DETERMINING IMPORTANT NOVEL MECHANISMS THAT REGULATE T-CELL-MEDIATED IMMUNE RESPONSES

A thesis submitted to The University of Manchester for the degree of Doctor of Philosophy in the Faculty of Biology, Medicine, and Health

2021

STEFANO ROSSI

SCHOOL OF BIOLOGICAL SCIENCES/ DIVISION OF IMMUNOLOGY, IMMUNITY TO INFECTION AND RESPIRATORY MEDICINE/ LYDIA BECKER INSTITUTE OF IMMUNOLOGY AND INFLAMMATION

List of contents

Word count: 34423

List of figures

List of tables

List of abbreviations

ADMIDAS	Adjacent to metal-ion-dependent adhesion site
ATAC-seq	Transposase-accessible chromatin with sequencing
BM	Bone marrow
BSA	Bovine serum albumin
CTLs	Cytotoxic T cells
cDC	Conventional dendritic cell
pDC	Plasmacytoid dendritic cell
DC	Dendritic cell
DEA	Differential expression analysis
ECM	Extracellular matrix
EM	Expectation maximisation
F-actin	Filamentous actin
FACS	Fluorescence-activated cell sorting
FCS	Foetal Calf Serum
GDI	GDP-dissociation inhibitor
GiGA	Graph-based iterative group
GO	Gene ontology
GTCF	Genomic Technologies Core Facility
HPRT	hypoxanthine phosphoribosyltransferase
IPA	Ingenuity pathway analysis
KEGG	Kyoto Encyclopaedia of Genes and Genomes
LAP	Latency associated peptide
LPS	Lipopolysaccharide
LTBP	Latent TGF-β binding protein
MACS	Magnetic-activated cell sorting

MHC	Major histocompatibility complex
MIDAS	Metal-ion-dependent adhesion site
mLNs	Mesenteric lymph nodes
MMP-14	Matrix metalloproteinase-14
p-Treg	Peripheral regulatory T cells
PBS	Phosphate-buffered saline
PC	Principal component
PCA	Principal component analysis
PI4,5-P2	Phosphatidylinositol 4,5-diphosphate
PIP5K	Phosphatidylinositol 4-monophosphate 5-kinase
PRRs	Pattern recognition receptors
PSI	Plexin-semaphorin-integrin
PTB	Phosphotyrosine-binding
PTK2	Protein tyrosine kinase 2
PTP	Protein tyrosine phosphatase
RA	Retinoic acid
REViGO	Reduce and Visualise Gene Ontology
RGD	Arg-Gly-Asp
RT-qPCR	Reverse transcription quantitative polymerase chain reaction
scRNA-seq	Single-cell RNA-seq
SLC	Small latency complex
SVA	Surrogate variable analysis
t-Treg	Thymic regulatory T cells
Tcm	Central memory T cells
TCR	T-cell receptor
Tem	Effector memory T cells

TLRs	Toll-like receptors
TMD	Transmembrane domain
Treg	Regulatory T cells
Trm	Tissue-resident memory T cells

Abstract

The immune system protects our body against foreign pathogens, being capable of promptly mounting a response against different insults. Nevertheless, immunity must be tightly regulated to prevent aberrant responses targeted against our tissues. TGF-β is a pleiotropic cytokine regulating immune responses, which the active cytokine moiety is non-covalently bound to the latency associated peptide (LAP). The activation state of latent TGF-β is regulated by the integrin αvβ8 – which engages with LAP to promote cytokine activation. On dendritic cells (DC), integrin αvβ8-mediated TGF-β activation is paramount – as conditional loss of this integrin correlates with severe inflammatory bowel disease in mice. Additionally, current studies in our lab indicate that expression of αvβ8 is high in effector memory T cells (Tem), with unpublished work indicating that αvβ8+ Tem suppress anti-viral immunity in mice. Importantly, members of the integrin family can act as signalling receptors, but pathways regulated by αvβ8 upon ligand engagement have not been well described to date, with no information available for how integrin αvβ8 signals in immune cells. In this PhD thesis, we used a transcriptomic approach with the aim of identifying important signalling pathways regulated by integrin αvβ8 in DC and T cells. Specifically, using RNA-seq analysis we found that expression of CXCR4 was downregulated in LAP-treated DC and that correlated with a reduced migrating behaviour toward the CXCR4-specific chemokine, CXCL12. These data might indicate that prior to TGF-β activation, signals propagated in DC limit their migrating behaviour. Also, transcriptome profiling indicates that αvβ8+ Tem marked a transcriptionally distinct cell population from the αvβ8- counterparts and are significantly enriched with anti-inflammatory pathways. These data complement previous findings in our lab, providing insights into why a newly identified subset of CD4+ memory T cells can inhibit anti-viral responses.

Together, these results indicate previously undetermined pathways that are regulated by αvβ8 in DC and T cell biology.

Declaration

I declare that no portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning

Copyright Statement

restriction declarations deposited in the University Library, the University Library's regulations (see http://www.library.manchester.ac.uk/about/regulations/) and in the University's policy on Presentation of Theses.

Acknowledgements

First, I wish to thank my supervisory team, Professor Mark Travis and Dr Jean-marc Schwartz, for guiding me throughout my PhD journey. Thanks to them, I like to think that I have acquired the right mind-set to face problems in and out the lab.

Mark, thanks for your thought-provoking discussions about science in general, when planning experiments, and reviewing articles during journal clubs. Your great support and expertise, clear willingness to help in all circumstances has been priceless. Jean-marc, your kindness and ability to gently introduce me to bioinformatics has been commendable. I have never imagined I was capable of mastering RNA-seq methodologies in such a short amount of time. Thanks both for welcoming me to your top-notch research groups.

My sincere appreciation to my funding body, the BBSRC DTP scheme, which made this experience possible.

To the Travis and Schwartz Labs, I could not have asked for a better place to spend this four years. Travis Lab, I will never forget the great time spent together – especially during the nights out that also helped me to discover the beautiful city of Manchester. Schwartz Lab, it has been a pleasure to be part of such an interdisciplinary and multicultural group – the perfect combination to learn and have fun during this period of time.

I like to acknowledge the solid support of my family. Thanks to my Italian friends and the great time spent together despite being many miles away. Finally, I wish to thank all the people I have met during these years of work in Manchester.

The Author

I earned my Bachelor's and Master's Degrees with the highest honours (110/110 cum laude) at the University of Naples 'Federico II' and ranked amongst the best students of my *alma mater*.

I have several years of wet lab experience and, to date, I have authored and co-authored five peer-reviewed papers – of which four are on Coeliac Disease. The latter is a severe condition brought on by an adverse response to gluten, a family of proteins contained in grain cereals, and affecting 1% of people worldwide. As an undergraduate, I produced a recombinant form of the main gluten component – i.e., the α-gliadin. This recombinant form was then used to trigger an intestinal immune response in a transgenic mouse model of Coeliac Disease. Also, I showed that enzymatic transamidation of the recombinant α-gliadin reduced inflammation. Then, biotechnology was adopted on pilot-scale testing for devising a cost-effective, low-calorie, and healthier strategy for innovative gluten-free products.

As BBSRC DTP PhD student, I improved my wet lab approaches and learned bioinformatics methodologies such as transcriptome profiling techniques. These techniques allowed me to expand my knowledge and help me to interrogate and navigate large datasets with ease. Furthermore, expertise acquired is giving me the opportunity to contribute to several joint projects within the faculty.

Dedication

In loving memory of Alfonso Maria Rossi

Blank Page

Chapter 1

Introduction

1.1 Integrins: structure and signalling

Integrins are members of the Type 1 receptor family, consisting of short cytoplasmic tails, a large ectodomain and one single-pass transmembrane domain (TMD) (Lau *et al.,* 2009). The first integrin molecule was characterised in 1986 as an integral membrane protein molecule linking the cytoskeleton to the extracellular matrix (ECM) (Tamkun *et al.,* 1986). Thirty-five years after their first identification and characterisation, integrins have been described to regulate many important biological processes as important cell adhesion receptors.

Indeed, integrins – apart from creating a physical connection between the ECM and the cytoskeleton – are receptors, too. Upon conformational changes, integrins can relay signalling cascades inside the cell following ligand binding and trigger pathways involved in cell differentiation, survival and signalling (Giancotti & Ruoslahti, 1999; Miranti & Brugge, 2002). Moreover, integrins play an important role in cell migration in different contexts such as during embryonic development. Integrins are also crucial in immunity; for example, via regulating leukocyte extravasation. Given its importance, alteration in integrin-mediated cell migration is linked to immunodeficiencies and cancer (Huttenlocher & Horwitz, 2011).

Given the important role played by the integrins in many biological contexts, major efforts have been made to better understand their biological properties.

1.1.1 Integrin structure

Each integrin molecule is composed of two non-covalently bound transmembrane glycoprotein subunits designated alpha (α) and beta (β). With 18 α and 8 β described subunits, 24 distinct assortments of $\alpha\beta$ dimers have been found in mammals – characterised by a distinct distribution throughout the body and specific binding properties (Figure 1.1).

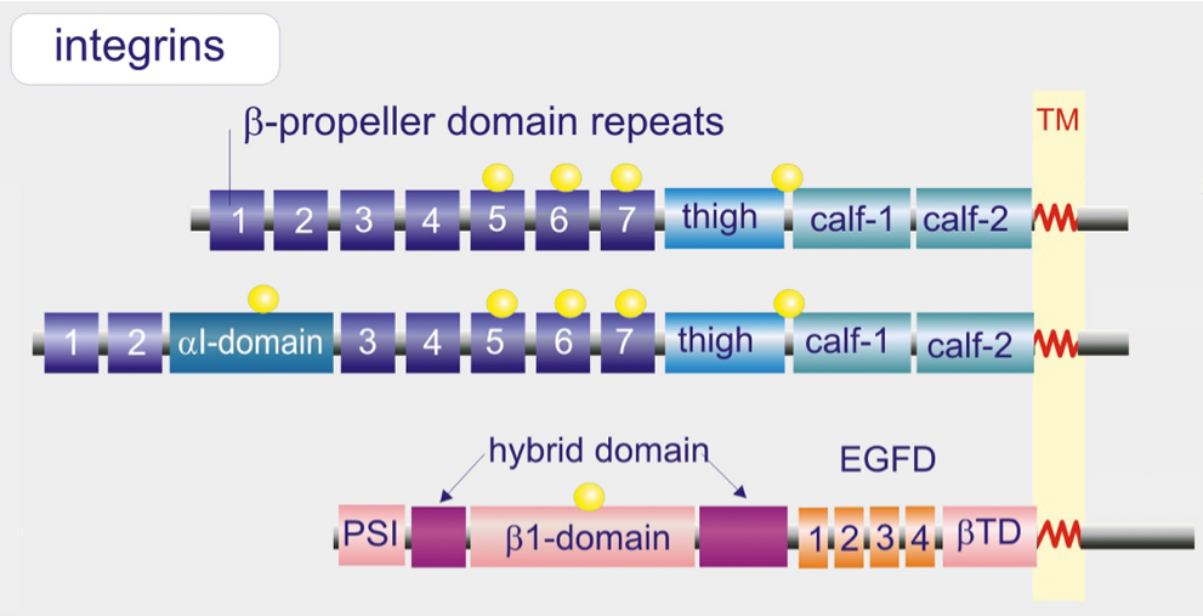


Figure 1.1 | Representative structure of an integrin molecule. The α subunit is characterised by a head, containing β-propeller domain repeats, a stalk, composed of one thigh and two calf regions, a transmembrane, and cytoplasmic domain. Some α chains also contain an αI-domain. In the extracellular part, the β chain is composed of a plexin-semaphorin-integrin (PSI) domain, a hybrid and β1 domain. The hybrid domain is key in regulating affinity for the ligand. The membrane-proximal portion of the extracellular molecule is composed of EGF-like domains and a β-tail domain. The β chain also contains a transmembrane and cytoplasmic region. The figure is modified from Kramer, 2016.

Given this variability and ligand specificity, integrins are classified in different groups – with the αv-, β1-, and β2-containing integrins the largest groups found (Figure 1.2). Pairing of heterodimers, occurring intracellularly, is believed to rely on the quantity of α subunit expressed in the cell – with the β subunits generally expressed in greater quantities (Barczyk et al., 2009; I. D. Campbell & Humphries, 2011). This is not true for αv integrins as several β subunits pair with a single α subunits (Figure 1.2).

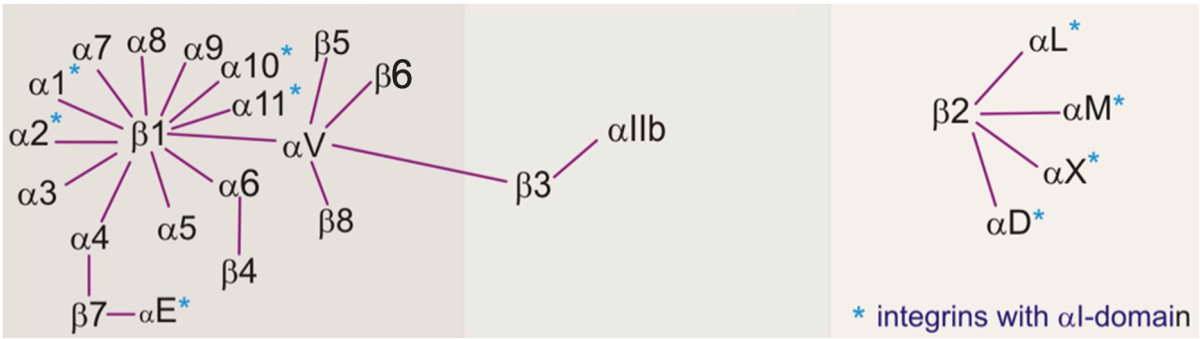


Figure 1.2 | Integrin subunits. Diagram showing known possible combinations of α and β subunits. Straight lines represent possible combinations between α and β subunits. The figure is modified from Kramer, 2016.

Each integrin subunit ranges in size from 750 to 1000 amino acidic residues – spanning the plasma membrane once. The extracellular N-terminal part of the integrin molecule is structured in different domains, whilst the intracellular part is composed of a small cytoplasmic tail (I. D. Campbell & Humphries, 2011). Despite a similar general structure, the α and β subunits are not homologous.

1.1.1.1 The α subunit

The α subunit is composed of a so-called 'leg' structure characterised by a 'thigh' and two 'calf' (calf-1 and calf-2) domains (Figure 1.1). The thigh and calf domains contain each 140 to 170 amino acidic residues organised in β-sandwich folds (I. D. Campbell & Humphries, 2011). Also, the leg supports a β propeller domain – structured in seven distinct blades (Humphries *et al.,* 2003). In 1989, Larson *et al*. found, in half of the α subunits, an extra structure called αI domain. The latter is placed between the second and third blade of the β propeller structure (Larson *et al.,* 1989). As reviewed by Campbell and Humphries, the α subunits also contain two interdomain regions: one is between the calf-1 and the thigh domains – called the knee; the second is a linker region between the thigh and β propeller (I. D. Campbell & Humphries, 2011).

1.1.1.2 The β subunit

Compared to the α subunits, the β chains have more domains – which are structured with a higher degree of complexity. Indeed, from the extracellular N-terminal end of the β subunit, the first domain is the PSI composed of two parts structured in an α/β folding (I. D. Campbell & Humphries, 2011). In the latter domain, a hybrid domain is inserted. Of note, in analogy with the I domain described in the α subunits, the β subunits have

a homologous I domain, too. As reviewed by Campbell and Humphries, this is inserted in the hybrid domain (Figure 1.1). More towards the plasma membrane, the β structures resolve in a β tail domain, more plastic than the α counterparts, which follows four cysteine-rich epidermal growth factor modules (Figure 1.1) (I. D. Campbell & Humphries, 2011).

Contrary to the α subunit, the cytoplasmic tail of the β subunits have a high degree of homology, with many sharing a highly conserved and membrane-proximal phosphorylatable tyrosine sequence – the NPxY (Asn-Pro-Xaa-Tyr) motif. Once the tyrosine is phosphorylated, this motif can bind proteins containing a phosphotyrosine binding (PTB) domain (Barczyk *et al.,* 2009). It is worth mentioning that the β8 cytoplasmic tails lack the latter motif, different and divergent from other β cytoplasmic tails (McCarty, 2020).

1.1.1.3 Transmembrane helices and membrane clasps

As mentioned earlier, each integrin subunit consists of a single-pass transmembrane domain – key in propagating signalling events downstream. Of note, experimental studies supported the hypothesis that a correlation between the α and the β TMD and integrin function exist (Ginsberg, 2014). As already mentioned, integrins switch their structural conformation which, in turn, reflect their receptor activity state. When receptor activity is low, the α- and β-TMD are associated, with the TMDs moving apart during high levels of receptor activity (Pagani & Gohlke, 2018). The resolved structure of the integrin αIIbβ3 identified two points of contact between the α- and β-subunit, referred to as the outer and inner membrane clasps (Ginsberg, 2014). The outer membrane clasp is found at the N-terminal end of the TMD and represents a point of interaction between a Gly-Xaa-Xaa-Xaa-Gly motif on the α chain and a Gly in position 708 on the β chain. The inner membrane clasp consists of an interaction between an

Arginine in position 995 on the α chain and an Aspartate in position 723 on the β chain (Ginsberg, 2014). Also, the latter clasp consists of two residues of Phenylalanine in positions 992 and 993 on the α chain interacting with the β subunit (Ginsberg, 2014). It is also worth mentioning that structural rearrangements occurring in the cytoplasmic tail of the integrins chain are transmitted to the ectodomains through the TMD (Kim *et al.,* 2011).

1.1.1.4 Ligand binding

Integrins can be classified in four main classes, reflecting their ligand specificity. One third of the integrins – including the integrin αvβ8 – recognise a highly conserved tripeptide Arg-Gly-Asp (RGD) sequence found on several molecules such as fibronectin. Importantly, the β2 subunit is mainly restricted to blood cells. Also, the β1 subunit binds laminin and collagen when it combines with some αI-domain-bearing α subunits. The combination of the β1 subunit with different α chains identifies another subclass of integrins, specialised in binding laminin molecules (Barczyk *et al.,* 2009) (Figure 1.3).

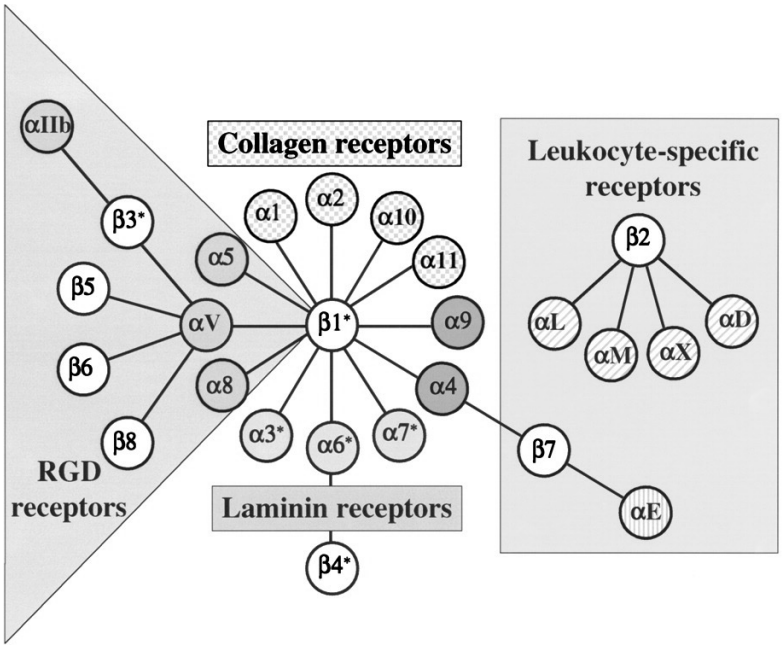


Figure 1.3 | Group classification of integrin molecules. The graph shows integrin subunits in mammals, their αβ association, and group specificity. Straight lines depict

heterodimer associations and how they are clustered into four subgroups. The figure is modified from Hynes, 2002 (Hynes, 2002).

The ligand-binding site is the junction between of the βI domain and β propeller – on the β and α chains, respectively. Evidence suggests that ligand engagement occurs at the interface between the β propeller in the α subunit and the βI domain and is directly dependent on divalent cations coordinating specific binding sites. Indeed, as reviewed by Campbell and Humphries, a divalent Mg2+ cation was described to regulate ligand binding in a motif called the metal-ion-dependent adhesion site (MIDAS), present in the α-I domain and β-I domain. The MIDAS in the β-I domain is flanked by two divalent Ca2+ cations, one per side. One of these sites, adjacent to MIDAS (ADMIDAS), coordinates an inhibitory divalent Ca2+ cation. A Mn2+ binding to ADMIDAS triggers integrin activation, upon structural rearrangement (I. D. Campbell & Humphries, 2011). The other mentioned flanking site is denoted as the synergistic metal ion binding site which acts as a positive regulator of ligand binding (Pan *et al.,* 2010).

The ligand specificity of an integrin has important implications for its function. If the ligand is part of a rigid structure – e.g., a component of the ECM, the integrin-ligand contact generates a force that change and remodel the ECM while having a role in cell contractility. The ability of integrins to engage with their ligands occur through an inducible process characterised by structural rearrangements (Arnaout *et al.,* 2005).

1.1.2 Structural changes and integrin activation

As mentioned, integrins were initially described as key linker molecules connecting the ECM with the cytoskeleton. Apart from their mechanical role, integrins propagate intracellular signalling cascades thus also act as signalling receptors. Relaying a signalling cascade intracellularly depends on the activation status of the integrin and its interaction with ligand.

It is commonly accepted that integrins oscillate between three distinct conformational states. The bent conformation generally reflects an inactive configuration with low binding affinity for

the ligand. A stretched, open, conformation is usually regarded as the active configuration (Askari *et al.,* 2009). Also, in this latter stage, engagement with a ligand triggers a high affinity state that defines this third structural change (Xiao *et al.,* 2004).

Different mechanisms orchestrate the transition from an inactive to active state, and this occurs through signals occurring in two directions across the plasma membrane, starting either from extracellular or intracellular stimuli. Indeed, signals generated from inside the cell induce a conformational change in the integrin molecule that promote a high binding affinity for the ligand (so-called inside-out signalling). During the inside-out activation process, the increased ligand affinity is thought to be due to structural rearrangements of the integrin chains at the transmembrane and cytosolic portion of the heterodimer (Askari *et al.,* 2009). Changes in the structure are then propagated to the extracellular portion of the integrin leading to its unbending. As reviewed by Askari *et al.,* a common step determining integrin activation is the interaction between the cytosolic tail of the β subunit and a molecule called talin, bearing the PTB domain. In the absence of stimuli, talin is found in an inactive closed conformation – with the PTB domain unable to bind the integrin molecule. Certain cues evoked intracellularly induce structural rearrangements leading to the interaction of talin with the integrin, leading to integrin activation.

Furthermore, if an integrin molecule engages with a specific ligand, this can also lead to propagation of intracellular signalling (so-called outside-in signalling). As mentioned for the inside-out signalling, conformational changes of the integrin are also triggered during outside-in signalling – resulting in a complete detachment of the α and β chains (Askari *et al.,* 2009; Shen *et al.,* 2012).

1.1.3 Integrin signalling

Integrins modulate their ligand binding affinity through a complex activation process requiring several steps. As mentioned earlier, the cytoplasmic tails of integrins come into contact with

intracellular adaptors – the latter inducing conformational changes in the whole integrin molecule leading to an increased ligand affinity.

Integrin signalling specificity depends on subunit pairs inducing binding interactions with specific intracellular signalling molecules (Romer *et al.,* 2006). Upon cell adhesion, this event drives integrins clustering and formation of focal adhesions (Romer *et al.,* 2006) – which is a supramolecular protein complex connecting the intracellular environment to the extracellular one (C. S. Chen *et al.,* 2003). As described below, talins and kindlins regulate key biological processes involved in integrin activation (H. Li *et al.,* 2017).

1.1.3.1 Talin

As mentioned earlier, the β cytoplasmic tail contains conserved motifs that are key in downstream signalling. Talin is a well-characterised integrin linker protein. Indeed, adherent fibroblast, microinjected with talin antibodies, inhibited formation of focal adhesions and thus fibroblast migration (Nuckolls *et al.,* 1992). Talin structure contains a FERM domain – modulating integrin adhesion.

As mentioned earlier, the inactive state of integrin molecules generally consists in a bent closed conformation – with the α and β TMD clasped. Therefore, talin brings about the dissociation of integrin chains by disrupting the salt bridge between α- and β-TMD – also involving the interaction with the NPxY motif in the membrane-proximal fraction of the β subunit with the FERM domain (I. D. Campbell & Humphries, 2011).

Also, talin has been described to connect integrins to actin cytoskeleton through its binding sites specific for the filamentous actin (F-actin) (Hemmings *et al.,* 1996) and other important signalling mediators. Concluding, talin is not only a key integrin activator, but also a protein involved in downstream signalling.

1.1.3.2 Kindlins

Kindlins are another class of molecules required for integrin signalling (Montanez *et al.,* 2008), with alterations associated to severe ailments such as Kindler syndrome (Siegel *et al.,* 2003).

Structurally, kindlins have a high degree of homology with all the molecules containing a PH domain – binding phosphoinositides. Moreover, and similar to talin, kindlins contain an atypical FERM domain but binding to a different, membrane-distal, NPxY target motif on the beta chain (H. Li *et al.,* 2017)

1.1.4 Inside-out signalling

As mentioned previously, integrins are a heterogeneous class of molecules with different biological roles. Integrin molecules cluster and connect the cytoskeleton with the ECM. Catalysis of GTPases RhoA and Rap1 to form focal adhesions occurs in many integrin molecules, upon ligand binding, activation, and interaction with cytoskeleton. Specifically, upon phosphorylation, tyrosine growth factor receptors recruit Crk adaptor molecule, binding to the nucleotide exchange factor RapGEF1, and activate Rap1 (Tanaka *et al.,* 1994). Following RAP1 activation, two effectors are then recruited.

Firstly, Rap-ligand is one signalling mediator interacting with the α cytoplasmic tails (Ij. M. Kramer, 2016). Secondly, Riam has been described to be another effector molecule that binds inositol lipids – i.e., phosphatidylinositol 4,5-diphosphate (PI4,5-P2) – on the plasma membrane where it can interact with Vasp, involved in the actin elongation process, and talin (Ij. M. Kramer, 2016; Torres-Gomez et al., 2020).

When RhoA is loaded with a GTP it is in an active state and catalyses the activation of the phosphatidylinositol 4-monophosphate 5-kinase (PIP5K). PIP5K catalyses the formation of PI4,5-P2 – with the latter being described in many downstream cellular pathways. Indeed, membrane sites enriched with such modified inositol lipids are required for talin attachment to the membrane and prerequisite for its activation (Atherton *et al.,* 2016).

Vinculin, another key component of focal adhesions, is also able to recognise PI4,5-P2 (Thompson *et al.,* 2017). Studies confirmed that vinculin recruitment to focal adhesions needs talin, which contains 11 vinculin binding sites (Atherton *et al.,* 2016). Indeed, once vinculin attaches to the inner modified leaflet of the plasma membrane, the overall structure changes (Gilmore & Burridge, 1996). Structural rearrangement of vinculin enables the molecule to interact with the α-actinin – associated with the F-actin – and talin (Ij. M. Kramer, 2016). Therefore, a strong junctional complex between the extra- and intracellular environment is created, with several other signalling mediators playing an important role in this process.

1.1.5 Outside-in signalling

As previously mentioned, outside-in mechanisms play a critical role in regulating several aspects of cell biology such as cell survival or apoptosis. Indeed, disruption of cell-matrix interactions induced apoptosis, in epithelial cells. Reliance of these cells upon their contact to specific surfaces relays an outside-in signalling, which prevents these cells from dying. In anchorage-dependent cells, this is a regulated phenomenon, regarded as 'anoikis' – i.e., homeless (Frisch & Francis, 1994).

Pathways controlling cell survival are a typical example of outside-in signalling. An example of such a pathway involves a scaffold molecule called paxillin binds the β cytoplasmic tail that, in turn, comes into contact with the focal adhesion protein tyrosine kinase 2 (PTK2). This leads to the auto-phosphorylation of PTK2, which then recruits the tyrosine kinase Src. Subsequently, Src catalyses the phosphorylation of other tyrosine residues on PTK2, promoting maximum catalytic activity (Ij. M. Kramer, 2016). PTK2 also contains a proline-rich region, which can act as a docking site for the SH3 domains found on Bcar1 (Lim *et al.,* 2004). The close contact allows the fully active PTK2 and Src to phosphorylate multiple tyrosine residues on Bcar1. Multiple phosphotyrosines on Bcar1 are therefore docking sites for SH2 domain-bearing proteins (Ij. M. Kramer, 2016).

Several proteins have been described to be the substrate of PTK2, including the phosphoinositide 3-kinase. The latter enzyme catalyses the phosphorylation in position 3 of inositol lipids, converting PI4,5-P2 to phosphatidylinositol-3,4,5-triphosphate (Carnero & Paramio, 2014). This generates a docking site for Pdk1 and Akt. Importantly, the close proximity between the two enzymes allows a Pdk1-mediated activation of Akt (Ij. M. Kramer, 2016). Upon Akt activation, this serine/threonine protein kinase regulates several biological processes – including gene repression of key molecules involved in apoptosis such as the FOXO signalling molecules (Zhang *et al.,* 2011).

The catalytic active PTK2 is also involved in the Ras-Erk pathway, where it recruits the adaptor molecules Grb2 and Sos1 involved in Ras activation (Mitra *et al.,* 2006). This outside-in mechanism of activation contributes in sustain an Erk-mediated signalling cascade. Indeed, a transient propagation of the Erk signalling cascade can be seen in fibroblasts incubated with growth factors and their regulation on cell cycle events (Assoian, 1997).

