTconv274 FWD



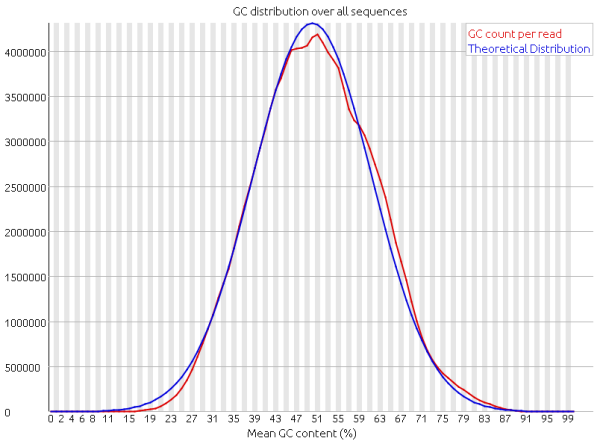
tTconv274 REV



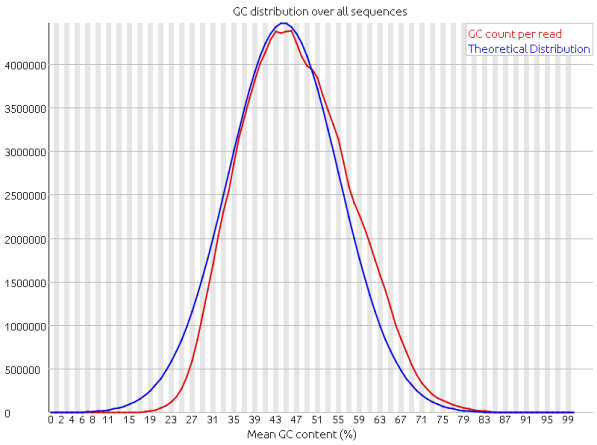
tTreg 276 REV



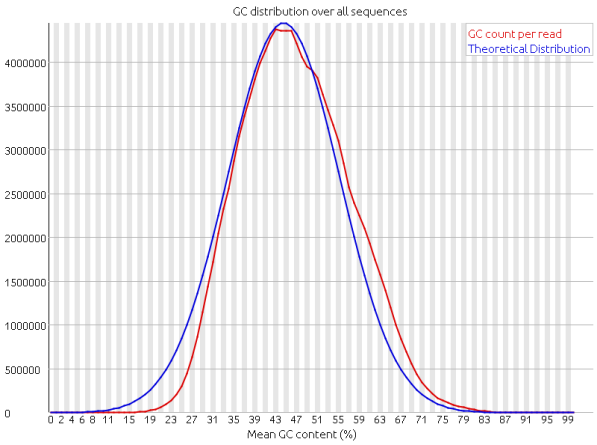
tTconv 276 REV



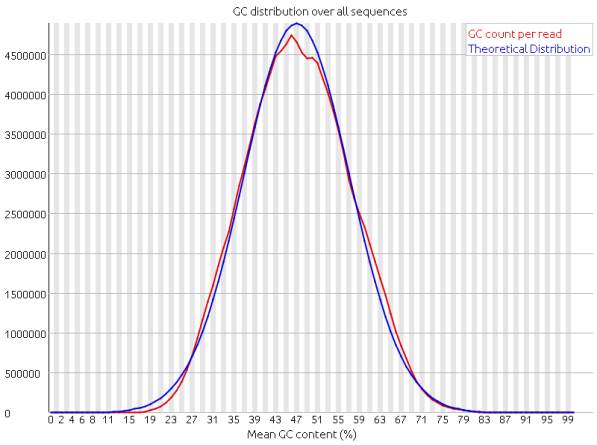
tTreg277 FWD



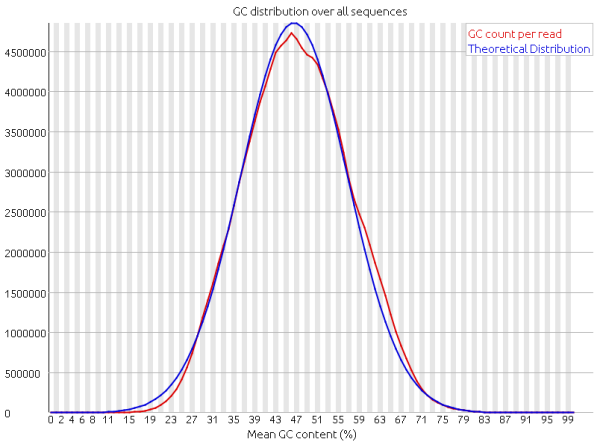
tTreg277 REV



tTconv277 FWD

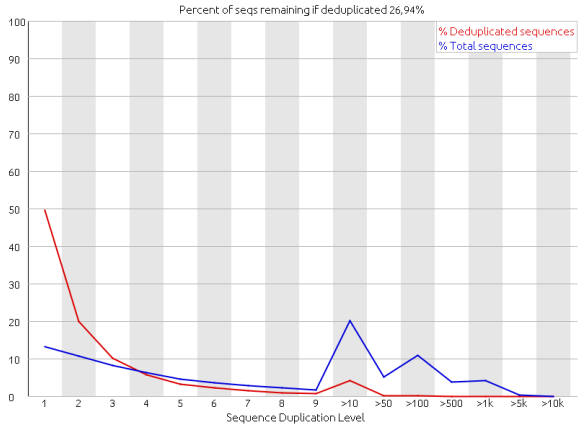


tTconv277 REV

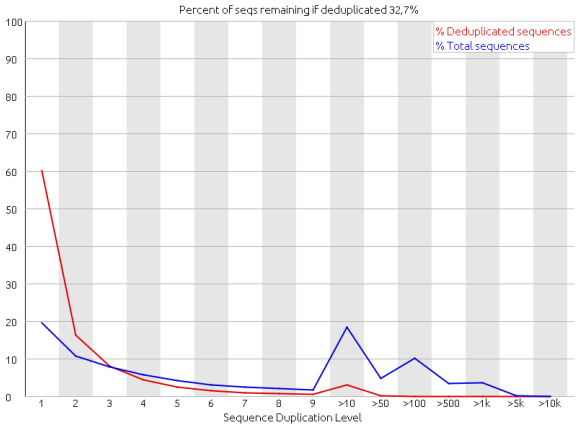


D

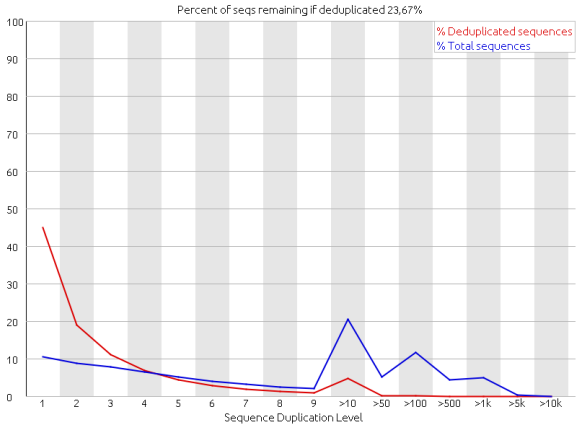
tTreg274 FWD



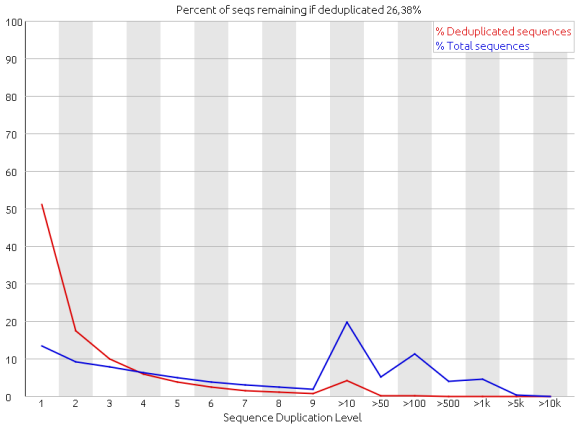
tTreg274 REV



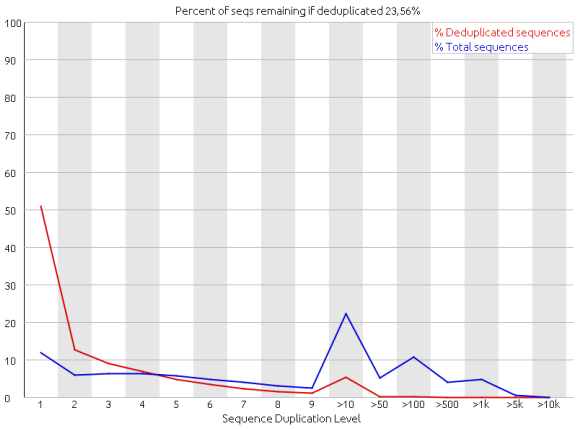
tTconv274 FWD



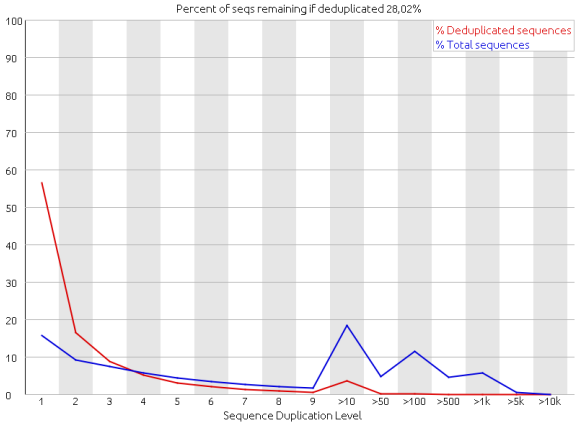
tTconv274 REV



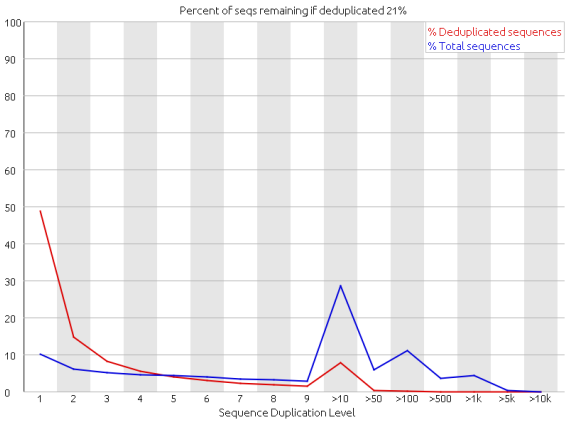
tTreg 276 REV



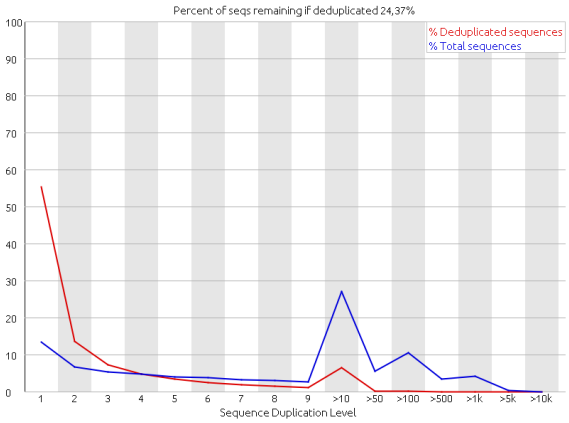
tTconv 276 REV



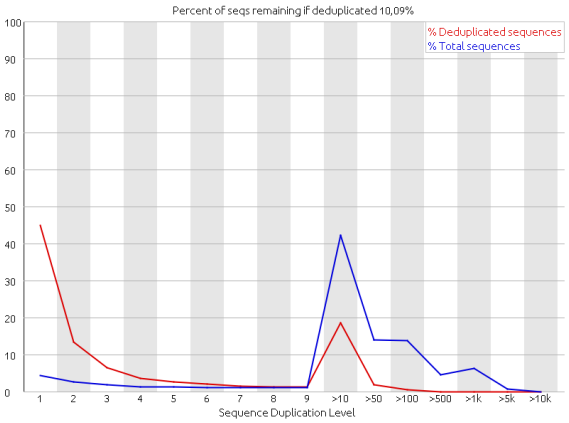
tTreg277 FWD



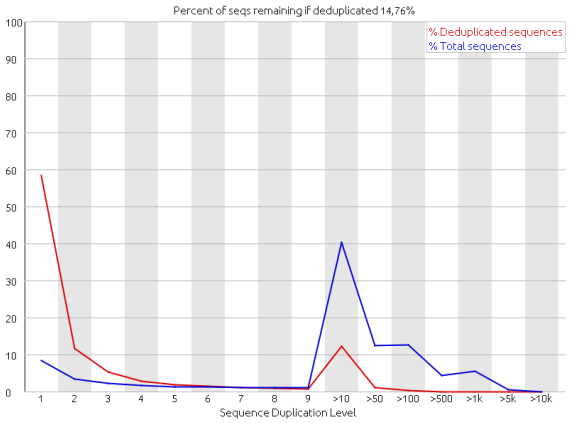
tTreg277 REV



tTconv277 FWD



tTconv277 REV



Annexe 3 Table with the 1047 genes significantly differentially expressed genes between tTreg and tTconv. Fold-change (FC) and FDR values are from edgeR tool. CPM, counts per million.

Gene ID	Ensembl ID	LogFC	FDR	FC	tTconv mean expression (normalized CPM)	tTreg mean expression (normalized CPM)
IL1R1	ENSG00000115594	8,54	7,79E-12	373,23	0,022	8,617
LINC01943	ENSG00000280721	7,22	2,11E-11	148,7	0,032	4,73
IL2RA	ENSG00000134460	6,78	7,71E-67	109,61	1,132	123,757
PCDH7	ENSG00000169851	6,63	6,92E-11	99,08	0,038	3,954
CLEC17A	ENSG00000187912	6,3	2,67E-18	79,05	0,105	8,61
AKR1C6P	ENSG00000151631	6,21	1,40E-09	73,79	0,094	6,894
RELN	ENSG00000189056	6,17	4,20E-13	71,9	0,091	6,507
LRRC32	ENSG00000137507	6,16	1,32E-27	71,7	0,382	27,699
BTNL8	ENSG00000113303	6,12	1,07E-11	69,38	0,074	5,266
SLC37A2	ENSG00000134955	6,1	9,56E-19	68,82	0,529	36,374
CLSTN2	ENSG00000158258	6,07	8,40E-18	67,09	0,078	5,357
CLNK	ENSG00000109684	6,03	7,49E-10	65,29	0,285	18,594
FAT3	ENSG00000165323	5,98	3,03E-11	63,05	0,265	16,75
PNOC	ENSG00000168081	5,82	1,66E-10	56,67	0,043	2,405
LIPN	ENSG00000204020	5,79	6,52E-11	55,42	0,049	2,832
CCR8	ENSG00000179934	5,77	2,08E-22	54,61	0,799	43,836
CPE	ENSG00000109472	5,7	3,78E-07	52,04	0,446	23,348
IL1RL1	ENSG00000115602	5,66	8,74E-11	50,73	0,279	14,224
ATP1A4	ENSG00000132681	5,64	7,35E-10	49,85	0,053	2,72
FN1	ENSG00000115414	5,62	2,55E-51	49,34	1,006	49,523
RP1-207H1.3	ENSG00000231150	5,55	4,92E-45	46,86	0,695	32,654
DNAH8	ENSG00000124721	5,54	1,35E-39	46,42	0,983	45,9
FOXP3	ENSG00000049768	5,52	2,23E-67	46	9,871	454,32
IRF5	ENSG00000128604	5,51	1,42E-40	45,43	0,761	34,641
RAB31	ENSG00000168461	5,46	2,29E-07	43,89	0,092	3,991
LAMA2	ENSG00000196569	5,44	1,62E-15	43,34	0,229	9,972
PIEZO2	ENSG00000154864	5,42	4,87E-10	42,74	0,056	2,419
CCL22	ENSG00000102962	5,34	2,00E-19	40,51	0,422	17,136
HNRNPA1P21	ENSG00000228168	5,33	1,84E-21	40,27	1,024	41,369
AC006042.6	ENSG00000227719	5,31	2,34E-05	39,73	0,251	10,034
TNFRSF11A	ENSG00000141655	5,3	1,96E-08	39,45	0,785	30,986
RP11-462G2.1	ENSG00000237643	5,27	9,78E-08	38,48	0,08	3,045
LRRC4B	ENSG00000131409	5,26	2,06E-11	38,23	0,154	5,88
LMNA	ENSG00000160789	5,16	2,07E-06	35,68	0,319	11,441
TNFRSF9	ENSG00000049249	5,14	1,88E-30	35,35	4,01	141,87
TMPRSS6	ENSG00000187045	5,11	2,38E-08	34,47	0,265	9,187
TNFRSF8	ENSG00000120949	5,09	1,51E-26	34	0,636	21,723
EBI3	ENSG00000105246	5,01	5,68E-13	32,16	0,419	13,527
FCRL3	ENSG00000160856	5,01	8,90E-24	32,15	4,972	160,013
RASGRP4	ENSG00000171777	5	1,39E-10	31,89	0,222	7,122
CXCR5	ENSG00000160683	4,99	3,40E-11	31,84	0,08	2,58
RYR1	ENSG00000196218	4,99	2,63E-10	31,72	0,74	23,57
DIRAS3	ENSG00000162595	4,97	7,96E-16	31,24	0,268	8,429
THSD7A	ENSG00000005108	4,96	2,17E-06	31,04	0,177	5,517
TNFRSF18	ENSG00000186891	4,94	4,78E-20	30,63	1,174	36,052
KCNS3	ENSG00000170745	4,91	1,28E-06	30,09	0,18	5,499
MRC1	ENSG00000260314	4,88	6,75E-20	29,35	0,801	23,611
SLITRK1	ENSG00000178235	4,87	1,35E-22	29,34	1,008	29,721
MAOB	ENSG00000069535	4,84	4,48E-25	28,6	1,934	55,442
PTPN3	ENSG00000070159	4,81	4,74E-29	28	1,794	50,377
CTD-2377D24.4	ENSG00000242407	4,77	1,45E-08	27,35	0,163	4,452
MIR4435-2HG	ENSG00000172965	4,77	9,48E-14	27,25	0,312	8,518

FAM198A	ENSG00000144649	4,72	2,94E-21	26,37	0,514	13,624
FANK1	ENSG00000203780	4,68	5,27E-06	25,64	0,188	4,87
NPW	ENSG00000183971	4,67	2,80E-10	25,43	0,138	3,59
PTCHD1	ENSG00000165186	4,64	1,09E-12	24,94	0,334	8,329
JAKMIP1	ENSG00000152969	4,62	3,27E-11	24,6	0,22	5,479
RNF175	ENSG00000145428	4,61	1,65E-11	24,35	0,119	2,921
SLC14A1	ENSG00000141469	4,59	1,26E-19	24,01	0,579	13,92
IL12RB2	ENSG00000081985	4,58	3,52E-14	23,89	0,765	18,303
CYTOR	ENSG00000222041	4,57	2,16E-24	23,68	0,941	22,372
RP11-1399P15.1	ENSG00000273445	4,55	6,94E-12	23,49	2,147	50,461
HTR4	ENSG00000164270	4,54	3,45E-06	23,22	0,434	10,149
ADPRH	ENSG00000144843	4,53	1,21E-12	23,16	0,447	10,407
PRICKLE1	ENSG00000139174	4,49	2,26E-09	22,5	0,615	13,908
FST	ENSG00000134363	4,49	2,78E-12	22,41	0,286	6,47
PTPRB	ENSG00000127329	4,48	5,33E-06	22,36	0,121	2,702
ICA1	ENSG00000003147	4,48	4,57E-27	22,26	6,671	148,555
ENOX1	ENSG00000120658	4,46	5,79E-08	22,03	0,16	3,562
NAT8L	ENSG00000185818	4,45	8,56E-12	21,87	0,702	15,429
CASQ1	ENSG00000143318	4,42	1,83E-12	21,45	0,188	4,054
TNFRSF13B	ENSG00000240505	4,42	3,70E-06	21,36	0,086	1,86
AKR1C1	ENSG00000187134	4,4	1,09E-05	21,11	2,551	53,885
MCF2L2	ENSG00000053524	4,36	3,35E-08	20,54	0,157	3,271
FCHO2	ENSG00000157107	4,36	5,27E-18	20,49	0,853	17,457
TIGIT	ENSG00000181847	4,32	1,45E-21	20,04	15,536	311,475
IL15RA	ENSG00000134470	4,3	3,14E-05	19,72	0,149	2,965
TOR4A	ENSG00000198113	4,26	4,54E-12	19,1	0,408	7,807
ACTN2	ENSG00000077522	4,24	2,76E-08	18,85	0,674	12,754
CD83	ENSG00000112149	4,24	3,61E-51	18,84	5,973	112,585
FAM46C	ENSG00000183508	4,2	1,44E-07	18,35	1,132	20,794
EDNRB	ENSG00000136160	4,18	1,32E-12	18,14	1,19	21,665
NEB	ENSG00000183091	4,18	9,97E-14	18,09	1,235	22,317
ZG16B	ENSG00000162078	4,17	2,17E-08	18,04	0,216	3,923
CHGB	ENSG00000089199	4,16	5,60E-20	17,82	4,237	75,577
MYO5C	ENSG00000128833	4,15	2,37E-16	17,79	1,58	28,157
ZNF662	ENSG00000182983	4,15	1,77E-29	17,78	1,762	31,337
LMCD1	ENSG00000071282	4,14	3,81E-27	17,66	1,966	34,776
HACD1	ENSG00000165996	4,14	1,16E-12	17,59	0,652	11,526
CTLA4	ENSG00000163599	4,11	7,22E-54	17,29	21,046	364,024
CHRNA6	ENSG00000147434	4,11	2,41E-06	17,27	0,602	10,435
ACTG2	ENSG00000163017	4,11	2,95E-08	17,27	0,278	4,85
CSF1	ENSG00000184371	4,11	0,000158	17,21	1,244	21,449
TNFRSF4	ENSG00000186827	4,1	1,50E-27	17,17	4,901	84,219
MAP3K8	ENSG00000107968	4,08	3,45E-10	16,97	0,198	3,393
XCL1	ENSG00000143184	4,08	1,92E-09	16,93	0,324	5,527
LINC00877	ENSG00000241163	4,08	3,31E-05	16,89	0,123	2,095
KIF26B	ENSG00000162849	4,07	4,01E-10	16,84	8,506	143,162
ZC3H12D	ENSG00000178199	4,04	2,59E-31	16,5	5,218	86,213
PTGS2	ENSG00000073756	4,03	5,55E-05	16,3	0,399	6,543
HMCN1	ENSG00000143341	3,99	4,47E-11	15,94	1,294	20,664
GBP5	ENSG00000154451	3,98	2,10E-25	15,78	11,875	187,513
CD79A	ENSG00000105369	3,97	5,07E-21	15,63	13,186	206,114
HGF	ENSG00000019991	3,97	7,60E-09	15,62	0,272	4,235
SPTBN2	ENSG00000173898	3,96	6,78E-23	15,53	1,984	30,876
CDK14	ENSG00000058091	3,96	6,86E-05	15,53	0,729	11,361
IKZF4	ENSG00000123411	3,95	1,58E-43	15,42	21,376	329,728
DUSP4	ENSG00000120875	3,9	4,56E-37	14,97	31,208	467,286
LAMB1	ENSG00000091136	3,89	7,39E-19	14,87	1,456	21,659
F5	ENSG00000198734	3,87	1,54E-24	14,67	2,323	34,081
PERP	ENSG00000112378	3,87	2,06E-11	14,66	0,602	8,857
ZC3H12C	ENSG00000149289	3,87	7,12E-30	14,59	5,061	73,873
FGFR3	ENSG00000068078	3,84	4,04E-05	14,27	0,18	2,587
CSMD1	ENSG00000183117	3,84	4,29E-08	14,27	0,166	2,37

MYO7A	ENSG00000137474	3,83	2,60E-14	14,22	2,603	37,094
NIPAL2	ENSG00000104361	3,8	1,25E-11	13,94	0,866	12,12
DGCR5	ENSG00000237517	3,78	1,10E-09	13,77	0,98	13,528
UNQ6494	ENSG00000237372	3,77	6,43E-09	13,6	0,591	8,083
HMOX1	ENSG00000100292	3,76	1,56E-06	13,59	0,84	11,448
HS3ST1	ENSG00000002587	3,76	8,79E-07	13,54	0,913	12,392
LAG3	ENSG00000089692	3,76	1,50E-10	13,52	1,352	18,325
ITGB8	ENSG00000105855	3,74	1,13E-05	13,39	0,252	3,385
ST6GALNAC3	ENSG00000184005	3,74	6,03E-26	13,36	2,912	38,942
ANKRD33B	ENSG00000164236	3,73	3,27E-11	13,3	1,073	14,319
OTOF	ENSG00000115155	3,73	1,27E-10	13,22	0,897	11,902
MEOX1	ENSG00000005102	3,72	1,60E-15	13,2	1,835	24,217
LPL	ENSG00000175445	3,72	2,79E-06	13,15	0,302	3,985
HILPDA	ENSG00000135245	3,7	2,63E-20	13,04	3,527	45,988
PIK3AP1	ENSG00000155629	3,66	5,85E-12	12,62	1,576	19,906
TBX21	ENSG00000073861	3,65	8,54E-06	12,57	0,214	2,713
CABLES1	ENSG00000134508	3,65	3,18E-06	12,56	0,336	4,247
LAPTM4B	ENSG00000104341	3,64	3,90E-28	12,47	12,584	156,933
FAM129A	ENSG00000135842	3,64	2,69E-12	12,44	6,993	87,056
CYP7B1	ENSG00000172817	3,64	4,65E-07	12,43	0,431	5,392
ITIH5	ENSG00000123243	3,62	2,93E-17	12,33	2,851	35,17
FABP5	ENSG00000164687	3,62	4,59E-09	12,31	0,491	6,052
PTGER2	ENSG00000125384	3,62	0,0007771	12,29	0,487	6,004
ADAM12	ENSG00000148848	3,58	5,05E-12	11,95	1,988	23,796
PRDM1	ENSG00000057657	3,56	5,81E-06	11,83	2,605	30,837
B3GNT5	ENSG00000176597	3,55	6,34E-07	11,75	0,203	2,414
METTL7A	ENSG00000185432	3,52	1,99E-20	11,51	3,757	43,286
S100A4	ENSG00000196154	3,52	7,76E-25	11,45	14,635	167,657
AKR1C2	ENSG00000151632	3,51	1,01E-05	11,42	0,485	5,561
TBC1D8-AS1	ENSG00000272902	3,51	2,47E-06	11,41	0,284	3,269
ICAM1	ENSG00000090339	3,48	8,33E-05	11,18	0,318	3,582
FBP1	ENSG00000165140	3,45	9,08E-10	10,93	0,933	10,226
RP4-673D20.4	ENSG00000234282	3,42	1,97E-05	10,67	0,142	1,526
WNT10A	ENSG00000135925	3,41	2,44E-09	10,65	1,409	15,044
CAV1	ENSG00000105974	3,41	3,34E-05	10,6	0,836	8,891
XXYLT1-AS2	ENSG00000230266	3,39	5,63E-08	10,49	0,381	4,011
PLXNB2	ENSG00000196576	3,38	1,71E-06	10,44	0,318	3,333
ABCA13	ENSG00000179869	3,38	5,89E-06	10,44	1,363	14,256
FAM124B	ENSG00000124019	3,34	1,46E-08	10,13	0,438	4,445
TUBA3D	ENSG00000075886	3,33	0,0002258	10,07	3,998	40,282
C20orf197	ENSG00000176659	3,32	1,63E-06	10	0,167	1,681
SUOX	ENSG00000139531	3,32	4,41E-25	9,97	5,665	56,531
TMCC3	ENSG00000057704	3,3	5,63E-08	9,84	0,605	5,984
ARAP3	ENSG00000120318	3,29	0,001215	9,75	0,281	2,77
CHRNA2	ENSG00000120903	3,29	3,88E-07	9,75	0,758	7,414
HDAC9	ENSG00000048052	3,27	0,0011618	9,62	0,54	5,221
CREB3L2	ENSG00000182158	3,27	1,42E-25	9,62	39,036	375,429
ATXN1	ENSG00000124788	3,25	1,01E-29	9,52	20,51	195,187
NCF4	ENSG00000100365	3,24	3,84E-08	9,43	1,229	11,605
SDC4	ENSG00000124145	3,23	6,12E-07	9,35	1,92	17,979
TMPRSS13	ENSG00000137747	3,22	0,002854	9,34	0,334	3,132
PHLDA1	ENSG00000139289	3,21	0,0001031	9,22	3,17	29,27
WHRN	ENSG00000095397	3,2	4,95E-08	9,17	1,547	14,203
HPSE	ENSG00000173083	3,18	3,95E-11	9,07	1,548	14,049
WWTR1	ENSG00000018408	3,17	3,45E-06	8,98	0,318	2,859
FGL2	ENSG00000127951	3,15	1,73E-17	8,9	5,458	48,544
PRDM8	ENSG00000152784	3,15	3,52E-10	8,85	1,589	14,098
SLC24A4	ENSG00000140090	3,13	0,0071098	8,78	0,239	2,09
SLIT1	ENSG00000187122	3,13	0,001181	8,73	0,819	7,157
VDR	ENSG00000111424	3,12	1,55E-19	8,72	4,327	37,749
RP11-799D4.1	ENSG00000267744	3,12	6,01E-06	8,69	0,445	3,893
GTSF1L	ENSG00000124196	3,12	8,38E-07	8,67	0,595	5,173

ENPP3	ENSG00000154269	3,11	7,67E-05	8,65	0,287	2,502
CACNB2	ENSG00000165995	3,11	1,37E-17	8,61	3,252	28,004
NCR3	ENSG00000204475	3,11	3,13E-09	8,61	1,622	13,992
OSM	ENSG00000099985	3,1	5,16E-06	8,6	0,233	2,011
SGPP2	ENSG00000163082	3,1	5,99E-10	8,56	1,516	13,012
HLA-DQB1	ENSG00000179344	3,1	4,15E-05	8,56	1,14	9,781
SLC41A2	ENSG00000136052	3,09	4,68E-09	8,5	1,237	10,533
RP11-239E10.2	ENSG00000236846	3,08	0,0008465	8,47	0,33	2,819
FAM43A	ENSG00000185112	3,07	2,79E-08	8,39	2,251	18,909
NEGR1	ENSG00000172260	3,07	3,29E-07	8,39	5,038	42,272
MIR155HG	ENSG00000234883	3,06	7,61E-12	8,37	1,833	15,364
NTRK1	ENSG00000198400	3,06	8,07E-05	8,34	0,379	3,168
MCAM	ENSG00000076706	3,04	0,0003264	8,25	0,649	5,371
IGFBP4	ENSG00000141753	3,04	8,79E-12	8,23	3,844	31,665
CTSZ	ENSG00000101160	3,02	3,85E-14	8,12	6,326	51,443
ZSCAN9	ENSG00000137185	3,02	0,0001773	8,12	0,421	3,444
IL18RAP	ENSG00000115607	3,01	7,36E-05	8,06	0,375	3,03
DUSP10	ENSG00000143507	3	5,68E-20	8,02	10,356	83,124
TSHR	ENSG00000165409	3	3,70E-06	8,02	0,262	2,119
MIR3142HG	ENSG00000253522	3	5,98E-08	7,98	1,387	11,071
LINC00891	ENSG00000281852	2,99	4,87E-12	7,97	3,274	26,086
METTL24	ENSG00000053328	2,99	4,78E-08	7,92	0,973	7,728
INPP5F	ENSG00000198825	2,96	1,60E-16	7,77	17,782	138,26
PRR5L	ENSG00000135362	2,95	8,26E-05	7,73	1,402	10,855
COL9A2	ENSG00000049089	2,95	3,45E-06	7,72	1,654	12,801
BIRC3	ENSG00000023445	2,95	7,47E-19	7,71	301,213	2321,358
TYMP	ENSG00000025708	2,94	6,85E-08	7,69	3,01	23,153
PHLDB2	ENSG00000144824	2,93	2,11E-09	7,63	2,407	18,384
RBMS3	ENSG00000144642	2,93	4,14E-21	7,62	5,815	44,312
C1QTNF6	ENSG00000133466	2,93	5,25E-13	7,61	7,732	58,829
SLC35F2	ENSG00000110660	2,92	8,72E-06	7,59	3,082	23,404
ST7	ENSG00000004866	2,91	1,07E-09	7,52	2,133	16,076
DAZL	ENSG00000092345	2,9	1,58E-05	7,45	0,387	2,904
PREX1	ENSG00000124126	2,9	1,26E-13	7,45	40,549	302,11
TBK1	ENSG00000183735	2,88	1,60E-05	7,39	1,675	12,392
TBC1D8	ENSG00000204634	2,88	5,57E-06	7,37	28,131	207,462
AC006042.8	ENSG00000233264	2,88	8,91E-06	7,36	0,411	3,037
AKAP5	ENSG00000179841	2,87	0,0001028	7,32	1,255	9,195
RASAL1	ENSG00000111344	2,87	1,06E-05	7,31	2,348	17,149
TMPRSS3	ENSG00000160183	2,84	1,43E-06	7,15	0,902	6,467
RP11-798M19.6	ENSG00000272870	2,83	9,59E-07	7,12	0,811	5,781
RP11-405M12.3	ENSG00000278607	2,83	0,0001649	7,12	0,301	2,132
GOLIM4	ENSG00000173905	2,83	5,46E-08	7,1	1,646	11,7
FNDC3B	ENSG00000075420	2,82	6,50E-15	7,08	5,789	40,961
PLN	ENSG00000198523	2,81	4,36E-08	7,03	4,223	29,713
SECTM1	ENSG00000141574	2,81	0,0004185	7,03	1,573	11,071
TSPAN5	ENSG00000168785	2,8	4,85E-29	6,96	17,559	122,161
IGLL3P	ENSG00000206066	2,8	6,08E-05	6,94	1,782	12,39
HAAO	ENSG00000162882	2,8	6,12E-07	6,94	1,048	7,285
CCDC122	ENSG00000151773	2,79	9,46E-05	6,94	0,413	2,876
FRMD6	ENSG00000139926	2,79	5,47E-08	6,91	1,704	11,774
BATF	ENSG00000156127	2,79	1,03E-16	6,9	8,281	57,174
GALNT8	ENSG00000130035	2,78	0,0181707	6,86	0,285	1,951
LINC01480	ENSG00000270164	2,77	2,12E-10	6,82	2,185	14,924
ARNTL2	ENSG00000029153	2,77	0,0001375	6,82	0,791	5,411
RP11-325F22.2	ENSG00000237513	2,76	1,28E-05	6,77	1,469	9,956
TNIP3	ENSG00000050730	2,76	2,06E-11	6,77	14,305	96,864
ALDH4A1	ENSG00000159423	2,75	6,31E-07	6,75	2,333	15,755
RP1-151F17.2	ENSG00000272341	2,75	1,61E-05	6,71	0,478	3,212
FAM129C	ENSG00000167483	2,74	0,000703	6,69	0,808	5,42
PRF1	ENSG00000180644	2,74	1,41E-12	6,67	11,19	74,691
SQLE	ENSG00000104549	2,73	1,12E-10	6,65	23,463	156,158