Concluding, bidirectional signalling pathways displayed by integrins play a central role in different biological context. Of note, conserved regions in the cytoplasmic tails of integrin subunits are pivotal in orchestrating signalling events. Importantly, the integrin αvβ8 does not contain the NPXY conserved motif on the β8 cytoplasmic domain, which in other integrins promotes interaction with FERM domain-bearing proteins such as talin (McCarty, 2020; S L Nishimura *et al.,* 1994). Therefore, likely binding partners interacting with the integrin αvβ8, and downstream signalling pathways are poorly understood.

1.2 The integrin αvβ8

The integrin β8 subunit heterodimerises solely with the αv subunit to form integrin αvβ8, which is classified as an RGD-binding integrin (McCarty, 2020). As reviewed by McCarty, containing the RGD conserved tripeptide sequence, the LAP – which is part of the latent form of the cytokine TGF-β – and vitronectin – a component of the ECM – bind to αvβ8 together with other

RGD containing ligands. Importantly, comparisons of affinity of αvβ8 to different ligands indicated that the latent TGF-β1 has high affinity binding for this integrin (Ozawa *et al.,* 2017). The integrin αvβ8 is mainly expressed in neurons, glial cells, fibroblasts, and epithelial cells, with studies also showing αvβ8 expression on immune cells such as T lymphocytes and DC (Stockis *et al.,* 2017).

1.2.1 Structural and functional role of the integrin αvβ8

Latent TGF-β – a pleiotropic cytokine with multiple regulatory roles, as described below – is a high affinity ligand for the integrin αvβ8 (Ozawa *et al.,* 2017). LAP, which is part of the latent TGF-β complex, specifically engages with the integrin αvβ8 through the recognition of the conserved RGD peptide sequence. In 2007, Yang *et al*. demonstrated that a single point mutation of the RGD site on the latent TGF-β converting the Aspartate to a Glutamate (RGE) alters integrin binding. Importantly, mice with such a mutation are phenotypically similar to TGF-β KO mice developing systemic inflammation and defect in vasculogenesis (Yang *et al.,* 2007). These findings suggested the importance of RGD-binding integrins as major activators of TGF-β in regulating several biological contexts, such as the immune system (Worthington *et al.,* 2011). Campbell *et al*. identified the mechanisms by which the integrin αvβ8 drives the activation of TGF-β, using cryo-electron microscopy. Specifically, this study elucidated a mechanism by which the integrin αvβ8 does not require actin cytoskeleton force to drive the cytokine activation (M. G. Campbell *et al.,* 2020) . Indeed, the latent TGF-β is bound to the integrin αvβ8 in the extended-closed conformation – accommodating ligand binding and regulating cytokine activation (M. G. Campbell *et al.,* 2020). As indicated by Campbell and colleagues, cytokine activation does not rely on release or diffusion mechanisms, based on cell-based experimental studies. Also, the mature cytokine does not diffuse and initiate a signalling cascade on the same cell expressing the latent TGF-β (M. G. Campbell *et al.,* 2020).

Importantly, rather than oscillating between different conformations, structural studies suggest that the αvβ8 is found partially active and in an extended-closed conformation, and does not adopt a bent, inactive conformation (McCarty, 2020). Also, integrin αvβ8 is less responsive to Mn2+, which can, at the most, increases affinity for LAP three-fold, whereas this cation can increase affinity of other integrins for their ligands approximately fifty-fold (J. Wang *et al.,* 2017). Mounting evidence suggests that integrin αvβ8 constitutively engages with ligand in an extended closed-conformation (Cormier *et al.,* 2018). Importantly, the canonical inside-out signalling pathways seen for other integrin molecules does not trigger ligand engagement by αvβ8 (McCarty, 2020). Apart for the integrin αvβ8, the integrin αvβ6 is another molecule described to be a TGF-β activator. Nevertheless, functional divergences exist between the two types of integrins.

1.2.1.1 Functional divergences in TGF-β activation by the integrin αvβ8 and αvβ6

In addition to integrin αvβ8, integrin αvβ6 can also bind to latent TGF-β. However, it is believed that the two integrins activate the cytokine in different ways. An inside-out mechanism promotes a conformational switch in αvβ6 leading from a low to high ligand binding affinity state transition. As mentioned earlier, rather than in an inactive bent-closed structure seen in the other integrin molecules such as the αvβ6 (Dong *et al.,* 2017), a constitutively active extended-closed configuration has been described for αvβ8. Indeed, a force generated from the cytoskeleton is not needed to generate radical conformational rearrangements in the αvβ8 heterodimer (M. G. Campbell *et al.,* 2020). Moreover, as previously mentioned, αv integrins such as αvβ6 have been described to bind several ligands (McCarty, 2020) whereas, in αvβ8 molecules, effective binding can only occur through a direct engagement with the latent TGF-β (M. G. Campbell *et al.,* 2020).

In 2002 Mu *et al*. demonstrated that the β8 cytoplasmic chain is dispensable for normal TGF-β activation (Mu *et al.,* 2002). On the other hand, for the integrin αvβ6, the β cytoplasmic chain is required for freeing the active cytokine moiety in the extracellular milieu (Munger *et al.,* 1999). Importantly, αvβ6 triggers a global conformational change upon TGF-β activation that is transmitted intracellularly (Dong *et al.,* 2017). Cryo-electron microscopy experiments elucidated mechanisms by which αvβ8-mediated TGF-β activation signals in the juxtacellular region and that cytokine diffusion is not required (M. G. Campbell *et al.,* 2020).

1.2.2 Signalling events triggered by the integrin αvβ8

As mentioned previously, the NPXY conserved motif in the β subunit cytoplasmic domain is required for propagating intracellular signalling by most integrins. However, the β8 subunit lacks this conserved region, suggesting that signal mediators are distinct in αvβ8. There is limited data on proteins that interact with the integrin β8 tail, which are described below.

The Rho GDP-dissociation inhibitor (GDI) 1 regulates the activation of the small GTPase Rho and interacts with αvβ8 at the leading-edge of cells in culture (Reyes *et al.,* 2013). The protein tyrosine phosphatase (PTP)-PEST positively regulates RhoGDI1. At the cell protrusions, activation of the GTPases Cdc42 and Rac1 is modulated by a Src-phosphorylated RhoGDI1. In turn, in complex with αvβ8, PTP-PEST stimulates RhoGDI1 detachment from the cell membrane with subsequent recruitment of GDP-loaded GTPases thereby controlling migrating cell behaviour in astrocytes (Shin *et al.,* 2015).

Also, studies have identified that Band 4.1B a likely interaction with β8 binding partners in regulating their structural switches. McCarty *et al*. showed that a 31-amino acid-long region of the β8 cytoplasmic tails is enriched with conserved motives that may serve as binding sites for the Band 4.1B proteins. It was then hypothesised that Band 4.1 could trigger an outside-in

signalling mechanism and orchestrate the activation of the integrin αvβ8. Indeed, astrocytes plated on the latent TGF-β induced colocalization of the integrin αvβ8 and Band 4.1B (McCarty *et al.,* 2005).

These experiments show evidence that the integrin αvβ8 interacts with intracellular signalling partners. Yet, it is still not clear whether functional signalling events are propagated inside cells upon integrin αvβ8 engagement with LAP. Moreover, given that the αvβ8 plays an important role in immunity through a direct regulation of TGF-β activity (McEntee *et al.,* 2020), it would be of interest to determine signalling pathways generated downstream of the integrin αvβ8 in immune cells as key regulators of TGF-β function.

Given the pleiotropic nature of TGF-β, biological role and modulation of its activity will be described below.

1.3 TGF-β and its biological role

1.3.1 TGF-β regulation

TGF-β proteins are a structurally-related family of proteins playing important roles in many biological contexts. TGF-β family members can be further divided in subfamilies, including TGF-β - consisting of three members – i.e., TGF-β1, 2, and 3 (if not stated otherwise, we will refer to the TGF-β1 as TGF-β). Studies confirmed that TGFβ isoforms elicit a similar function *in vitro*. Nevertheless, mice lacking individual isoforms have a distinct phenotype. These data confirm that, *in vivo*, each isoform has non-redundant biological roles (Annes *et al.,* 2003).

Whilst TGF-β receptors are ubiquitously expressed throughout the body, the latent ligand cannot directly bind to its receptor. So, the functional role of TGF-β is regulated at the level of its activation (Worthington *et al.,* 2011). Indeed, TGF-β molecules are produced as inactive complexes and released by the cell as non-biologically active molecules (Worthington *et al.,*

2012) (Figure 1.4). A single product is encoded by TGF-β genes consisting of a pro-peptide fragment called LAP and the portion corresponding to the biologically active TGFβ cytokine – at the N- end C-terminal end, respectively (Worthington *et al.,* 2012). A dimer is formed, then a Golgi-based enzyme called furin cleaves between LAP and the active TGF-β fragment (McEntee *et al.,* 2020). Nevertheless, LAP interacts via a non-covalent association with the cytokine moiety in a so-called 'straightjacket' conformation preventing the active moiety engaging with its receptor and forming the small latent complex (Worthington *et al.,* 2012). Also, LAP and TGF-β forms the large latent complex upon association with the latent TGF-β binding protein (LTBP, Figure 1.4); This complex is tethered to the ECM thus localising TGF-β in the extracellular environment (Worthington *et al.,* 2011).

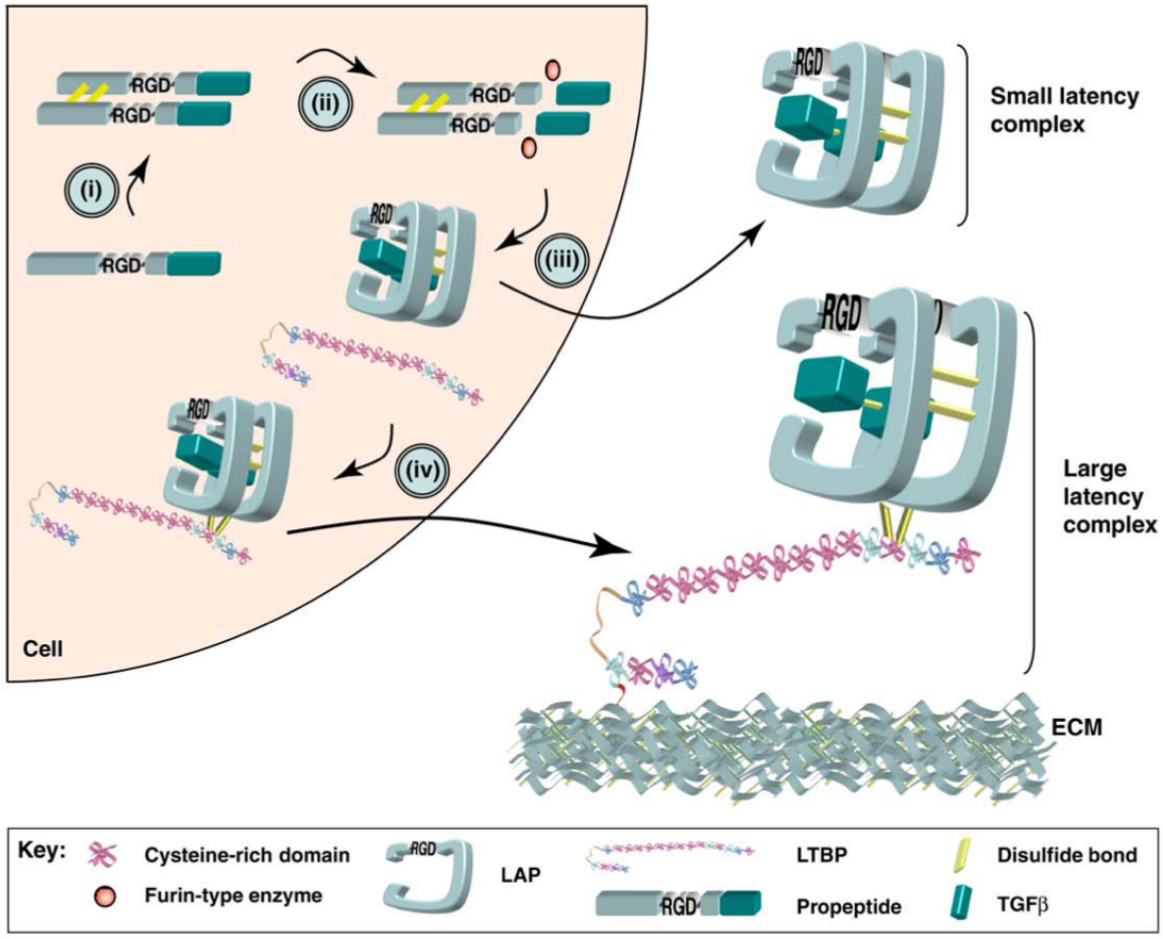


Figure 1.4 | TGF-β synthesis. (i) The just-translated TGF-β molecule, as a pro-peptide, homodimerises. (ii) A furin-type enzyme catalyses a proteolytic reaction resulting in the maturation of the pro-peptide. (iii) The biologically active TGF-β molecule is retained in an inactive state upon non-covalent binding with the LAP molecule – forming the small latency complex (SLC). Cells can secrete SLC. (iv) Alternatively, LTBP can interact with the SLC,

through a disulphide bond formation, and form the large latency complex which can also be release by the cells. The latter tethers the ECM as a ready-made product available to other cells. The figure was taken from a review by Worthington *et al*. (Worthington *et al.,* 2011).

TGF-β activation can occur through physical and chemical mechanisms such as heat, low pH, reactive oxygen species, and catalytic activity by proteases. However, as mentioned earlier, studies have highlighted a key role for αv integrins in regulating TGF-β activity (McEntee *et al.,* 2020).

1.3.2 Integrin-mediated TGF-β activation

As mentioned already, certain integrins bind latent TGF-β via the RGD tripeptide sequence present on LAP. Different integrins appear to activate latent TGF-β in different ways (Worthington *et al.,* 2011). Cell cultures experiments confirmed that integrin αvβ5- and αvβ6-mediated TGF-β activation occurs independently from proteases (Munger *et al.,* 1999; Wipff *et al.,* 2007). These results indicate that αvβ5 and αvβ6 do not require proteolytic cleavage to regulate the activation status of TGF-β. Markedly, αvβ6 engaging with the latent TGF-β does not release the active moiety in culture medium (Munger *et al.,* 1999). Therefore, evidence suggests that integrin αvβ6 drives a conformational change enabling the active cytokine to bind to the TGF-β receptor (Worthington *et al.,* 2011).

Cytochalasin D, an actin polymerization inhibitor, reduced the activation of TGF-β in αvβ5- and αvβ6-expressing cells. These results indicated that these two integrin molecules require cell contraction to regulate biological activity of TGF-β (Munger *et al.,* 1999; Wipff *et al.,* 2007). It is worth noting that LTBP1 is required for the αvβ6-mediated TGF-β activation where it targets the inactive cytokine to the ECM (Annes *et al.,* 2004). Therefore, a pulling force is proposed to be propagated by actin filaments linked to the cytoplasmic tails of the LAP-engaged αvβ5 and αvβ6. Also, a holding force in the opposite direction is generated by the LTBP1 (Worthington *et al.,* 2011). The joint activity of these forces frees the active TGF-β that is able to bind to its cognate receptor (Worthington *et al.,* 2011).

1.3.2.1 αvβ8-mediated TGF-β activation

Integrin αvβ8-mediated TGF-β activation was initially proposed to be regulated in a different way compared to αvβ5 and αvβ6. Indeed, this process was proposed not to rely on cell contraction (Worthington *et al.,* 2011). Instead, initial studies suggested that αvβ8 requires metalloprotease activity to activate TGF-β as this activation is prevented by specific metalloprotease inhibitors (Mu *et al.,* 2002). Moreover, Mu *et al*. demonstrated that TGF-β engagement with αvβ8 resulted into a matrix metalloproteinase-14 (MMP-14)-dependent release of the active cytokine (Mu *et al.,* 2002). Nevertheless, a MMP-14-dependent activation of TGF-β has been tested in the respiratory tract and in human monocytes (Araya *et al.,* 2007; Mu *et al.,* 2002).

1.3.3 Classical TGF-β signalling

Upon activation, TGF-β binds to its receptor and propagates a signalling cascade inside the target cell. TGF-β receptor is made of two subunits – type I and II receptor subunits – eliciting both a serine/threonine and tyrosine kinase activity (Heldin & Moustakas, 2016). Ligand engagement triggers receptor dimerization of the type II receptor and induces the type II receptor cytoplasmic tail to catalyse phosphorylation of the type I receptor cytoplasmic tail on the GS domain – a region enriched with glycine and serine residues (Heldin & Moustakas, 2016). This phosphorylation event causes activation of the type I receptor catalytic activity.

The catalytically active type I receptor phosphorylates serine and threonine residues of the receptor-activated-Smad members, Smad2 and 3 (Feng & Derynck, 2005). The protein Smad4 then associates with phosphorylated receptor-activated-Smads to form a trimeric complex that enters the nucleus and regulates expression of specific genes (Y. Shi & Massagué, 2003). A negative feedback mechanism also exists to shut down TGF-β signalling, with the cascade promoting expression of the negative regulator Smad7. Once expressed, Smad7 is retained into the nucleus bound to Smurf2, a ubiquitin ligase (Heldin & Moustakas,

2016). Following TGFβ signalling, the Smad7/Smurf2 complex attaches to the plasma membrane and ubiquitinates the type I receptor. The ubiquitinated target is then degraded, after having entered the proteasomal degradation pathway (Kavsak *et al.,* 2000), thus dampening TGFβ signalling.

TGF-β signalling can also occur through Smad-independent routes, indicating that several signalling pathways are evoked in different settings (Worthington *et al.,* 2012). Therefore, as mentioned above, regulation of TGF-β activity is strictly controlled in different biological contexts such as immune responses.

1.4 TGF-β and its role in controlling immune responses

For an appropriate immunological response to be mounted, a tight regulation is needed to ensure a proper production of required lymphocyte subsets in appropriate number (David & Massagué, 2018). Depending on the specific context, TGF-β can elicit an anti- or pro-inflammatory role (Worthington *et al.,* 2012). Therefore, TGF-β cytokines can orchestrate several biological roles in immunity (Sanjabi *et al.,* 2017).

1.4.1 TGF-β and T lymphocytes

Completely differentiated T cells that leave the thymus in search of their cognate antigen are regarded as naïve T cells. Naïve T cells continuously recirculate throughout the body where these cells access, from the bloodstream, the secondary lymphoid tissues thus reaching the white pulp or T cell zones – in the spleen and lymph nodes, respectively (Jameson & Masopust, 2018). As it will be described below, upon recognition of its specific antigen presented by the major histocompatibility complex (MHC), in the context of co-stimulation, naïve T cells can proliferate and differentiate into effector T cells. There are several classes of effector T cells, each specialised in eliciting a specific task. Markedly, TGF-β acts in an articulated and context-dependent manner in T cell homeostasis.

1.4.1.1 TGF-β regulation of CD8+ T cells

Upon recognition of a specific antigen presented on MHC-I and in the presence of co-stimulation, CD8+ T cells can differentiate into cytotoxic T cells (CTLs) and subsequently kill infected target cells. In mice expressing a dominant negative form of type II TGF-β receptor in all T cells – blunting TGF-β signalling – the CD8+ T cell population was expanded (Lucas *et al.,* 2000). Also, mice in which TGF-β signalling is altered in T cells showed a dysregulation of CD8+ T cell development (M. O. Li *et al.,* 2006). The effector function of CTLs in provoking cell death of target cells is also regulated by TGF-β. For example, TGF-β reduces the expression of perforin (Smyth *et al.,* 1991) and suppresses IFN-γ production by repressing the expression of the transcription factor T-bet (Ahmadzadeh & Rosenberg, 2005).

1.4.1.2 TGF-β regulating T helper 1 and 2-mediated responses

In contrast to CD8+ T cells, CD4+ T cells can be divided in several, functionally distinct, subsets. Activation of effector T cells relies on antigen recognition, presented by MHC-II by antigen presenting cells such as DC. Specifically, T cell differentiation is a complex mechanism that classically requires three distinct stimulations. As will be described below, the first two depend on signalling cascades propagated by the T cell receptor (TCR) and costimulatory molecules (Korn *et al.,* 2009) upon encounter of DC. Then, various cytokines can generate the third signal required to trigger the appropriate cell fate differentiation (Korn *et al.,* 2009).

As reviewed by Zhu and colleagues, CD4+ T cells, in the presence of IL-12 and IFN-γ, undergo T helper 1 cell differentiation – described to play a key role in defending the host against intracellular pathogens. TGF-β suppresses T helper 1 differentiation. Indeed, type 2 TGF-β receptor KO mice on T cells develop a T helper 1-cell-mediated aberrant inflammatory disease (M. O. Li *et al.,* 2006; Marie *et al.,* 2006). TGF-β can

directly suppress T helper 1 cell development through a negative regulation of T-bet – the master regulator of T helper 1 cell differentiation (Gorelik *et al.,* 2002).

Furthermore, T helper 2 cell fate specification is evoked in the presence of IL-4; these cells have been described to mount an immune reaction against large extracellular pathogens such as helminths (Jinfang Zhu, 2018). Crossing mice lacking type 2 TGF-β receptor molecule – the phenotype of which is an aberrant T helper 1-mediated multiorgan inflammation – with mice lacking the expression of the T helper 1 master regulator – i.e., T-bet – generated an offspring developing a T helper 2-mediated inflammation (M. O. Li *et al.,* 2006). Studies show that TGF-β inhibits the expression of GATA3 – the T helper 2 cell master regulator (Gorelik *et al.,* 2000). Indirectly, TGF-β can also regulate T helper 2 cell differentiation through the regulation of Sox4 – suppressing GATA3 transcriptional activity – and repressing the release of IL-5 (Kuwahara *et al.,* 2012). Together, these results highlight the importance of TGF-β in limiting T helper 1 and 2 cell responses.

1.4.1.3 TGF-β orchestrating regulatory T cells (Treg) / T helper 17 axis

T helper 17 cells make up another important group of immune cells, important in protecting host cells against fungi and extracellular bacteria. T helper 17 cell polarisation can occur in presence of TGF-β and IL-6 (Korn *et al.,* 2009). As reviewed by Korn *et al.,* expression of RORγt – the master regulator of T helper 17 cells – is enhanced in the presence of IL-6 and IL-1β. In a context where the pro-inflammatory cytokine IL-6 is missing, Foxp3 – the master regulator of Treg – is induced (Korn *et al.,* 2009). Noteworthy, IL-6 KO mice show defects in T helper 17 cell development in favour of a much higher proportion of Treg (Korn *et al.,* 2007). These studies confirm that a finely regulated axis exists between Treg and T helper 17 cell development.

1.4.1.3.1 TGF-β regulating T helper 17

T helper 17 cells were regarded as a unique subset of effector T cells during organ-specific autoimmunity. In fact, in this setting, a subset of T helper cells mounted responses against autoantigens triggering tissue inflammation (Korn *et al.,* 2009). Mice with blunted TGF-β signalling (expressing a type 2 TGF-β receptor under the control of the CD4 promoter) have a reduced amount of T helper 17 cells following induction of experimental autoimmune encephalomyelitis (Veldhoen *et al*., 2006). Other cytokines are involved in T helper 17 development such as IL-1β and IL-23. Studies conducted by Ghoreschi *et al*. indicated that TGF-β could be dispensable for T helper 17 development as RORγt was induced by a concerted activity of IL-6, IL-1β, and IL-23 (Ghoreschi *et al.,* 2010). More work is needed to elucidate the functional importance of TGF-β in regulating T helper 17 cells.

1.4.1.3.2 TGF-β regulating Treg

TGF-β promotes differentiation of Treg by increasing Foxp3 expression. Indeed, the phenotype of mice lacking TGF-β expression mirrors that of mice with functional loss of Foxp3 (Brunkow *et al.,* 2001). Treg falls into two, developmentally-distinct, categories: thymic (t-Treg) and peripheral Treg (p-Treg). As the name suggests, t-Treg complete their development within the thymus and represent the largest fraction of the Treg pool in the body (Sanjabi *et al.,* 2017). Liu and colleagues showed that conditional loss of the type I TGF-β receptor consistently leads to a significant reduction of thymocytes marked positive for Foxp3 expression, in 3- to 5-day old mice. This indicates that TGF-β controls early stages of t-Treg development (Liu *et al.,* 2008).

p-Treg are induced in peripheral tissues upon induction of Foxp3 in CD4+T cells. Of note, p-Treg induction occurs in concerted action of IL-2 and retinoic acid (RA) that, together with TGF-β, support expression of Foxp3 in differentiated naïve T cells. This differentiation relies on the classical Smad signalling pathway – as CD4+ T cells lacking the expression of Smad2 and 3 were not able to induce Foxp3 in presence of TGF-β (Takimoto *et al.,* 2010). Several studies suggested the importance of p-Treg in immune regulation at peripheral sites. Indeed, inflammation was generated in the intestine and the lung in mice where the induction of p-Treg was abrogated (Josefowicz *et al.,* 2012).

TGF-β is not only crucial during Treg development. Indeed, in fully differentiated Treg, TGF-β can also contribute to the immunosuppressive function of such cells. A transfer model of colitis – involving the transfer of Treg and naïve CD4+ T cells in mice lacking lymphocytes- has been used to elucidate the role played by TGF-β in Treg function (Powrie *et al.,* 1996). As reviewed by Kelly and colleagues, severe colitic inflammation was generated upon transferring naïve T cells into lymphocyte-deficient, whereby Treg were able to inhibit development of this inflammation (Kelly *et al.,* 2017). Moreover, Fahlen *et al*. found that in a transfer model Treg were not able to suppress inflammation if the ability of CD4+ T cells to sense TGF-β was reduced (Fahlén *et al.,* 2005).

In 2008, Zhou *et al*. identified that TGF-β induces Foxp3 and RORγt. This suggests that the discriminating factor is the environmental context in regulating cell fate differentiation. Indeed, Treg are induced in the presence of IL-2 and TGF-β, whereas IL-6, and IL-1β plus TGF-β and IL-21 drive T helper 17 differentiation (Horwitz *et al.,* 2008; L. Zhou *et al.,* 2008). Interestingly, Foxp3 can inhibit IL-17 production by cells.

This shifts the differentiation fate more toward the Treg phenotype by directly binding to RORγt (L. Zhou *et al.,* 2008). Finally, it has been hypothesised that the level of TGF-β is key in determining cell fate. Indeed, naïve T cells undergo T helper 17 differentiation when levels of TGF-β are relatively low; on the contrary, Foxp3 expression is upregulated by high concentrations of TGF-β which induces a regulatory phenotype (L. Zhou *et al.,* 2008).

1.4.1.4 TGF-β and immunological memory

A primary cell-mediated immune response involves activation of naïve T cells, in concert with an increased proliferation rate and differentiation into effector cells. In parallel, a primary response triggers the formation of a long-lived cell pool referred to as memory T cells. Surviving the absence of their antigens, the memory cells are poised to mount a vigorous immune reaction upon reencounter of the infective pathogen (MacLeod *et al.,* 2009). TGF-β has been described to have a role in orchestrating memory T cell response.

1.4.1.4.1 Central memory T cells (Tcm) and Tem biology

As previously mentioned, CCR7 – as well as expression of CD62L – is key in regulating homing to secondary lymphoid tissues. Studies conducted by Sallusto *et al*. identified CCR7 as an important marker to separate two functionally distinct subsets of human memory T cells. Cells marked negative for CCR7 expression are referred to as Tem (Sallusto *et al.,* 1999). Expressing high levels of chemokine receptors, indicating capacity of these cells to migrate towards inflamed sites (Sallusto *et al.,* 2004), Tem have been also described to express β1 and β2 integrins (Sallusto *et al.,* 1999). On the contrary, Tcm are marked positive for CCR7 expression (Sallusto *et al.,* 1999), being these more prone to recirculate as naïve T cells. As reviewed by MacLeod *et al.,* the CD45

gene undergoes a process of alternative splicing producing several isoforms including CD45RA and CD45R0 (MacLeod *et al.,* 2009) – the latter of which characterises the Tcm subset (Sallusto *et al.,* 2004). CD4+ and CD8+ T cells can either differentiate into functionally distinct Tem or Tcm (Sallusto *et al.,* 2004). Tem are capable of releasing IL-4, IL-5, and IFN-γ upon antigen reencounter; also, the largest amount of perforin is brought by CD8+ Tem due to their expression of CD45RA (Sallusto *et al.,* 2004). Finally, Tcm – through the upregulation of CD40L – are more sensitive to antigen stimulation. Upon TCR engagement, Tcm release IL-2. Furthermore, upon proliferation and differentiation to effector cells, Tcm produce IL-4 or IFN-γ. Therefore, Tem act immediately after re-infection, whereas Tcm provide more prolonged protection (Lanzavecchia & Sallusto, 2005).

Studies showed that TGF-β seems to play an inhibitory effect on the Tcm subset. Indeed, expression of CD62L and CD44 increased following blockade of TGF-β, with the use of a blocking antibody, or after inhibition of a type I TGF-β receptor kinase, in splenic CD8+ T cells. Data support the hypothesis that TGF-β elicits a suppressive function upon Tcm precursors by acting in autocrine fashion (Ma & Zhang, 2015; Takai *et al.,* 2013). Importantly, TGF-β inhibits the Tcm phenotype even in fully differentiated cells. Indeed, CD62L was lost and accompanied by a reduction in IFN-γ production, when Tcm were incubated with TGF-β (Takai *et al.,* 2013). On the contrary, very little is known about TGF-β, its activation, and Tem biology.

1.4.1.4.2 Tissue-resident memory T cells (Trm) biology

Given the lack of CCR7 expression, it was not clear how Tem could egress from non-lymphoid tissues and recirculate between these, and blood and

lymphatic circulation (Schenkel & Masopust, 2014). As reviewed by Schenkel and Masopust, seminal works suggested that memory T cells were, in fact, able to reside in non-lymphoid tissues. Indeed, Trm precursors arrive and permanently remain in peripheral tissues during immune cell responses. Phenotypic changes induced in the Trm groups are triggered soon after migration to tissues. For example, following a viral infection in mice, effector T cells transiently increased integrin $\alpha4\beta7$ expression – required for these cells to reach the small intestine (Masopust *et al.,* 2010). In mice, the arrival to small intestine by CD8+ effector cells mark $\alpha4\beta7$ repression. It is worth mentioning that bioinformatic analysis indicated that Trm are transcriptionally distinct from naïve T cells, Tcm and Tem (Mackay *et al.,* 2013).