HECW2	ENSG00000138411	2,73	0,0003215	6,63	1,94	12,878
KANK3	ENSG00000186994	2,73	8,02E-05	6,63	0,683	4,545
AHNAK	ENSG00000124942	2,72	1,02E-12	6,6	54,646	360,415
LSR	ENSG00000105699	2,72	3,68E-06	6,58	2,672	17,614
CSF2RB	ENSG00000100368	2,71	0,0011126	6,54	0,796	5,223
IGHM	ENSG00000211899	2,71	0,0003715	6,54	1,329	8,71
IL18R1	ENSG00000115604	2,7	2,75E-13	6,52	11,389	74,259
PYHIN1	ENSG00000163564	2,7	3,82E-23	6,52	11,287	73,583
FAM135A	ENSG00000082269	2,68	8,52E-25	6,4	12,394	79,308
TTC7B	ENSG00000165914	2,67	6,55E-06	6,38	0,776	4,964
LAIR2	ENSG00000167618	2,67	1,72E-07	6,37	4,36	27,804
CXCR3	ENSG00000186810	2,66	0,0016495	6,32	0,851	5,394
STX11	ENSG00000135604	2,66	6,52E-12	6,32	4,429	28,011
OXER1	ENSG00000162881	2,65	0,0025656	6,28	0,471	2,966
BHLHE40	ENSG00000134107	2,65	6,54E-11	6,27	9,762	61,233
SLC22A15	ENSG00000163393	2,64	0,0004962	6,25	0,451	2,832
RP11-179A10.1	ENSG00000254401	2,63	0,020424	6,21	0,339	2,122
RAB30	ENSG00000137502	2,63	1,51E-26	6,2	26,657	165,322
LINC01671	ENSG00000225431	2,61	8,76E-05	6,12	0,39	2,386
IZUMO4	ENSG00000099840	2,61	6,00E-05	6,1	1,029	6,29
IL2RB	ENSG00000100385	2,6	8,35E-21	6,05	153,119	927,102
TLR7	ENSG00000196664	2,59	0,001738	6,03	1,759	10,622
CLDN16	ENSG00000113946	2,59	7,86E-05	6	0,391	2,362
HHIP	ENSG00000164161	2,57	0,001792	5,92	0,623	3,701
MAF	ENSG00000178573	2,56	7,46E-05	5,89	6,167	36,352
MYO1F	ENSG00000142347	2,55	0,0001323	5,86	1,386	8,137
RGS16	ENSG00000143333	2,55	6,01E-08	5,84	2,967	17,337
PTPN14	ENSG00000152104	2,54	1,84E-10	5,82	8,769	51,026
TLR2	ENSG00000137462	2,52	0,0024299	5,74	0,388	2,237
ERICD	ENSG00000280303	2,52	0,0001017	5,72	0,568	3,254
NCF2	ENSG00000116701	2,5	0,0108323	5,66	0,519	2,948
RP11-568N6.1	ENSG00000260101	2,5	0,0011308	5,65	0,542	3,077
ATP1B1	ENSG00000143153	2,49	2,21E-08	5,62	7,904	44,401
LINC01281	ENSG00000235304	2,48	2,28E-12	5,57	6,467	36,074
NETO2	ENSG00000171208	2,47	0,0002695	5,55	0,987	5,489
SSH3	ENSG00000172830	2,46	8,94E-05	5,51	0,914	5,047
FLVCR2	ENSG00000119686	2,45	0,0001028	5,48	0,929	5,102
KIAA1614	ENSG00000135835	2,45	8,90E-06	5,45	1,745	9,519
TNFRSF1B	ENSG00000028137	2,44	9,25E-19	5,42	193,521	1048,509
C3AR1	ENSG00000171860	2,44	0,0029655	5,42	0,797	4,322
HAPLN3	ENSG00000140511	2,43	3,20E-16	5,4	15,162	81,895
ACOT9	ENSG00000123130	2,42	8,63E-09	5,37	3,101	16,656
METRNL	ENSG00000176845	2,41	0,0249914	5,31	0,847	4,511
A2MP1	ENSG00000256069	2,41	0,0034441	5,31	0,54	2,867
SESN1	ENSG00000080546	2,4	3,52E-10	5,29	114,886	607,607
TSPAN13	ENSG00000106537	2,39	0,0003345	5,24	0,8	4,197
AC079610.2	ENSG00000196096	2,38	0,0064338	5,22	0,379	1,988
A2M	ENSG00000175899	2,38	0,0003668	5,21	0,856	4,469
AC074289.1	ENSG00000225889	2,37	0,0001264	5,18	0,593	3,069
FBN1	ENSG00000166147	2,37	1,53E-05	5,18	2,767	14,355
MYO1C	ENSG00000197879	2,37	3,13E-06	5,18	3,059	15,828
REEP2	ENSG00000132563	2,37	8,10E-05	5,16	1,157	5,975
LACC1	ENSG00000179630	2,36	0,0044225	5,15	0,795	4,11
GNG8	ENSG00000167414	2,36	1,44E-06	5,14	15,545	79,91
RASGRF1	ENSG00000058335	2,36	0,0008569	5,13	0,32	1,649
MAP7	ENSG00000135525	2,35	7,68E-07	5,09	3,652	18,619
GFI1	ENSG00000162676	2,35	3,22E-06	5,08	11,051	56,185
RAP1GAP2	ENSG00000132359	2,35	3,48E-06	5,08	2,27	11,548
TRIB1	ENSG00000173334	2,34	4,60E-12	5,07	23,684	120,03
GPR161	ENSG00000143147	2,33	0,0042935	5,04	0,494	2,506
LINC00963	ENSG00000204054	2,33	3,64E-09	5,03	7,782	39,171
CXorf21	ENSG00000120280	2,31	1,43E-06	4,95	2,731	13,54

PHTF1	ENSG00000116793	2,3	2,57E-05	4,92	1,651	8,117
TRAF1	ENSG00000056558	2,28	8,19E-15	4,87	21,485	104,589
GAB1	ENSG00000109458	2,28	0,0008641	4,84	0,451	2,183
DUSP16	ENSG00000111266	2,27	4,03E-19	4,83	84,478	408,095
NFKBIZ	ENSG00000144802	2,27	9,22E-11	4,82	24,397	117,501
MARCH3	ENSG00000173926	2,27	1,22E-07	4,81	3,693	17,763
DTX1	ENSG00000135144	2,25	9,16E-09	4,77	5,437	25,93
SLC35G3	ENSG00000164729	2,25	0,0060202	4,75	0,454	2,167
ADRB1	ENSG00000043591	2,25	0,0104798	4,74	0,965	4,592
PZP	ENSG00000126838	2,24	0,0171602	4,73	0,747	3,533
ZBP1	ENSG00000124256	2,23	2,18E-05	4,68	2,379	11,154
ADGRV1	ENSG00000164199	2,23	0,0377221	4,68	0,443	2,071
PARVB	ENSG00000188677	2,22	0,0010152	4,67	1,671	7,813
MICAL3	ENSG00000243156	2,21	0,0003869	4,63	1,268	5,875
MPST	ENSG00000128309	2,2	0,0002409	4,61	1,552	7,159
LFNG	ENSG00000106003	2,2	5,81E-16	4,59	19,359	88,905
HLA-L	ENSG00000243753	2,2	0,0056769	4,59	0,741	3,399
NRIP1	ENSG00000180530	2,2	8,08E-14	4,59	139,669	641,109
LINC00426	ENSG00000238121	2,2	9,78E-08	4,58	3,753	17,187
SMS	ENSG00000102172	2,18	6,08E-10	4,54	43,76	198,481
LGALSL	ENSG00000119862	2,18	0,0037654	4,52	0,555	2,517
SETBP1	ENSG00000152217	2,16	1,55E-08	4,48	4,517	20,247
SMPD3	ENSG00000103056	2,16	2,54E-08	4,48	26,903	120,473
DAB2IP	ENSG00000136848	2,15	8,82E-05	4,43	2,506	11,118
FAS	ENSG00000026103	2,15	0,0004209	4,43	4,913	21,761
PMAIP1	ENSG00000141682	2,14	0,0005072	4,4	10,807	47,567
KCNQ3	ENSG00000184156	2,12	0,0010003	4,34	0,871	3,786
GLCCI1	ENSG00000106415	2,1	1,64E-13	4,29	82,816	355,17
GADD45G	ENSG00000130222	2,08	7,12E-06	4,22	7,252	30,622
IRAK3	ENSG00000090376	2,07	0,0105133	4,2	1,235	5,188
SLCO3A1	ENSG00000176463	2,07	7,54E-06	4,19	3,113	13,042
GCNT2	ENSG00000111846	2,07	5,55E-05	4,19	2,152	9,018
FABP5P7	ENSG00000234964	2,07	0,0067714	4,19	1,674	7,021
UBE2QL1	ENSG00000215218	2,06	0,0043751	4,18	0,492	2,066
GBP4	ENSG00000162654	2,06	5,29E-14	4,18	26,58	111,089
TRPC1	ENSG00000144935	2,06	0,0015986	4,18	0,662	2,769
RP11-20D14.6	ENSG00000249790	2,06	0,0057278	4,17	0,87	3,637
MAP3K5	ENSG00000197442	2,05	5,87E-06	4,15	6,44	26,707
BISPR	ENSG00000282851	2,05	4,16E-05	4,14	3,486	14,441
FAM172BP	ENSG00000175841	2,04	0,0025762	4,12	1,231	5,075
SELP	ENSG00000174175	2,04	0,0206034	4,12	0,721	2,967
PDE4A	ENSG00000065989	2,04	0,0427353	4,11	1,222	5,036
SLC43A1	ENSG00000149150	2,04	3,24E-05	4,11	2,634	10,827
ALCAM	ENSG00000170017	2,04	0,0001487	4,1	2,072	8,508
LINC00598	ENSG00000215483	2,03	0,000321	4,08	1,427	5,831
CPXM1	ENSG00000088882	2,02	1,87E-05	4,05	11,547	46,822
ADCY3	ENSG00000138031	2,02	8,57E-11	4,05	11,829	47,916
SYT11	ENSG00000132718	2,01	4,53E-05	4,02	12,496	50,276
CEP170B	ENSG00000099814	2,01	0,0004011	4,02	1,965	7,904
POU2F2	ENSG00000028277	2	1,45E-07	4,01	9,401	37,722
TCEAL3	ENSG00000196507	2	0,0052594	4,01	1,286	5,15
GRIN3A	ENSG00000198785	2	0,0369914	3,99	0,511	2,038
CYTH3	ENSG00000008256	1,99	6,24E-09	3,98	16,921	67,349
IRF8	ENSG00000140968	1,99	1,88E-06	3,96	6,102	24,17
SCN9A	ENSG00000169432	1,98	0,0011438	3,94	3,431	13,503
PLAGL1	ENSG00000118495	1,97	3,17E-05	3,93	3,585	14,09
NTNG2	ENSG00000196358	1,95	0,0280754	3,87	0,883	3,427
DUSP5	ENSG00000138166	1,95	7,37E-06	3,86	6,745	26,063
CCDC81	ENSG00000149201	1,94	0,004871	3,85	0,528	2,04
ZC2HC1A	ENSG00000104427	1,94	0,0111661	3,85	1,067	4,106
IKZF2	ENSG00000030419	1,94	2,60E-09	3,85	324,394	1247,316
TIGD2	ENSG00000180346	1,94	3,70E-12	3,84	19,404	74,602

HES4	ENSG00000188290	1,94	0,0090946	3,84	0,576	2,224
CTTNBP2NL	ENSG00000143079	1,94	0,0119336	3,84	3,033	11,663
TBC1D4	ENSG00000136111	1,94	3,54E-09	3,83	88,975	340,664
SNX33	ENSG00000173548	1,94	0,000169	3,82	2,247	8,6
EXD3	ENSG00000187609	1,93	0,0430549	3,82	0,871	3,335
PAM	ENSG00000145730	1,93	1,90E-07	3,82	6,381	24,35
RP5-1031D4.2	ENSG00000232591	1,93	0,0041323	3,81	1,166	4,444
BCL2A1	ENSG00000140379	1,93	4,25E-05	3,8	4,318	16,434
KRT17P8	ENSG00000256937	1,93	0,0143039	3,8	0,53	2,012
LGALS9	ENSG00000168961	1,92	5,89E-08	3,78	29,787	112,739
CXXC5	ENSG00000171604	1,91	1,41E-08	3,77	8,089	30,488
MDGA1	ENSG00000112139	1,91	0,0032342	3,75	0,865	3,246
DOCK5	ENSG00000147459	1,91	0,00903	3,75	1,204	4,515
GNB4	ENSG00000114450	1,9	1,63E-08	3,74	10,302	38,566
GADD45B	ENSG00000099860	1,9	3,29E-05	3,73	20,157	75,247
IL10RA	ENSG00000110324	1,9	2,13E-07	3,72	35,495	132,085
STAM	ENSG00000136738	1,89	1,41E-11	3,72	35,151	130,632
HLA-DPB1	ENSG00000223865	1,89	0,001499	3,71	2,165	8,031
NABP1	ENSG00000173559	1,88	5,16E-08	3,69	14,12	52,036
GK	ENSG00000198814	1,88	3,53E-06	3,67	5,625	20,669
CDH1	ENSG00000039068	1,88	0,0006365	3,67	3,201	11,751
LAYN	ENSG00000204381	1,87	1,57E-13	3,67	32,481	119,087
MAGEH1	ENSG00000187601	1,86	1,22E-12	3,63	131,178	476,075
NDNF	ENSG00000173376	1,85	1,51E-07	3,61	58,167	210,05
THEM5	ENSG00000196407	1,85	0,0015449	3,6	1,77	6,382
KCNK5	ENSG00000164626	1,84	0,0001031	3,57	12,641	45,174
TLR5	ENSG00000187554	1,82	3,21E-05	3,54	7,605	26,9
RP11-367G6.3	ENSG00000272053	1,81	0,0078896	3,52	1,121	3,94
RGS1	ENSG00000090104	1,81	1,80E-06	3,52	301,646	1060,523
CPM	ENSG00000135678	1,81	0,020766	3,51	2,328	8,18
SPATS2L	ENSG00000196141	1,81	3,34E-09	3,51	25,636	89,946
NRSN2	ENSG00000125841	1,81	0,0002592	3,51	2,867	10,062
SWAP70	ENSG00000133789	1,81	1,45E-11	3,5	26,708	93,603
AC011893.3	ENSG00000226806	1,81	4,53E-05	3,5	3,667	12,82
PCBD1	ENSG00000166228	1,8	0,0054936	3,48	1,215	4,237
ZBTB38	ENSG00000177311	1,79	3,47E-07	3,47	53,649	185,906
GRK3	ENSG00000100077	1,79	7,19E-12	3,46	45,621	157,787
LTA	ENSG00000226979	1,78	8,06E-09	3,44	16,202	55,758
HVCN1	ENSG00000122986	1,78	8,86E-06	3,43	8,871	30,425
DLGAP1-AS1	ENSG00000177337	1,78	0,00327	3,43	3,457	11,857
MERTK	ENSG00000153208	1,78	0,0083471	3,43	0,718	2,466
RAB33A	ENSG00000134594	1,78	6,52E-08	3,43	14,96	51,296
AL450992.2	ENSG00000234614	1,77	0,0027165	3,42	1,443	4,942
GCLC	ENSG00000001084	1,76	5,64E-09	3,4	19,358	65,776
ST8SIA1	ENSG00000111728	1,76	0,0464281	3,39	2,981	10,117
TNIP1	ENSG00000145901	1,76	4,37E-12	3,39	103,734	351,591
AJ006998.2	ENSG00000229425	1,76	0,0045569	3,39	1,31	4,439
TFRC	ENSG00000072274	1,76	5,02E-06	3,38	70,502	238,415
HECTD2	ENSG00000165338	1,76	0,0012616	3,38	2,479	8,381
KLHDC7B	ENSG00000130487	1,75	0,0041313	3,37	3,279	11,049
ARHGAP18	ENSG00000146376	1,75	4,82E-05	3,36	5,602	18,832
PAOX	ENSG00000148832	1,73	0,0004325	3,33	2,954	9,835
ETNK2	ENSG00000143845	1,73	0,0230926	3,33	0,505	1,687
ABCB10	ENSG00000135776	1,73	4,37E-12	3,33	51,359	170,8
GABARAPL1	ENSG00000139112	1,73	6,65E-08	3,31	9,656	31,974
CCDC171	ENSG00000164989	1,73	0,0002236	3,31	8,493	28,116
LINC00888	ENSG00000240024	1,72	0,0117523	3,3	0,947	3,128
GADD45A	ENSG00000116717	1,72	4,24E-05	3,3	8,151	26,875
IQCH-AS1	ENSG00000259673	1,72	0,0091364	3,29	1,218	4,014
KLF6	ENSG00000067082	1,72	0,000321	3,29	253,904	835,749
LINC01619	ENSG00000257242	1,71	9,69E-05	3,28	4,473	14,655
MFSD2A	ENSG00000168389	1,71	0,0069715	3,27	1,458	4,77

P2RY1	ENSG00000169860	1,71	8,47E-09	3,27	19,737	64,463
CD8A	ENSG00000153563	1,7	0,0090946	3,25	19,469	63,318
ANXA4	ENSG00000196975	1,7	0,0006443	3,25	4,09	13,272
ERRFI1	ENSG00000116285	1,69	0,013076	3,23	1,015	3,277
KIAA1211L	ENSG00000196872	1,68	0,0214807	3,2	1,165	3,73
DLGAP1-AS2	ENSG00000262001	1,68	0,02319	3,2	0,783	2,5
PRRT3	ENSG00000163704	1,67	0,0047329	3,19	5,855	18,696
ANKRD6	ENSG00000135299	1,65	0,0128446	3,14	1,615	5,071
BST2	ENSG00000130303	1,65	0,0013909	3,14	31,507	98,927
PON2	ENSG00000105854	1,65	0,0010196	3,13	3,528	11,047
SLC9B2	ENSG00000164038	1,64	0,0030581	3,13	4,131	12,915
RASGRF2	ENSG00000113319	1,64	1,77E-08	3,12	31,31	97,754
RP11-712B9.2	ENSG00000245552	1,64	0,0071826	3,12	2,171	6,779
SLC25A29	ENSG00000197119	1,64	0,0267499	3,12	0,566	1,77
EFHC1	ENSG00000096093	1,64	0,0007746	3,11	3,4	10,59
FAM110A	ENSG00000125898	1,63	6,13E-06	3,1	14,118	43,773
AF127936.9	ENSG00000235609	1,63	0,0136498	3,1	1,369	4,248
CD79B	ENSG00000007312	1,62	5,45E-06	3,08	19,497	59,976
DGKH	ENSG00000102780	1,61	3,37E-10	3,06	73,715	225,381
C1RL-AS1	ENSG00000205885	1,61	0,0358647	3,05	1,991	6,069
SPINT1	ENSG00000166145	1,6	0,0158429	3,04	1,093	3,32
SYP	ENSG00000102003	1,6	0,0008492	3,03	3,904	11,816
IGSF3	ENSG00000143061	1,59	0,0297061	3,02	1,536	4,644
PHLPP1	ENSG00000081913	1,59	0,0001301	3	9,863	29,604
RP11-1348G14.4	ENSG00000251417	1,58	0,0343501	3	0,665	1,996
SIRPB3P	ENSG00000223750	1,58	0,0383359	3	1,287	3,863
KLRG1	ENSG00000139187	1,58	1,26E-05	2,99	13,746	41,157
SH3TC1	ENSG00000125089	1,58	0,0041313	2,98	33,763	100,703
DHRS7	ENSG00000100612	1,57	5,00E-06	2,97	16,751	49,775
REL	ENSG00000162924	1,57	1,51E-05	2,96	57,698	171,052
EOGT	ENSG00000163378	1,57	0,007424	2,96	3,081	9,129
HIP1	ENSG00000127946	1,56	0,0257483	2,95	1,921	5,66
CHST2	ENSG00000175040	1,56	1,06E-07	2,95	159,596	470,084
ARRDC3	ENSG00000113369	1,55	6,52E-08	2,94	136,878	402,149
PRDM11	ENSG00000019485	1,55	0,0091586	2,94	3,825	11,226
NEK3	ENSG00000136098	1,55	0,0149031	2,93	2,678	7,854
CD74	ENSG00000019582	1,55	1,67E-07	2,93	294,867	864,07
FCHSD1	ENSG00000197948	1,55	0,0006365	2,93	7,536	22,073
KIAA1324	ENSG00000116299	1,53	0,0011929	2,89	9,971	28,822
WASF1	ENSG00000112290	1,52	1,05E-05	2,87	13,675	39,216
PPP1R12B	ENSG00000077157	1,52	8,24E-08	2,86	35,241	100,82
OTUD7B	ENSG00000264522	1,52	0,0003671	2,86	5,478	15,66
IRF4	ENSG00000137265	1,51	2,08E-08	2,86	96,699	276,11
MAN1A1	ENSG00000111885	1,51	1,07E-07	2,85	44,469	126,584
IKZF3	ENSG00000161405	1,51	4,63E-07	2,84	194,438	551,975
GPR155	ENSG00000163328	1,5	6,32E-08	2,82	35,361	99,736
CLIP2	ENSG00000106665	1,5	0,0001753	2,82	8,193	23,119
SESN3	ENSG00000149212	1,5	3,62E-06	2,82	169,333	477,577
ABAT	ENSG00000183044	1,5	4,32E-06	2,82	30,901	87,123
SLC9A7	ENSG00000065923	1,49	7,96E-06	2,81	12,294	34,596
MEGF6	ENSG00000162591	1,49	0,0071826	2,81	11,796	33,145
B4GALT5	ENSG00000158470	1,49	0,0002792	2,8	9,834	27,586
MTSS1	ENSG00000170873	1,48	1,62E-07	2,8	44,006	123,084
CH507-338C24.1	ENSG00000277991	1,48	0,0442572	2,79	0,695	1,944
SLC43A2	ENSG00000167703	1,48	0,001403	2,79	5,465	15,264
MEI1	ENSG00000167077	1,48	0,0001264	2,79	8,154	22,739
INPP1	ENSG00000151689	1,47	0,0066739	2,78	3,065	8,514
BCL3	ENSG00000069399	1,47	0,0073176	2,78	11,545	32,048
RGPD2	ENSG00000185304	1,47	0,0002976	2,77	15,04	41,659
CD58	ENSG00000116815	1,47	0,0249914	2,76	4,885	13,512
USP51	ENSG00000247746	1,47	0,0055706	2,76	3,616	10,002
REC8	ENSG00000100918	1,46	0,0006443	2,76	10,679	29,453

BTN2A3P	ENSG00000124549	1,46	0,0244597	2,76	1,632	4,502
SLC2A13	ENSG00000151229	1,46	0,0297998	2,75	2,006	5,525
LYST	ENSG00000143669	1,46	4,89E-07	2,75	162,071	445,046
TJP2	ENSG00000119139	1,45	7,44E-05	2,74	13,039	35,734
DRP2	ENSG00000102385	1,45	0,0064576	2,74	4,681	12,816
NFKB2	ENSG00000077150	1,45	6,20E-09	2,73	39,62	108,251
TP53INP1	ENSG00000164938	1,45	2,42E-07	2,73	42,722	116,531
C2orf16	ENSG00000221843	1,45	0,0479844	2,73	1,092	2,979
CFLAR	ENSG00000003402	1,44	6,90E-06	2,71	130,609	353,909
PHEX	ENSG00000102174	1,44	0,02189	2,7	5,562	15,053
NR4A3	ENSG00000119508	1,43	8,55E-06	2,7	47,351	127,735
ZNF563	ENSG00000188868	1,43	0,0011561	2,7	4,96	13,378
HSPA1B	ENSG00000204388	1,43	0,0290568	2,69	31,951	86,037
FUCA2	ENSG00000001036	1,43	0,0199151	2,69	3,278	8,825
FBXO2	ENSG00000116661	1,42	0,0430932	2,68	0,918	2,469
MGST2	ENSG00000085871	1,42	0,0270784	2,68	4,505	12,092
FGFRL1	ENSG00000127418	1,42	0,0026838	2,68	10,114	27,094
CTTN	ENSG00000085733	1,42	0,0024431	2,68	8,546	22,893
PTGIR	ENSG00000160013	1,42	0,0004175	2,68	20,131	53,898
SLC29A1	ENSG00000112759	1,42	0,0116798	2,68	3,48	9,316
RORA	ENSG00000069667	1,42	0,0006283	2,68	55,716	149,046
ATP2C1	ENSG00000017260	1,42	1,46E-07	2,67	85,227	227,557
PHC1	ENSG00000111752	1,41	1,28E-07	2,67	77,8	207,371
NFKBIE	ENSG00000146232	1,41	0,0015128	2,66	6,196	16,466
PHKB	ENSG00000102893	1,4	6,52E-08	2,65	95,692	253,249
APLP2	ENSG00000084234	1,4	1,11E-06	2,65	62,108	164,292
NAALADL1	ENSG00000168060	1,4	0,0189283	2,64	2,695	7,124
IER5	ENSG00000162783	1,4	0,0018709	2,64	38,98	102,951
ZDHHC23	ENSG00000184307	1,4	0,0480625	2,64	2,151	5,675
PHC1P1	ENSG00000179899	1,39	3,62E-05	2,62	12,266	32,198
LCOR	ENSG00000196233	1,38	1,52E-07	2,61	68,821	179,533
ARHGEF12	ENSG00000196914	1,38	0,0041313	2,61	10,191	26,575
ENOX2	ENSG00000165675	1,37	5,23E-07	2,59	29,979	77,583
B3GNT2	ENSG00000170340	1,36	3,30E-05	2,57	69,397	178,459
AHR	ENSG00000106546	1,36	5,32E-05	2,57	15,611	40,116
DNPH1	ENSG00000112667	1,36	6,65E-07	2,57	26,916	69,158
NFIA	ENSG00000162599	1,35	0,0401724	2,54	2,992	7,608
PDLIM7	ENSG00000196923	1,35	0,0088036	2,54	4,238	10,781
GPA33	ENSG00000143167	1,35	0,0227203	2,54	30,841	78,4
M6PR	ENSG00000003056	1,34	1,50E-05	2,53	139,647	353,551
VASH1	ENSG00000071246	1,34	0,0060677	2,53	5,311	13,446
RAB37	ENSG00000172794	1,34	3,71E-05	2,53	23,704	59,984
GPRIN3	ENSG00000185477	1,34	6,03E-06	2,53	94,588	239,295
STAT4	ENSG00000138378	1,34	0,023572	2,53	10,786	27,254
LAMP3	ENSG00000078081	1,33	6,01E-06	2,52	22,652	57,068
TMEM120A	ENSG00000189077	1,33	0,0038743	2,52	14,153	35,601
CTPS2	ENSG00000047230	1,32	0,001499	2,5	6,635	16,616
TTC7A	ENSG00000068724	1,32	5,93E-07	2,5	80,116	200,219
IL4R	ENSG00000077238	1,32	0,0001858	2,5	109,469	273,212
PTMS	ENSG00000159335	1,32	0,0012928	2,49	7,937	19,792
CTD-3092A11.2	ENSG00000270055	1,31	0,0006579	2,48	13,141	32,63
MTHFD2	ENSG00000065911	1,31	0,0003798	2,48	11,02	27,296
TIFA	ENSG00000145365	1,3	0,0017311	2,47	22,869	56,504
LRIG1	ENSG00000144749	1,3	5,87E-06	2,47	29,026	71,611
MAP3K14	ENSG00000006062	1,3	0,0001143	2,46	28,454	70,087
PELI1	ENSG00000197329	1,29	4,47E-07	2,45	90,348	221,095
ARSD	ENSG00000006756	1,29	0,0154813	2,45	5,824	14,245
PLEKHG1	ENSG00000120278	1,29	2,13E-05	2,44	29,524	71,951
SERPINB9	ENSG00000170542	1,28	4,50E-06	2,44	94,271	229,589
FOXO1	ENSG00000150907	1,28	1,10E-05	2,43	175,761	427,106
LINC-PINT	ENSG00000231721	1,28	0,0481879	2,42	2,524	6,117
CHST11	ENSG00000171310	1,28	0,0012593	2,42	65,525	158,646

NAGA	ENSG00000198951	1,27	0,0001806	2,41	16,097	38,779
TGIF1	ENSG00000177426	1,26	1,77E-05	2,4	18,686	44,863
GBP2	ENSG00000162645	1,26	1,36E-06	2,4	187,434	449,702
TMEM64	ENSG00000180694	1,26	0,0007405	2,4	11,852	28,44
FLNA	ENSG00000196924	1,26	2,47E-06	2,4	307,646	737,93
TNFSF13B	ENSG00000102524	1,26	0,0007322	2,4	9,86	23,633
ZC3H12A	ENSG00000163874	1,26	0,0014975	2,39	11,592	27,729
CAPN2	ENSG00000162909	1,25	0,0002316	2,39	24,023	57,312
CCDC50	ENSG00000152492	1,24	4,85E-06	2,37	27,873	65,939
ADAT2	ENSG00000189007	1,24	0,0010953	2,36	10,63	25,069
CRTAP	ENSG00000170275	1,24	5,91E-06	2,36	30,646	72,214
FANCL	ENSG00000115392	1,23	0,0407367	2,35	4,07	9,565
PCED1B	ENSG00000179715	1,22	0,0009929	2,34	17,813	41,635
SAMSN1	ENSG00000155307	1,22	0,0002702	2,34	30,734	71,789
CACNA2D2	ENSG00000007402	1,22	0,0052256	2,32	7,219	16,77
DST	ENSG00000151914	1,21	0,0222325	2,32	10,154	23,53
PIM2	ENSG00000102096	1,21	3,10E-05	2,31	141,985	328,086
ZC3H7A	ENSG00000122299	1,21	1,61E-05	2,31	127,714	294,541
TMED8	ENSG00000100580	1,2	0,0024182	2,31	56,093	129,307
SHMT2	ENSG00000182199	1,2	1,55E-05	2,3	60,992	140,28
NEAT1	ENSG00000245532	1,2	0,0321987	2,3	22,586	51,911
RGPD1	ENSG00000187627	1,2	0,0043054	2,29	19,176	44,002
CD82	ENSG00000085117	1,2	8,33E-05	2,29	38,185	87,533
RP11-717F1.2	ENSG00000274333	1,2	0,0108994	2,29	6,057	13,874
UST	ENSG00000111962	1,19	0,0394452	2,29	4,401	10,057
SEMA7A	ENSG00000138623	1,19	0,0447323	2,28	4,257	9,726
RAB11FIP1	ENSG00000156675	1,19	0,0003372	2,28	18,174	41,485
ARHGAP19	ENSG00000213390	1,19	3,72E-05	2,27	38,99	88,67
FHL3	ENSG00000183386	1,19	0,0146047	2,27	7,162	16,289
PLAGL2	ENSG00000126003	1,18	0,0003425	2,27	42,771	97,076
WHAMMP3	ENSG00000276141	1,18	0,0190203	2,27	7,986	18,116
IFNAR2	ENSG00000159110	1,18	0,0001001	2,26	26,245	59,386
PPP1R3F	ENSG00000049769	1,17	0,0350939	2,26	7,993	18,043
NFE2L3	ENSG00000050344	1,17	0,011054	2,26	13,522	30,503
EPHX2	ENSG00000120915	1,17	3,84E-05	2,25	106,12	239,172
CASK	ENSG00000147044	1,17	4,57E-05	2,25	25,943	58,456
PLXNC1	ENSG00000136040	1,17	0,023357	2,25	8,889	20,016
RELB	ENSG00000104856	1,17	0,0002872	2,25	16,682	37,539
ITGB7	ENSG00000139626	1,17	0,0058009	2,25	13,188	29,669
IL21R	ENSG00000103522	1,16	0,0002257	2,24	70,253	157,171
C7orf50	ENSG00000146540	1,16	0,0321987	2,23	10,404	23,218
ITFG1	ENSG00000129636	1,15	8,28E-05	2,23	44,725	99,513
RP11-632K20.7	ENSG00000223509	1,15	0,0108846	2,22	8,841	19,58
BCL2L1	ENSG00000171552	1,15	7,44E-05	2,21	28,903	63,984
CPEB3	ENSG00000107864	1,15	0,0441821	2,21	4,5	9,961
SGPP1	ENSG00000126821	1,14	0,0006578	2,21	70,913	156,782
PSD	ENSG00000059915	1,14	0,0034682	2,21	10,394	22,955
TCEAL4	ENSG00000133142	1,14	0,0177074	2,21	8,521	18,797
SETD7	ENSG00000145391	1,14	0,0024037	2,2	12,299	27,092
ABCA2	ENSG00000107331	1,14	0,0066739	2,2	30,516	67,142
TNFSF12	ENSG00000239697	1,13	0,0029627	2,19	13,333	29,147
TXNDC11	ENSG00000153066	1,13	6,88E-05	2,18	53,969	117,776
HABP4	ENSG00000130956	1,13	0,031272	2,18	8,159	17,807
SLAMF6	ENSG00000162739	1,13	0,000296	2,18	67,55	147,347
CD27	ENSG00000139193	1,12	0,0001844	2,17	125,062	271,799
HEATR5A	ENSG00000129493	1,12	0,0398871	2,17	6,655	14,42
IDH1	ENSG00000138413	1,11	0,035468	2,16	5,697	12,302
CH507-24F1.2	ENSG00000275496	1,11	0,0333809	2,16	5,462	11,799
ATOX1	ENSG00000177556	1,11	0,0013366	2,15	17,535	37,764
TNFAIP3	ENSG00000118503	1,1	0,0001038	2,15	842,792	1812,297
NINJ1	ENSG00000131669	1,1	0,0013206	2,15	15,843	34,047
DNMBP	ENSG00000107554	1,1	0,0335551	2,14	12,061	25,86