TGF-β has been observed to induce upregulation of CD103, the αE integrin subunit that forms integrins αEβ7, which binds to the adhesion molecule E-cadherin found on epithelia. As reviewed by Kelly and colleagues, many studies indicated that TGF-β is a central factor inducing CD8+ CD103+ Trm at barrier sites such as skin, lung, and gut. Importantly, keratinocytes have been described to express $\alpha v\beta6$ and $\alpha v\beta8$, regulating TGF-β activation, and drive Trm localization to the skin during a viral infection (Mohammed *et al.,* 2016). Also, Mani and colleagues found that TGF-β is involved in conditioning naïve CD8+ T cells, from lymph nodes and during homeostasis, to differentiate into Trm. This conditioning effect was dependent on the expression of αv-integrins on DC (Mani *et al.,* 2019). Importantly, many Trm that do not express CD103 home to several tissues throughout the body (Mueller & Mackay, 2016; Steinert et al., 2015). This evidence indicates that some biological mechanisms regulating Trm homing to non-lymphoid tissues needs to be clarified.

1.4.2 DC and the role played by TGF-β

Generation of an appropriate adaptive immune response requires the activation of DC. DC are bone marrow (BM)-derived leukocytes patrolling the body and can recognise pathogens and present antigen to T cells via MHC proteins. (Merad *et al.,* 2013). Cytologically, these cells are provided with long processes resembling neural dendrites – from which their name is derived.

Importantly, DC respond to several harmful insults leveraging their vast repertoire of innate receptors – which are able to recognise molecular patterns associated with microbial signatures – i.e., pathogen-associated molecular patterns – or damage-associated molecular patterns exposed by damaged cells. These molecular cues are the target of the so-called pattern recognition receptors (PRRs), including the Toll-like receptors (TLRs) family, which is possessed by cells of the innate immunity such as DC (Reis e Sousa, 2004). Engagement of PRRs with cognate ligands induce the maturation process in DC (Soto *et al.,* 2020). This is a phenotypic change consisting in a much greater expression of MHC-peptide complexes and costimulatory molecules CD80, CD86, and CD40. Also, DC maturation consists in the upregulation of the chemokine receptor CCR7 – allowing DC to migrate to lymphoid from peripheral tissues (Cabeza-Cabrerizo *et al.,* 2021).

Therefore – thanks to ICAM-1, ICAM-2, and CD58 expressed on the surface of DC – phenotypically mature DC can interact with circulating naïve T cell through an efficient interaction of LFA-1 and CD2, respectively. This interaction is crucial for allowing naïve T cells to eventually recognise the MHC-peptide complex that induce a conformational rearrangement of LFA-1 thereby increasing ligand binding (Murphy *et al.,* 2012).

As mentioned previously, three distinct signals are propagated by DC to naïve T cells for inducing clonal expansion and cell differentiation (Kapsenberg, 2003). An activation signal is relayed downstream upon TCR engagement with MHC-peptide complex and further interaction with CD4 or CD8 co-receptors. Then, a co-stimulatory signal is also induced, for example by CD80 and CD86 costimulatory molecules expressed on DC surface binding to

CD28 expressed on T cells. As mentioned previously, a third differentiation co-stimulatory signal is relayed inside naïve T cells upon exposure to cytokines in the environment which shapes the phenotypic role played by effector T cells (Kapsenberg, 2003).

This process induces naïve T cells to enter G1 phase and express high-affinity IL-2 receptor. Therefore, engagement of IL-2 receptor with its cognate ligand induces activated T cells to progress through the cell cycle and clonally expand (M. Shi *et al.,* 2009). Of note, ICOSL is an important molecule expressed on DC that interacts with ICOS (Greenwald *et al.,* 2005). Signals propagated upon this interaction do not induce IL-2 expression, key to induce class switching and induce helper cell function – in B and T cells, respectively. Also, bi-directional signalling is generated by the interaction between CD40, on DC, and its cognate ligand CD40L on T cells. This signalling pathway also relays costimulatory signals to the target T cell (Kapsenberg, 2003). Finally, the interaction between OX40, expressed on T cells, and its ligand – i.e., OX40L – has been described to enhance cell proliferation and survival (Sun *et al.,* 2018). Therefore, DC are key in regulating several aspects of immune responses. Given the crucial role played in immunity, these cells are functionally distinct in different subsets. Therefore, in line with the scope of our project, a general classification of DC is described below.

1.4.2.1 Plasmacytoid and conventional dendritic cells (cDC)

DC are classically divided into 2 major groups: plasmacytoid (pDC) and cDC. As reviewed by Merad *et al.,* pDC – mainly populating the blood and lymph nodes – have been described to express TLRs 7 and 9 – designated to the recognition of viral RNA and DNA. After sensing foreign nucleic acid, pDC release type-1 interferon and increase their ability of antigen presenting cells (Merad *et al.,* 2013).

On the other hand, cDC can be found in lymphoid and non-lymphoid tissues where they can sample tissue antigens and, thanks to their enhanced ability to process and

present molecules (Villadangos & Schnorrer, 2007), prime T lymphocytes (Banchereau & Steinman, 1998).

DC express CD11c, MHC-II, and CD45. As some of the hemopoietic markers such as CD11c are shared with other cell subsets – i.e., macrophages – cDC can be discerned from the latter for their differential expression of the tyrosine kinase receptor Flt3 (Merad *et al.,* 2013). Also, cDC specifically express the transcription factor Zbtb46 and the homing receptor CCR7 (Förster *et al.,* 1999; Meredith *et al.,* 2012).

There are two major subtypes of cDC, referred to as cDC1 and cDC2. CD8α+ DC and CD103+ DC belong to the cDC1, whereas cDC2 express CD11b (Soto *et al.,* 2020). As reviewed by Cabeza-Cabrerizo *et al.,* in lymphoid tissues, cDC1 express the CD8αα marker, whereas this marker is missing on equivalent cells in nonlymphoid organs. Also, in the periphery, cDC1 have an increased expression of CD103 or CD24. Furthermore, cDC1 express XCR1 and DNGR-1 – a chemokine and C-type lectin receptor, respectively (Cabeza-Cabrerizo *et al.,* 2021).

Markedly, cDC1 are IRF8- and Batf3-dependent, whereas cDC2 are IRF4-dependent (Reynolds & Haniffa, 2015). Also, Flt3 is more expressed in cDC1 compared to cDC2, with the latter expressing higher levels of CD11b and Sirpα (Cabeza-Cabrerizo *et al.,* 2021). Irf8 has a specific, nonredundant, role in cDC1 development, acting to inhibit the development of granulocytes while promoting differentiation of myeloid cells (Cabeza-Cabrerizo *et al.,* 2021; Merad *et al.,* 2013). Indeed, Irf8-/- mice are characterised by a myeloproliferative disease and this is phenocopied by a point mutation in the Irf8 locus of BHX2 mice, a mouse strain phenotypically similar to Irf8-/- mouse – with altered CD8α+ DC development (Tailor *et al.,* 2008).

cDC2, are the other major cDC subset – which has been described to be more heterogeneous and abundant than cDC1. cDC2 are marked positive for CD4, Sirpα,

and CD11b expression. The signalling protein Traf6, Irf2, and Irf4 can all regulate cDC2 development (Cabeza-Cabrerizo *et al.,* 2021). For example, mice lacking the expression of Irf4 have an impaired number of CD4+ DC with no developmental alterations of the CD8α+ DC subset (Merad *et al.,* 2013).

1.4.2.2 Functional roles played by cDC

Of note, cDC are strategically located throughout the body to sample antigens. For example, long projections are extended by cDC1 into the conducting airways to detect any potential foreign airborne pathogen (Guilliams *et al.,* 2013). Markedly, intestinal luminal pathogens are detected by long interdigitations of cDC1 found in the epithelial cell layer (Chieppa *et al.,* 2006).

Double-stranded, viral, RNA engages with the endosomal TLR3 – which is highly expressed on cDC1 (Schulz *et al.,* 2005). As reviewed by Cabeza-Cabrerizo *et al.,* this appears to be key in anti-viral reactions. Engagement with TLR11 and TLR12 ligands – recognising profilin-like molecule – triggers massive production of IL-12 (Cabeza-Cabrerizo *et al.,* 2021). Damaged cells release extracellularly the F-actin recognised as DAMP by high expression of DNGR-1 on cDC1 (Cabeza-Cabrerizo *et al.,* 2021) – which is key during cytotoxic and memory responses (Iborra *et al.,* 2012, 2016). As previously mentioned, CD8α+ DC are an important class of cDC1 – marked by their expression of CD8αα – and residing in primary and secondary lymphoid organs, where they make up 20-40% the total DC fraction (Merad *et al.,* 2013). Importantly, Flt3-/- mice have a significantly reduced amount of CD8α+ DC, indicating this is important in the expansion of this cell subset (McKenna *et al.,* 2000). Splenic CD8α+ DC are strategically located in the marginal zone, where they sample blood antigens. In the lymph nodes, CD8α+ DC are in the subcapsular sinus, encountering antigens drained from peripheral tissues, at the entry site of the afferent lymphatics (Merad *et al.,* 2013). Markedly, CD103+ DC are another cell subset grouped in the cDC1 compartment and

expressed in many tissues such as the gut and lungs (Soto *et al.,* 2020). As reviewed by Soto *et al.,* CD103+ and CD8α+ DC share important surface markers such as CD24, CD205, and langerin. On the contrary, while CD8α+ DC can initiate T helper 1 responses due to secretion of IL-12p70 – following cell activation, CD103+ DC are more prone to induce T helper 17 response upon release of IL-1β and IL-6 (Soto *et al.,* 2020). Studies confirmed that CD8α+ and CD103+ DC have an enhanced ability activate CD8+ T cells due to their higher cross-presentation and cell priming potential compared to cDC2 (Merad *et al.,* 2013).

On the other hand, cDC2 exclusively express TLR7 – which recognises exogenous single-stranded RNA (Cabeza-Cabrerizo *et al.,* 2021). These cells are described in lymphoid and nonlymphoid tissues. In sharp contrast with cDC1, cDC2 do not cross-present antigens in an efficient way but rather induce T helper 17 and T helper 2 responses (Soto *et al.,* 2020).

Several studies indicate the central role played by TGF-β in regulating cDC function. Indeed, upon encountering apoptotic T cells, cDC can release TGF-β and induce Treg (Perruche *et al.,* 2008). Also, mounting evidence suggests that regulation of DC is context-dependent. Indeed, *in vivo* studies indicated that DC can induce a tolerogenic phenotype by releasing TGF-β upon exposure of apoptotic DC cells – a process that was repressed in LPS-activated apoptotic DC (Notley *et al.,* 2015). Also, infected apoptotic cells induced a pro-inflammatory T helper 17-mediated response (Torchinsky *et al.,* 2009). Importantly, TGF-β signalling orchestrates in an indirect way cDC biology by preventing T-cell mediated autoimmune reactions (Kelly *et al.,* 2017). Indeed, functional loss of type 2 TGF-β receptor in murine CD11c+ cells correlated with multiorgan inflammation and death (Ramalingam *et al.,* 2012). Also, Ramalingam *et al*. discovered that this phenotype is linked with an increased proportion of migratory DC marked positive for CCR7 and directed to lymphoid tissues. Of note, Ramalingam and colleagues also determined that this was consistent with a pro-inflammatory

phenotype seen in cDC – which expressed a higher proportion of pro-inflammatory cytokines (Ramalingam *et al.,* 2012). It is worth mentioning that epithelial cell-derived TGF-β is proposed to be important in inducing a tolerogenic phenotype in cDC with greater potential to regulate Treg (Kelly *et al.,* 2017). Evidence suggests that CD103 might localise cDC to intestinal epithelial cells and that the interaction with the CD103 ligand E-cadherin might induce the tolerogenic phenotype upon release of intestinal epithelial cells-derived mediators, such as TGF-β and RA (Iliev, Mileti, *et al.,* 2009; Iliev, Spadoni, *et al.,* 2009). Therefore, intestinal epithelial cell-conditioned tolerogenic cDC were seen to promote Treg differentiation and, upon adoptive transfer, were able to protect mice from colitis (Iliev, Mileti, *et al.,* 2009). It is worth noting that E-cadherin is expressed on pro-inflammatory DC with an increased expression on cDC in TGF-β1-/- mice. A hypothesis suggests that TGF-β may orchestrate intestinal homeostasis by repressing expression of E-cadherin to allow cDC to directly bind intestinal epithelia (Kelly *et al.,* 2017).

1.5 The role of the integrin αvβ8-mediated TGF-β activation in immunity

As previously mentioned, αvβ8 is a major activator of TGF-β – with this regulation found to be crucial in immunity. One evidence of this derived from experimental results obtained by Lacy-Hulbert *et al*. when studying functional importance of Itgav – encoding the αv integrin subunit to which the β8 chain exclusively binds. Interestingly, mutant mice – lacking immune cell αv subunit – showed a reduced proportion of Treg leading to colitis, autoimmunity and colorectal adenocarcinomas (Lacy-Hulbert *et al.,* 2007).

Furthermore, Travis *et al*. identified that αvβ8 is crucial in regulating immune responses. Indeed, mice lacking the expression of αvβ8 specifically on all leukocytes developed a severe form of bowel inflammation and autoimmune response (Travis *et al.,* 2007). The group found that such inflammatory phenotype was mirrored in mice lacking the expression of αvβ8 on DC

(Travis *et al.,* 2007). Also, *in vitro* studies suggested that DC lacking β8 expression were less able to induce Treg, and this was due to reduced ability of the integrin αvβ8 KO DC to activate latent TGF-β (Worthington *et al.,* 2011). In humans, DC also express integrin αvβ8 and enhances their ability to activate TGF-β. Indeed, Fenton *et al*. identified that high amount of integrin αvβ8 are expressed on human DC (Fenton *et al.,* 2017). Inflammatory bowel disease patients showed a much higher proportion of αvβ8 compared to the control groups (Fenton *et al.,* 2017). This suggested that pro-inflammatory stimuli may induce the expression of this integrin on DC (Fenton *et al.,* 2017). Indeed, work identified that TLR4 engagement upregulated integrin αvβ8 expression and the ability of human DC to activate TGFβ (Fenton *et al.,* 2017). These findings indicate that αvβ8 can be regulated by environmental stimuli such as an inflammatory condition.

Importantly, our group also determined that Treg express high levels of αvβ8. Higher integrin expression resulted in a superior ability to activate TGF-β (Worthington *et al.,* 2015). It is worth mentioning that αvβ8 was not required by Treg to maintain cell homeostasis because deletion of αvβ8 on Foxp3+ effector cells did not induce inflammation (Worthington *et al.,* 2015). Yet, an ongoing inflammatory process could not be reverted by Treg with functional deletion of αvβ8 (Worthington *et al.,* 2015). These results indicate that αvβ8 is key to suppress effector T cell function but not to maintain T cell homeostasis (Worthington *et al.,* 2015). Interestingly, Treg express a transmembrane protein named glycoprotein A repetitions predomain that has been showed to promote TGF-β activation in concerted action with αvβ8 (McCarty, 2020).

Also, αvβ8 has a role during parasite infections. Furthermore, our group found that loss of αvβ8 on DC protected mice from *Trichuris muris* infection – a process that did not require Treg. On the contrary, conditional loss of αvβ8 on DC resulted in an increased production of protective T helper 2-derived cytokines (Worthington *et al.,* 2013). Also, Steel *et al*. determined that *Trichinella spiralis* infection triggered a T helper 17-cell mediated response in the small

intestine – which increased intestinal contractility. Conditional loss of αvβ8 on DC reduced host ability to expel worms (Steel *et al.,* 2019). These results indicated that αvβ8 is required in regulating worm expulsion, activity that is promoted by a T helper 17 environment which increase the smooth muscle contractility, necessary for parasite expulsion (Steel *et al.,* 2019).

Altogether, these results elucidate the crucial role played by integrin αvβ8 in regulating several biological processes in the context of immune responses. Furthermore, given its role in promoting the activation of TGF-β, this cytokine is a suitable therapeutic target. Specifically, the RGD binding domain sequence could be used to design function-blocking antibodies or specific inhibitors to treat ailments, such as fibrosis (McCarty, 2020). On the contrary, blocking αvβ8 could recapitulate autoimmune conditions seen in mice with functional loss of αvβ8 (McCarty, 2020). Importantly, identification of other αvβ8 downstream therapeutic targets might be important for regulating immune responses. Therefore, leading-edge technologies in the field of proteomics and transcriptomics could help determining identification of novel target molecules.

1.6 Transcriptome analysis

To study the effects elicited downstream of β8 upon ligand engagement, one solution would be to investigate transcriptional changes induced by LAP in DC. Likewise, gene expression profiling approaches may also be an appealing tool to describe differences at transcriptional level between different immune cells expressing β8 or not, as discussed below. Transcriptome analysis studies the whole set of transcripts expressed under specific conditions by the cell. Alongside the cost reduction needed to run this analysis, this technique has become increasingly used in different fields such as immunology. Success of these approaches over other methods such as microarray resides in an in-depth and more accurate determination of RNA concentration of samples analysed – harnessing high-throughput sequencing and transcriptome quantification (Z. Wang *et al.,* 2009). RNA-seq is mainly used to perform

differential expression analysis (DEA) – in which gene expression changes are measured in two or more group conditions to determine the magnitude of a change in gene expression (Costa-Silva *et al.,* 2017).

Upon RNA extraction from cultured cells or tissues, a library of cDNA is generated. Adaptors are attached at both ends of each cDNA fragment and each molecule is sequenced. Depending on whether high-throughput sequencing is performed at one or both ends of the cDNA fragments, a single- or paired-end sequencing is done respectively (Corley *et al.,* 2017). Single-end sequencing is cost-effective but less accurate than paired-end one. Indeed, with the latter method, a distance between the reads pair is determined and this would ensure a finer alignment against the reference in the downstream steps of the pipeline (Corley *et al.,* 2017).

1.6.1 Quality check of sequencing reads

Illumina sequenced reads are saved as FASTQ files (Cock *et al.,* 2009), a plain text format containing – for each sequenced nucleotide – a sequence quality score. The first step is, in fact, to check raw reads for their quality by looking at parameters such as the Phred quality score in Illumina machines – i.e., a value referring to the base call accuracy. Should the general quality of FASTQ files be scarce, the sequencing files are subjected to a trimming step aimed at removing low quality inputs. Importantly, towards the 3' end of reads, read quality decreases and base pairs should be removed to increase reads alignment (Conesa *et al.,* 2016). Subsequently, trimmed reads are mapped to a known reference (Stark *et al.,* 2019). Percentage of mapping is an important quality parameter – being this an indicator of DNA contamination and other quality parameters such as sequencing accuracy (Conesa *et al.,* 2016).

1.6.2 Pseudo-alignment with the Kallisto tool

Estimation of gene expression is the most common function of RNA-seq (Conesa *et al.,* 2016). This occurs through a process called alignment. Nevertheless, this process generates estimates of how many times a read enriches a specific region of the given reference but is relatively time-consuming and computationally expensive (Bray *et al.,* 2016) due to several technical steps needed to account for introns and exons in the reference. Therefore, samples can alternatively be quantified – circumventing the alignment step – by deconstructing sample reads into short nucleotide fragments called κ-mers (Bray *et al.,* 2016). Indeed, technological advances have significantly improved the computational time and accuracy needed to perform the alignment process. Alignment-free tools associate reads contained in the FASTQ files with transcripts. For example, the Kallisto tool computes transcript counts with high level of accuracy by significantly reducing the whole computational time needed to process sequenced samples (Bray *et al.,* 2016). As discussed by Bray *et al.,* the pseudo-alignment is a novel method that identifies where a read is originated from rather than accounting for their exact mapping position on the reference – e.g,, the transcriptome. Importantly, to perform the pseudo-alignment process, the latter method requires an index – created by processing a species-specific transcriptome file (Bray *et al.,* 2016). Specifically, κ-mers originates from previously shredded transcripts to construct a reference graph, a path. Therefore, each sample is also decomposed into κ-mers to generate a compatibility matrix between the reads and the reference (Bray *et al.,* 2016). To speed-up the approach, the compatibility matrix is converted into equivalence classes, indicating how many sequenced reads are contained in each class. This step ensures a smoother calculation aimed at maximising the likelihood function (Bray *et al.,* 2016). Counts estimate are therefore generated.

1.6.3 DEA with DESeq2

The output of these alignment-free tools such as Kallisto are converted into count estimates that are compared among experimental samples in a process called differential gene expression, modelled with tools such as DESeq2 (Love *et al.,* 2014). These tools estimate

which genes significantly changed and their level of expression in a given experimental comparison (Stark *et al.,* 2019). Briefly, DESeq2 normalises the raw reads counts, estimates the variability between replicates, and calculates gene changes between conditions (Love *et al.,* 2014). The DESeq2 package uses the negative binomial distribution to model RNA-seq count data distribution – aimed at testing the null hypothesis that gene changes between two separate treatments or conditions is equal to zero (Love *et al.,* 2014).

1.6.4 Outlook

RNA-seq conducted from bulk cells or tissues helped to clarify important questions in several field of biology – becoming one of the gold standard in transcriptome analysis (Conesa *et al.,* 2016). Nowadays, interest is mainly shifting in understanding transcriptional profile at single-cell level by harnessing a method referred to as single-cell RNA-seq (Stark *et al.,* 2019). As reviewed by Conesa *et al.,* the major advantage of that is aimed at the identification of uncharacterised tissue-specific cell subsets that may play a role in a certain experimental setting. Importantly, this approach made it possible to explain how two identical cells, in different contexts, may have a different transcriptional profile (Conesa *et al.,* 2016).

RNA-seq technologies are developing fast given their role in revolutionising how the transcriptome is quantified (Z. Wang *et al.,* 2009). For this reason, we decided to harness this technique for our project.

1.7 Thesis aims and hypotheses

Integrin $\alpha v\beta 8$ plays a crucial role in regulating cell homeostasis in immunity with DC expressing high levels of $\alpha v\beta 8$ (Travis *et al.,* 2007). As mentioned previously, $\alpha v\beta 8$ regulates the activation of TGF-β and expression of this integrin on DC contributes to maintaining a tolerogenic phenotype through the induction of Treg (Travis *et al.,* 2007). Despite the role played by $\alpha v\beta 8$ in promoting TGF-β activation, whether a signalling cascade is propagated

downstream of the integrin αvβ8 upon ligand engagement has not been described yet. We, therefore, hypothesised that αvβ8 stimulation may induce immunologically important transcriptional changes in DC. Additionally, unpublished data from the lab has shown that integrin αvβ8 is expressed by a subset of CD4+ Tem, but the importance of this expression to cellular phenotype is not understood. Thus, we speculated that αvβ8 expression could provide transcriptionally distinct properties to the latter cell subset. Therefore, in this thesis, the aims are to address the following questions:

What signalling cascades are evoked upon ligand engagement with αvβ8 in DC?

This aim is addressed in chapters 3 and 4. Firstly, we performed an RNA-seq analysis on a previously generated dataset containing DC expressing or not β8, using different bioinformatic approaches to interrogate the effects of ligand engagement with LAP in regulating the DC transcriptome. Secondly, a new RNA-seq experiment and analysis was performed, using different time points post-ligand engagement by DC.

Are β8+ Tem transcriptionally different from β8- Tem?

Previous work in the lab identified high expression of β8 on Tem. Results in chapter 5 contain RNA-seq experiments and pathway analyses of CD4+ Tem and Treg expressing or not β8 to test whether the immunosuppressive role seen *in vitro* in β8+ Tem could entail a more similar transcriptional profile to Treg.

Chapter 2

Material and Methods

2.1 Animals

Female and male mice were used between 6 and 14 weeks old. C57BL/6 (Charles River, UK), CD11c Cre+Itgb8fl/fl, and CD11c Cre-Itgb8fl/fl mice (Travis *et al.,* 2007) were kept in specific pathogen-free conditions, in line with the provisions of the Institution and UK Home Office guidelines in the Biological Services Facility at The University of Manchester. Schedule 1 procedures were conducted in accordance with the Home Office Scientific Procedures Act (1986) under licence and/or supervision.

2.2 Murine cell isolation

2.2.1 CD11c+ enrichment of splenocytes via Magnetic-Activated Cell Sorting (MACS)

Spleens were removed from Schedule 1-killed mice. At both ends, spleens were injected with a Collagenase D solution – containing 1mg/mL Collagenase D (Roche) and 50U/mL DNase I (Roche), in serum free RPMI (Sigma Aldrich) – and incubated at 37ºC for 30 minutes. Spleens were then chopped up and passed through a 40µm filter in order to get rid of cell clumps. Cells were centrifuged at 300 x g for 10 minutes and resuspended in culture medium (RPMI 1640 with 10% FCS, 10mM HEPES, 50µM 2-Mercaptoethanol, 1% penicillin/streptavidin, 2mM L-Glutamine, 1% non-essential amino acids (Sigma Aldrich)). Erythrocytes contained in the cell suspension were lysed with red blood cell lysis buffer (Sigma Aldrich) for 1 minute at room temperature. To quench lysis buffer, 8 mL of culture medium was added. Cells were counted using a haemocytometer and cells resuspended in culture medium at $2x10^8$cells/mL. To reduce non-specific binding, cell suspensions were incubated with Fc-blocking antibody (eBioscience) at a concentration of 5µg/mL and kept on ice for 5 minutes. CD11c+ splenocytes were positively selected using CD11c MicroBeads UltraPure (Miltenyi Biotec) as per manufacturer's instructions. In brief, 30µL of CD11c microbeads with 5µg/mL of DNAse I were added to splenocyte suspensions for every $10x10^9$ cells and kept on ice for 30 minutes. Cell were washed, centrifuged at 300 x g for 10 minutes, and resuspended in 500µL of MACS

buffer – containing phosphate-buffered saline (PBS) (Sigma Aldrich), 2% foetal calf serum and 1mM EDTA – per $1x10^8$ total cells. Splenocytes suspensions were applied onto a previously calibrated MACS LS column (Miltenyi Biotec) to separate CD11c+ cells via magnetic separation, as per protocol.

2.2.2 Fluorescence-activated cell sorting (FACS) separation of previously enriched CD11c+ splenocytes

Enriched CD11c+ splenocytes were stained in PBS with Zombie Aqua viability dye (BioLegend) and incubated for 30 minutes at room temperature, in the dark, at a dilution of 1:500 for $5x10^6$ cells. FACS buffer (PBS, 2%FCS, 2mM EDTA) was used to wash CD11c+ cells. In an alternative panel, DAPI was used as cell death marker. Cells were centrifuged and resuspended in 400µL FACS buffer at 5µg/mL of Fc-blocking antibody (eBioscience) and kept for 20 minutes, in the dark, at 4ºC. PE/Cy7- or APC-conjugated anti-MHCII antibody (Clone MS/114.15.2, BioLegend) and APC- or FITC-conjugated anti-CD11c antibody (Clone N418, BioLegend) was added at 1.5µg/mL, in the dark, for 30 minutes at 4ºC. Stained cells were washed via centrifugation in 1mL FACS buffer. Cells were resuspended in 200µL Sort Buffer (Ca2+- and Mg2+-free PBS, 25mM HEPES and 2.5mM EDTA) and finally filtered through 50µm sieve. Live CD11c+ MHCII+ cells were sorted using a BD FACS Aria Fusion or BD Influx cell sorter. Data analysis was performed using the FlowJo software (Treestar).

2.2.3 CD11c+ MHCII+ splenocytes incubation with LAP

MACS-enriched and/or FACS sorted CD11c+ MHCII+ splenocytes from WT mice were resuspended at a concentration of $1x10^6$ cells/mL in pre-warmed culture medium and distributed in either 48- or 96-well plates. Cells were incubated with bovine serum albumin (BSA, Sigma Aldrich) or LAP (R&D System) at 5µg/mL for 2 or 6 hours at 37ºC.

2.3 RNA

2.3.1 RNA extraction

Following incubation with LAP or BSA, CD11c+ MHCII+ cells were centrifuged at 300 x g for 10 minutes. Supernatant was removed and the cell pellet resuspended with 350µL of RLT buffer (Qiagen), adding 10µL of 14.3M 2-mercaptoethanol per 1mL RLT buffer. Cell lysates were stored at -80ºC for later use. RNA extraction was performed using the RNeasy Micro Kit (Qiagen) as per manufacturer's instructions. In brief, 70% ethanol equal to the volume of the cell lysates was added, samples added to spin columns and centrifuged for 1 minute at >10,000g in a microcentrifuge. Columns were washed twice with 350µL RW1 buffer. Columns were then washed with 500µL RPE buffer, then with 500µL of 80% ethanol. A drying step was then performed by centrifuging the column with no liquid, before and elution of the column by addition of 21µL RNase-free water followed by centrifugation. RNA purity was either measured with a NanoDrop 2000c (Thermo Fisher Scientific) or Agilent 4200 TapeStation system to measure RNA integrity.

2.3.2 Reverse transcription of RNA into cDNA

RNA was reverse transcribed to cDNA by using a SuperScript III reverse transcriptase kit (Invitrogen). In brief, a master mix containing 500nM dNTPs (Invitrogen), 500ng oligo(dT)$_{12\text{-}18}$ primers, the whole amount of extracted RNA, and UltraPure DNase/RNase-Free Distilled Water (Thermo Fisher Scientific) was prepared. Using the 2720 thermal cycler (Applied Biosystem), Oligo(dT) primers were annealed to RNA for 5 minutes at 65ºC, then samples kept on ice for at least 1 minute. Next, a mix of reaction buffer, 10mM DTT, 40units/µL RNaseOUT and 200units/µL of SuperScript III reverse transcriptase was prepared and added to samples. Reverse transcription was performed by incubating samples at 50ºC for 60 minutes, then 70ºC for 15 minutes, then reaction terminated by reducing the temperature to 4ºC. cDNA was measured on the nanodrop, used soon thereafter or stored at -20ºC.

2.3.3 Reverse transcription quantitative polymerase chain reaction (RT-qPCR)

Expression of candidate genes (see Table 2.1 for mouse primer list) was measured using a StepOnePlus Real-Time PCR System (Thermo Fisher Scientific) or QuantStudio 12K Flex System (Applied Biosystems). RT-qPCR was performed with either Fast SYBR Green Master Mix (Applied Biosystems) or the PowerUp SYBR Green Master Mix (Applied Biosystems). The reaction mix contained 1µl of template, UltraPure DNase/RNase-Free Distilled Water (Thermo Fisher Scientific), and 0.5µM of gene-specific forward and reverse primers. Samples were subjected to the following cycles: 95ºC for an initial 2 minutes, then 95ºC for 5 seconds and 60ºC for 30 seconds, repeated 40 times. The expression of the hypoxanthine phosphoribosyltransferase (HPRT) housekeeping gene was used to normalise gene expression levels using the $2^{-\Delta\Delta}$Ct Method (Livak & Schmittgen, 2001).

Primer	Forward (5'-3')	Reverse (5'-3')
HPRT	TCAGTCAACGGGGGACATAAA	GGGGCTGTACTGCTTAACCAG
CXCR4	GACTGGCATAGTCGGCAATG	AGAAGGGGAGTGTGATGACAAA
SMAD3	CACGCAGAACGTGAACACC	GGCAGTAGATAACGTGAGGGA
RHO-GDI2	GGTCACTCGTTACATGGACTTC	CCATTAAGCTAGATGCCAGGG
SKIL	AGGCAGAGACAAGTAAGTCCA	CGTCTGGGTAAGACACTGTTTTT
CXCL16	CCTTGTCTCTTGCGTTCTTCC	TCCAAAGTACCCTGCGGTATC

Table 2.1 | Forward and reverse primers for candidate murine genes selected for RT-qPCR

2.4 Quality check of fastq files

2.4.1 Quality control of sequenced samples

Paired- or singled-end sequenced samples, in FASTQ file format, were checked for quality of their reads with the FastQC software version 0.11.8 (Andrews *et al.,* 2015). If the trimming

step was required, then FastQC reports were generated again and files from all samples were collated by using the MultiQC software (Ewels *et al.,* 2016).

2.4.2 Adapter trimming and filtering of low-quality reads

Reads quality trimming was performed to remove adapter contamination and filter out fragments of reads with poor quality score. In brief, the latter parameter is calculated as the negative 10 times base-10 logarithm of the probability of a base calling error – i.e., $-10 \log_{10} P$ – and defined as Phred score.

2.5 Pseudoalignment with the Kallisto software

The Kallisto software (Bray *et al.,* 2016) was used to pseudoalign single- or paired-end sequenced fastq files, with default parameters, against the reference mouse transcriptome (GRCm38 or GRCm39) – ensuring relatively fast and accurate results to quantify transcript abundances (Figure 2.1).

$$L(\alpha) \propto \prod_{f \in F} \sum_{t \in T} \mathbf{y}_{f,t} \frac{\alpha_t}{l_t} = \prod_{e \in E} \left(\sum_{t \in e} \frac{\alpha_t}{l_t} \right)^{c_e}$$

Figure 2.1 | Quantification step used by Kallisto. L is the likelihood function. F is and T are the set of fragments and transcripts, respectively. y$_{f,t}$ is a matrix – containing information about fragment and transcript compatibility. αt are parameters of interest – being abundances. Equivalence classes – containing reads matching the same sets of transcripts – are generated factorizing the compatibility matrix. The Expectation-Maximization (EM) algorithm is applied to increase likelihood function and stop when, for each subsequent iteration, for every transcript $t \in$ T, αtN > 0.01 – being N the whole amount of fragments – is measured to change < 1%.

2.5.1 Index building

Firstly, the Kallisto software creates an index to allow estimation of count abundances through the formation of a De Bruijin graph, where each node is a κ-mer – i.e., a nucleotide sequence of length κ. Mapping information about κ-mers and original transcripts are stored.

2.5.2 κ-compatibility classes, equivalence classes, and optimisation through the EM algorithm

The index is used to analyse sequenced sample reads by estimating κ-compatibility classes – from which sets of transcripts of origin are determined. Information stored in the κ-compatibility classes were summarised into equivalence classes to make the analysis smoother. The EM algorithm is finally used to optimise the RNA-seq likelihood function – calculated as the highest likelihood from which a read may have originated from (Figure 2.1). The counts table was finally used for the downstream statistical analysis.

2.6 Gene-level analysis

2.6.1 Summarisation of transcript-level abundances to gene-level matrices

Transcript counts generated with the Kallisto software were aggregated to gene-level counts with the Tximport package (release 1.16.1) – converted into a matrix that can be imported for the downstream analysis.

2.6.2 DEA

DEA, statistical analysis on normalised gene counts, was performed with the DESeq2 package (version 1.28.1). Principal component analysis (PCA) plots generated with the DESeq2 package were visualised with the ggplot2 package (version 3.3.2). Supervised hierarchical clustering and volcano plots were generated with the Pheatmap package (version 1.0.12) and the Enhanced Volcano package (version 1.6.0), respectively.

2.7 Batch effect correction

The surrogate variable analysis (SVA) (version 3.36.0) was used to control and remove unwanted source of bias affecting chosen RNA-seq datasets. PCA plots were generated to measure the level of clustering in RNA-seq experiments.

2.8 Pathway and network analysis

2.8.1 Ingenuity Pathway Analysis (IPA), functional annotation, and the Reduce and Visualise Gene Ontology (REViGO) tool

Lists of outputs generated with the DESeq2 package were used as input files for the IPA tool (QIAGEN, http://www.ingenuity.com). Specifically, canonical pathways were calculated using default settings and represented as the negative logarithm of the P value. Also, the upstream analysis was used to determine common regulators that were significantly enriched in the list of differentially expressed genes.

The DAVID tool (D. W. Huang *et al.,* 2009) and the Kyoto Encyclopaedia of Genes and Genomes (KEGG) database (Kanehisa & Goto, 2000) were used to assess the function of differentially expressed genes. Specifically, the functional annotation method shows each cluster in which each gene falls into, ranked by an enrichment score and indicating proportion of significantly expressed genes in each category. Gene ontology (GO) terms were also identified using the DAVID tool and entries from biological processes and molecular function were further analysed using the REViGO tool to reduce pathways redundancy.

2.8.2 Graph-based iterative group (GiGA) analysis

An R-based version of the GiGA analysis, extended to RNA-seq experiments, was used to determine putative functional interactions between differentially expressed genes (Breitling *et al.,* 2004). In brief, the GiGA algorithm identifies significantly regulated subset of genes in an experimental setting – creating ‘evidence graphs’ (Breitling *et al.,* 2004). As discussed in Breitling *et al.,* in this context, evidence is a common functional annotation or could be referred

to a shared physiological process associated with annotated genes. A list of differentially expressed genes is also provided and genes are ranked based on their expression values (Breitling *et al.,* 2004). Network nodes with a lower rank than all their connections in the graph are then identified and process finishes until all nodes are covered (Breitling *et al.,* 2004).

2.9 Functional validation

2.9.1 Gene expression analysis in LAP-treated DC

To candidate genes that were seen to be differentially expressed by RNA-seq, MACS- and FACS-sorted CD11c+ MHCII+ splenocytes were incubated with either 5µg/mL BSA, 5µg/mL LAP, 10ng/mL active TGF-β (Peprotech), 100µg/mL anti-TGF-β blocking antibody (clone 1D11, BioXcell) – hereinafter referred as anti-TGFβ antibody – or 5µg/mL LAP + 10ng/mL free active TGF-β. Gene expression was tested via RT-qPCR.

2.9.2 Migration assay

CD11c+ splenocytes were incubated at a concentration of $1x10^6$ cells/mL with 5µg/mL LAP or BSA for 4 hours in the presence of 500ng/mL LPS at 37ºC. Cells were loaded onto the upper chamber of a transwell apparatus (Corning) and exposed to the presence – in the lower chamber of the apparatus – of 200ng/mL CXCL12 (Peprotech) or serum-free RPMI as negative control, for 2 hours at 37ºC. Cells were then collected from the upper and lower chambers in separate vials and total events were counted via acquisition on an LSR-II flow cytometer (BD Biosciences). Percentage of migration was expressed as number of events measured in the lower chamber divided by the total events acquired in both the lower and upper chambers, multiplied by 100.

2.9.3 Microscopy analysis

FACS-sorted CD11c+ MHCII+ splenocytes were incubated for six hours with 5µg/mL LAP or BSA at 37ºC, in culture medium at a concentration of $1x10^6$ cells/mL. During the incubation time, cells were seeded on 8 well slides (Thermo Fisher Scientific) coated with poly-L-Lysin (Sigma Aldrich) and fibronectin (Sigma Aldrich). To fix the cells, 1 volume of 4% paraformaldehyde (Thermo Fisher Scientific) was added to the media for 20 minutes. Cell media was aspirated and cells washed three times with PBS. Cells were then incubated for 5 minutes with 0.01% Triton x-100 (Sigma Aldrich) in PBS. Cells were blocked overnight at 4ºC with 3% BSA in PBS. The day after, cells were stained with phalloidin conjugated with Alexa Fluor 555 (Thermo Fisher Scientific) at 3 U/ml in 3% BSA for 30 minutes at room temperature, and cells imaged with a confocal microscope SP8 (Leica) using a 63x/NA 1.40 objective lens. Cell area and perimeter was measured using the Fiji software (Schindelin *et al.,* 2012).

2.10 Statistical analysis

Unless otherwise stated, all statistical analysis was performed using the GraphPad Prism 9 software (GraphPad Software, Inc). Data are expressed as the mean ± standard deviation. Data sets were tested for normality using the Shapiro-Wilk Test. In case of normally distributed data, a parametric test was performed. Non-parametric test was used for non-normally distributed data. For RNA-seq analysis, the Wald test was performed, followed by the Benjamini-Hochberg correction.

Chapter 3

Determining transcriptional changes induced by the integrin αvβ8 in DC

3.1 Introduction

Integrins are transmembrane glycoproteins, composed of two non-covalently linked subunits (α and β) spanning the cell membrane. Apart from their role as cell-matrix adhesion molecules, integrins are also signalling receptors. Therefore, they can relay intracellular signalling upon ligand binding (Kechagia *et al.,* 2019). This signalling cascade is propagated through interactions of the cytoplasmic domain of the integrin molecules with cytoplasmic signalling mediators. Yet, the primary sequence of the β8 molecule diverges from the other integrin β subunits (Moyle *et al.,* 1991). In particular, the β8 cytoplasmic domain lacks a highly conserved peptide motif (McCarty, 2020) that, in other β subunits, interacts with talin (Shen *et al.,* 2013) – the role of which is to connect integrins with the actin cytoskeleton. Therefore, given that this mechanical link is missing in β8, it is worth considering whether this subunit might propagate different signalling pathways from those relayed inside the cells by other integrins. Therefore, understanding whether and how signalling events are relayed inside the cell upon ligand-receptor engagement – in different contexts – may help elucidate which effector proteins are involved in the integrin αvβ8 signalling pathway.

Integrin αvβ8 plays a pivotal role in regulation of the immune system, via activation of the latent form of TGF-β (Mu *et al.,* 2002). Successful engagement of the integrin αvβ8 by LAP releases the active moiety of TGF-β that is subsequently able to elicit its biological function (Stephen L. Nishimura, 2009). Furthermore, studies on β8 conditional KO mice, lacking the expression of the integrin αvβ8 on DC, showed that – in this cell subset – the integrin αvβ8 is an activator of the TGF-β cytokine and an important regulator of intestinal homeostasis (Travis *et al.,* 2007). These findings suggest that the αvβ8-mediated TGF-β activation is required by DC for orchestrating important functions in the context of immunity. Nevertheless, if and how the integrin αvβ8 plays any role in intracellular signalling in DC, to date, is not understood.

In this chapter, we aimed to identify gene expression changes induced by the integrin αvβ8 engagement of LAP in cells marked positive for CD11c – which classically defines DC population – using RNA-seq and pathway analysis approaches.

3.2 Developing methods to purify wildtype and αvβ8 KO DC for treatment with LAP

The primary purpose of this part of my thesis was to optimise an efficient workflow to analyse bioinformatics datasets with subsequent identification of αvβ8-dependent signalling pathways. More specifically, the overarching goal of this project is to investigate whether ligand engagement by the integrin αvβ8 can drive changes in gene expression regulating DC function. In this chapter, we focus our attention on bioinformatic analysis of genes identified to be directly altered by αvβ8 upon its engagement with LAP on cells expressing the CD11c molecule, commonly used to define DC (Wu *et al.,* 2018) and isolated from murine mesenteric lymph nodes (mLNs). Initially, this was achieved by performing analysis of an RNA-seq experimental dataset that was previously generated in our lab. In order to rule out any non-β8-specific signals on CD11c+ cells, conditional KO mice lacking the expression of αvβ8 on CD11c+ cells were used as an experimental group (Figure 3.1A).

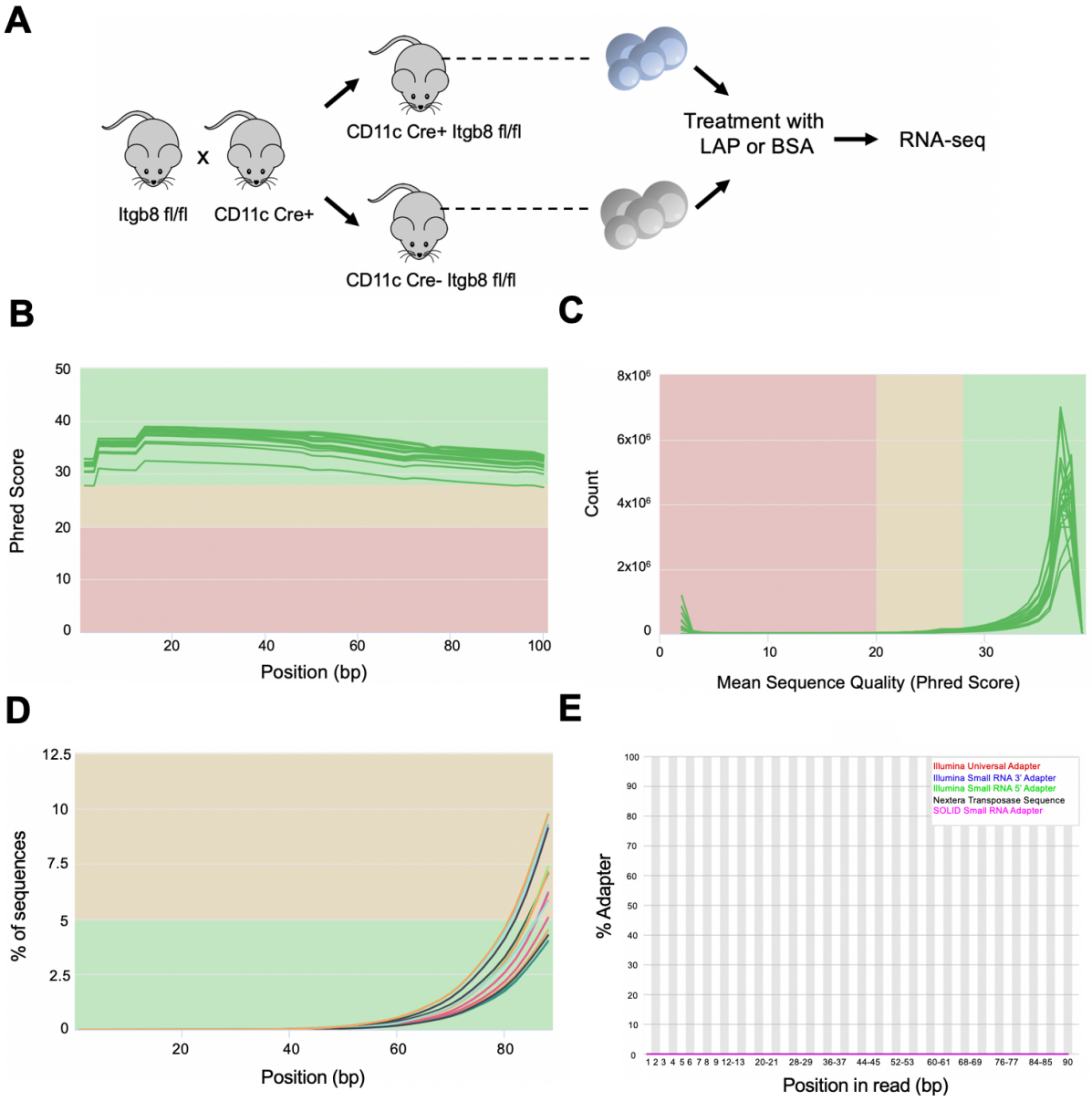


Figure 3.1 | Sequence quality assessment of raw FASTQ files obtained from high throughput sequencing protocols. A) Schematic representation of the CD11c Cre+ Itgb8 fl/fl or CD11c Cre- Itgb8 fl/fl used in the RNA-sequencing experiment – that is, conditional KO mice lacking the expression of the integrin αvβ8 on CD11c+ cells and their controls, respectively. CD11c+ Cre+ mice, in which the Cre recombinase enzyme expression was regulated by the CD11c promoter, were mated with Itgb8 fl/fl mice, in which the Itgb8 gene is floxed and thus substrate of the Cre recombinase. CD11c Cre+ Itgb8 fl/fl or CD11c Cre- Itgb8 fl/fl cells isolated from murine mLNs were sequenced, after a six-hour long incubation with 5µg/mL LAP or BSA (credit: Ellie Sherwood). In B and C) The FastQC tool was used to estimate the quality of 12 paired-end sequencing reads coming from the Illumina sequencer. Results were collected and analysed with the MultiQC tool. B) Quality check of sequenced reads performed by plotting the phred quality score, on the Y-axis, against each base of sequencing reads, on the X-axis. Phred quality score is inversely correlated with wrong estimation of base calling probability. C) Distribution of phred quality scores where the quality sequence average is on the X-axis and sequence counts, for each fastq file, are on the Y-axis. D) Proportion of adapter contamination, on the Y-axis, at each position, on the X-axis. Coloured lines are different samples. E) Representative plot showing fraction of adapter

contaminants after the trimming step. Percentage of adapter is on the Y-axis and base pair position on the X-axis. N=12.

After cell culture, RNA was extracted from LAP- or BSA-treated CD11c+ cells from mice lacking β8 expression on CD11c+ cells and wildtype ones – each of which was treated with LAP or BSA for six hours – was paired-end sequenced with the HiSeq 2500 Illumina machine (Figure 3.1A). Fastq files generated by the Illumina analyser retained information including nucleotide sequence of the starting material. Importantly, each nucleotide contained in the fastq file was associated with a phred quality score. The latter parameter is an estimate of a base calling error, given by the sequencing machine upon detection of a nucleotide-specific light signal. Thus, the probability of a wrong base being called is directly related to the strength of the signal at each cycle of the sequencing process. Importantly, given that technical issues – at each step of the pipeline – have a consistent impact in the further steps of the analysis, the starting quality of sequenced samples was assessed.

Quality check was performed on raw reads – contained in the fastq files – through a series of tests run by the FastQC tool (Figure 3.1). Next, these results were collated with the MultiQC analysis software. Of note, amongst different tests conducted, an in-depth scan of the nucleotide-specific phred quality score across all reads was performed for all samples and visualised with the MultiQC tool (Figure 3.1B). We found that distribution of reads was skewed towards the upper quality range of the graph (Figure 3.1C). These results suggest an overall good starting quality of fastq reads, despite minor adapter contamination affected sequenced samples (Figure 3.1D).

Therefore, in order to increase reliability in the further steps of the analysis, Illumina adapters were deleted from raw reads and low-quality reads filtered out. In this regard, a trimming step was performed with the Trimmomatic tool. In short, at the beginning and end of the sequenced reads, Trimmomatic cropped fragments in case phred quality score were less than 10 – which

corresponds to an estimated accuracy of base calling fewer than 90%. Moreover, the Trimmomatic tool also trimmed off low-quality bases whereby an average score, calculated within a four-base sliding window, fell below 15. Finally, in order to avoid ambiguous targeting to the reference, reads shorter than 25 bases were dropped. As a result, at the end of the trimming step, all the samples contained <0.1% adapter contaminants (Figure 3.1E). In this way, quality-filtered reads were ready for the next step of the pipeline.

3.3 Wildtype murine CD11c+ cells upon incubation with LAP show minor transcriptional changes compared to the β8 KO counterparts

Once the quality check and trimming were completed, a process called pseudoalignment was performed. Briefly, the pseudoalignment (or pseudomapping) step involves identifying sets of the reference transcriptome segments, from which a sequenced read – contained in the fastq files and therefore from each sample – may have originated. The Kallisto tool was shown to quickly and accurately perform the pseudoalignment step because it does not require information about the exact matching position of a specific read against its reference transcriptome, in contrast to other tools (Bray *et al.,* 2016).

Kallisto firstly processed the reference mouse transcriptome (GRCm38) through the construction of an index (Bray *et al.,* 2016). Next, the Kallisto tool could pseudoalign the previously trimmed paired-end fastq files (Figure 3.2). Of note, four samples showed a relatively low pseudoaligned rate ranging from 51.9 to 62%, likely linked to issues at any of the upstream steps of the pipeline such as contamination by ribosomal RNA during sample preparation.

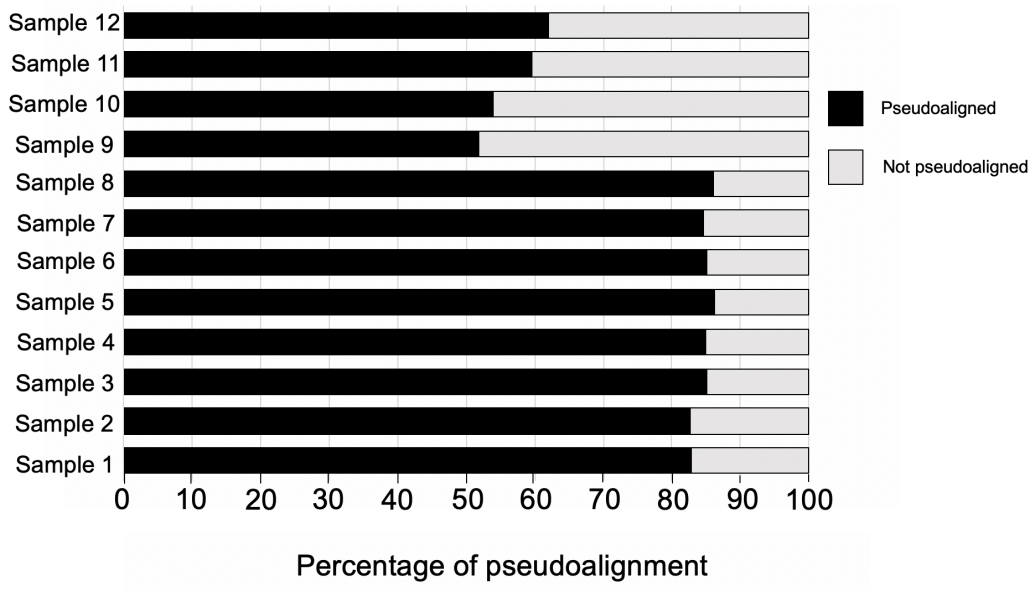


Figure 3.2 | Percentage of successfully pseudoaligned reads. The stacked bar chart depicts – on the X-axis – percentage of successfully pseudoaligned, paired-end trimmed fastq files against the mouse index transcriptome performed with the Kallisto tool. Samples are vertically listed on the Y-axis. Proportion of successfully pseudoaligned reads are coloured in black; coloured in grey are reads, for which the pseudoaligned was not successful.

After the pseudomapping step, transcript read counts estimated by the Kallisto tool were cumulated to the gene level for downstream analysis. This process was achieved through the aggregation of the pseudocounts generated for all the transcript variants to their cognate gene. Finally, data was ready for being statistically inferred in the further stage of the analysis.

Therefore, pseudoaligned counts were used for determining which genes significantly changed when wildtype CD11c+ cells were compared with β8 KO counterparts and/or incubated with LAP. The DESeq2 package was used for this purpose. In the first place, library-related technical variables (e.g., differences in depth of sequencing) were normalised in the DESeq2 package through the estimation of scaling factors for each gene in the dataset (Love *et al.,* 2014). Secondly, DESeq2 calculated gene-wise dispersion values returning estimates about gene variability across different biological replicates. Thirdly, the negative binomial distribution was used to model gene counts and hypothesis testing was performed to identify differentially expressed genes. Then, diagnostic plots were used to determine correlation of samples clustering.

PCA is a widely-used approach in bioinformatics aimed at summarising information gathered from RNA-seq samples. In brief, in a PCA, vectors generated account for a certain share of experimental variance and are hierarchically ranked depending on the amount of variation they account for. In general terms, the first vector – i.e., the principal component (PC) 1 – is the one that captures the highest variation of the experiment, followed by PC 2 in order of level of variance. Conventionally, PC 1 is plotted against PC 2 in a PCA plot, which is a graphical approach used to visualise similarities and differences across different samples. Noteworthy, samples that segregate together show a high degree of correlation, meaning that their transcriptional profile is similar.

Therefore, in order to determine how wildtype and β8 KO CD11c+ cells differed upon a six-hour-long incubation with LAP, PCA plots were generated in DESeq2 (Figure 3.3, N=3). In the first analysis, samples showed a poor within-group correlation due to unknown reasons (Figure 3.3A). To measure how different experimental groups were in our experiments, a common approach is to calculate sample distance from the gene count matrix and represent these results as a heatmap. Indeed, these results were confirmed by a sample-to-sample heatmap that identifies three major clusters that were not related to experimental groups (Figure 3.3B). Nevertheless, we would have expected a higher degree of correlation from, at least, groups of the same genetic background – should the effect of LAP treatment be minimal. Given that the source of bias was not definable, the SVA package, a tool used to remove unmodeled factors in the analysis, was performed to correct the variability affecting the whole dataset (Leek & Storey, 2007).

Therefore, after having applied the SVA correction, an increased level of clustering was measured across the two different genotype conditions – i.e., β8 KO CD11c+ cells and the wildtype group (Figure 3.3C). On the other hand, we found that LAP-dependent effects were minimal whereas the major transcriptional changes were due to Itgb8 gene suppression – which encodes the integrin β8 subunit. To validate that experimental variation was mainly due

to β8 gene suppression, we studied sample similarities in a sample-to-sample heatmap to see how samples clustered after the SVA correction. We found that two major clusters, representing the β8 KO and the wildtype groups, were identified – regardless of any big effect on the phenotype induced by LAP (Figure 3.3D).

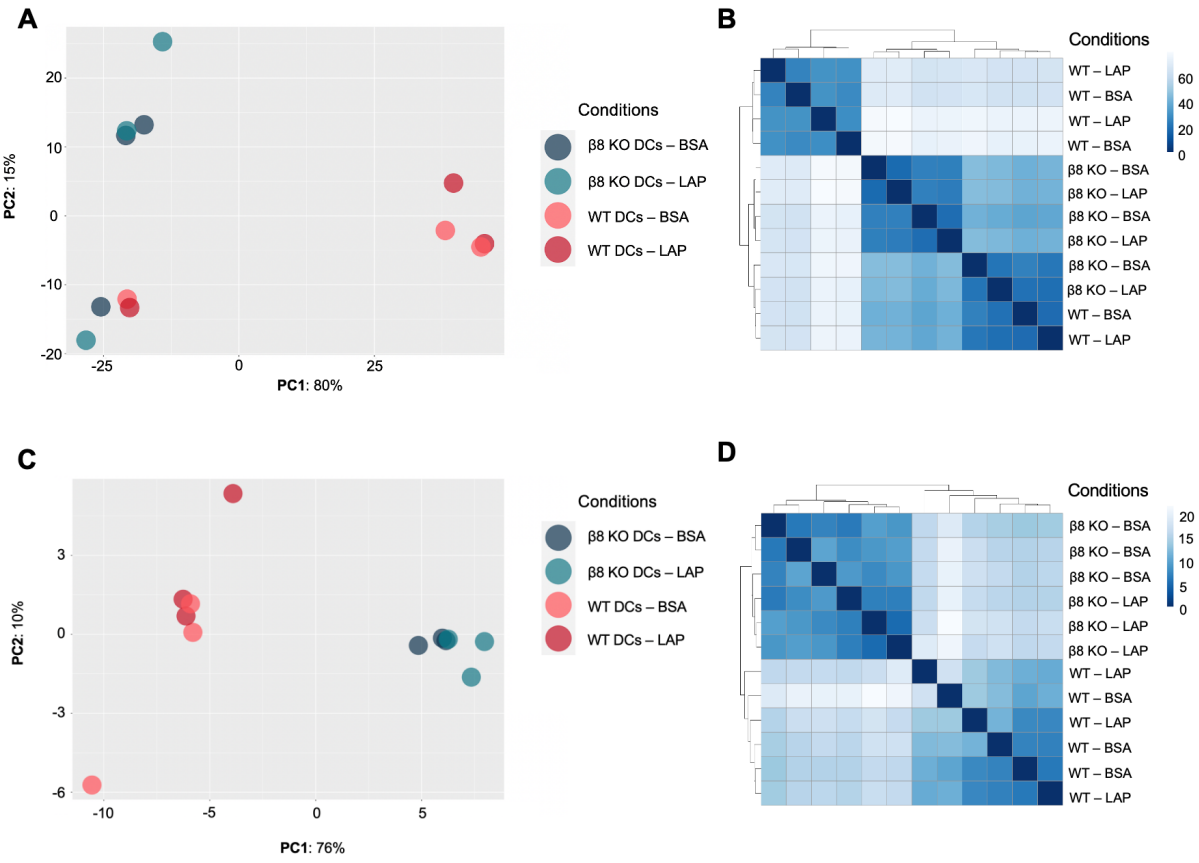


Figure 3.3 | PCA plot reveals that sequenced samples are affected by a source of bias further rescued by the SVA. A) PCA plot representing sample correlation when CD11c+ cells were treated with LAP or BSA, if the SVA tool is not applied. B) Sample-to-sample heatmap describing sample similarities in the uncorrected dataset. C) PCA plots showing clustering achieved after performing the SVA algorithm. D) Sample-to-sample similarities represented as a heatmap after the SVA approach was applied. A) and C) Blue and light-blue dots are β8 KO CD11c+ cells incubated with BSA or LAP, respectively; pink and red dots are wildtype CD11c+ cells incubated with BSA or LAP, respectively. PC 1 – on the X-axis – and PC 2 – on the Y-axis – explain the highest variation of the experiment. B) and D) Degree of similarity is high if colour is dark-blue. Each dot or tile is one biological replicate. N=3.