APOBEC3C	ENSG00000244509	1,1	0,0010953	2,14	24,179	51,787
SLC6A6	ENSG00000131389	1,1	0,0001863	2,14	27,563	58,934
IVNS1ABP	ENSG00000116679	1,1	0,0002389	2,14	131,015	279,882
SP140	ENSG00000079263	1,09	0,0005582	2,13	23,211	49,479
EFHD2	ENSG00000142634	1,09	0,000454	2,13	33,183	70,733
NFKBIA	ENSG00000100906	1,09	0,0006563	2,13	195,446	416,377
HAVCR2	ENSG00000135077	1,09	0,0492807	2,12	8,952	19,002
VOPP1	ENSG00000154978	1,09	0,0001495	2,12	104,432	221,661
PTPRJ	ENSG00000149177	1,08	0,0003822	2,12	173,005	366,495
PRR5	ENSG00000186654	1,08	0,0401202	2,11	9,355	19,775
SEC14L1	ENSG00000129657	1,08	5,55E-05	2,11	66,766	140,733
P2RY10	ENSG00000078589	1,07	0,0049936	2,1	22,415	47,178
TP53I11	ENSG00000175274	1,07	0,0062645	2,1	17,813	37,482
TRAF2	ENSG00000127191	1,07	0,0013132	2,1	21,481	45,098
PCK2	ENSG00000100889	1,07	0,00154	2,1	20,66	43,361
ZNF704	ENSG00000164684	1,07	0,0103505	2,1	14,877	31,22
IFIT2	ENSG00000119922	1,07	0,0008384	2,09	23,089	48,362
EEF2K	ENSG00000103319	1,06	0,0094277	2,09	12,406	25,923
CBX7	ENSG00000100307	1,06	0,0053307	2,08	22,61	46,993
CMTM7	ENSG00000153551	1,05	0,0145516	2,07	74,243	153,859
LRRC8D	ENSG00000171492	1,05	0,0028697	2,07	131,614	272,46
FBLN7	ENSG00000144152	1,05	0,009338	2,07	48,595	100,367
E2F3	ENSG00000112242	1,04	0,0290567	2,06	13,008	26,76
STARD4	ENSG00000164211	1,04	0,0047411	2,05	17,772	36,511
STX7	ENSG00000079950	1,04	0,0044048	2,05	23,068	47,293
ASNS	ENSG00000070669	1,03	0,0145986	2,05	11,528	23,587
SAT1	ENSG00000130066	1,03	0,0136454	2,04	54,82	111,936
ITGAL	ENSG00000005844	1,02	0,00027	2,03	320,669	652,459
KSR1	ENSG00000141068	1,02	0,0118492	2,03	10,233	20,784
HS3ST3B1	ENSG00000125430	1,02	0,0043129	2,03	23,556	47,778
LY75	ENSG00000054219	1,02	0,0094989	2,02	31,124	62,952
PHTF2	ENSG00000006576	1,02	0,0129243	2,02	35,307	71,407
BAZ2B	ENSG00000123636	1,01	0,0011809	2,02	56,278	113,718
DLD	ENSG00000091140	1,01	0,0002769	2,01	71,448	143,545
LDLRAD4	ENSG00000168675	1,01	0,0002316	2,01	190,316	382,288
FRMD4B	ENSG00000114541	1,01	0,0382262	2,01	10,117	20,307
KIAA1161	ENSG00000164976	1	0,0333072	2,01	8,342	16,725
GLB1	ENSG00000170266	1	0,0007789	2	26,492	53,119
RHOBTB2	ENSG00000008853	-1	0,0385365	-2	22,535	11,243
PITPNC1	ENSG00000154217	-1	0,0001801	-2,01	159,542	79,509
DENND2D	ENSG00000162777	-1,01	0,0001517	-2,01	583,227	290,555
TLDC1	ENSG00000140950	-1,01	0,0056956	-2,01	50,595	25,185
DGKD	ENSG00000077044	-1,01	0,001741	-2,01	95,654	47,537
RARRES3	ENSG00000133321	-1,02	0,000558	-2,02	65,915	32,583
TSEN54	ENSG00000182173	-1,02	0,0018486	-2,02	57,155	28,225
SAE1	ENSG00000142230	-1,03	0,0001084	-2,04	178,321	87,611
STMN3	ENSG00000197457	-1,03	0,0122098	-2,04	161,006	78,843
LRP12	ENSG00000147650	-1,03	0,0125796	-2,05	20,494	10,019
TBXAS1	ENSG00000059377	-1,03	0,0411742	-2,05	12,437	6,076
PCNX2	ENSG00000135749	-1,03	0,0003939	-2,05	86,924	42,435
ZHX3	ENSG00000174306	-1,04	0,0064328	-2,05	31,853	15,54
SATB1-AS1	ENSG00000228956	-1,04	0,0001683	-2,06	151,057	73,483
CAMK4	ENSG00000152495	-1,04	0,0023626	-2,06	800,051	389,156
TTC39B	ENSG00000155158	-1,04	0,0008755	-2,06	97,875	47,464
VCL	ENSG00000035403	-1,05	0,027826	-2,07	28,912	13,992
PARP8	ENSG00000151883	-1,05	8,99E-05	-2,07	134,308	64,977
BACH2	ENSG00000112182	-1,05	0,0007373	-2,07	497,035	240,419
MAP2K1	ENSG00000169032	-1,05	0,0001077	-2,07	121,155	58,542
ACER3	ENSG00000078124	-1,06	0,00187	-2,08	38,484	18,487
ADGRL1	ENSG00000072071	-1,06	0,0020338	-2,08	65,012	31,212
LINC01550	ENSG00000246223	-1,06	0,0118307	-2,08	60,109	28,855
CUBN	ENSG00000107611	-1,06	0,0104241	-2,08	27,4	13,148

SDCBP	ENSG00000137575	-1,06	0,0001918	-2,09	227,81	109,252
DDIT4	ENSG00000168209	-1,06	0,0272784	-2,09	219,545	105,196
CARMIL1	ENSG00000079691	-1,06	0,0010077	-2,09	33,942	16,252
MAML2	ENSG00000184384	-1,06	9,09E-05	-2,09	176,257	84,324
KCNQ1	ENSG00000053918	-1,07	0,0125796	-2,09	34,522	16,493
GPR183	ENSG00000169508	-1,07	0,0002979	-2,1	106,125	50,421
LINC00865	ENSG00000232229	-1,07	0,0318688	-2,11	14,982	7,118
ZNF185	ENSG00000147394	-1,08	0,0439785	-2,11	11,145	5,276
CISD3	ENSG00000277972	-1,08	0,0178148	-2,12	16,158	7,621
HELB	ENSG00000127311	-1,09	7,89E-05	-2,13	93,774	43,97
CD55	ENSG00000196352	-1,1	0,0062909	-2,14	122,099	57,087
ARRB1	ENSG00000137486	-1,1	0,038504	-2,14	26,07	12,184
NAA16	ENSG00000172766	-1,1	0,0004925	-2,14	92,298	43,108
SVIL	ENSG00000197321	-1,1	0,0044362	-2,15	28,169	13,108
GIMAP8	ENSG00000171115	-1,11	7,36E-05	-2,15	75,493	35,078
PLXND1	ENSG00000004399	-1,11	5,55E-05	-2,15	76,62	35,594
FRY	ENSG00000073910	-1,11	0,0005389	-2,16	85,442	39,633
MEST	ENSG00000106484	-1,11	0,02189	-2,16	29,082	13,446
MFHAS1	ENSG00000147324	-1,12	0,0026571	-2,17	84,542	38,902
SERINC5	ENSG00000164300	-1,12	0,0006365	-2,18	113,347	51,976
MPP7	ENSG00000150054	-1,13	0,0019762	-2,19	54,064	24,733
ADARB1	ENSG00000197381	-1,13	1,19E-05	-2,19	96,717	44,192
PLAC8	ENSG00000145287	-1,13	0,0065098	-2,19	57,321	26,174
ARHGAP32	ENSG00000134909	-1,14	0,0001085	-2,2	101,348	46,002
SORL1	ENSG00000137642	-1,15	7,86E-05	-2,21	372,206	168,304
HPGD	ENSG00000164120	-1,15	4,89E-05	-2,22	100,005	45,061
MGAT4A	ENSG00000071073	-1,16	0,0006046	-2,23	137,015	61,405
SFMBT1	ENSG00000163935	-1,17	0,0092225	-2,24	22,373	9,971
ITGA5	ENSG00000161638	-1,17	1,65E-05	-2,25	87,229	38,748
ANKS6	ENSG00000165138	-1,17	0,0022848	-2,25	29,387	13,041
RNF144A	ENSG00000151692	-1,18	0,0002315	-2,27	129,815	57,303
HOOK1	ENSG00000134709	-1,18	0,0001128	-2,27	58,063	25,609
RP1-47M23.3	ENSG00000280135	-1,18	0,0014643	-2,27	56,214	24,755
RXRA	ENSG00000186350	-1,18	0,0017313	-2,27	24,533	10,801
TMSB10	ENSG00000034510	-1,19	9,67E-06	-2,28	2489,063	1091,085
CASS4	ENSG00000087589	-1,19	0,0439785	-2,28	35,504	15,543
MFSD12	ENSG00000161091	-1,19	0,0055903	-2,29	57,559	25,184
FUT8	ENSG00000033170	-1,2	5,28E-06	-2,31	89,525	38,838
FAR2	ENSG00000064763	-1,2	0,0104592	-2,31	17,737	7,694
JAML	ENSG00000160593	-1,21	4,07E-06	-2,31	82,409	35,607
NSG1	ENSG00000168824	-1,21	0,0171537	-2,32	14,148	6,093
MPP6	ENSG00000105926	-1,22	0,0021369	-2,33	36,188	15,544
KBTBD11	ENSG00000176595	-1,22	0,00018	-2,33	35,587	15,262
PLEKHF1	ENSG00000166289	-1,23	0,0056607	-2,35	95,22	40,568
APBA2	ENSG00000034053	-1,23	0,0015119	-2,35	65,375	27,785
AKTIP	ENSG00000166971	-1,24	0,0005764	-2,36	45,49	19,253
AJUBA	ENSG00000129474	-1,24	0,0141256	-2,37	15,026	6,352
RP11-330A16.1	ENSG00000229646	-1,25	0,0194412	-2,37	9,571	4,037
CD101	ENSG00000134256	-1,25	0,0003822	-2,38	62,335	26,224
SCML4	ENSG00000146285	-1,25	0,0223821	-2,38	351,225	147,367
WDR27	ENSG00000184465	-1,25	0,0019897	-2,39	21,631	9,065
RUNX1	ENSG00000159216	-1,26	2,12E-05	-2,39	318,901	133,618
MICAL2	ENSG00000133816	-1,26	5,26E-05	-2,39	209,122	87,586
FAM65B	ENSG00000111913	-1,26	1,74E-06	-2,39	433,846	181,262
CYSLTR1	ENSG00000173198	-1,26	0,0004617	-2,39	41,53	17,347
GCSAM	ENSG00000174500	-1,26	0,0134167	-2,4	18,026	7,505
SLC35F6	ENSG00000213699	-1,27	0,0007887	-2,4	41,565	17,289
TMEM229B	ENSG00000198133	-1,27	8,42E-06	-2,4	67,209	27,958
SOX8	ENSG00000005513	-1,28	0,0464517	-2,42	10,027	4,144
CEBPB	ENSG00000172216	-1,29	0,0268299	-2,45	37,338	15,24
BCAT1	ENSG00000060982	-1,3	9,64E-05	-2,46	38,208	15,554
ADPRM	ENSG00000170222	-1,3	3,30E-05	-2,46	41,403	16,814