These results confirmed that the whole experiment was affected by an unidentifiable source of variability that could be controlled through the implementation of surrogate variables. Secondly, following LAP incubation, results suggested that αvβ8 might relay an intracellular

signalling cascade that evokes less transcriptome changes than transcriptional alterations induced by the absence of β8 gene expression in the absence of LAP treatment. Therefore, a robust approach was needed for controlling the unwanted technical noise and minimal transcriptional changes evoked by LAP in this experimental setting.

3.4 Outlier removal increases degree of clustering of wildtype and β8 KO CD11c+ cells confirming minimal effects triggered by LAP

In view of results in figure 3.3A and B, four biological samples (samples 1 to 4) – one per each biological replicate – that did not correlate with their cognate group were excluded from the analysis. These samples were deemed as outliers based on the prior assumption that replicates from the same genotype group should be transcriptionally similar. Therefore, two replicates per group were re-analysed.

After having scaled down the experiment by reducing the number of biological replicates from three to two, samples were then re-analysed and PCA was performed again on the remaining samples (Figure 3.4A). PCA plots showed that samples segregated based on the expression of β8 rather than their treatment with LAP. In fact, a LAP-driven effect on phenotype was not explained by variations measured on PC 1 and PC 2.

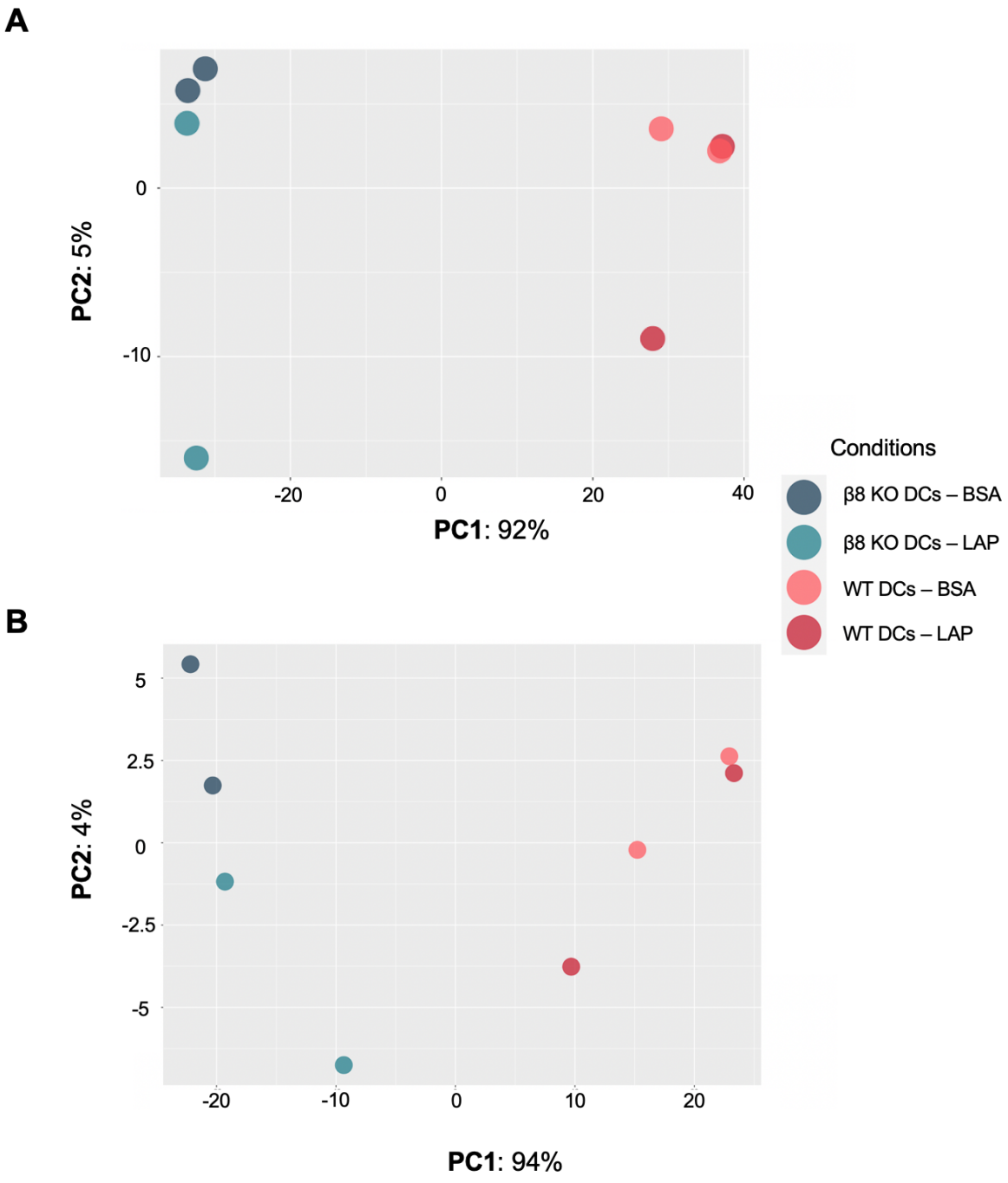


Figure 3.4 | Outliers removal increases clustering of sequenced samples. A) PCA plot performed on a dataset where CD11c+ cells were incubated with the LAP versus a control group, and after removal of four biological points deemed as outliers. The highest amount of experimental variance is explained by the PC 1 and PC 2 – on the X- and Y-axis, respectively. B) PCA plot of samples corrected with the surrogate variable analysis tool. β8 KO CD11c+ cells and treated with BSA are coloured in blue; coloured in light-blue are dots representing β8 KO CD11c+ cells treated with LAP; wildtype CD11c+ cells treated with BSA or LAP, are coloured in pink and red, respectively. N=2.

Next, we sought to determine whether the SVA approach could be extended to this dataset with only two replicates to see if the unknown source of variation was still present after outliers were removed (Figure 3.4B). After applying the SVA algorithm, differences between LAP- and BSA-treated β8 KO CD11c+ cells were distinguishable in two, albeit not completely discrete, clusters. On the other hand, transcriptome profiling of wildtype groups showed high degree of correlation suggesting that these samples, incubated with LAP or BSA, were transcriptionally similar even after SVA correction. These results confirmed that wildtype CD11c+ cells never showed a clearly distinct transcriptome profile upon LAP incubation either when samples were corrected and/or outliers were removed.

We conclude that the experimental data analysed indicate that there were minimal effects triggered by LAP on CD11c+ cells, with only small differences in transcriptomes observed in cells treated or not with LAP. In order to determine the impact of these differences across the four groups analysed, and visualise the impact on the transcriptome due to the absence of integrin αvβ8 expression and LAP treatment, supervised hierarchical clustering of the top 500 variable genes was performed on uncorrected samples (Figure 3.5, N=2). Euclidean distance – used to determine similarities across samples – confirmed that the transcriptome profile of wildtype and β8 KO CD11c+ cells were separated in only two major clusters rather than four, expected if LAP treatment would have induced an overt effect on the phenotype.

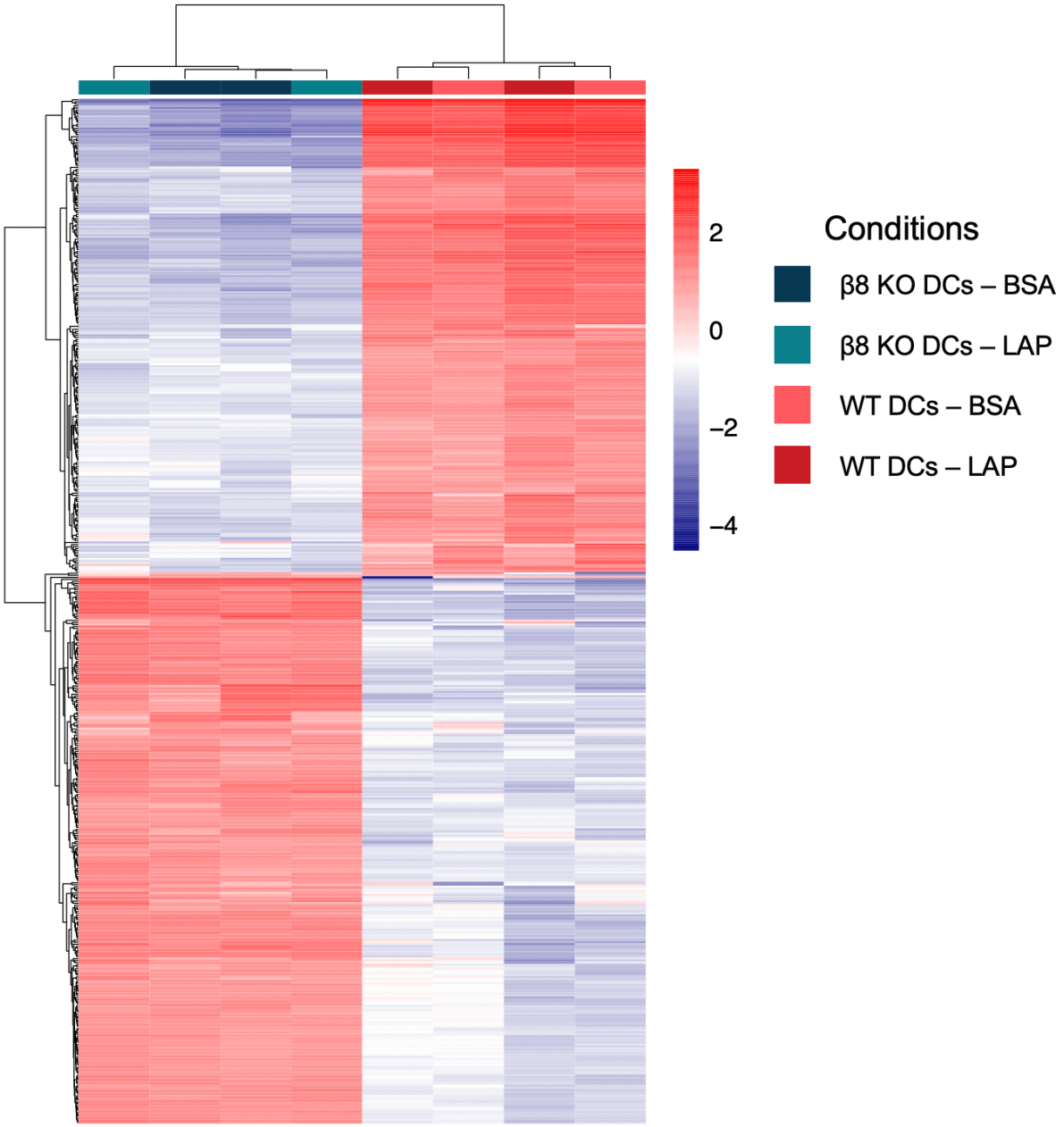


Figure 3.5 | Supervised hierarchical clustering of the top 500 variable genes compared across wildtype CD11c+ cells and their αvβ8 KO counterparts suggests minimal effect of LAP. Figure shows heatmap of the top most variable genes in the experiment. Columns and rows represent samples and genes, respectively. Regularised logarithmic transformation was used to stabilise sample variance. Values are centred on the mean. Euclidean distance was calculated for both row and column dendrograms. Up- and down-regulated genes are shown in red and blue, respectively. N=2.

Furthermore, these results were further confirmed when two MA plots were produced (Figure 3.6). Briefly, an MA plot is a graphical visualisation in which the base mean, that is the average of normalised counts generated with the DESeq2 package, is plotted against the log2 fold change values. In this way, gene dispersion reflects how much a certain gene varies as a function of its abundance – between selected comparisons. When comparison was made between wildtype CD11c+ cells treated with LAP versus BSA, subtle differences in gene expression were found, given the fact that, only few genes were significantly above the log2 fold change cut off of 1.5 (Figure 3.6A). This suggested that, upon incubation with LAP, the two groups analysed have a similar transcriptome profile. On the other hand, comparison between wildtype versus β8 KO CD11c+ cells resulted in major transcriptional differences (Figure 3.6B).

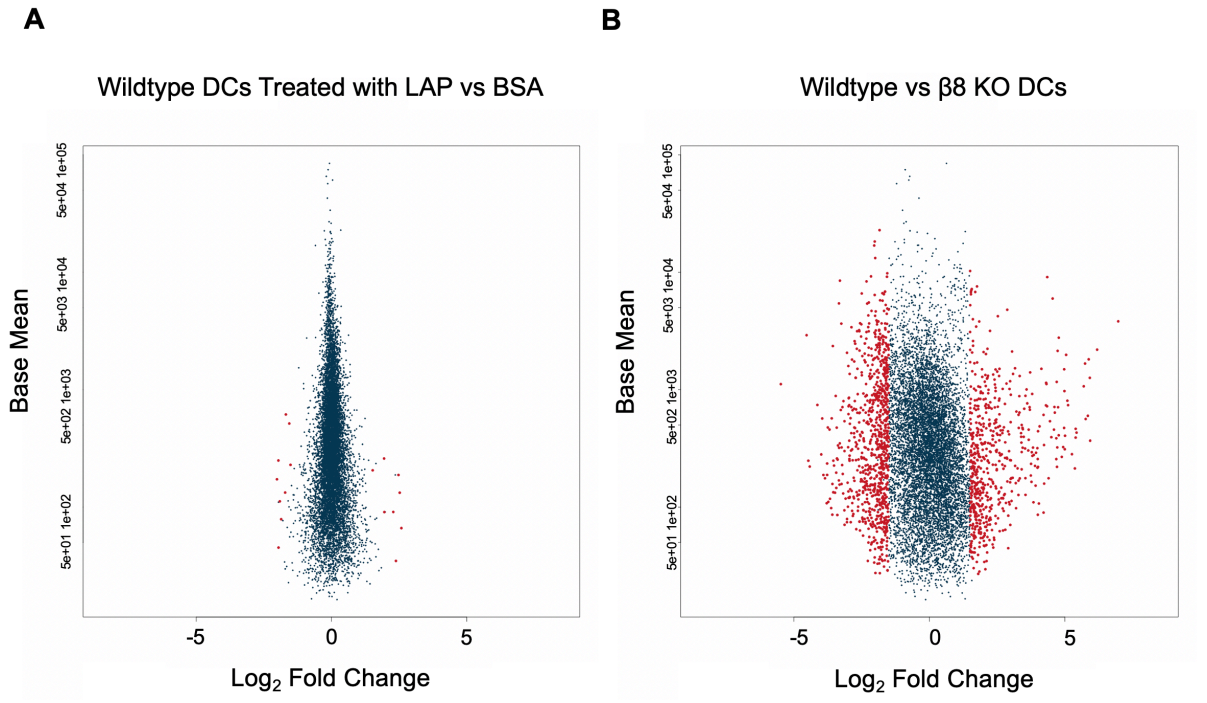


Figure 3.6 | Transcriptional profile induced by the differential expression of β8 subunit on CD11c+ cells show more pronounced phenotypic changes than those elicited by the sole treatment of LAP. A) MA-plot illustrating distribution of genes, represented as dots, when wildtype CD11c+ cells were treated with LAP versus a control condition. B) MA plot showing gene dispersion, when wildtype CD11c+ cells are compared with β8 KO ones. X- and Y-axis refers to the log2 fold change and base mean, respectively. Data are gathered from experiment where the SVA method was not applied. Red-coloured dots are genes that reached an adjusted p-value < 0.05. N=2.

Therefore, these results finally indicate that – in this experiment – effects evoked by LAP are minimal. Moreover, conditional loss of the integrin β8 on CD11c+ cells result in bigger phenotypic differences when these cells are compared to wildtype ones. Given the high experimental noise and minor biological effects induced by LAP, it was necessary to combine different bioinformatic approaches to retrieve potentially useful biological information about which genes were directly regulated by the LAP-αvβ8 signalling axis.

3.5 DEA, performed at different degrees of corrections and statistical power, reveals potential αvβ8-dependent genes in murine CD11c+ cells

As previously mentioned, the SVA approach provided an estimate of surrogate variables to be used in the downstream analysis, after having captured experimental heterogeneity arising from the detection of unmodelled factors that could skew final observations. Therefore, with the ultimate aim of finding αvβ8-dependent genes in CD11c+ cells, we decided to extend the usage of the SVA algorithm for DEA. Indeed, given the high experimental variability, covariates constructed with the SVA algorithm were used for statistical analysis. In fact, in the uncorrected dataset only one gene reached significance when wildtype CD11c+ cells were incubated with LAP. On the other hand, if one considered the SVA-corrected results only, in case of an over adjustment, a likely share of false observations might be introduced for downstream validation. Therefore, we decided to apply a more stringent approach.

As a result, Venn diagrams were used to extrapolate biological information regarding genes involved in the LAP-αvβ8 signalling cascade while controlling for confounding variables that were likely to occur – in case the SVA algorithm would have over fitted the experimental observations (Figure 3.7A). Specifically, Venn diagrams were drawn to describe a potential

common gene signature virtually retained across different conditions analysed and for the same type of comparison. In view of these observations, genes shared between different methods would then be selected for experimental validation.

For the same comparison – that is, wildtype CD11c+ cells treated with LAP or not – DEA was therefore performed four times at different degrees of SVA-correction and/or statistical power (Figure 3.7A). Specifically, the experiment contained a list of genes, corrected or not with the SVA package, from the full dataset and intersected with outputs generated after outliers' removal. Results showed that differentially expressed genes, in which outliers were removed from the experiment showed similar results, indicating that the biggest experimental variation was due to outliers and not by the impact of the correction. Therefore, we identified 12 potential candidate targets that were shared across the four sets (Table 3.1, N=3).

Gene name	Log2 fold change
Cxcr4	-0.78
Ski	-0.68
Zfp950	1.65
Smad7	-0.86
Skil	-1.09
Lss	0.98
Lpxn	-0.79
Itgae	-1.38
Pmepa1	-1.38
Zfp518a	2.22
Fmnl3	-1.24
Zc3h13	-1.65

Table 3.1 | Genes of interest, resulting from the intersection of different bioinformatic approaches (see text) when wildtype CD11c+ cells were incubated with LAP for 6 hours. Log2 fold change values refer to results obtained after having applied the SVA method to the full experiment. Threshold cut off was log2 fold change > log 1.5 and padj < 0.05. N=3.

Theoretically, gene changes seen in table 3.1 may be due to an $\alpha v\beta 8$-dependent or independent signalling cascade. We then applied the same approach as before to perform

DEA in a context when DC lacked expression of integrin αvβ8. Genes represented in both the lists, with the same expression changes, were considered to be integrin β8-independent and therefore excluded from the analysis. To this end, DEA was performed on β8 KO CD11c+ cells after treatment with LAP. After this analysis, samples were intersected in a three-way Venn diagram (Figure 3.7B). Indeed, if the dataset was not corrected with the SVA tool, comparisons between LAP versus control group in β8 KO CD11c+ cells returned no results.

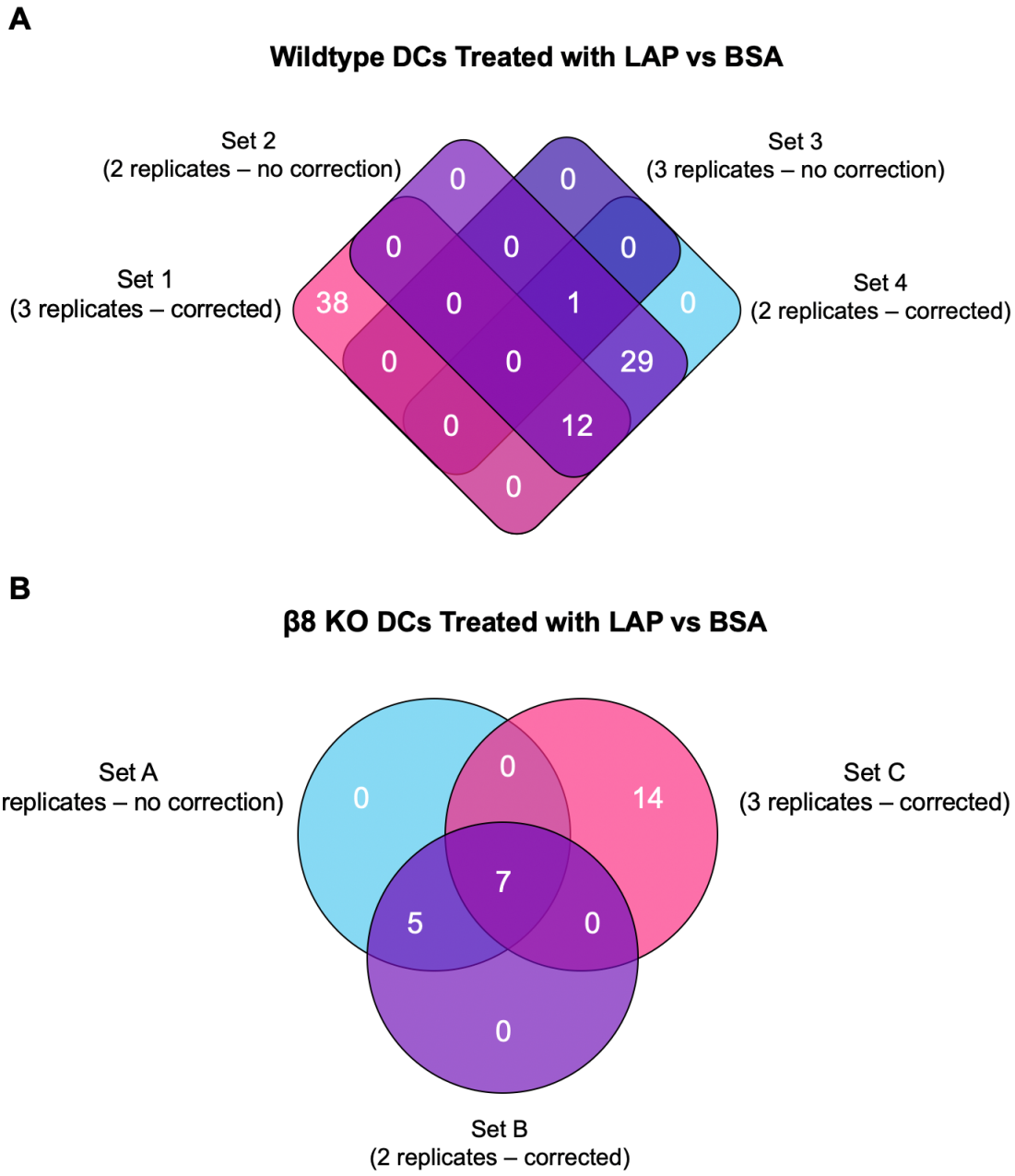


Figure 3.7 | Venn diagrams comparing differential gene expression in CD11c+ cells treated with BSA versus LAP after different methods of correction/analysis of data. A) Four-way Venn diagram comparing LAP- versus BSA-treated wildtype CD11c+ cells at different degrees of correction and/or statistical power. Set 1, in pink, contains 3 replicates and shows number of differentially expressed genes reaching significance when dataset was corrected with the SVA method. Set 3 (violet) is an uncorrected set of significantly differentially expressed genes, with Set 2 (dark blue) being the same analysis but with outliers removed. Set 4 (light-blue) is the set of differentially expressed genes after the SVA correction method with 2 replicates (outliers removed). B) Three-way Venn diagram showing genes statistically different when β8 KO CD11c+ cells were treated with LAP versus BSA. Set A, in light-blue, is the set of genes when experiments were performed after outliers removal (2 replicates) and no correction was applied. Set B, in pink, contains differentially expressed genes when no outliers were removed, corrected with the SVA method. Set C, in violet, includes SVA-corrected results when β8 KO CD11c+ cells were analysed with outliers removed in duplicates. In all the experiment significance threshold was denoted with a padj<0.05 and fold change ≥ log 1.5.

Uncorrected and SVA-corrected gene lists after outlier removal were qualitatively similar. Next, these sets were intersected with outputs generated after DEA performed on SVA-corrected dataset derived from the full experiment. Finally, seven genes resulted from this intersection (Figure 3.7B).

To exclude αvβ8-independent genes from the analysis, a list of candidate genes differentially expressed in wildtype CD11c+ cells, treated with LAP versus BSA, were filtered against the same comparison in β8 KO CD11c+ cells. Results showed that two genes –Ski and Smad7 – were shared between these two experimental comparisons and were removed from the analysis as molecules independently regulated by integrin αvβ8.

Despite the rigorous and conservative approach adopted, given the large experimental variability and low number of observations generated, we decided it was a better option repeating the whole RNA-seq experiment, analysing two different time points post-ligand treatment.

3.6 Discussion

3.6.1 Transcriptome profiling of DC from wildtype or conditional KO background used to determine αvβ8-dependent genes

The role played by the integrin αvβ8 in immunity is paramount (Travis *et al.,* 2007). To date, whether signalling cascades are evoked downstream of integrin αvβ8 expressed by DCs upon ligand binding is not yet determined. To address this, CD11c+ cells from mLNs – described to express high amount of αvβ8 (Païdassi *et al.,* 2011) – were treated for six hours with LAP – the αvβ8-specific ligand. Thanks to transcriptome profiling techniques, we aimed to determine genes induced or repressed after LAP treatment. Nevertheless, we could not rule out likely indirect responses upon LAP treatment, not directly due to the integrin αvβ8 signalling.

In fact, having used LAP at saturating concentrations (Mark Travis, personal communication) without its active moiety, TGF-β-related genes could also be altered. This because LAP and the latent form of TGF-β (LAP+TGF-β), present in variable quantities in the FCS used for cell culture media preparation, compete for the same binding site of the integrin αvβ8. Therefore, in this setting, it would be unlikely that the latent TGF-β is activated by the integrin αvβ8 to evoke its biological effect. In the light of this, gene changes seen in the RNA-seq experiment could not be directly linked to a direct effect of LAP.

Lastly, in this experimental setting, should a change be TGF-β-independent, transcriptional effects triggered upon LAP incubation could be due to either an αvβ8-dependent or independent signalling cascade. Thus, in order to exclude changes not due to αvβ8, CD11c+ cells lacking the expression of the integrin αvβ8 were incubated with LAP.

3.6.2 RNA-seq analysis required bias detection and correction

Analysis of this previously generated dataset was characterised by an undefined source of noise possibly due to lack of metadata such as information regarding experimental batches. Diagnostic plots showed that samples were not correlating within each of the four groups

considered. Lack of within-group correlation may lead to the wrong assumption that samples are different, even if these were subjected to an identical treatment or sharing the same genetic background.

Since this dataset showed great variation in part due to an unknown source of bias, best practice is to rely on tools aimed at detecting the background noise skewing final observations. To this end, the SVA method was used. Through the use of surrogate variables integrated in the experimental design, the SVA package removed the unwanted experimental variability.

In order to control for bias, while avoiding unnecessary technical over adjustment, the same comparison – i.e., LAP- versus BSA-treated CD11c+ cells – was analysed at different degrees of correction and statistical power. Next, results obtained were intersected and studied in a four-way Venn Diagram with the ultimate aim of defining a common gene structure shared across different methodologies applied. Only one gene differentially expressed was found in the full, uncorrected, experiment and found to be a false observation given the fact it was not shared with other correction methods. Minor corrections were performed by the SVA package when four outliers were removed from the dataset. We found that LAP elicited a minimal effect on CD11c+ cells as only twelve genes were identified with the aforementioned method.

Consequently, to exclude β8-independent gene changes, candidate genes were further matched against genes differentially expressed in β8 KO CD11c+ cells when treated with LAP. As a result of this process, two genes were removed not being the latter ones due to a direct β8 signalling cascade – according to the method adopted.

As the dataset was affected by a level of background noise, this made it difficult to have clear evidence about whether a biologically important intracellular signalling pathway could be evoked when αvβ8 engaged with LAP. Therefore, we decided to repeat the RNA-seq experiment, as will be presented in the next chapter.

3.7 Conclusion

In this chapter, adopting several bioinformatic approaches to analyse a previously generated dataset, we demonstrated that LAP incubation induced minor transcriptional changes, in CD11c+ murine cells. Given that we could not rule out that part of the experimental variation seen in this dataset could be due to technical issues during sample preparation, we decided it was a safer option repeat the RNA-seq experiment. In this way, we could control for any unwanted technical variables, as explored in the next chapter.

Chapter 4

Determining genes differentially expressed in splenic DC upon LAP engagement

4.1 Introduction

In chapter 3, we performed analysis of a previously generated RNA-seq dataset in our lab to determine the signalling role of integrin αvβ8 in DC by studying genes differently regulated in primary cells following a six-hour long incubation with LAP. We detected and attempted to remove variation affecting the dataset by introducing surrogate variables in our statistical analysis. Even with dataset correction, the effect of LAP on DC was minimal. Furthermore, we noticed that the major source of variation was given by differences between β8-KO versus wildtype – regardless of whether these samples were incubated with LAP or not.

Building on these observations, we decided it was a better option to repeat the RNA-seq experiment – as discussed in the following paragraphs.

4.2 Devising a protocol to isolate high-purity CD11c+ MHCII+ cells for RNA-seq with high yield

Results from the previous chapter suggest a high degree of experimental variability that may have caused poor sample segregation upon LAP incubation. Another possible explanation is that LAP could induce transcriptional changes in DC relatively early after incubation. Therefore, we decided to perform another RNA-seq experiment to cover a window of early and later gene expression changes, upon incubation with LAP. This time, DC were isolated from spleen rather than mLNs, given its ease of extraction and tissue manipulation; choosing spleen would have lowered the number of experimental mice needed – due to the higher expected yield of RNA, compared to other tissues.

RNA was extracted and sequenced from DC after a two and six hour-long incubation with LAP (Figure 4.1A). Specifically, the CD11c+ population of splenocytes were initially enriched by positive selection. In brief, magnetic beads were used to label CD11c+ cells, subsequently fractionated based on the different magnetic charge carried by this cell population. To get a

highly pure and homogeneous DC population, the enriched eluate was stained for CD11c and MHCII markers and sorted by flow cytometry (Figure 4.1B). The average purity measured after the first enrichment was 87.7% (±2.6), reaching 99.4% (±0.3) after cell sorting (Table 4.1).

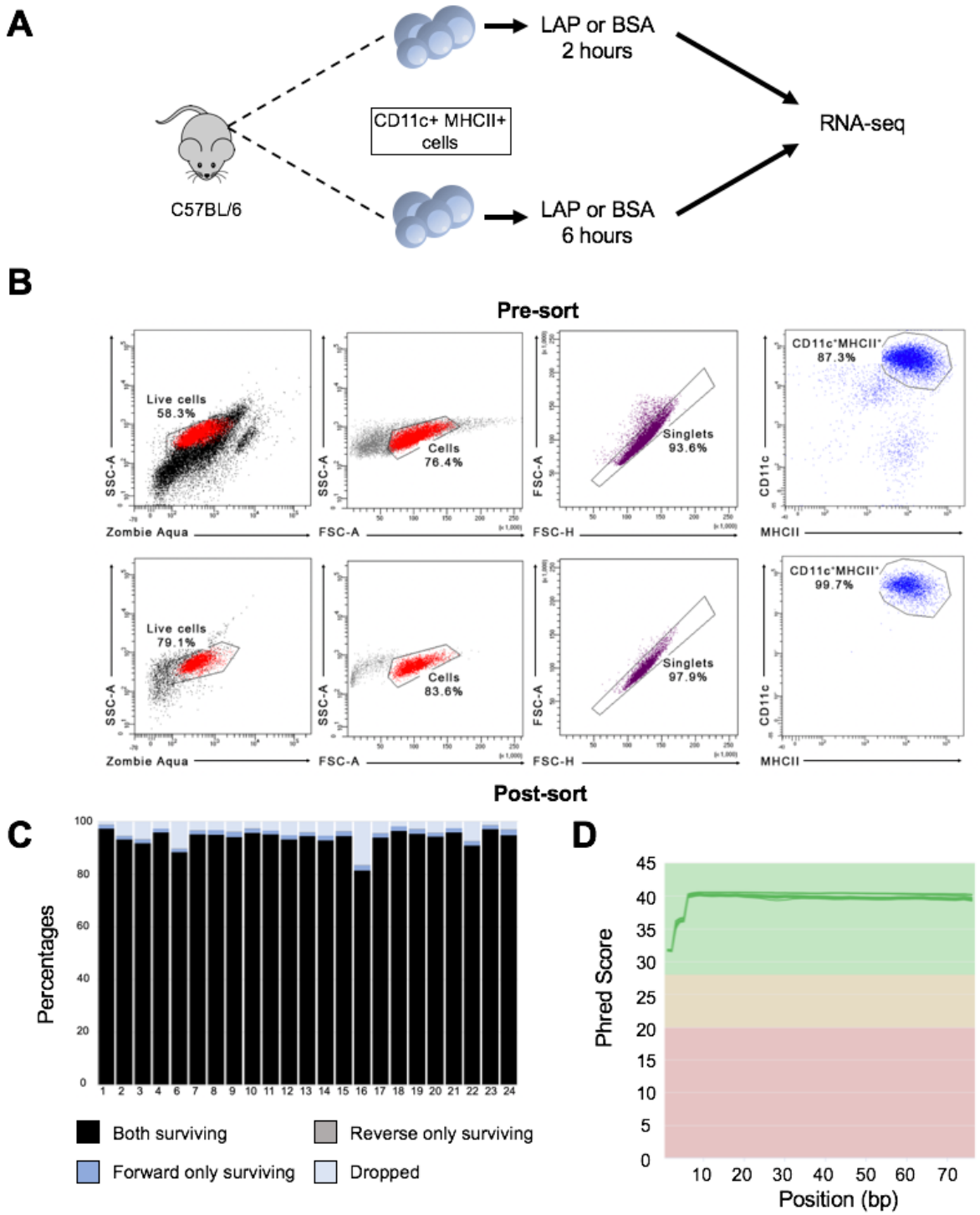


Figure 4.1 | Sample preparation and quality check of sequenced reads. A) Schematic figure of the pipeline adopted. Spleens were removed from C57BL/6 mice and CD11c+ MHCII+ splenocytes were treated with LAP or BSA (5µg/mL) for two or six hours. RNA was extracted, sequenced and RNA-seq performed. B) Representative FACS plots comparing

purity of CD11c+ MHCII+ cell population after the sole enrichment with CD11c-specific magnetic beads (upper panel) or after further purification step by flow cytometry (lower panel). Live cell population was gated on Zombie Aqua positive population and side scatter area, and on forward and side scatter area (first and second dot plots highlighted in red). Single cells were gated on forward scatter height and side scatter area (third plots in violet). Areas of double positive CD11c+ MHCII+ population were drawn using unstained controls. Data is representative of six independent experiments in each of which five biological samples were pooled together after cell sorting. C) Histogram showing percentages of reads – on the Y-axis – surviving the trimming step with the Trimmomatic tool for each sequenced sample – listed on the X-axis. Black bars represent percentage of reads surviving at both ends of the fastq files; blue and light grey bars refer to reads surviving, respectively, on the forward or reverse strand only; light blue is the percentage of reads that were dropped from the analysis. D) Quality check of sequenced reads analysed with the FastQC and MultiQC tools. Plot shows average quality score – on the Y-axis – per base pair position – on the X-axis – of trimmed reads.

Subsequently, DC were treated with LAP or BSA (5µg/mL) at 1x10^6cells/mL and incubated at 37ºC. In order to obtain enough RNA for sequencing, for each experimental replicate, DC were obtained by pooling five mice housed together in the same cage (N=6).

	Post Enrichment					Post Sorting
	Mouse 1	Mouse 2	Mouse 3	Mouse 4	Mouse 5	
Batch 1	89.7	88.4	89	86.2	85	99.2
Batch 2	89.3	87.8	89	91	87.3	98.9
Batch 3	86.9	87.3	87.3	87.8	89	99.7
Batch 4	88.1	88.5	91.7	87.4	84.9	99.6
Batch 5	81	89.7	84.3	84.7	87.6	99.2
Batch 6	89.8	85.1	86.8	84.7	87.3	99.5

Table 4.1 | The percentage purity of cells post enrichment and after flow cytometric cell sorting

Following the incubation time, RNA was extracted, and its concentration measured by automated RNA electrophoresis in the Genomic Technologies Core Facility (GTCF), University of Manchester.

Next, the RNA-seq library was prepared at the GTCF, and pair-ended sequenced to increase reliability in the downstream steps of the protocol. Only one sample failed to make an RNA-seq library – that is, one biological replicate of DC treated for 2 hours with BSA.

Concluding, combining two purification steps in sequence, a highly pure population of DC was obtained from murine spleens. Apart from one sample, the RNA extracted from all the remaining ones was deemed to be acceptable for RNA-seq library preparation.

4.3 Quality check and pseudoalignment of sequenced RNA samples extracted from primary DC

Fastq files – i.e., outputs generated by the high throughput sequencer – contained all the information required for the downstream analysis. Yet, these files needed further verification steps to determine likely sequencing issues such as adapter contamination. In the first place, raw fastq files were assessed for quality by using the FastQC tool and all the information were then collected by the MultiQC software. Despite the majority of raw fastq files showing good quality, some files showed a quality score < 30 in their first base-pair positions – that is, an error rate of base calling of at least 1 in 1000. Moreover, some overrepresented sequences – i.e., contaminants accounting for more than 0.1% of the total transcript length – were also filtered out. Therefore, raw sequencing reads were processed with the Trimmomatic tool in order to increase quality of all samples while avoiding issues in further steps of the pipeline (Figure 4.1C). The QC was performed again on trimmed samples which showed a significant reduction in percentage of overrepresented sequences – i.e., <0.01% – and an increased average phred quality score at the beginning of the sequencing reads (Figure 4.1D). Following this preliminary step, the Kallisto tool was used to pseudoalign trimmed reads to the mouse transcriptome (GRCm38) where the percentage of pseudoalignment ranged between 84.0% and 90.2% (Figure 4.2).

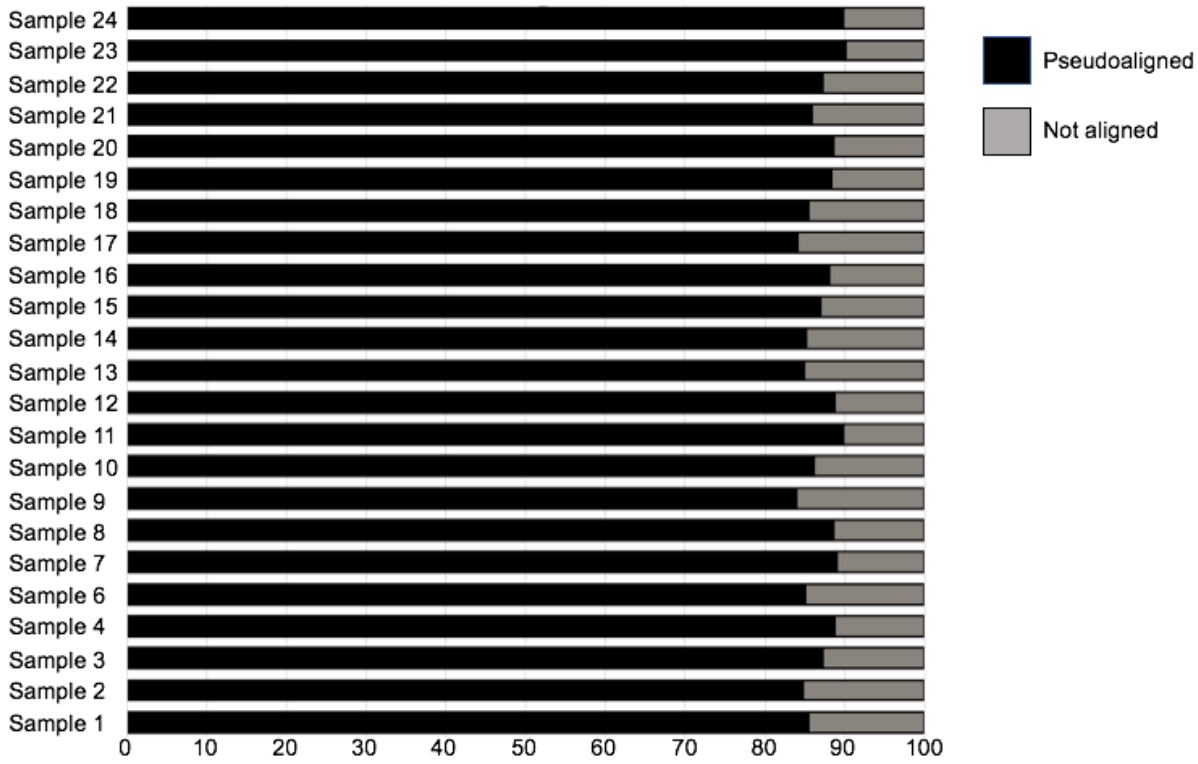


Figure 4.2 | Percentage of trimmed sequencing reads pseudoaligning against the mouse transcriptome. Pair-ended trimmed fastq files were pseudoaligned against the mouse transcriptome using the Kallisto tool. In each sample – on the Y-axis – the horizontal stacked bar chart shows, in black, the percentage of reads – on the X-axis – that successfully pseudomapped the reference transcriptome; in grey is the proportion of reads that were not pseudoaligned.

In conclusion, results showed that raw fastq files generated by the Illumina platform were generally good quality – excluding problems potentially occurred in the initial steps of the pipeline. Subsequently, reads were trimmed to enrich for high quality sequences and remove source of contamination. These steps ensured high reliability in the pseudoalignment step performed with the Kallisto tool.

4.4 PCA plots show a time-dependent effect of LAP incubation on DC gene expression

Gene counts, generated after the pseudomapping step with the Kallisto tool (Bray *et al.,* 2016), were used as input files for DEA. In brief, statistical analysis was performed in R with the DESeq2 package (Love *et al.,* 2014) to determine the number of differentially expressed genes

and magnitude of their change. As mentioned previously, a PCA plot is a diagnostic plot used to determine within-group correlation of samples and their degree of similarity when compared to other groups. These plots are also useful to identify potential problems requiring further investigation – e.g., poor sample correlation, prior to the analysis.

Results showed that the transcriptional effects elicited by LAP on CD11c+ MHCII+ murine splenocytes were secondary to the incubation time the cells were exposed to. Regardless of their treatment, 75% of experimental variation on the PC 1 was explained by segregation of two major clusters – which referred to their experimental incubation time and not to an effect on the transcriptome elicited by LAP (Figure 4.3A, N=5-6). Given that the overarching scope of the project was to identify LAP-dependent transcriptional differences, it was necessary to see whether PCA could also provide insights on gene changes when comparing LAP- versus BSA-treated samples – at the same time point. Therefore, PCA was repeated by selecting, from the whole experiment, groups of samples incubated at two and six hours (Figure 4.3B, D). In both the PCA plots, samples are not correlating according to their specific treatments. Importantly, segregation appeared to be batch dependent. Specifically, samples treated with LAP or BSA on the same day were more phenotypically similar than other group replicates – gathered on different days.

In an attempt to address this, the Limma's removeBatchEffect (Ritchie *et al.,* 2015) was used to see if control of batch-related noise would have ameliorated the within-group clustering. Results showed that minor differences were obtained by comparing samples treated for two hours with LAP both prior to and after batch effect correction, suggesting that few changes were induced by LAP at this time point (Figure 4.3C). On the other hand, a more defined group clustering, albeit not completely discrete, was achieved for batch-effect corrected samples treated at six hours compared to the untreated group (Figure 4.3E). These results indicated that transcriptional effects in CD11c+ MHCII+ murine splenocytes elicited by a six hour-long treatment with LAP could be determined when samples were corrected for batches.

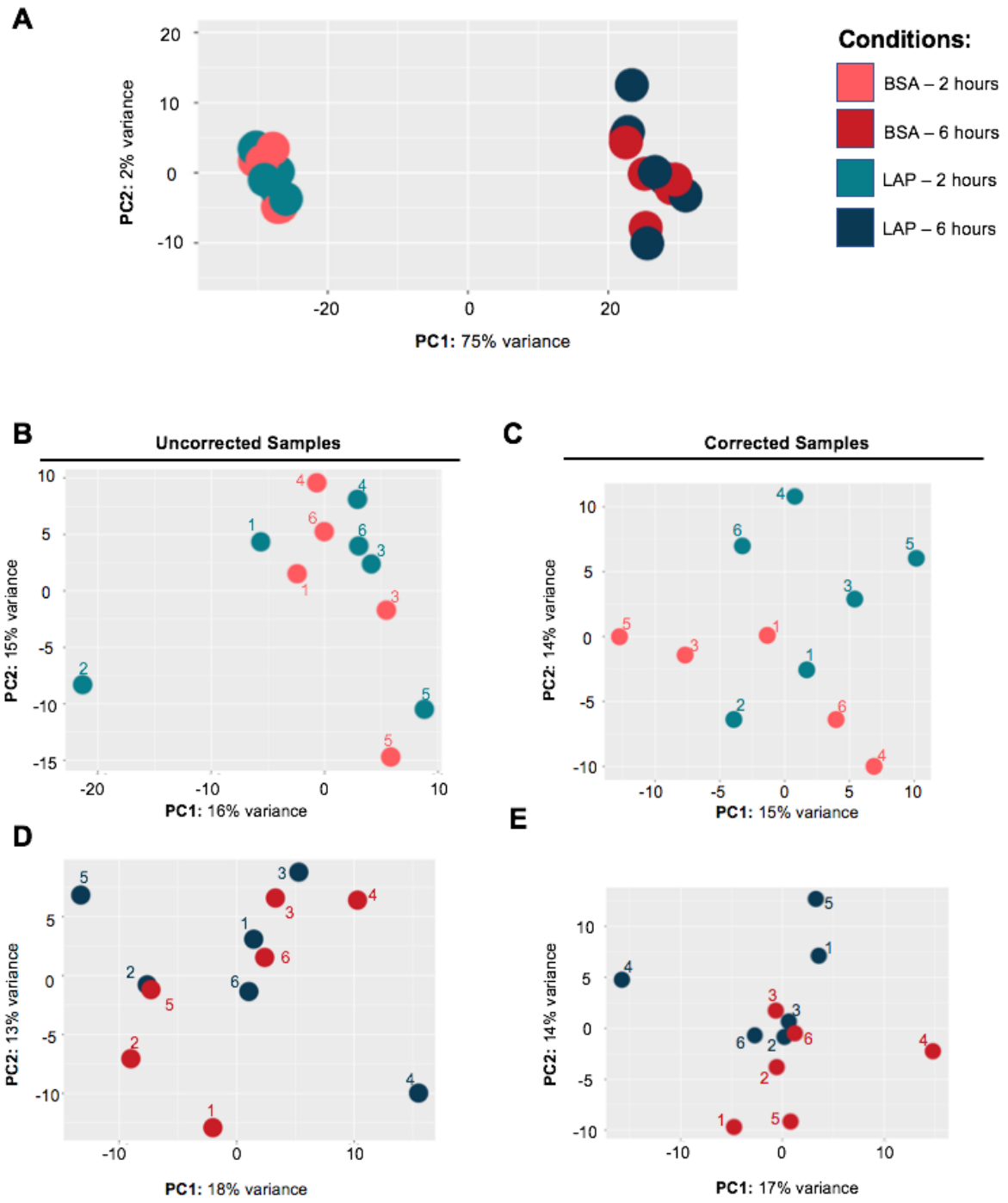


Figure 4.3 | PCA plots show minor transcriptional effects induced by LAP. PCA plots show how samples correlate within each group. A) General PCA plot of DC treated with BSA or LAP at two or six hours indicate the presence of two distinct clusters mainly separating on the PC 1. Plots in B) and C) show segregation of DC at two hours post-incubation with LAP or BSA – before and after batch effect correction, respectively. D) and E) are PCA plots of DC sequenced after a six hour-long incubation with LAP or BSA – controlled or not for batch effect, respectively. Pink and green dots are DC treated for 2 hours with BSA or LAP, respectively; samples treated with BSA or LAP for 6 hours are represented as red and blue dots, respectively. From B) to E), numbers in the PCA plots indicate day of experiment. Each biological replicate is the pool of five different mice, in which splenic CD11c+ MHCII+ cells were equally distributed in each of the four different treatment groups. N=5-6.

In the light of these observations, treatment at two hours was considered a time frame not long enough to detect any transcriptional difference of biological interest. Therefore, analysis of CD11c+ MHCII+ murine splenocytes treated at two hours with LAP were excluded from further downstream analysis. On the other hand, samples incubated at six hours were used for further investigation – as will be discussed in the following paragraphs.

4.5 RNA-seq reveals genes differentially expressed following six hour-long incubation with LAP in splenic DC

After batch correction, samples treated at six hours with LAP or BSA were re-analysed with the DESeq2 package to identify significantly altered genes. In the first place, we found high fluctuations for genes with low counts. Indeed, if genes with very low counts have high fold changes this may be due to expression changes that are not biologically important but rather due to a technical noise. Therefore, we considered genes with a base mean – that is, the average normalised gene counts with an arbitrary cut-off greater than 200 – which were less prone to technical fluctuations, according to our analysis (Figure 4.4A). Therefore, the list of differentially expressed genes was imported into the IPA tool, a pathway analysis software, and enriched pathways examined. These analyses indicated that a number of differentially expressed genes were involved in the TGF-β signalling pathway, actin cytoskeleton, and γ-linolenate biosynthesis (Figure 4.4B). Moreover, hints provided by the pathway analysis tool were used to focus on genes potentially linked to the pathways of interest. Apart from genes enriched in the IPA tool – i.e., Fads 1 and 2 involved in γ-linolenate biosynthesis (Sergeant *et al.,* 2016) and TGF-β related genes – we found that Vasp, involved in cytoskeleton regulation (Krause *et al.,* 2003) and involved in trafficking of TGF-β receptor (Yakymovych *et al.,* 2018), was -downregulated upon a six-hour-long LAP incubation. Also, the contact between the GTPase Rab11, involved in recycling molecules on the cell surface, and the type II TGF-β receptor is mediated by VASP (Yakymovych *et al.,* 2018). Importantly, we found that Rab11fip4 – described as a molecule interacting with Rab11 (Wallace *et al.,* 2002) – was also

suppressed upon LAP incubation. Furthermore, Rasgrp3 is an exchange factor described to modulate the activation of small GTPases (Rebhun *et al.,* 2000) was found to be downregulated – suggesting modulation of its activity following LAP incubation. Therefore, normalised counts of selected genes from the DESeq2 tool were exported (Figure 4.4C, ***, $p < 0.001$; ****, $p < 0.0001$).

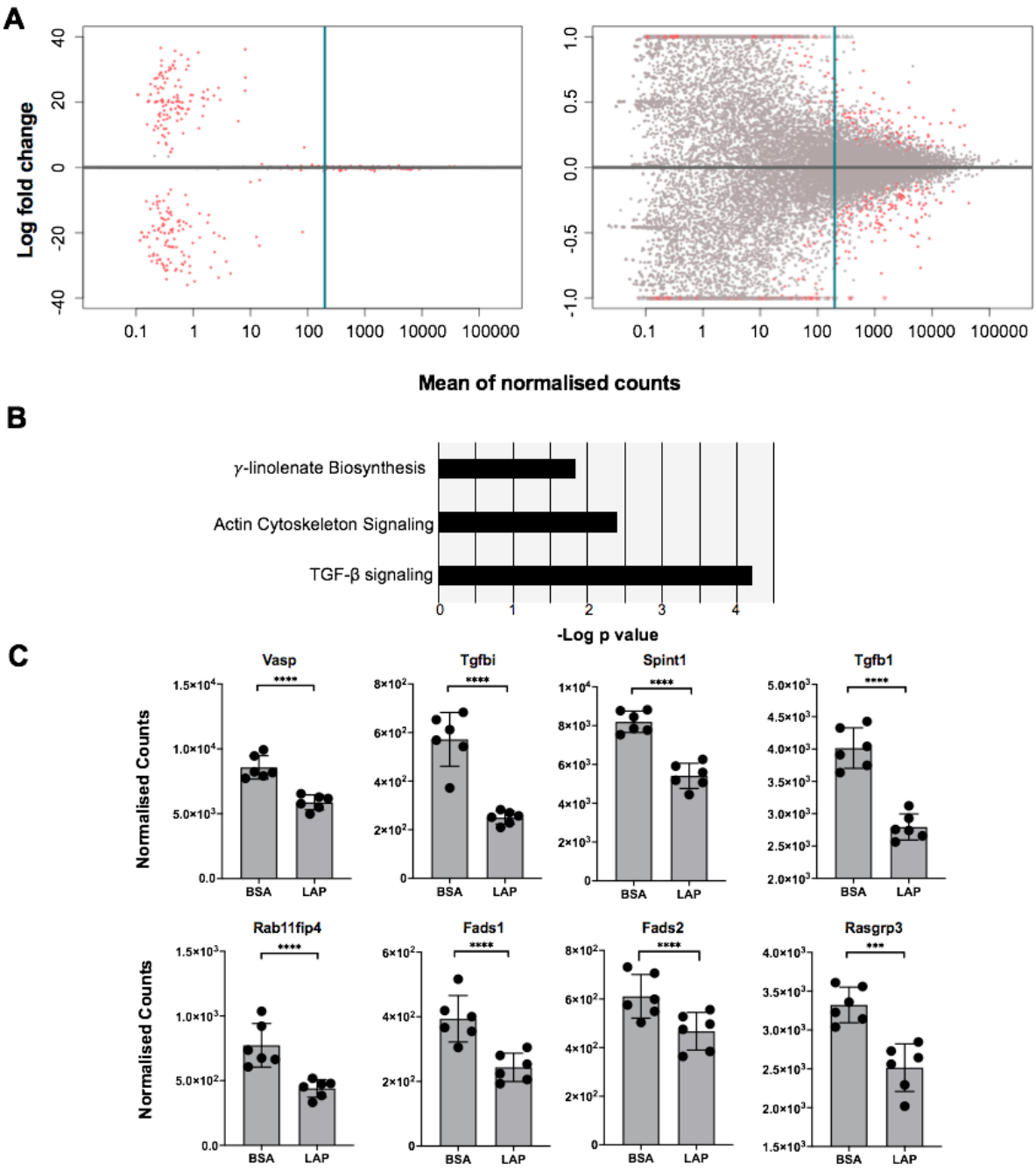


Figure 4.4 | Transcriptome profiling of DC incubated with LAP or BSA for six hours shows enrichment of TGF-β signalling-related proteins, and regulation of genes

involved in actin-cytoskeleton rearrangement and fatty acids metabolism. A) MA plots showing dispersion of genes regulated when CD11c+ MHCII+ murine splenocytes were treated with LAP for six hours. Specifically, the average gene counts – on the X-axis – are plotted against the log2 fold change – on the Y-axis. MA plots are from the same experiment but reported at two different Y-axis intervals. Blue vertical bar is the cut off applied for filtering average normalised gene counts with high dispersion values – that is, greater than 200. Red dots are genes significantly changed. B) Supervised analysis of significantly enriched pathways generated with the IPA tool and represented as a bar chart. On the X-axis is the negative logarithm of the p value calculated with the IPA tool where a positive correlation exists between the height of bars and their statistical significance. C) Normalised counts generated with the DESeq2 package of candidate genes found significantly changed in an experiment where LAP-treated CD11c+ MHCII+ splenocytes are compared to the control group. Error bars are mean with standard deviation. Asterisks represent a p value computed with the DESeq2 package where the expression values differ significantly (***, $p < 0.001$; ****, $p < 0.0001$) using the Wald test. (N=6).

In conclusion, despite some technical background noise, efforts were made to account for this by setting a stringent threshold cut off to focus our attention on gene changes that may potentially be important for biological validation. Consequently, pathway enrichment analysis of top significant genes differentially expressed identified potential pathways that appeared affected upon LAP treatment. The potential functional importance of this will be discussed below.

4.6 Incubation with LAP downregulates CXCR4 in DC and reduces migrating behaviour in CD11c+ murine splenocytes

In this dataset, the gene encoding the chemokine receptor CXCR4 was significantly downregulated upon incubation with LAP for six hours (Figure 4.5A, left panel, N=6; ****, $p < 0.0001$). Briefly, CXCR4 is the most widely expressed chemokine receptor, regulating important biological processes such as cell migration (Bianchi & Mezzapelle, 2020). Therefore, RT-qPCR was performed to validate these results. In brief, RNA was extracted from FACS-sorted CD11c+ MHCII+ murine splenocytes and incubated with BSA or LAP (5µg/mL), at 37ºC. After RNA reverse transcription to cDNA, RT-qPCR was performed and CXCR4 expression values normalised to the level of HPRT1. Results confirmed that, upon LAP engagement, CD11c+ MHCII+ splenocytes downregulate CXCR4 gene expression (Figure 4.5A, right panel, N=4; *, $p < 0.05$).

The chemokine receptor CXCR4 was described to play a key role in orchestrating DC migration to peripheral tissues (Ricart *et al.,* 2011). Therefore, to determine whether a LAP-dependent gene repression of CXCR4 in CD11c+ MHCII+ murine splenocytes impacted their migrating behaviour, we established a transwell migration experiment. In short, CD11c+ enriched splenocytes were incubated with BSA or LAP (5µg/mL) in presence of LPS (0.5µg/mL) for four hours at 37ºC (Figure 4.5B, upper panel). Then, cells were loaded onto the upper layer of a transwell and exposed to the presence of a chemoattractant, CXCL12, for 2 hours in the lower chamber of the plate (Figure 4.5B, lower panel). At the end of the experiment, the whole number of cells were acquired by flow cytometery, in both upper and lower wells. Results show that LAP-treated CD11c+ splenocytes migrated less compared to the BSA-treated group, when cells were exposed to CXCL12 – a CXCR4-specific chemoattractant (Figure 4.5C; N=3; *, $p < 0.05$).