HRH2	ENSG00000113749	-1,3	0,0001094	-2,47	35,786	14,518
PNP	ENSG00000198805	-1,31	0,0119336	-2,48	26,625	10,754
LZTFL1	ENSG00000163818	-1,31	8,49E-06	-2,48	124,984	50,447
DENND5A	ENSG00000184014	-1,32	1,41E-06	-2,5	95,943	38,423
DUSP14	ENSG00000276023	-1,33	0,0193228	-2,51	11,735	4,679
NOSIP	ENSG00000142546	-1,33	4,53E-05	-2,51	115,867	46,083
RARG	ENSG00000172819	-1,34	0,0014334	-2,53	18,796	7,435
CD200R1	ENSG00000163606	-1,34	0,0002119	-2,53	43,937	17,362
DLG4	ENSG00000132535	-1,34	3,61E-05	-2,54	33,241	13,108
ARRB2	ENSG00000141480	-1,34	9,74E-06	-2,54	106,215	41,84
MBOAT2	ENSG00000143797	-1,35	0,0060463	-2,55	23,981	9,418
AC145124.2	ENSG00000255495	-1,35	0,0343501	-2,55	16,188	6,34
ZBTB18	ENSG00000179456	-1,36	4,18E-07	-2,56	158,722	62,033
CD109	ENSG00000156535	-1,36	0,0001164	-2,56	39,297	15,361
TNFSF11	ENSG00000120659	-1,36	0,0002075	-2,56	26,287	10,252
CTC-463A16.1	ENSG00000280047	-1,36	0,0336027	-2,57	5,998	2,331
KIAA0355	ENSG00000166398	-1,38	3,19E-05	-2,6	153,897	59,249
PKIG	ENSG00000168734	-1,38	0,0189894	-2,6	6,418	2,465
LINC01128	ENSG00000228794	-1,38	1,16E-06	-2,6	70,318	27,006
SYTL2	ENSG00000137501	-1,39	9,67E-06	-2,61	223,072	85,394
GLB1L3	ENSG00000166105	-1,39	0,0030478	-2,62	13,45	5,125
ACVR2A	ENSG00000121989	-1,39	0,0108846	-2,63	8,147	3,102
RP11-664D1.1	ENSG00000256862	-1,39	0,0186259	-2,63	28,056	10,668
CDC14A	ENSG00000079335	-1,39	0,0002202	-2,63	62,592	23,797
KIAA1671	ENSG00000197077	-1,4	3,15E-05	-2,65	113,25	42,786
FAM213A	ENSG00000122378	-1,41	2,71E-06	-2,65	44,295	16,705
AMIGO1	ENSG00000181754	-1,43	1,43E-06	-2,69	50,19	18,677
ZNF774	ENSG00000196391	-1,43	0,0166582	-2,7	8,396	3,114
AXIN2	ENSG00000168646	-1,43	1,07E-07	-2,7	67,581	25,034
RSAD2	ENSG00000134321	-1,44	5,16E-08	-2,71	173,898	64,283
CIAPIN1	ENSG00000005194	-1,44	0,0020872	-2,71	183,638	67,866
PRKCA	ENSG00000154229	-1,44	0,0008543	-2,71	68,627	25,327
MRC2	ENSG00000011028	-1,44	5,60E-05	-2,72	52,871	19,441
RFLNB	ENSG00000183688	-1,45	3,68E-08	-2,73	297,25	108,73
ECT2	ENSG00000114346	-1,46	0,0336027	-2,75	9,336	3,397
SMARCA2	ENSG00000080503	-1,46	2,98E-08	-2,75	418,076	152,111
SPEG	ENSG00000072195	-1,46	0,019978	-2,75	12,046	4,381
TNFRSF10D	ENSG00000173530	-1,46	2,87E-05	-2,75	48,438	17,618
CTD-2562J17.6	ENSG00000279117	-1,46	0,0178148	-2,75	10,707	3,889
JAG2	ENSG00000184916	-1,46	0,0434043	-2,76	2,562	0,929
CTDSPL	ENSG00000144677	-1,46	0,0002344	-2,76	19,03	6,896
ATP6V0E2	ENSG00000171130	-1,47	0,0006386	-2,77	124,912	45,167
C4orf32	ENSG00000174749	-1,47	0,0001512	-2,77	66,232	23,913
CDK18	ENSG00000117266	-1,47	0,029362	-2,77	7,731	2,789
KIF24	ENSG00000186638	-1,47	0,0308279	-2,77	2,662	0,958
ADGRB2	ENSG00000121753	-1,47	0,0089899	-2,77	17,385	6,263
FAM86B1	ENSG00000186523	-1,47	0,0326178	-2,77	4,102	1,478
COL6A3	ENSG00000163359	-1,47	7,76E-07	-2,78	178,571	64,275
RAB27B	ENSG00000041353	-1,48	0,0151559	-2,78	11,441	4,111
SLC6A20	ENSG00000163817	-1,49	0,0229399	-2,8	6,782	2,419
PRKD3	ENSG00000115825	-1,49	8,08E-05	-2,81	190,679	67,93
TRDV1	ENSG00000211804	-1,5	0,0280841	-2,83	17,118	6,055
TMEM220	ENSG00000187824	-1,5	0,040173	-2,83	6,956	2,46
DPPA4	ENSG00000121570	-1,5	6,89E-05	-2,83	29,75	10,513
MKI67	ENSG00000148773	-1,5	0,0052802	-2,83	27,19	9,595
DBH-AS1	ENSG00000225756	-1,51	1,13E-05	-2,84	79,495	27,994
AC093642.4	ENSG00000226423	-1,52	0,0070399	-2,86	9,246	3,236
DDR1	ENSG00000204580	-1,52	3,08E-05	-2,86	49,314	17,226
CYP2J2	ENSG00000134716	-1,52	0,02319	-2,86	5,389	1,878
SLC30A4	ENSG00000104154	-1,52	0,0006055	-2,87	18,196	6,347
C3orf52	ENSG00000114529	-1,52	0,0369914	-2,88	4,01	1,396
SULT1B1	ENSG00000173597	-1,53	3,01E-07	-2,88	99,496	34,546

HOPX	ENSG00000171476	-1,53	0,01708	-2,89	8,845	3,062
PANK1	ENSG00000152782	-1,55	0,0034909	-2,92	12,428	4,251
IMPA2	ENSG00000141401	-1,55	0,0200887	-2,92	3,382	1,156
CECR1	ENSG00000093072	-1,55	6,03E-08	-2,92	244,272	83,541
NLRP6	ENSG00000174885	-1,55	0,0006903	-2,93	19,199	6,558
FZD6	ENSG00000164930	-1,55	0,0016163	-2,94	19,799	6,741
FHL2	ENSG00000115641	-1,56	0,0051419	-2,94	7,441	2,531
CEP112	ENSG00000154240	-1,57	0,0268428	-2,96	3,03	1,021
SSX2IP	ENSG00000117155	-1,57	8,52E-06	-2,97	34,525	11,632
RLN1	ENSG00000107018	-1,57	0,0267499	-2,97	3,136	1,058
SMIM3	ENSG00000256235	-1,59	0,0073176	-3	6,905	2,3
UBXN10-AS1	ENSG00000225986	-1,59	0,0006361	-3,01	14,287	4,745
YES1	ENSG00000176105	-1,59	1,67E-08	-3,02	58,069	19,22
TRAT1	ENSG00000163519	-1,61	5,02E-06	-3,04	363,227	119,319
NDFIP1	ENSG00000131507	-1,61	3,76E-06	-3,04	1138,236	373,85
OAF	ENSG00000184232	-1,61	0,0407097	-3,05	2,156	0,705
EPAS1	ENSG00000116016	-1,61	0,0023303	-3,05	16,587	5,435
CASP10	ENSG00000003400	-1,62	4,31E-07	-3,08	48,17	15,652
CELF5	ENSG00000161082	-1,62	0,0318588	-3,08	2,251	0,73
LIPC	ENSG00000166035	-1,63	0,0277239	-3,09	2,823	0,909
GOLGA7B	ENSG00000155265	-1,63	0,0003242	-3,1	13,443	4,341
ADAMTS17	ENSG00000140470	-1,63	0,0025493	-3,1	8,993	2,897
CHI3L2	ENSG00000064886	-1,64	3,29E-09	-3,12	498,021	159,78
PALD1	ENSG00000107719	-1,65	0,0006825	-3,13	17,463	5,577
TUSC3	ENSG00000104723	-1,65	0,030725	-3,13	8,961	2,856
CPNE7	ENSG00000178773	-1,65	0,0439141	-3,14	2,608	0,829
RAD54B	ENSG00000197275	-1,66	0,0278798	-3,15	5,154	1,638
RP11-61O1.1	ENSG00000259097	-1,66	0,0469744	-3,15	5,585	1,769
PFKFB2	ENSG00000123836	-1,66	0,0129799	-3,16	11,333	3,579
SLC20A1	ENSG00000144136	-1,67	2,06E-11	-3,17	102,423	32,27
TMEM71	ENSG00000165071	-1,68	3,37E-06	-3,2	29,856	9,339
EEPD1	ENSG00000122547	-1,68	3,29E-05	-3,2	20,883	6,515
AC069513.4	ENSG00000229178	-1,68	0,0140521	-3,2	2,396	0,746
MCC	ENSG00000171444	-1,69	0,0062645	-3,22	30,441	9,449
AC010731.2	ENSG00000228577	-1,72	0,0137912	-3,29	4,582	1,39
AC112721.2	ENSG00000222032	-1,72	0,0145516	-3,3	7,067	2,142
CTD-3203P2.1	ENSG00000259940	-1,73	0,0214807	-3,31	1,755	0,53
MTSS1L	ENSG00000132613	-1,73	0,0190409	-3,32	10,486	3,162
SH3BP5	ENSG00000131370	-1,74	0,0050054	-3,33	3,662	1,097
LINC02132	ENSG00000268804	-1,74	0,0129922	-3,34	4,86	1,455
RIPK2	ENSG00000104312	-1,75	0,0036971	-3,35	74,53	22,225
ANKRD55	ENSG00000164512	-1,75	0,0012211	-3,36	107,757	32,033
EPPK1	ENSG00000261150	-1,75	0,0153224	-3,37	26,692	7,912
PCSK5	ENSG00000099139	-1,75	3,66E-08	-3,37	45,215	13,399
CD1B	ENSG00000158485	-1,76	0,0122922	-3,38	11,411	3,376
CXADR	ENSG00000154639	-1,76	0,0079804	-3,39	3,9	1,147
AOAH	ENSG00000136250	-1,77	9,51E-07	-3,41	83,029	24,327
DLL1	ENSG00000198719	-1,78	0,0229399	-3,42	5,038	1,468
RP11-161M6.3	ENSG00000260496	-1,79	0,0101482	-3,46	2,217	0,643
PRTFDC1	ENSG00000099256	-1,79	0,0267174	-3,47	2,088	0,604
ITGA6	ENSG00000091409	-1,82	4,87E-12	-3,53	213,386	60,434
BUB1B	ENSG00000156970	-1,83	0,0367004	-3,55	3,063	0,861
EDAR	ENSG00000135960	-1,83	6,04E-07	-3,56	86,477	24,323
RP11-403I13.5	ENSG00000232721	-1,85	0,0076351	-3,6	2,83	0,788
DMTN	ENSG00000158856	-1,85	0,0260018	-3,61	2,429	0,67
PIK3R6	ENSG00000276231	-1,86	0,0015119	-3,62	6,324	1,746
RIMS3	ENSG00000117016	-1,86	1,90E-06	-3,63	21,586	5,958
CR2	ENSG00000117322	-1,87	0,0042654	-3,66	9,99	2,737
ALS2CL	ENSG00000178038	-1,88	0,0010014	-3,67	16,103	4,387
IER3	ENSG00000137331	-1,88	0,0032435	-3,68	5,652	1,533
RHOU	ENSG00000116574	-1,88	0,0003693	-3,69	10,451	2,827
ELFN1	ENSG00000225968	-1,89	0,0175532	-3,7	3,941	1,062

MND1	ENSG00000121211	-1,9	0,0486784	-3,72	2,647	0,707
SEC14L2	ENSG00000100003	-1,9	0,0001846	-3,72	12,931	3,469
ULK2	ENSG00000083290	-1,9	1,67E-08	-3,73	42,021	11,272
MB21D2	ENSG00000180611	-1,9	0,0021739	-3,73	4,427	1,187
AHRR	ENSG00000063438	-1,91	0,0030536	-3,76	8,298	2,2
CCR12P	ENSG00000238241	-1,92	1,36E-06	-3,78	54,026	14,305
RP11-382A20.3	ENSG00000166503	-1,93	0,0003066	-3,81	11,85	3,105
AKAP6	ENSG00000151320	-1,93	0,0059769	-3,82	2,908	0,762
CD226	ENSG00000150637	-1,96	1,88E-09	-3,88	148,976	38,417
CAMKK1	ENSG00000004660	-1,96	3,62E-05	-3,89	20,5	5,273
ADAMTS7P1	ENSG00000274376	-1,96	0,0071826	-3,9	17,269	4,424
GRB10	ENSG00000106070	-1,97	2,94E-05	-3,91	19,549	4,999
FSBP	ENSG00000265817	-1,98	0,0096761	-3,94	4,449	1,128
MAGI3	ENSG00000081026	-1,99	1,43E-09	-3,96	59,33	14,982
PTCH1	ENSG00000185920	-1,99	0,0003928	-3,97	41,607	10,491
MDS2	ENSG00000197880	-1,99	0,0003242	-3,97	7,671	1,934
SMAD1	ENSG00000170365	-1,99	3,48E-06	-3,98	14,176	3,564
STX3	ENSG00000166900	-1,99	0,000576	-3,98	11,738	2,955
CENPF	ENSG00000117724	-2	1,19E-08	-4	48,573	12,144
FRRS1	ENSG00000156869	-2,01	0,0024431	-4,03	2,235	0,553
RP11-61O1.2	ENSG00000258511	-2,02	3,46E-05	-4,07	16,075	3,954
RNF157	ENSG00000141576	-2,03	0,0005225	-4,08	6,043	1,483
ZNF280B	ENSG00000275004	-2,03	0,0008789	-4,08	4,793	1,177
MAP1A	ENSG00000166963	-2,03	0,012257	-4,08	6,559	1,608
SEPT5	ENSG00000184702	-2,03	3,61E-05	-4,09	17,419	4,262
RP11-412B14.1	ENSG00000255101	-2,03	0,0052059	-4,09	7,697	1,886
CAMSAP2	ENSG00000118200	-2,03	6,57E-06	-4,1	55,835	13,634
RHPN1-AS1	ENSG00000254389	-2,04	0,0173366	-4,11	1,972	0,477
ACTN1	ENSG00000072110	-2,04	7,00E-14	-4,11	417,761	101,592
TC2N	ENSG00000165929	-2,05	4,37E-12	-4,13	329,572	79,805
CLDN1	ENSG00000163347	-2,06	7,47E-08	-4,17	297,552	71,27
RP11-18H21.1	ENSG00000245954	-2,06	4,89E-07	-4,18	27,857	6,665
FZD1	ENSG00000157240	-2,07	0,0012928	-4,19	35,726	8,529
SLC15A3	ENSG00000110446	-2,07	0,0021443	-4,2	2,138	0,506
NLRP14	ENSG00000158077	-2,07	0,0032362	-4,21	7,617	1,808
FAM26F	ENSG00000188820	-2,08	0,0072262	-4,22	4,394	1,037
LINC01891	ENSG00000231682	-2,08	0,0016998	-4,23	3,132	0,742
NEURL1	ENSG00000107954	-2,09	0,0013223	-4,25	8,693	2,046
ADGRE4P	ENSG00000268758	-2,1	0,0232052	-4,28	4,032	0,94
GP5	ENSG00000178732	-2,1	5,34E-05	-4,28	7,602	1,776
PGLYRP2	ENSG00000161031	-2,1	0,0019787	-4,29	1,76	0,411
STOX1	ENSG00000165730	-2,1	0,0011026	-4,29	6,761	1,573
BBC3	ENSG00000105327	-2,12	8,33E-05	-4,33	18,668	4,309
AC079922.3	ENSG00000237753	-2,12	0,0304135	-4,34	3,644	0,842
EPHA4	ENSG00000116106	-2,13	0,0020419	-4,36	10,76	2,465
RBM11	ENSG00000185272	-2,13	0,0024592	-4,38	3,426	0,778
LL21NC02-1C16.2	ENSG00000272825	-2,14	0,0175506	-4,4	2,201	0,501
SORBS3	ENSG00000120896	-2,16	0,001766	-4,48	4,577	1,021
NPAS2	ENSG00000170485	-2,17	0,00081	-4,49	21,16	4,71
KIAA1211	ENSG00000109265	-2,17	0,0003118	-4,49	5,363	1,196
ABCB1	ENSG00000085563	-2,18	6,43E-14	-4,53	77,439	17,101
LDB2	ENSG00000169744	-2,2	3,89E-05	-4,61	12,786	2,772
CSGALNACT1	ENSG00000147408	-2,21	7,44E-08	-4,61	71,871	15,568
LZTS1	ENSG00000061337	-2,22	0,0005584	-4,67	7,761	1,659
IGSF9B	ENSG00000080854	-2,22	0,0006277	-4,67	5,3	1,133
Y_RNA	ENSG00000201635	-2,24	0,0093784	-4,73	1,7	0,363
TRGV2	ENSG00000233306	-2,26	0,0048369	-4,8	3,734	0,779
KIRREL	ENSG00000183853	-2,27	6,60E-05	-4,81	10,475	2,179
TKTL1	ENSG00000007350	-2,28	0,000158	-4,84	4,095	0,846
KIAA1522	ENSG00000162522	-2,28	0,0012728	-4,84	15,176	3,136
CD40LG	ENSG00000102245	-2,29	1,12E-15	-4,89	250,809	51,316
RP5-1184F4.7	ENSG00000277301	-2,31	0,0022348	-4,96	2,707	0,547