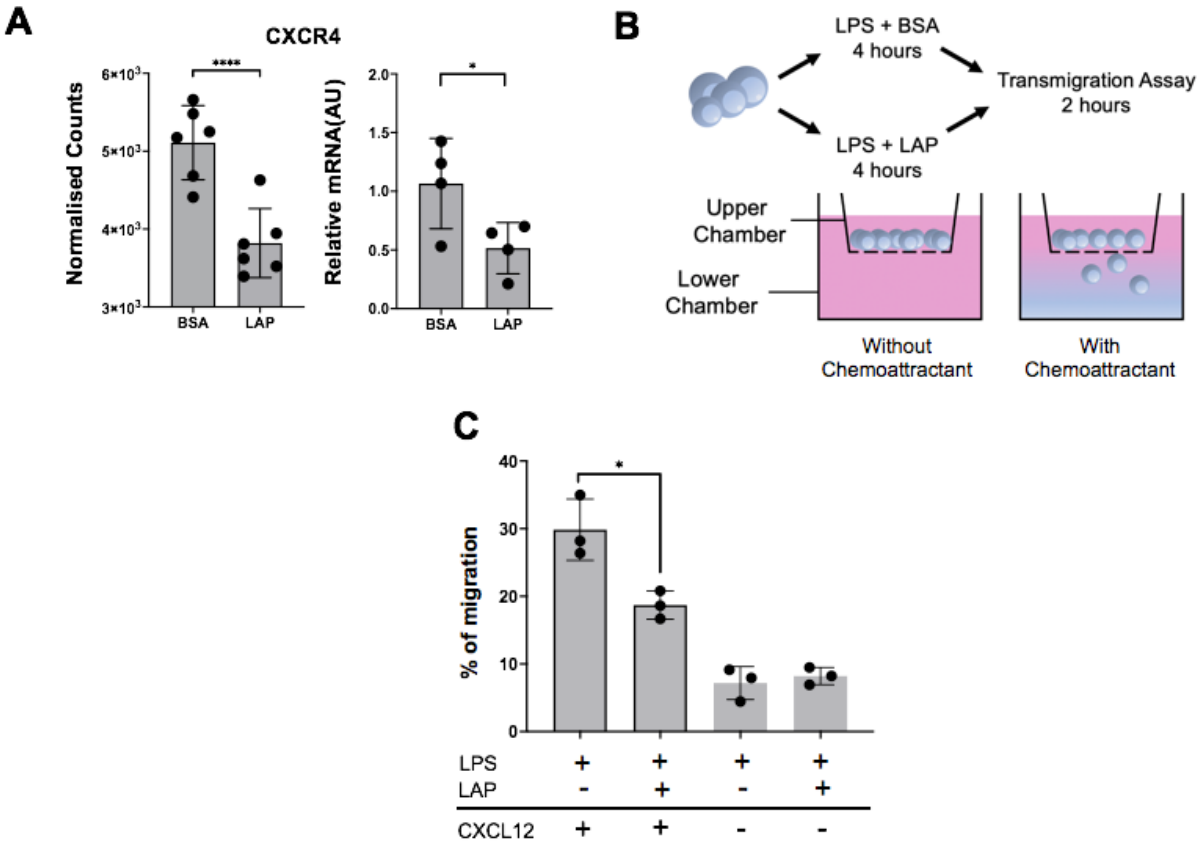


Figure 4.5 | LAP represses CXCR4 expression in murine DC and decreases CXCR4-specific migration in CD11c+ splenocytes. In A) HPRT1 was used to normalise CXCR4

expression levels. Results – expressed in arbitrary units (AU) on the Y-axes – were calculated with the 'delta-delta Ct' method. A) Between-group comparisons of CXCR4 normalised counts generated with the DESeq2 package (left panel) and RT-qPCR expression values (right panel) after incubation with LAP, in murine splenic CD11c+ MHCII+ cells. B) (Upper panel) schematic representation of the workflow used to perform the migration assay: cells were co-incubated with LPS (0.5µg/mL) and BSA or LAP (5µg/mL) for four hours at 37ºC; then, transmigration experiment was performed for 2 hours. (Lower panel) schematic representation of the transmigration experiment: cells loaded onto the upper chamber of a transwell apparatus were exposed (right well) or not (left well) to the presence of chemoattractants, in the lower chamber. C) Percentage of migration of CD11c+ cells exposed to CXCL12 (200ng/mL) versus a control. The first two bars are experimental conditions and last two are negative controls – in which cells were exposed to CXCL12 or not, respectively. On the X-axis is the treatment group. The '+' or '-' symbols indicates the presence or absence of a certain reagent used, respectively. Y-axes report the percentage of migration measured as cell in the lower chamber divided by the total amount of cell and multiplied by 100. Asterisks represent differences that are statistically different (*, $p < 0.05$; ****, $p < 0.0001$). For RT-qPCR and migration experiment a paired T-test was used. The Wald test with the Benjamini-Hochberg multiple testing correction was used to test for differences between normalised counts. (N=3-6).

Overall, incubation with LAP suppressed CXCR4 transcription and reduced a CXCR4-dependent migration in CD11c+ splenocytes.

4.7 Incubation of splenic DC with LAP downregulates expression of the TGF-β signalling protein Smad3

Smad3 is a key transcriptional regulator involved in the TGF-β signalling pathway (Datto *et al.,* 1999). Of note, results from RNA-seq experiment showed that CD11c+ MHCII+ splenocytes significantly increased Smad3 expression upon LAP incubation (Figure 4.6A).

To test the hypothesis that upregulation of Smad3 was a direct effect of integrin αvβ8 upon ligand binding, the level of Smad3 expression was measured on DC isolated from spleen of conditional KO mice lacking αvβ8 expression on CD11c+ cells. Results showed no difference in Smad3 expression when LAP-treated β8 KO mice were compared to their littermate control (Figure 4.6B). In addition, to determine whether Smad3 transcript expression was dependent on TGF-β, we investigated its regulation in splenic DC – isolated from a wildtype background – and incubated with BSA, LAP, the TGF-β blocking antibody (i.e., 1D11), free active TGF-β,

and in presence of LAP and TGF-β. Results confirm Smad3 upregulation upon LAP treatment (N=3-6; *, p < 0.05). Also,when DC were incubated with 1D11, TGF-β, and LAP+TGF-β, effects on Smad3 expression was not significantly different from the control group (Figure 4.6C).

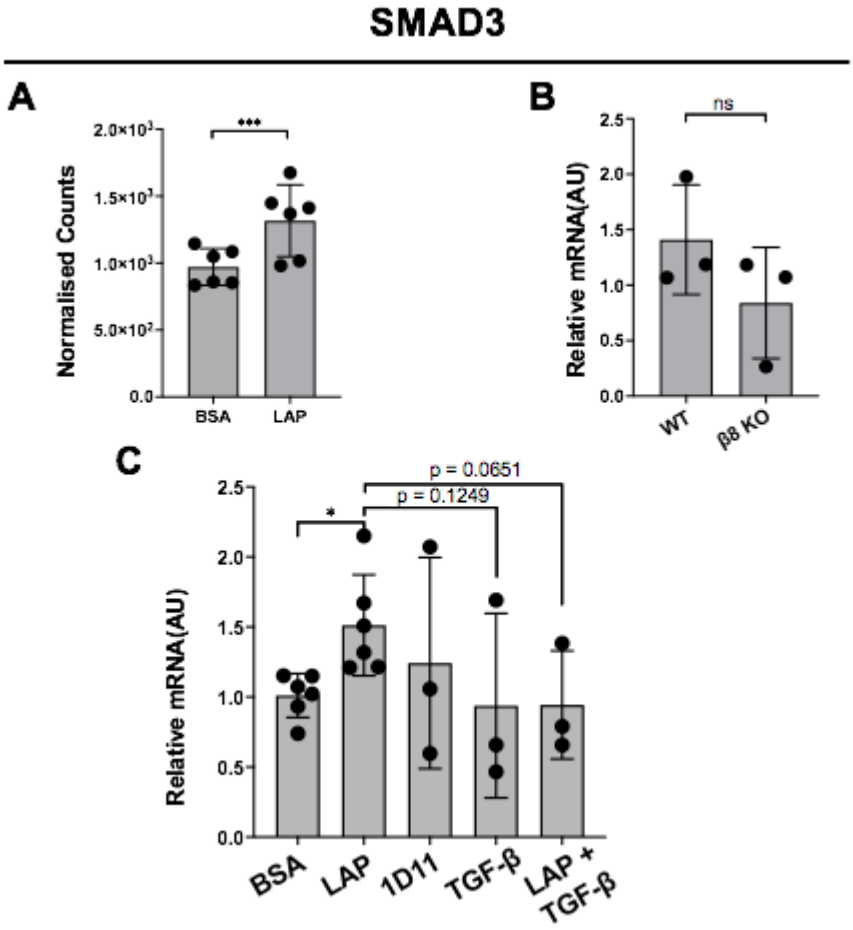


Figure 4.6 | Smad3 expression levels are regulated by LAP. A) Smad3 normalised values from bioinformatic experiment – in which murine CD11c+ MHCII+ splenocytes were incubated with LAP or BSA for 6 hours. B) Smad3 expression values where murine CD11c+ MHCII+ splenocytes, from wildtype mice versus mice lacking the expression of αvβ8 on CD11c+ cells – i.e., CD11c+ cre + itgb8 fl/fl, were incubated with LAP. C) RT-qPCR results of Smad3 in an experiment where CD11c+ MHCII+ splenocytes were incubated with BSA or LAP (5µg/mL), TGF-β blocking antibody (clone 1D11, 40µg/mL), active TGF-β (10ng/mL), and LAP (5µg/mL) + TGF-β (10ng/mL). In RT-qPCR experiments, Smad3 expression values were normalised to the level of HPRT1 and expressed as arbitrary units (AU) following the application of the $2Ct^{\Delta treated}/2Ct^{\Delta untreated}$ method. A response that differs significantly is marked with an asterisk (*, p < 0.05; ***, p < 0.001). Statistical significance between normalised counts was calculated using the Wald test and paired T-test in RT-qPCR experiments (N=3-6).

In conclusion, results indicate that, in this experiment, Smad3 expression is enhanced by LAP treatment in DC. Our preliminary results could anticipate that this might occur in an integrin αvβ8-dependent manner, as lack of β8 expression resulted in no increase in Smad3 expression following LAP treatment versus wildtype DC. Due to the small sample size and the small increase measured in the RNA-seq experiment, we could not conclude that LAP engagement with integrin αvβ8 may regulate the TGFβ signalling pathway.

4.8 Incubation of splenic DC with LAP downregulates genes involved in actin cytoskeleton rearrangement

The IPA tool indicated that proteins involved in actin cytoskeleton signalling were altered by LAP treatment and that genes involved in small GTPases regulation were significantly repressed (Figure 4.4B and C). Yet, amongst tested genes –including Rab11fip4 and Rasgrp3 (data not shown) – only a LAP-dependent downregulation of Rho-GDI2 gene was confirmed via RT-qPCR (Figure 4.7A; N=4-6; *, $p < 0.05$; ***, $p < 0.001$). Markedly, studies have shown a key role elicited by the Rho small GTPases in reorganising the actin cytoskeleton, with GDI molecules regulating their biological activity (Rivero *et al.,* 2002). Importantly, we confirmed downregulation of Skil and Cxcl16, upon LAP incubation, seen in the RNA-seq experiment (Figure 4.7B; N=3-6; *, $p < 0.05$; ***, $p < 0.001$, ****, $p < 0.0001$). Of note, previous studies showed correlation of these two gene products in actin cytoskeleton rearrangements, too (Caligaris *et al.,* 2015; Singh *et al.,* 2016). Therefore, given that the actin cytoskeleton signalling pathway was significantly enriched in the pathway analysis tool used (Figure 4.4B), with other genes found to be differentially regulated and linked to this biological pathway (Figure 4.7C; N=6; ***, $p < 0.001$; ****, $p < 0.0001$), we sought to determine whether αvβ8 integrin activation was involved in regulating actin cytoskeleton remodelling.

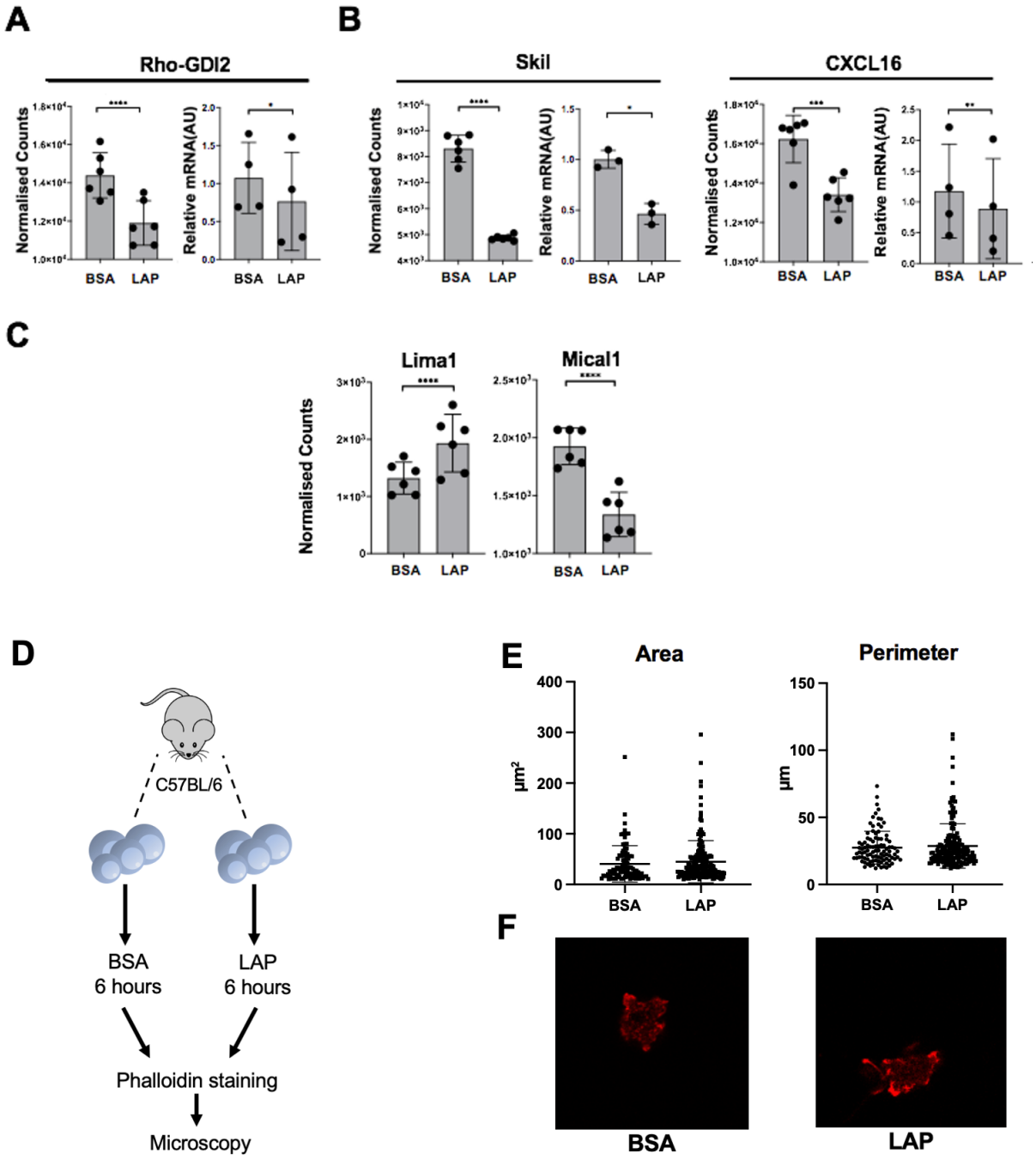


Figure 4.7 | Functional validation of target genes differentially expressed with RNA-seq experiment confirms no phenotypical differences upon LAP incubation. A) and B) Correlation between normalised counts – computed in the DESeq2 package – and RT-qPCR expression values of target genes. C) Normalised counts of target genes involved in cytoskeleton rearrangement. D) Schematic graphics representing protocol used for studying actin cytoskeleton changes. In brief, highly pure population of CD11c+ MHCII+ murine splenocytes were incubated with LAP or BSA for six hours and stained with phalloidin, prior to the microscopy analysis. E) Microscopy analysis where area (micrometre, μm^2) and perimeter (micron, µm) are compared between groups analysed. Bar is the mean. F) Representative figures showing phalloidin staining of CD11c+ MHCII+ splenocytes upon incubation with BSA or LAP. A response that significantly marks a difference between conditions is denoted with an asterisk (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ****, $p < 0.0001$). Statistical analysis for

normalised counts was assessed with the Wald test and paired T-test in RT-qPCR experiments. Microscopy analysis was tested using the Mann-Whitney non-parametric test. Bars are the mean with standard deviation. (N=3-6).

With the final aim to see if measurable structural differences could be induced in DC upon $\alpha v\beta 8$ engagement with its ligand, a microscopy experiment was performed (Figure 4.7D). Briefly, FACS-sorted CD11c+ MHCII+ murine splenocytes were incubated with BSA or LAP (5µg/mL) and stained with phalloidin – highly specific for actin filaments, after being seeded on slides coated with poly-L-Lysin and fibronectin. Following microscopy analysis, cell area and perimeter was measured and compared between LAP- and BSA-treated groups. Results showed no difference in cell area and perimeter, suggesting no correlation between transcriptional changes and actin cytoskeleton rearrangements, as also suggested by qualitative analysis of phalloidin staining (Figure 4.7E, F).

These results exclude any structural rearrangement induced upon a six-hour-long treatment with LAP, in CD11c+ MHCII+ splenic murine DC. Nevertheless, the contribution of target genes in other alternative cell functions cannot be ruled out – given their multiple effector roles elicited downstream.

4.9 Discussion

4.9.1 Minor time-dependent transcriptional changes induced by LAP, in primary splenic murine DC

In this chapter, we revealed that LAP induced minor transcriptional changes compared to the control group. In order to achieve this goal, we firstly tried to determine in which time frame an incubation with 5µg/mL LAP could determine an overt transcriptional response in splenic DC. Therefore, CD11c+ MHCII+ splenic cells were exposed to LAP or BSA at two different time-points – that is, two and six hours – to pinpoint changes occurring at an earlier and later stage of the transcriptional process. For the RNA to be successfully sequenced and given the

experimental setting where each biological replicate had to be split in four treatment groups, enough amount of starting nucleic acid was required. Therefore, CD11c+ MHCII+ splenocytes were isolated form five murine spleens and pooled together to be equally divided in each experimental group to ensure good RNA yield. Thus, for technical reasons it was not feasible to start and complete the whole experiment in one day. Instead, six biological replicates were split into six separate batches. Inevitably, this introduced a batch effect that needed to be corrected during DEA.

Importantly, we did not find evidence to support that ligand engagement with integrin $\alpha v\beta 8$ caused biologically distinct phenotypes. It is worth mentioning that, as evidenced by PCA plots in figure 4.3A, the major variable determining separation of samples is the lapse of time at which each sample was exposed to – regardless of this being LAP or BSA. Indeed, samples were mainly separating along the first PC, clustering in a two and a six-hour group – rather than creating four distinct clusters as per treatment groups. Pairwise combinations of all principal components never clustered into four distinct groups (data not shown), suggesting that effects of the variable of interest is not contained in any of these PCs. This reflects minor effects that LAP elicited in our experiment.

As a result, poor correlation was confirmed for samples measured within each time point. First, samples treated at two hours showed no direct response to LAP treatment – despite a batch effect correction tool being used. These results suggest that transcriptional changes triggered at two hours by LAP were not enough to induce a distinct phenotype from the control group, in splenic DC. On the other hand, samples treated at six hours showed a more evident batch effect-related issue. Indeed, accounting for batch effect, at this time point, a better clustering could be achieved, as shown in Figure 4.3D and E. A sample segregation achieved suggested that LAP incubation evoked a more prominent response compared to the one measured at two hours. Despite a filtering step was performed using the DESeq2 package, genes with low counts were inflated by high fold changes values, as shown in figure 4.4A. This issue may

lead to false observations if not controlled for adequately. Specifically, low statistical power associated with low counts relates high fold changes to poor confidence values. On the other hand, genes with higher counts were less affected by counts oscillation and thus considered more consistent. Therefore, only the latter were included in our experiments. Yet, amongst these genes differentially regulated, none of these were found above the commonly accepted cut off threshold of 1.5 as log2 fold change. These results suggest that – despite PCA plot indicates that a marginal clustering was achieved for a six-hour treatment with LAP – only a small number of genes showed differential expression indicative of an important functional effect and therefore taken forward for validation.

4.9.2 Transcriptional changes elicited by LAP induce few biological differences

A list of genes differentially expressed, filtered for low average gene counts, were used for downstream bioinformatic analysis. Pathways significantly enriched with IPA were mainly related to fatty acid metabolism, actin cytoskeleton signalling and the TGF-β signalling pathway. Of note, expression values of high counts-genes changed by a multiplicative factor less than 3 when LAP treatment is compared to the control group (i.e., a log2 fold change less than 1.5). Therefore, correlation with such changes rarely reached significance in an RT-qPCR setting as the number of biological replicates, needed for rejecting the null hypothesis when fold changes are relatively small, should be consistently high.

4.9.2.1 LAP reduces migration of splenic DC but does not induce cytoskeleton rearrangement

In our experimental setting, we measured that CXCR4 was significantly downregulated upon LAP treatment. Results support the hypothesis that treatment of CD11c+ splenocytes with LAP inhibited their migratory behaviour. Whether a causal link exists between LAP treatment and the outcome measured has yet to be determined. Indeed, in preliminary studies we found that, despite migration towards FCS – containing a

complex mix of chemoattractants – was significantly reduced in a FACS-sorted population of LAP-treated DC, we did not measure a consistent difference between groups when these cells were exposed to CXCL12 (data not shown). Therefore, whether these observations are biologically relevant require further investigation. Reduction in migrating behaviour upon LAP stimulation might indicate CD11c+ cells are maintained for prolonged time where the latent TGF-β is encountered, to better respond to TGF-β.

Importantly, reduction in cell migration is related with differences in cytoskeleton rearrangements that was supported by the finding that several genes associated with cytoskeletal rearrangement were differentially regulated in our RNA-seq experiment. Thus, we tested if differences in cell area or perimeter were induced upon incubation with LAP in murine splenic DC. Our results confirm that incubation for six hours with LAP does not induce changes in area and perimeter of these cells. One could hypothesise that, despite LAP regulating expression of genes involved in cytoskeleton rearrangements, such as small GTPases, other signalling cascades – not indicated in our pathway analysis – could possibly be involved.

4.9.3 Regulation of TGF-β signalling-related genes after incubation with LAP

We functionally tested several genes found to be differentially regulated in our RNA-seq experiment that were described to have a role in regulating TGF-β signalling pathway – that, is PMEPA1, TGF-β-induced, TGF-β1, and SKI (data not shown). However, we found nonsignificant differences in LAP- versus BSA-treated groups for these genes, in RT-qPCR validation experiments.

Bioinformatic analysis indicated that the TGF-β signalling related gene SMAD3 was significantly upregulated, a result that was validated *in vitro* in one experiment. Yet, preliminary data do not indicate that such regulation is integrin-specific rather than TGF-β-dependent.

Furthermore, LAP incubation did not significantly changed expression of SMAD3 in splenic DC isolated from wildtype versus conditional KO mice. Therefore, more work is needed to determine whether SMAD3, or other TGF-β-related genes, is responsive to LAP incubation in a TGF-β independent fashion. One way to better characterise an effect elicited by LAP on DC would be to either select cells with a higher documented β8 expression (Païdassi *et al.,* 2011) and/or increase LAP concentration. An appealing working hypothesis is that encountering of the latent form of TGF-β may prepare DC for the arrival of the active cytokine via integrin αvβ8-mediated activation by upregulating genes involved in the TGF-β signalling pathway. Indeed, if LAP, through the αvβ8, may propagate signalling pathways regulating expression of TGF-β-related genes needs further transcriptional studies on several DC – isolated from peripheral tissues where β8 expression is significantly higher than those isolated from spleen (Païdassi *et al.,* 2011).

Also, a better understanding on potential signalling mediators involved in regulating integrin αvβ8 signalling could be addressed by coupling transcriptome analysis with Western blot studies. In this way, a finer association between transcription and translational processes could be determined – through the use of antibodies specific for targets of interest, transcription of which was preliminary found to be up- or down-regulated. Importantly, one could hypothesise that an overt signalling pathway evoked downstream of αvβ8 can be studied in other DC – where the expression of αvβ8 is higher. If these results are confirmed, these may support the hypothesis that some TGF-β-related intracellular pathways are triggered upon LAP engagement with αvβ8, preparing the cell for responding to TGF-β.

4.9.4 Limitations of our current approaches

As mentioned already, β8 expression is expressed at minimal level in splenic DC (Païdassi *et al.,* 2011). Therefore, to determine a direct effect on DC behaviour upon LAP incubation – different cell subsets could be used in future experiments. Indeed, results from RNA-seq

experiment showed that only a few genes changed between LAP-treated versus untreated group of splenic DC at two and six hours. Moreover, such gene alterations – apart from a big experimental noise that we attempted to control – were small, indicating a mild effect elicited by LAP in this experimental setting. Therefore, trying to validate *in vitro* such small gene expression changes by RT-qPCR is technically difficult for low number of replicates. Indeed, most of the candidate genes chosen from the RNA-seq experiment did not show significant LAP-dependent regulation on gene expression. Also, some of the candidate genes showing statistically significant differences in one experiment – e.g., CXCR4, SKIL (data not shown) – were not consistently changed in other experimental settings. These limitations reinforce the idea that other cell candidates must be used for further experimental analysis to study effects elicited upon LAP incubation.

It is worth mentioning that ligand concentration could also be increased in future prospective experiments. We used a concentration that previous work in the lab identified as optimal – i.e., of 5μg/mL – with higher amounts being cytotoxic on primary cultures of DC (Eleanor Sherwood, unpublished data). Nevertheless, LAP was also used at higher concentrations – 10, 30, or 50μg/mL – in *in vitro* stimulation experiments where its biological effect was dose-dependent, in T cell cultures (Nakamura *et al.,* 2004). Unfortunately, we could not rely on internal controls as readout of appropriate cell stimulation, being $\alpha v\beta 8$-related gene expression changes in DC – to date – unknown.

In addition, another viable approach is to expand the time range at which LAP incubation is performed. In our experiments, we incubated primary DC at two and six hours with LAP versus a control with the hope to detect substantial transcriptional regulation downstream from $\alpha v\beta 8$. Ideally, one potential strategy could result in shorter and longer incubation times. As an exploratory analysis, one could reduce the number of biological replicates while increasing the time points of cells incubated with LAP versus their cognate controls. Anyways, even reducing

number of replicates needed for an RNA-seq experiment, current sequencing cost represent a bottleneck to this study.

Therefore, future efforts could be made to repeat bulk RNA-seq experiments by sorting cells with consistently higher expression of β8, increasing LAP concentration, and covering a broader time range. Another constraint of transcriptome analysis is the lack of correlation that sometimes occur when studying transcripts that are not converted into protein products. Therefore, a better understanding of potential signalling pathways propagated downstream of αvβ8 in primary cultures of DC can integrate bulk RNA-seq analysis with proteomic approaches, such as mass spectrometry. This approach could then be used to study other cell types expressing high levels of αvβ8 and what could be the biological impact of its activation.

4.10 Conclusion

In this chapter, we demonstrated that LAP induced a few transcriptional differences after an incubation of 6 hours, with no biologically important changes at the 2-hour time-point. I*n vitro* validation of target pathways revealed minor effect on the phenotypes. Indeed, computational analysis confirmed that, in our experiment, LAP did not elicit an orchestrated regulation of a wide proportion of gene patterns for experimental validation. We then decided to study the role played by αvβ8 in other cell contexts, as described in the Chapter 5.

Chapter 5

Determining transcriptional profile of Tem expressing the integrin αvβ8

5.1 Introduction

In chapter 3 and 4, we sought to determine signalling pathways downstream of the integrin αvβ8 in DCs through the use of RNA-seq analysis. Results showed that integrin αvβ8 ligand LAP triggers minor transcriptional differences and biological effects. Specifically, the transcriptional profile of wildtype DC versus integrin β8 KO DC incubated with LAP showed that the major transcriptional differences seen were due to β8 expression (Chapter 3). Moreover, when DC were incubated with LAP for two or six hours, as indicated by the PCA plots, bigger differences were seen in transcriptional profiles due to the time in culture rather than the time of exposure to the ligand (Chapter 4).

Next, to address the overarching goal of our project to understand how integrin αvβ8 expression on immune cells regulates immunity, we decided to study the effects of αvβ8 on a different cell population our lab found to be able to activate latent TGF-β, given to the high expression of the receptor. Specifically, we found that a population of αvβ8+ Tem could reverse an exacerbated secondary T cell response mounted against respiratory viruses, in mice.

5.2 Expression of αvβ8 on Tem marks a transcriptionally distinct subset from the αvβ8-negative counterparts

Previous work in the lab suggested that the latent form of TGF-β can be activated by a population of CD4+ Tem, due to their high expression of the integrin αvβ8. Also, following a secondary memory recall response, our lab measured an expansion of influenza specific CD8+ T cells in mice lacking the expression of αvβ8 on CD4+ cells. This anti-viral immune response was not seen in mice lacking β8 under transcriptional control of the Foxp3 promoter. For this reason, we sought to determine if the phenotype seen in the Tem αvβ8+ shared transcriptional similarities with Treg and whether the transcriptional profile of αvβ8+ and αvβ8- Tem was distinct. Therefore, spleens were isolated from Itgb8 reporter mice – expressing a fluorescent

dtTomato reporter molecule under transcriptional regulation of the Itgb8 promoter region (Nakawesi *et al.,* 2021) – and splenocytes sorted as TCRβ+ CD4+ Foxp3- CD44+ CD62L- CXCR5- and further divided into β8+ or β8- populations (Figure 5.1A, B, credit: Craig McEntee). Fastq sequenced files that were assessed for quality and trimmed afterwards (Figure 5.1C). Quality control indicates that the trimming step was needed to remove adapter contamination and overrepresented sequence from samples. All sequenced reads had a phred score always greater than 30, indicating high sample quality in all base pair positions. Also, following the pseudoalignment step with the Kallisto package (Bray *et al.,* 2016), diagnostic plots were generated with the DESeq2 tool (Love *et al.,* 2014). PCA plots indicated that Tem αvβ8+ and αvβ8- segregated into distinct clusters and that differences between these clusters represented the major difference between cells (PC 1 value of 55%) (Figure 5.1D, N=6).