FHIT	ENSG00000189283	-2,32	3,68E-06	-4,99	47,493	9,523
HBEGF	ENSG00000113070	-2,33	9,46E-05	-5,02	2,992	0,594
IL7R	ENSG00000168685	-2,34	5,40E-13	-5,08	3704,678	729,207
AIF1L	ENSG00000126878	-2,35	0,0062356	-5,11	2,512	0,495
ALDH7A1	ENSG00000164904	-2,36	0,0032413	-5,13	4,012	0,784
DTX4	ENSG00000110042	-2,36	0,0003425	-5,14	4,514	0,882
RP11-747H7.3	ENSG00000260711	-2,36	0,0012348	-5,15	4,528	0,88
BAG3	ENSG00000151929	-2,36	7,00E-14	-5,15	78,265	15,196
CRTAM	ENSG00000109943	-2,37	1,46E-05	-5,17	12,878	2,49
RNF144A-AS1	ENSG00000228203	-2,38	0,0001299	-5,2	7,021	1,352
ADGRE1	ENSG00000174837	-2,4	7,66E-12	-5,29	84,502	15,968
TSPAN9	ENSG00000011105	-2,41	9,69E-05	-5,33	3,746	0,704
RASL11A	ENSG00000122035	-2,42	0,0001427	-5,34	3,724	0,696
ELAVL4	ENSG00000162374	-2,42	0,0010108	-5,34	2,872	0,536
CLEC4A	ENSG00000111729	-2,42	0,0004499	-5,36	3,539	0,657
FGFBP2	ENSG00000137441	-2,43	5,79E-09	-5,37	18,36	3,419
DEPTOR	ENSG00000155792	-2,43	0,0017576	-5,4	2,918	0,54
RP5-997D24.3	ENSG00000225632	-2,49	0,0008522	-5,61	5,7	1,018
LEF1-AS1	ENSG00000232021	-2,49	1,82E-10	-5,62	19,943	3,548
SLCO4C1	ENSG00000173930	-2,49	6,73E-08	-5,63	18,486	3,278
ADAM23	ENSG00000114948	-2,49	2,19E-11	-5,64	161,734	28,694
ENC1	ENSG00000171617	-2,5	6,11E-09	-5,66	44,241	7,82
GPR65	ENSG00000140030	-2,5	5,88E-10	-5,67	63,914	11,268
GIPC3	ENSG00000179855	-2,51	1,41E-07	-5,71	12,352	2,164
RASSF8	ENSG00000123094	-2,52	5,54E-09	-5,72	64,329	11,245
CCR9	ENSG00000173585	-2,52	5,28E-11	-5,73	470,202	82,059
MYOM2	ENSG00000036448	-2,52	0,0012398	-5,75	16,747	2,909
FCGRT	ENSG00000104870	-2,52	1,24E-15	-5,75	148,911	25,881
WWC2	ENSG00000151718	-2,55	0,0155611	-5,85	2,654	0,453
SPRY2	ENSG00000136158	-2,55	0,0001773	-5,86	18,522	3,155
PLXDC1	ENSG00000161381	-2,56	1,85E-12	-5,91	41,267	6,978
CATSPERB	ENSG00000133962	-2,59	4,19E-05	-6,04	2,515	0,414
ENPP2	ENSG00000136960	-2,59	0,0004342	-6,04	12,795	2,113
RBFOX2	ENSG00000100320	-2,6	1,46E-05	-6,05	7,095	1,171
PTAFR	ENSG00000169403	-2,62	8,28E-05	-6,16	10,798	1,751
DIP2C	ENSG00000151240	-2,66	2,42E-06	-6,31	31,676	5,018
HCAR1	ENSG00000196917	-2,66	1,23E-06	-6,34	9,078	1,429
CHD7	ENSG00000171316	-2,67	2,85E-15	-6,36	128,643	20,221
ENPP5	ENSG00000112796	-2,67	5,21E-06	-6,37	12,917	2,031
LMO4	ENSG00000143013	-2,69	1,59E-08	-6,43	11,852	1,838
TMIE	ENSG00000181585	-2,69	8,94E-05	-6,47	4,298	0,668
NET1	ENSG00000173848	-2,7	5,27E-18	-6,49	72,657	11,195
EPHA1-AS1	ENSG00000229153	-2,7	0,0004222	-6,5	4,819	0,741
LPCAT2	ENSG00000087253	-2,71	4,06E-08	-6,52	21,614	3,31
DNAH6	ENSG00000115423	-2,71	2,06E-06	-6,56	12,435	1,893
AUTS2	ENSG00000158321	-2,72	2,47E-06	-6,59	329,517	49,969
MIR4697HG	ENSG00000280237	-2,73	7,36E-06	-6,62	5,745	0,864
NELL2	ENSG00000184613	-2,74	3,01E-20	-6,66	78,811	11,835
RP11-117D22.2	ENSG00000230138	-2,76	0,0003242	-6,79	1,804	0,266
SIPA1L2	ENSG00000116991	-2,79	1,47E-05	-6,93	4,595	0,66
TRABD2A	ENSG00000186854	-2,82	1,47E-22	-7,06	186,632	26,441
PPFIBP2	ENSG00000166387	-2,82	1,84E-06	-7,06	5,112	0,721
RASSF8-AS1	ENSG00000246695	-2,86	4,99E-05	-7,25	5,174	0,715
GNG7	ENSG00000176533	-2,86	2,63E-06	-7,26	12,567	1,732
QPCT	ENSG00000115828	-2,86	3,89E-11	-7,27	29,74	4,092
PLEK2	ENSG00000100558	-2,87	6,27E-07	-7,34	11,25	1,536
CACNA1I	ENSG00000100346	-2,89	0,0023015	-7,41	7,127	0,959
ZNF462	ENSG00000148143	-2,92	4,64E-05	-7,57	4,403	0,582
ATP6V0E2-AS1	ENSG00000204934	-2,93	1,55E-06	-7,61	26,303	3,453
CD300A	ENSG00000167851	-2,93	2,47E-06	-7,64	4,954	0,645
AC022182.1	ENSG00000254777	-2,94	6,47E-06	-7,65	2,192	0,284
ST6GALNAC1	ENSG00000070526	-2,94	1,41E-05	-7,67	4,868	0,637

CLMN	ENSG00000165959	-2,95	5,23E-08	-7,7	6,382	0,827
SHTN1	ENSG00000187164	-2,95	0,0015301	-7,75	2,255	0,29
ACSL6	ENSG00000164398	-2,95	1,98E-06	-7,75	8,129	1,05
VIPR1	ENSG00000114812	-2,98	1,58E-05	-7,88	9,083	1,152
CD1C	ENSG00000158481	-3,03	0,0001328	-8,18	5,383	0,659
SCN3A	ENSG00000153253	-3,06	2,76E-14	-8,33	38,158	4,583
ZNF467	ENSG00000181444	-3,13	1,19E-09	-8,74	19,692	2,251
CR1	ENSG00000203710	-3,16	0,0001172	-8,92	1,832	0,207
CDHR1	ENSG00000148600	-3,17	1,40E-07	-9,03	25,478	2,824
EPHA1	ENSG00000146904	-3,18	6,14E-10	-9,07	25,344	2,795
KRT2	ENSG00000172867	-3,19	0,0001037	-9,12	3,138	0,343
GABRD	ENSG00000187730	-3,19	4,31E-07	-9,13	4,885	0,535
MPZL2	ENSG00000149573	-3,2	3,31E-06	-9,16	4,22	0,464
RHOBTB1	ENSG00000072422	-3,21	2,13E-09	-9,26	10,37	1,12
LRRN3	ENSG00000173114	-3,23	1,61E-23	-9,37	216,689	23,136
TPST1	ENSG00000169902	-3,23	6,99E-07	-9,41	2,407	0,255
ANK3	ENSG00000151150	-3,32	2,09E-18	-9,98	34,558	3,464
ANXA1	ENSG00000135046	-3,35	8,06E-14	-10,18	19,213	1,888
TSKU	ENSG00000182704	-3,36	5,00E-07	-10,24	11,807	1,158
SIAH3	ENSG00000215475	-3,37	2,33E-10	-10,35	13,052	1,263
AC013264.2	ENSG00000231621	-3,42	2,74E-10	-10,69	11,834	1,105
LINC01146	ENSG00000258867	-3,43	1,14E-05	-10,8	1,771	0,163
RGMB	ENSG00000174136	-3,47	5,46E-06	-11,11	4,462	0,399
KRT73-AS1	ENSG00000257495	-3,48	4,33E-05	-11,16	8,751	0,779
RCN3	ENSG00000142552	-3,48	3,29E-07	-11,18	8,907	0,795
PLLP	ENSG00000102934	-3,52	0,0004947	-11,47	16,436	1,435
DACT1	ENSG00000165617	-3,54	6,29E-14	-11,62	70,124	6,037
EVA1A	ENSG00000115363	-3,55	2,65E-06	-11,71	10,108	0,861
DAPK2	ENSG00000035664	-3,62	5,89E-08	-12,32	2,137	0,173
AC020571.3	ENSG00000229056	-3,63	8,63E-09	-12,4	19,21	1,548
TMEM200A	ENSG00000164484	-3,68	2,73E-16	-12,81	30,169	2,354
CELA1	ENSG00000139610	-3,77	5,12E-10	-13,69	4,669	0,341
FBLN2	ENSG00000163520	-3,78	4,55E-09	-13,73	8,937	0,649
GNAI1	ENSG00000127955	-3,8	2,43E-10	-13,94	24,052	1,729
WNT5A	ENSG00000114251	-3,83	2,02E-07	-14,27	4,256	0,3
GZMA	ENSG00000145649	-3,85	7,72E-13	-14,44	12,327	0,854
PI16	ENSG00000164530	-3,92	8,93E-06	-15,11	5,264	0,349
AMPH	ENSG00000078053	-3,95	4,26E-13	-15,5	30,273	1,952
SLC18A2	ENSG00000165646	-3,99	4,33E-08	-15,89	40,295	2,537
EPGN	ENSG00000182585	-4,07	0,0001334	-16,81	17,079	1,016
HTR2B	ENSG00000135914	-4,12	3,01E-10	-17,36	7,764	0,449
CDCP1	ENSG00000163814	-4,16	8,46E-10	-17,93	13,416	0,743
HKDC1	ENSG00000156510	-4,17	1,47E-12	-18,03	7,886	0,439
EPHA2	ENSG00000142627	-4,17	2,33E-07	-18,03	2,898	0,164
RP11-735G4.1	ENSG00000226409	-4,18	5,82E-07	-18,18	4,182	0,232
LRRC4	ENSG00000128594	-4,19	9,79E-19	-18,22	16,912	0,925
ITGA1	ENSG00000213949	-4,19	2,70E-34	-18,26	53,021	2,899
TGFA	ENSG00000163235	-4,28	1,25E-12	-19,37	12,018	0,621
LINC00494	ENSG00000235621	-4,28	1,47E-12	-19,37	19,186	0,992
AFAP1L1	ENSG00000157510	-4,29	2,46E-11	-19,58	4,55	0,232
ADD2	ENSG00000075340	-4,35	4,30E-19	-20,42	20,249	0,99
RNASE6	ENSG00000169413	-4,36	1,24E-19	-20,52	26,61	1,299
NDST3	ENSG00000164100	-4,4	6,42E-18	-21,08	28,599	1,362
PDZD4	ENSG00000067840	-4,4	1,39E-11	-21,1	9,364	0,444
SVOPL	ENSG00000157703	-4,68	1,98E-09	-25,66	4,164	0,164
IGFBP5	ENSG00000115461	-5,08	4,54E-12	-33,79	12,723	0,375
SPRY4	ENSG00000187678	-5,09	2,06E-11	-34,05	31,146	0,915
WNT10B	ENSG00000169884	-5,1	1,49E-08	-34,24	1,378	0,043
S100B	ENSG00000160307	-5,24	6,29E-14	-37,76	4,981	0,128
NTRK3	ENSG00000140538	-5,42	5,23E-14	-42,76	30,072	0,704
ADAMTS1	ENSG00000154734	-5,7	2,33E-09	-52,03	9,238	0,173
CUX2	ENSG00000111249	-5,79	5,27E-18	-55,23	9,611	0,172

LINC01871	ENSG00000235576	-6,75	1,63E-18	-107,77	6,151	0,055
PRMT8	ENSG00000111218	-7,29	1,00E-16	-156,8	19,408	0,129
DCHS2	ENSG00000197410	-11,3	3,49E-19	-2519,55	3,359	0

Annexe 4 Table with the 196 uniquelly expressed genes in tTreg subset. Fold-change (FC) and FDR values are from edgeR tool. CPM, counts per million.

Gene ID	Ensembl ID	LogFC	FDR	FC	tTconv mean expression (normalized CPM)	tTreg mean expression (normalized CPM)
FN1	ENSG00000115414	5,625	2,55E-51	49,338	1,006	49,523
DNAH8	ENSG00000124721	5,537	1,35E-39	46,420	0,983	45,900
CCR8	ENSG00000179934	5,771	2,08E-22	54,609	0,799	43,836
HNRNPA1P21	ENSG00000228168	5,332	1,84E-21	40,269	1,024	41,369
SLC37A2	ENSG00000134955	6,105	9,56E-19	68,818	0,529	36,374
IRF5	ENSG00000128604	5,506	1,42E-40	45,428	0,761	34,641
RP1-207H1.3	ENSG00000231150	5,550	4,92E-45	46,856	0,695	32,654
TNFRSF11A	ENSG00000141655	5,302	1,96E-08	39,451	0,785	30,986
SLITRK1	ENSG00000178235	4,875	1,35E-22	29,343	1,008	29,721
LRRC32	ENSG00000137507	6,164	1,32E-27	71,697	0,382	27,699
MRC1	ENSG00000260314	4,875	6,75E-20	29,345	0,801	23,611
RYR1	ENSG00000196218	4,987	2,63E-10	31,723	0,740	23,570
CPE	ENSG00000109472	5,702	3,78E-07	52,040	0,446	23,348
CYTOR	ENSG00000222041	4,566	2,16E-24	23,685	0,941	22,372
TNFRSF8	ENSG00000120949	5,087	1,51E-26	33,996	0,636	21,723
CLNK	ENSG00000109684	6,029	7,49E-10	65,291	0,285	18,594
IL12RB2	ENSG00000081985	4,578	3,52E-14	23,891	0,765	18,303
FCHO2	ENSG00000157107	4,357	5,27E-18	20,489	0,853	17,457
CCL22	ENSG00000102962	5,340	2,00E-19	40,506	0,422	17,136
FAT3	ENSG00000165323	5,978	3,03E-11	63,051	0,265	16,750
NAT8L	ENSG00000185818	4,451	8,56E-12	21,874	0,702	15,429
ANKRD33B	ENSG00000164236	3,734	3,27E-11	13,303	1,073	14,319
IL1RL1	ENSG00000115602	5,665	8,74E-11	50,731	0,279	14,224
SLC14A1	ENSG00000141469	4,586	1,26E-19	24,015	0,579	13,920
PRICKLE1	ENSG00000139174	4,492	2,26E-09	22,503	0,615	13,908
FAM198A	ENSG00000144649	4,721	2,94E-21	26,368	0,514	13,624
DGCR5	ENSG00000237517	3,783	1,10E-09	13,768	0,980	13,528
EBI3	ENSG00000105246	5,007	5,68E-13	32,164	0,419	13,527
ACTN2	ENSG00000077522	4,236	2,76E-08	18,848	0,674	12,754
HS3ST1	ENSG00000002587	3,759	8,79E-07	13,537	0,913	12,392
NIPAL2	ENSG00000104361	3,801	1,25E-11	13,935	0,866	12,120
OTOF	ENSG00000115155	3,725	1,27E-10	13,224	0,897	11,902
HACD1	ENSG00000165996	4,137	1,16E-12	17,590	0,652	11,526
HMOX1	ENSG00000100292	3,765	1,56E-06	13,594	0,840	11,448
LMNA	ENSG00000160789	5,157	2,07E-06	35,677	0,319	11,441
CDK14	ENSG00000058091	3,957	6,86E-05	15,526	0,729	11,361
CHRNA6	ENSG00000147434	4,111	2,41E-06	17,274	0,602	10,435
ADPRH	ENSG00000144843	4,533	1,21E-12	23,158	0,447	10,407
FBP1	ENSG00000165140	3,450	9,08E-10	10,926	0,933	10,226
HTR4	ENSG00000164270	4,538	3,45E-06	23,224	0,434	10,149
AC006042.6	ENSG00000227719	5,312	2,34E-05	39,731	0,251	10,034
LAMA2	ENSG00000196569	5,437	1,62E-15	43,335	0,229	9,972
TMPRSS6	ENSG00000187045	5,107	2,38E-08	34,467	0,265	9,187
CAV1	ENSG00000105974	3,406	3,34E-05	10,600	0,836	8,891
PERP	ENSG00000112378	3,874	2,06E-11	14,660	0,602	8,857
IL1R1	ENSG00000115594	8,544	7,79E-12	373,228	0,022	8,617
CLEC17A	ENSG00000187912	6,305	2,67E-18	79,051	0,105	8,610
MIR4435-2HG	ENSG00000172965	4,768	9,48E-14	27,254	0,312	8,518
DIRAS3	ENSG00000162595	4,966	7,96E-16	31,244	0,268	8,429
PTCHD1	ENSG00000165186	4,640	1,09E-12	24,938	0,334	8,329
UNQ6494	ENSG00000237372	3,766	6,43E-09	13,601	0,591	8,083
TOR4A	ENSG00000198113	4,255	4,54E-12	19,097	0,408	7,807

METTL24	ENSG00000053328	2,986	4,78E-08	7,921	0,973	7,728
CHRNA2	ENSG00000120903	3,285	3,88E-07	9,750	0,758	7,414
HAAO	ENSG00000162882	2,795	6,12E-07	6,941	1,048	7,285
SLIT1	ENSG00000187122	3,126	0,00118097	8,729	0,819	7,157
RASGRP4	ENSG00000171777	4,995	1,39E-10	31,892	0,222	7,122
AKR1C6P	ENSG00000151631	6,205	1,40E-09	73,794	0,094	6,894
PTGS2	ENSG00000073756	4,027	5,55E-05	16,303	0,399	6,543
RELN	ENSG00000189056	6,168	4,20E-13	71,899	0,091	6,507
FST	ENSG00000134363	4,486	2,78E-12	22,413	0,286	6,470
TMPRSS3	ENSG00000160183	2,838	1,43E-06	7,152	0,902	6,467
IZUMO4	ENSG00000099840	2,609	6,00E-05	6,099	1,029	6,290
FABP5	ENSG00000164687	3,622	4,59E-09	12,308	0,491	6,052
PTGER2	ENSG00000125384	3,620	0,00077711	12,295	0,487	6,004
TMCC3	ENSG00000057704	3,299	5,63E-08	9,845	0,605	5,984
LRRC4B	ENSG00000131409	5,257	2,06E-11	38,234	0,154	5,880
RP11-798M19.6	ENSG00000272870	2,833	9,59E-07	7,125	0,811	5,781
AKR1C2	ENSG00000151632	3,514	1,01E-05	11,421	0,485	5,561
XCL1	ENSG00000143184	4,082	1,92E-09	16,935	0,324	5,527
THSD7A	ENSG00000005108	4,956	2,17E-06	31,043	0,177	5,517
KCNS3	ENSG00000170745	4,911	1,28E-06	30,092	0,180	5,499
NETO2	ENSG00000171208	2,473	0,00026955	5,551	0,987	5,489
JAKMIP1	ENSG00000152969	4,620	3,27E-11	24,595	0,220	5,479
FAM129C	ENSG00000167483	2,743	0,00070296	6,695	0,808	5,420
ARNTL2	ENSG00000029153	2,770	0,00013754	6,823	0,791	5,411
CXCR3	ENSG00000186810	2,660	0,0016495	6,322	0,851	5,394
CYP7B1	ENSG00000172817	3,636	4,65E-07	12,434	0,431	5,392
MCAM	ENSG00000076706	3,044	0,00032642	8,246	0,649	5,371
CLSTN2	ENSG00000158258	6,068	8,40E-18	67,086	0,078	5,357
BTNL8	ENSG00000113303	6,117	1,07E-11	69,384	0,074	5,266
CSF2RB	ENSG00000100368	2,710	0,00111263	6,541	0,796	5,223
HDAC9	ENSG00000048052	3,266	0,00116183	9,621	0,540	5,221
GTSF1L	ENSG00000124196	3,116	8,38E-07	8,667	0,595	5,173
FLVCR2	ENSG00000119686	2,455	0,00010277	5,483	0,929	5,102
SSH3	ENSG00000172830	2,462	8,94E-05	5,509	0,914	5,047
TTC7B	ENSG00000165914	2,673	6,55E-06	6,380	0,776	4,964
FANK1	ENSG00000203780	4,680	5,27E-06	25,639	0,188	4,870
ACTG2	ENSG00000163017	4,110	2,95E-08	17,268	0,278	4,850
LINC01943	ENSG00000280721	7,216	2,11E-11	148,699	0,032	4,730
ADRB1	ENSG00000043591	2,246	0,01047984	4,745	0,965	4,592
KANK3	ENSG00000186994	2,729	8,02E-05	6,630	0,683	4,545
METRNL	ENSG00000176845	2,409	0,0249914	5,311	0,847	4,511
A2M	ENSG00000175899	2,381	0,00036685	5,208	0,856	4,469
CTD-2377D24.4	ENSG00000242407	4,774	1,45E-08	27,353	0,163	4,452
FAM124B	ENSG00000124019	3,341	1,46E-08	10,135	0,438	4,445
C3AR1	ENSG00000171860	2,437	0,00296547	5,415	0,797	4,322
CABLES1	ENSG00000134508	3,651	3,18E-06	12,563	0,336	4,247
HGF	ENSG00000019991	3,965	7,60E-09	15,619	0,272	4,235
TSPAN13	ENSG00000106537	2,389	0,00033445	5,238	0,800	4,197
LACC1	ENSG00000179630	2,365	0,00442251	5,151	0,795	4,110
ZC2HC1A	ENSG00000104427	1,943	0,01116608	3,846	1,067	4,106
CASQ1	ENSG00000143318	4,423	1,83E-12	21,453	0,188	4,054
XXYLT1-AS2	ENSG00000230266	3,390	5,63E-08	10,485	0,381	4,011
RAB31	ENSG00000168461	5,456	2,29E-07	43,888	0,092	3,991
LPL	ENSG00000175445	3,717	2,79E-06	13,153	0,302	3,985
PCDH7	ENSG00000169851	6,631	6,92E-11	99,080	0,038	3,954
RP11-367G6.3	ENSG00000272053	1,814	0,00788956	3,517	1,121	3,940
ZG16B	ENSG00000162078	4,173	2,17E-08	18,038	0,216	3,923
RP11-799D4.1	ENSG00000267744	3,119	6,01E-06	8,688	0,445	3,893
KCNQ3	ENSG00000184156	2,116	0,00100032	4,335	0,871	3,786
HHIP	ENSG00000164161	2,566	0,00179197	5,922	0,623	3,701
RP11-20D14.6	ENSG00000249790	2,061	0,00572782	4,173	0,870	3,637
NPW	ENSG00000183971	4,668	2,80E-10	25,427	0,138	3,590

ICAM1	ENSG00000090339	3,482	8,33E-05	11,177	0,318	3,582
ENOX1	ENSG00000120658	4,461	5,79E-08	22,026	0,160	3,562
PZP	ENSG00000126838	2,242	0,01716024	4,731	0,747	3,533
ZSCAN9	ENSG00000137185	3,021	0,00017726	8,118	0,421	3,444
NTNG2	ENSG00000196358	1,954	0,02807535	3,874	0,883	3,427
HLA-L	ENSG00000243753	2,199	0,00567688	4,591	0,741	3,399
MAP3K8	ENSG00000107968	4,084	3,45E-10	16,965	0,198	3,393
ITGB8	ENSG00000105855	3,743	1,13E-05	13,387	0,252	3,385
EXD3	ENSG00000187609	1,934	0,04305489	3,820	0,871	3,335
PLXNB2	ENSG00000196576	3,384	1,71E-06	10,442	0,318	3,333
SPINT1	ENSG00000166145	1,602	0,01584291	3,036	1,093	3,320
ERRFI1	ENSG00000116285	1,691	0,01307604	3,228	1,015	3,277
MCF2L2	ENSG00000053524	4,361	3,35E-08	20,543	0,157	3,271
TBC1D8-AS1	ENSG00000272902	3,512	2,47E-06	11,406	0,284	3,269
ERICD	ENSG00000280303	2,516	0,00010168	5,720	0,568	3,254
MDGA1	ENSG00000112139	1,907	0,00323416	3,750	0,865	3,246
RP1-151F17.2	ENSG00000272341	2,746	1,61E-05	6,710	0,478	3,212
NTRK1	ENSG00000198400	3,061	8,07E-05	8,344	0,379	3,168
TMPRSS13	ENSG00000137747	3,223	0,002854	9,336	0,334	3,132
LINC00888	ENSG00000240024	1,723	0,01175227	3,302	0,947	3,128
RP11-568N6.1	ENSG00000260101	2,499	0,00113081	5,653	0,542	3,077
AC074289.1	ENSG00000225889	2,374	0,00012644	5,182	0,593	3,069
RP11-462G2.1	ENSG00000237643	5,266	9,78E-08	38,479	0,080	3,045
AC006042.8	ENSG00000233264	2,879	8,91E-06	7,359	0,411	3,037
IL18RAP	ENSG00000115607	3,011	7,36E-05	8,061	0,375	3,030
C2orf16	ENSG00000221843	1,446	0,0479844	2,725	1,092	2,979
SELP	ENSG00000174175	2,041	0,02060335	4,115	0,721	2,967
OXER1	ENSG00000162881	2,651	0,00256561	6,280	0,471	2,966
IL15RA	ENSG00000134470	4,302	3,14E-05	19,720	0,149	2,965
NCF2	ENSG00000116701	2,502	0,01083228	5,663	0,519	2,948
RNF175	ENSG00000145428	4,606	1,65E-11	24,349	0,119	2,921
DAZL	ENSG00000092345	2,898	1,58E-05	7,452	0,387	2,904
CCDC122	ENSG00000151773	2,795	9,46E-05	6,940	0,413	2,876
A2MP1	ENSG00000256069	2,408	0,00344412	5,307	0,540	2,867
WWTR1	ENSG00000018408	3,167	3,45E-06	8,980	0,318	2,859
LIPN	ENSG00000204020	5,792	6,52E-11	55,420	0,049	2,832
SLC22A15	ENSG00000163393	2,644	0,0004962	6,250	0,451	2,832
RP11-239E10.2	ENSG00000236846	3,082	0,0008465	8,469	0,330	2,819
ARAP3	ENSG00000120318	3,286	0,00121504	9,753	0,281	2,770
TRPC1	ENSG00000144935	2,062	0,00159855	4,175	0,662	2,769
ATP1A4	ENSG00000132681	5,639	7,35E-10	49,846	0,053	2,720
TBX21	ENSG00000073861	3,652	8,54E-06	12,568	0,214	2,713
PTPRB	ENSG00000127329	4,483	5,33E-06	22,362	0,121	2,702
FGFR3	ENSG00000068078	3,835	4,04E-05	14,273	0,180	2,587
CXCR5	ENSG00000160683	4,993	3,40E-11	31,845	0,080	2,580
LGALSL	ENSG00000119862	2,176	0,00376537	4,517	0,555	2,517
GPR161	ENSG00000143147	2,332	0,00429348	5,036	0,494	2,506
ENPP3	ENSG00000154269	3,113	7,67E-05	8,653	0,287	2,502
DLGAP1-AS2	ENSG00000262001	1,676	0,02319005	3,196	0,783	2,500
FBXO2	ENSG00000116661	1,425	0,04309316	2,685	0,918	2,469
MERTK	ENSG00000153208	1,777	0,00834707	3,428	0,718	2,466
PIEZO2	ENSG00000154864	5,418	4,87E-10	42,744	0,056	2,419
B3GNT5	ENSG00000176597	3,555	6,34E-07	11,751	0,203	2,414
PNOC	ENSG00000168081	5,824	1,66E-10	56,669	0,043	2,405
LINC01671	ENSG00000225431	2,614	8,76E-05	6,124	0,390	2,386
CSMD1	ENSG00000183117	3,835	4,29E-08	14,272	0,166	2,370
CLDN16	ENSG00000113946	2,586	7,86E-05	6,003	0,391	2,362
TLR2	ENSG00000137462	2,520	0,00242986	5,736	0,388	2,237
HES4	ENSG00000188290	1,943	0,0090946	3,844	0,576	2,224
GAB1	ENSG00000109458	2,275	0,00086412	4,841	0,451	2,183
SLC35G3	ENSG00000164729	2,247	0,00602023	4,747	0,454	2,167
RP11-405M12.3	ENSG00000278607	2,832	0,00016488	7,119	0,301	2,132

RP11-179A10.1	ENSG00000254401	2,634	0,020424	6,209	0,339	2,122
TSHR	ENSG00000165409	3,003	3,70E-06	8,016	0,262	2,119
LINC00877	ENSG00000241163	4,078	3,31E-05	16,894	0,123	2,095
SLC24A4	ENSG00000140090	3,134	0,00710978	8,779	0,239	2,090
ADGRV1	ENSG00000164199	2,226	0,03772209	4,679	0,443	2,071
UBE2QL1	ENSG00000215218	2,064	0,00437509	4,180	0,492	2,066
CCDC81	ENSG00000149201	1,943	0,00487096	3,846	0,528	2,040
GRIN3A	ENSG00000198785	1,997	0,03699136	3,991	0,511	2,038
KRT17P8	ENSG00000256937	1,926	0,01430387	3,800	0,530	2,012
OSM	ENSG00000099985	3,104	5,16E-06	8,600	0,233	2,011
RP11-1348G14.4	ENSG00000251417	1,584	0,03435009	2,999	0,665	1,996
AC079610.2	ENSG00000196096	2,383	0,00643376	5,218	0,379	1,988
GALNT8	ENSG00000130035	2,777	0,01817069	6,856	0,285	1,951
CH507-338C24.1	ENSG00000277991	1,482	0,04425718	2,793	0,695	1,944
TNFRSF13B	ENSG00000240505	4,417	3,70E-06	21,356	0,086	1,860
SLC25A29	ENSG00000197119	1,641	0,02674993	3,119	0,566	1,770
ETNK2	ENSG00000143845	1,734	0,02309259	3,326	0,505	1,687
C20orf197	ENSG00000176659	3,321	1,63E-06	9,996	0,167	1,681
RASGRF1	ENSG00000058335	2,358	0,00085693	5,125	0,320	1,649
RP4-673D20.4	ENSG00000234282	3,415	1,97E-05	10,667	0,142	1,526

Annexe 5 Table with the 46 transcription factors identified in tTreg over-expressed genes. Fold change (FC) and false discovery rate (FDR) values are from edgeR tool.CPM, counts per million.

Gene ID	Ensembl ID	LogFC	FDR	FC	tTconv mean expression (normalized CPM)	tTreg mean expression (normalized CPM)
FOXP3	ENSG00000049768	5,52	2,23E-67	46	9,87	454,32
IRF5	ENSG00000128604	5,51	1,42E-40	45,43	0,76	34,64
ZNF662	ENSG00000182983	4,15	1,77E-29	17,78	1,76	31,34
IKZF4	ENSG00000123411	3,95	1,58E-43	15,42	21,38	329,73
MEOX1	ENSG00000005102	3,72	1,60E-15	13,2	1,84	24,22
TBX21	ENSG00000073861	3,65	8,54E-06	12,57	0,21	2,71
PRDM1	ENSG00000057657	3,56	5,81E-06	11,83	2,61	30,84
CREB3L2	ENSG00000182158	3,27	1,42E-25	9,62	39,04	375,43
PRDM8	ENSG00000152784	3,15	3,52E-10	8,85	1,59	14,1
VDR	ENSG00000111424	3,12	1,55E-19	8,72	4,33	37,75
ZSCAN9	ENSG00000137185	3,02	1,77E-04	8,12	0,42	3,44
BATF	ENSG00000156127	2,79	1,03E-16	6,9	8,28	57,17
ARNTL2	ENSG00000029153	2,77	1,38E-04	6,82	0,79	5,41
BHLHE40	ENSG00000134107	2,65	6,54E-11	6,27	9,76	61,23
MAF	ENSG00000178573	2,56	7,46E-05	5,89	6,17	36,35
GFI1	ENSG00000162676	2,35	3,22E-06	5,08	11,05	56,19
SETBP1	ENSG00000152217	2,16	1,55E-08	4,48	4,52	20,25
POU2F2	ENSG00000028277	2	1,45E-07	4,01	9,4	37,72
IRF8	ENSG00000140968	1,99	1,88E-06	3,96	6,1	24,17
PLAGL1	ENSG00000118495	1,97	3,17E-05	3,93	3,59	14,09
IKZF2	ENSG00000030419	1,94	2,60E-09	3,85	324,39	1247,32
TIGD2	ENSG00000180346	1,94	3,70E-12	3,84	19,4	74,6
HES4	ENSG00000188290	1,94	9,09E-03	3,84	0,58	2,22
CXXC5	ENSG00000171604	1,91	1,41E-08	3,77	8,09	30,49
ZBTB38	ENSG00000177311	1,79	3,47E-07	3,47	53,65	185,91
KLF6	ENSG00000067082	1,72	3,21E-04	3,29	253,9	835,75
REL	ENSG00000162924	1,57	1,51E-05	2,96	57,7	171,05
IRF4	ENSG00000137265	1,51	2,08E-08	2,86	96,7	276,11
IKZF3	ENSG00000161405	1,51	4,63E-07	2,84	194,44	551,98
NFKB2	ENSG00000077150	1,45	6,20E-09	2,73	39,62	108,25
ZNF563	ENSG00000188868	1,43	1,16E-03	2,7	4,96	13,38
NR4A3	ENSG00000119508	1,43	8,55E-06	2,7	47,35	127,74
RORA	ENSG00000069667	1,42	6,28E-04	2,68	55,72	149,05
LCOR	ENSG00000196233	1,38	1,52E-07	2,61	68,82	179,53
AHR	ENSG00000106546	1,36	5,32E-05	2,57	15,61	40,12
NFIA	ENSG00000162599	1,35	4,02E-02	2,54	2,99	7,61
STAT4	ENSG00000138378	1,34	2,36E-02	2,53	10,79	27,25
FOXO1	ENSG00000150907	1,28	1,10E-05	2,43	175,76	427,11
TGIF1	ENSG00000177426	1,26	1,77E-05	2,4	18,69	44,86
PLAGL2	ENSG00000126003	1,18	3,43E-04	2,27	42,77	97,08
NFE2L3	ENSG00000050344	1,17	1,11E-02	2,26	13,52	30,5
RELB	ENSG00000104856	1,17	2,87E-04	2,25	16,68	37,54
SP140	ENSG00000079263	1,09	5,58E-04	2,13	23,21	49,48
ZNF704	ENSG00000164684	1,07	1,04E-02	2,1	14,88	31,22
E2F3	ENSG00000112242	1,04	2,91E-02	2,06	13,01	26,76
BAZ2B	ENSG00000123636	1,01	1,18E-03	2,02	56,28	113,72