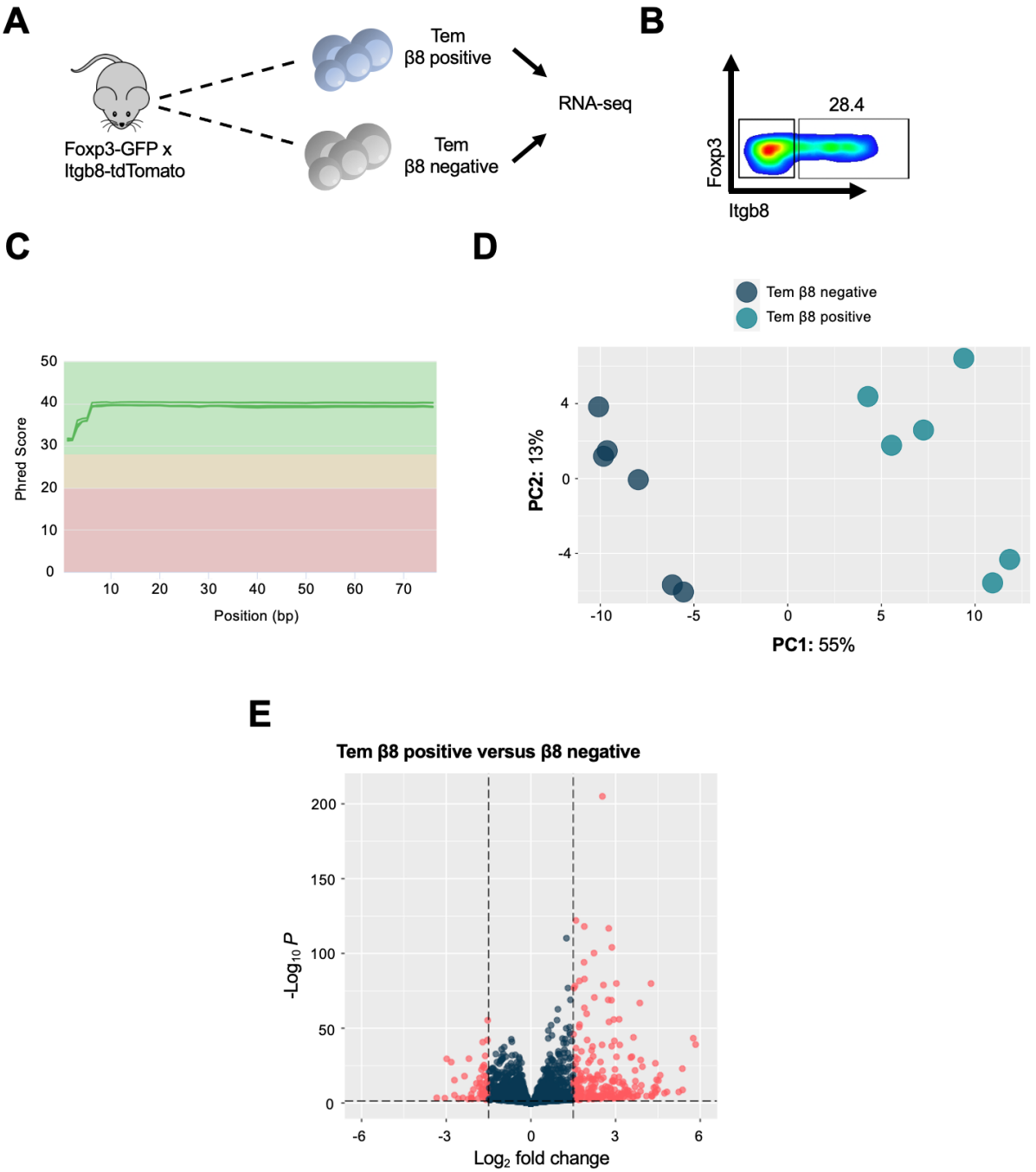


Figure 5.1 | β8 positive Tem are transcriptionally distinct from β8 negative counterparts. A) Representative figure of the adopted experimental setting. In brief, β8 positive or negative Tem were isolated from murine spleen of Foxp3-GFP x Itgb8-tdTomato mice. RNA was extracted and sequenced in two separate batches. B) Representative FACS plot of the gating strategy used, in which the Tem population was gated as single, CD4+, TCR-beta+, CD44Hi, CD62L-, CXCR5-. Gating on Foxp3-, cells were sorted based on their differential expression of β8. C) Quality check of trimmed fastq files from the full experiment. On the X-axis, is the base pair position. On the Y-axis, is the phred quality score assigned to each sample. D) A plot comparing sample segregation of β8 positive Tem (in blue) versus β8 negative counterparts (in light blue). On the X axis, is the PC 1 which accounts for the largest experimental variability. On the Y-axis is the PC 2 – accounting for the second major source of variance. E) Volcano plot representing differentially expressed genes when the Tem β8 positive group is compared to the Tem β8 negative one. On the X-axis is the log2 fold change. On the Y-axis

is the negative base 10 logarithm of the P-value. Vertical and horizontal lines marks chosen threshold cut off. Genes differentially expressed are coloured in red. (N=6).

Moreover, a Volcano plot summarising information from the DEA shows that the vast majority of genes was upregulated when Tem αvβ8+ were compared to Tem αvβ8- group (Figure 5.1E). Specifically, 51 genes were downregulated and 213 were upregulated – using threshold cut-off values of log2 fold change ≥ 1.5 and padj ≤ 0.05. These data suggest that integrin αvβ8 expression expressed on Tem is associated with a transcriptionally distinct pattern versus cells that do not express integrin αvβ8.

Next, functional annotation of the gene significantly altered was performed using the DAVID tool and the KEGG pathway analysis performed. Of note, the "Cytokine-cytokine receptor" module was the most significantly enriched pathway in the analysis (Figure 5.2). Amongst several modules identified with this approach (data not shown), transcription of genes – the product of which has been described to elicit a regulatory or co-stimulatory function – was enhanced in the αvβ8+ Tem. Specifically, in our analysis we found that CD30 and its cognate ligand, CD30L, were differentially upregulated in αvβ8+ Tem versus the αvβ8- counterpart. CD30 is a member of the Tumour necrosis factor receptor family and, together with its cognate ligand, is expressed on activated T cells. CD30 signalling pathway modulates T-cell survival and reduce their cytolytic activity (A. C. Zhou *et al.,* 2017). Noteworthy, we determined that OX40 was significantly upregulated in αvβ8+ Tem compared to αvβ8- cells. Of note, previous work indicated that OX40 and CD30 have common signalling cascades, where OX40-dependent signalling cascade in OX40-deficient mice was compensated by CD30-dependent pathways (Gaspal *et al.,* 2005). IL-10, a key anti-inflammatory molecule, was significantly upregulated when Tem expressed the integrin αvβ8. In fact, these results are in line with experimental findings that were previously generated in the lab where αvβ8+ Tem showed an immune suppressive phenotype compared to the αvβ8- counterparts, *in vivo* (Craig McEntee, unpublished data).

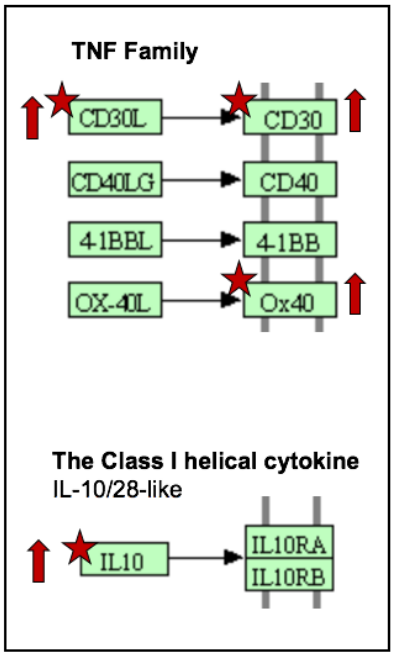


Figure 5.2 | Functional annotation of genes with the DAVID tool of the differentially expressed genes when Tem αvβ8+ are compared to αvβ8- ones. List of differentially expressed genes was analysed using the DAVID tool and KEGG pathway interrogated. The figure shows selected genes from the top enriched KEGG pathway term named "Cytokine-cytokine Receptor Interaction" enriched with a P-value corrected with the Benjamini-Hochberg method of 1.4E-2. Green boxes contain gene product names, described to activate (black arrow) a designated molecule. A red star indicates the gene was found differentially expressed in Tem αvβ8+ versus αvβ8-. Red arrows were added afterwards and refers to an upregulation measured in the abovementioned comparison.

Therefore, given bioinformatics results and experimental validation data suggesting that αvβ8+ Tem display a suppressive phenotype, we next questioned whether the transcriptional profile of these cells was similar to another crucial immunosuppressive cell type, the Foxp3+Treg.

5.3 Comparative analysis of two separate experiments indicate that αvβ8+ Tem are transcriptionally distinct from Treg

A first attempt to determine whether a subset of Treg – isolated from mLNs and marked positive or negative for the αvβ8 expression – was transcriptionally similar to αvβ8+ Tem, was

made by examining a previously generated dataset in a lab at the Cancer Research Centre of Lyon. In brief, Treg were sorted from murine mLNs of Itgb8 reporter mice and single end sequenced afterwards (Figure 5.3A, N=3). A quality check of sequenced reads was performed and identified no significant issues in trimmed reads analysed with the FastQC tool, collated with the MultiQC software (Figure 5.3B). Nevertheless, a PCA plot did not detect two distinct groups, with each group instead having one outlier present (Figure 5.3C). A volcano plot showed that the largest number of genes differentially expressed in a significant way – i.e., Log2 fold change ≥ 1.5 and padj ≤ 0.05 – were upregulated (Figure 5.3D).

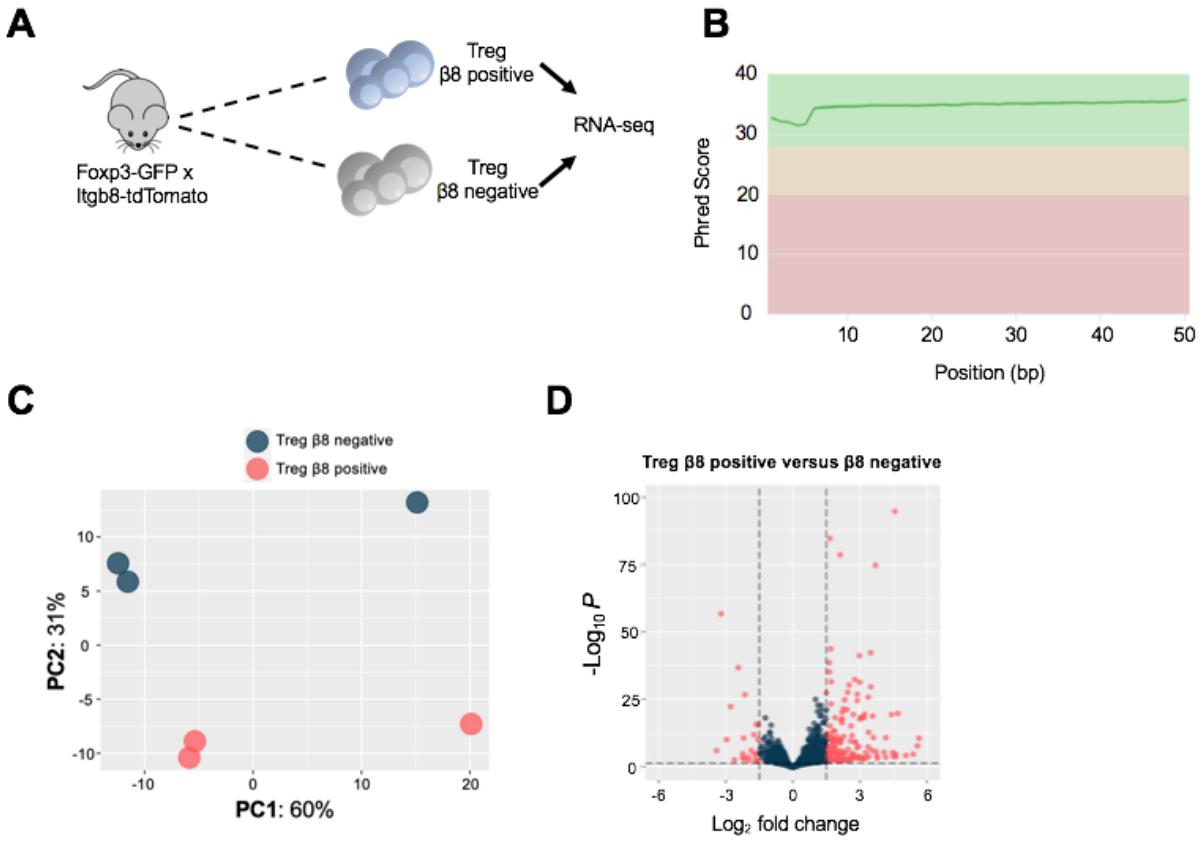
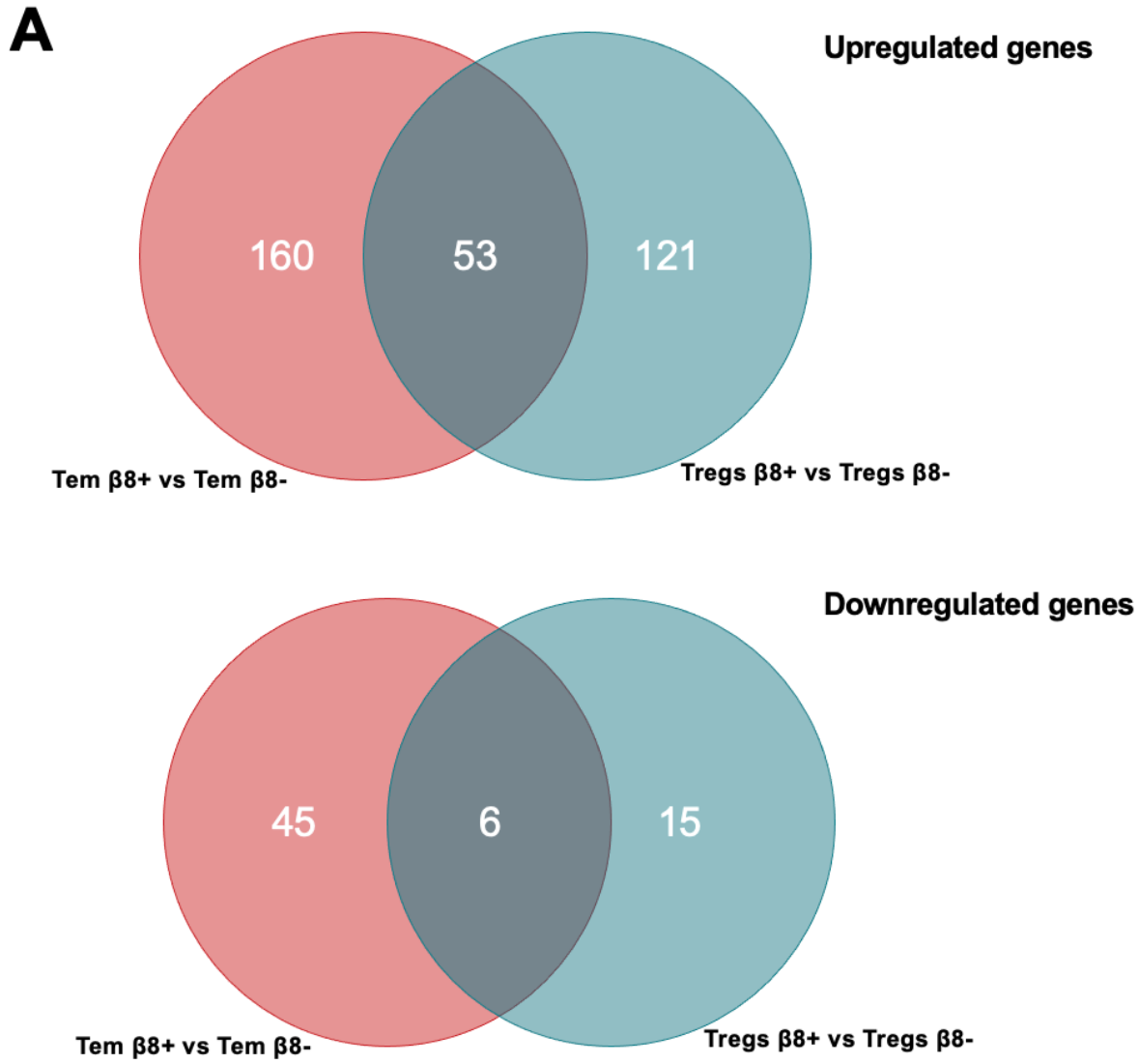


Figure 5.3 | Expression of β8 Treg correlates with a distinct transcriptional profile compared to β8 negative Treg. A) Schematic figure of the experiments. In brief, Treg were isolated from mLNs, in Foxp3-GFP x Itgb8-tdTomato. Then, extracted RNA was single-ended sequenced. B) Quality control of trimmed reads used for the downstream analysis. C) PCA plot of β8 positive and negative Treg. Blue dots are β8 negative Treg; pink dots are β8 positive Treg. PC 1 and 2 – explaining the major source of experimental variability – are on the X- and Y- axis, respectively. D) Volcano plot showing genes that changed significantly, in red, when β8 positive Treg are compared to β8 negative Treg. Log2 fold change values are shown on the X-axis and the P-values – transformed as the negative base 10 logarithm – are shown on the Y-axis. Vertical and horizontal lines are used to draw cut off. (N=3).

A list of genes differentially expressed from our analysis of Tem αvβ8+ versus Tem αvβ8- subsets was then intersected to the list generated when comparing Treg αvβ8+ versus Treg αvβ8- counterparts, to see if a common gene signature between the two experiments could be drawn-out. A log2 fold change higher or equal to 1.5 and P-value adjusted with the Benjamini Hochberg correction method less or equal to 0.05 was chosen for both the experiments as a significance cut off. The two-way Venn Diagrams indicated a gene overlap of around 22% in the Tem group versus the Treg one (Figure 5.4A). Furthermore, it is worth mentioning that all the shared genes showed a similar trend of regulation – i.e., upregulation or downregulation. This indicated that, for genes significantly altered in this experimental setting, the integrin αvβ8 is correlated with a similar expression pattern in both the cell subsets – as also indicated for the candidate genes selected (Figure 5.4B; N=3-6; ****, p adjust ≤ 0.0001). Specifically, the Itgb8 gene, encoding the β8 subunit, was upregulated in αvβ8+ Treg and Tem compared to their αvβ8- counterparts. These results indicate a correlation between β8 expression and its gene regulation in both cell subsets. Importantly, we found that expression of αvβ8 in Treg and Tem was linked with upregulation of genes involved in anti-inflammatory pathways or with an immune suppressive function. Also, Lag3 expression – described as an important immune suppressor (C.-T. Huang *et al.,* 2004) – was upregulated in both αvβ8+ Tem and Treg compared to αvβ8- ones. Importantly, Tigit was another key molecule significantly upregulated in αvβ8+ Treg and Tem relative to the αvβ8- counterparts. This molecule has been shown to regulate T cell immunity through inhibition of T cell proliferation and regulation of cytokine production (Joller *et al.,* 2014).


A
Upregulated genes
160
53
121
Tem β8+ vs Tem β8-
Tregs β8+ vs Tregs β8-
Downregulated genes
45
6
15
Tem β8+ vs Tem β8-
Tregs β8+ vs Tregs β8-

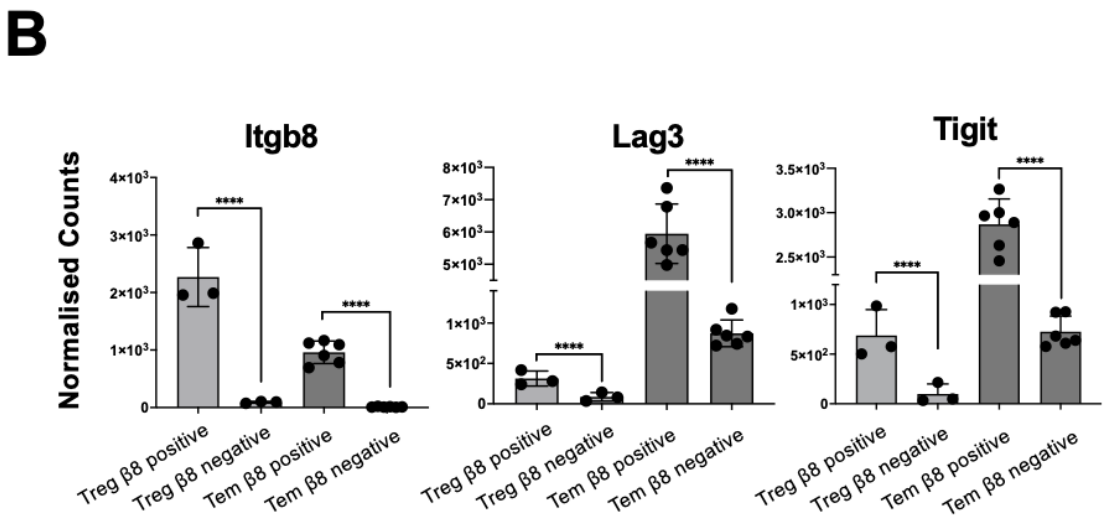

B
Itgb8
Lag3
Tigit
Normalised Counts
Treg β8 positive
Treg β8 negative
Tem β8 positive
Tem β8 negative
****

Figure 5.4 | β8+ Tem show a gene overlap of 22% when compared to β8+ Treg. A) DEA was performed with the DESeq2 package and genes were deemed as significant based on a log2 fold change cut off greater or equal to 1.5 and P value adjusted with the Benjamini Hochberg method less or equal to 0.05. Venn Diagrams show intersections of common up-regulated (upper panel) and downregulated (lower panel) genes significantly changed when Tem β8+ were compared to Tem β8- (pink circle) and Treg β8+ were compared to Treg β8- (light blue circle). B) Normalised counts generated in both the experiments for candidate genes selected. Asterisks refer to a difference that changed significantly (****, p adjust ≤ 0.0001) in the DESeq2 package, where Wald test was performed to calculate p values adjusted with the Benjamini-Hochberg correction. (N=3-6).

To gain a better insight into cell populations analysed and to avoid false observations due to likely batch effects, we decided it was a better option to repeat the RNA-seq analysis by including the four groups of interest in the same experimental design.

5.4 A new RNA-seq experiment indicates that αvβ8+ Tem mark a new, transcriptionally distinct, cell subset

To have a clear indication of whether expression of αvβ8 on Tem could entail a distinct phenotype, we decided to perform transcriptional profiling on Tem and Treg subsets isolated from the same location, from the same mouse. Therefore, primary Treg and Tem – each marked positive or negative for β8 and Foxp3 expression – were isolated from spleen of Itgb8 reporter mouse (Figure 5.5A), and sorted as single, live, TCRβ+, CD4+, CD8-, CD44Hi, CD62L-, CXCR5- (Figure 5.5B).

A quality check of trimmed fastq files increased overall sample quality, as shown by the phred score measured for each read in the experiment (Figure 5.5C). Moreover, overrepresented sequences – consisting of sequencing contaminants – made up less than 1% of trimmed samples compared to a higher contamination seen in the untrimmed samples (data not shown). Upon the pseudomapping step, a PCA plot was generated using the DESeq2 package (Figure 5.5D; N=3). Data indicated that all samples segregated from each other on the two major PCs, forming four clear clusters. This confirmed that expression of αvβ8 by Tem marks a transcriptionally distinct subset compared to αvβ8- counterparts and Treg. To address the overarching

goal of our project and thus directly compare population of Tem and Treg, expressing the integrin or not, DEA was performed for relevant comparisons, and genes dispersion represented in Volcano plots (Figure 5.5E).

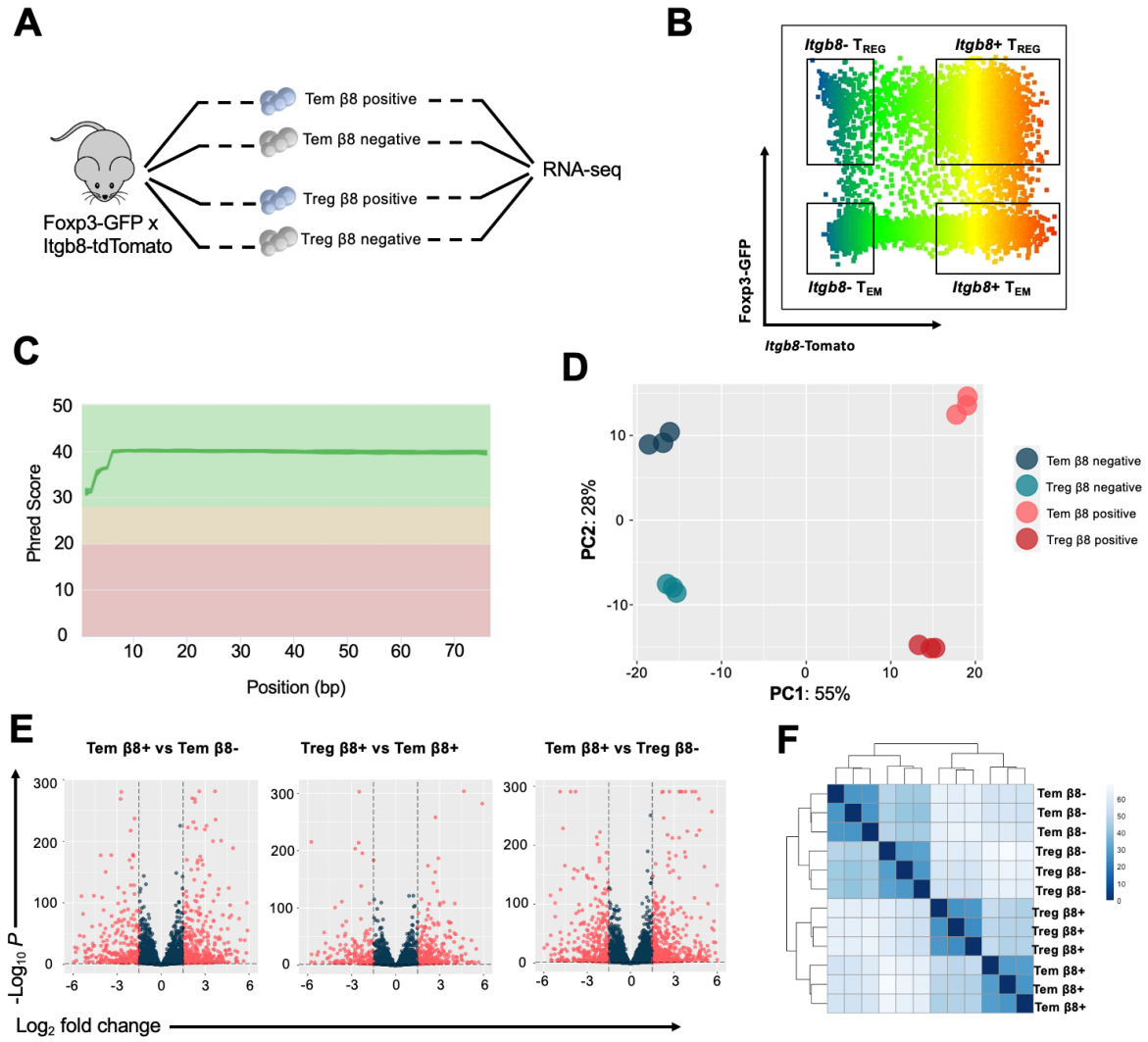


Figure 5.5 | RNA-seq experiment on a newly generated dataset indicates that expression of β8 on Tem or Treg triggers phenotypically distinct changes and that β8 positive Tem share key genes with β8 positive Treg. A) Schematic representation of the adopted experimental design. Briefly, Tem or Treg were extracted from Foxp3-GFP x Itgb8-tdTomato and RNA, extracted in one single batch, was paired-end sequenced. B) Representative FACS plot of isolated populations (credit: Dr Craig McEntee). C) Quality control of trimmed sequenced reads. Green areas indicate a phred quality score, on the Y-axis, greater than 30 – suggesting a correct base calling > 99.9%. On the X-axis is the base pair position of processed reads. D) PCA plot representing clustering of the four groups analysed. Pink and red dots are Tem and Treg β8 positive, respectively; blue and light blue dots are Tem and Treg β8 negative, respectively. PC 1 and 2 reflects the biggest variability in this experiment and are reported on the X- and Y-axis – respectively. E) Volcano plots representing pairwise comparisons of interest. Fold change values are represented on the X-axis and negative log transformed P values

are reported on the Y-axis. Red dots are genes differentially expressed. F) Sample to sample distance of log transformed counts processed with the regularised logarithmic transformation. Darker colour indicates a higher degree of similarity between samples analysed. (N=3).

As expected from PCA results and confirmed in the Volcano plots, the highest number of differentially expressed genes were seen when comparing Tem β8+ versus Treg β8- (Figure 5.5E) and Treg β8+ versus Tem β8- (data not shown), indicating that differences were the highest for these pairwise comparisons. These observations were also confirmed when measuring Euclidean distances measured on the log-transformed values to determine sample-to-sample correlation (Figure 5.5F). Of note, the sample-to-sample heatmap reported in figure 5.5F suggested that groups expressing the integrin αvβ8 were more transcriptionally related compared to samples who were marked negative for the latter receptor. Together, these results suggest that presence or absence of the integrin αvβ8 is required to segregate the four conditions analysed, each into a separate cluster.

5.5 Comparative analysis of Treg and Tem identifies regulatory genes enriched in β8+ subsets

In the light of these results, we then sought to address transcriptional differences between β8+ and β8- Treg. To achieve this goal, we first merged datasets from the Treg experiment generated at the Cancer Research Centre in Lyon (Figure 5.3) with β8+ and β8- Treg generated in our lab (Figure 5.5). In this way, we increased the power of our analysis while indirectly determining whether populations of Treg extracted from different sites – i.e., spleen and mLNs – were transcriptionally homogeneous. Indeed, batch-effect controlled PCA plot showed that Treg mainly separated based on the expression of the integrin αvβ8 (Figure 5.6A, N=6).

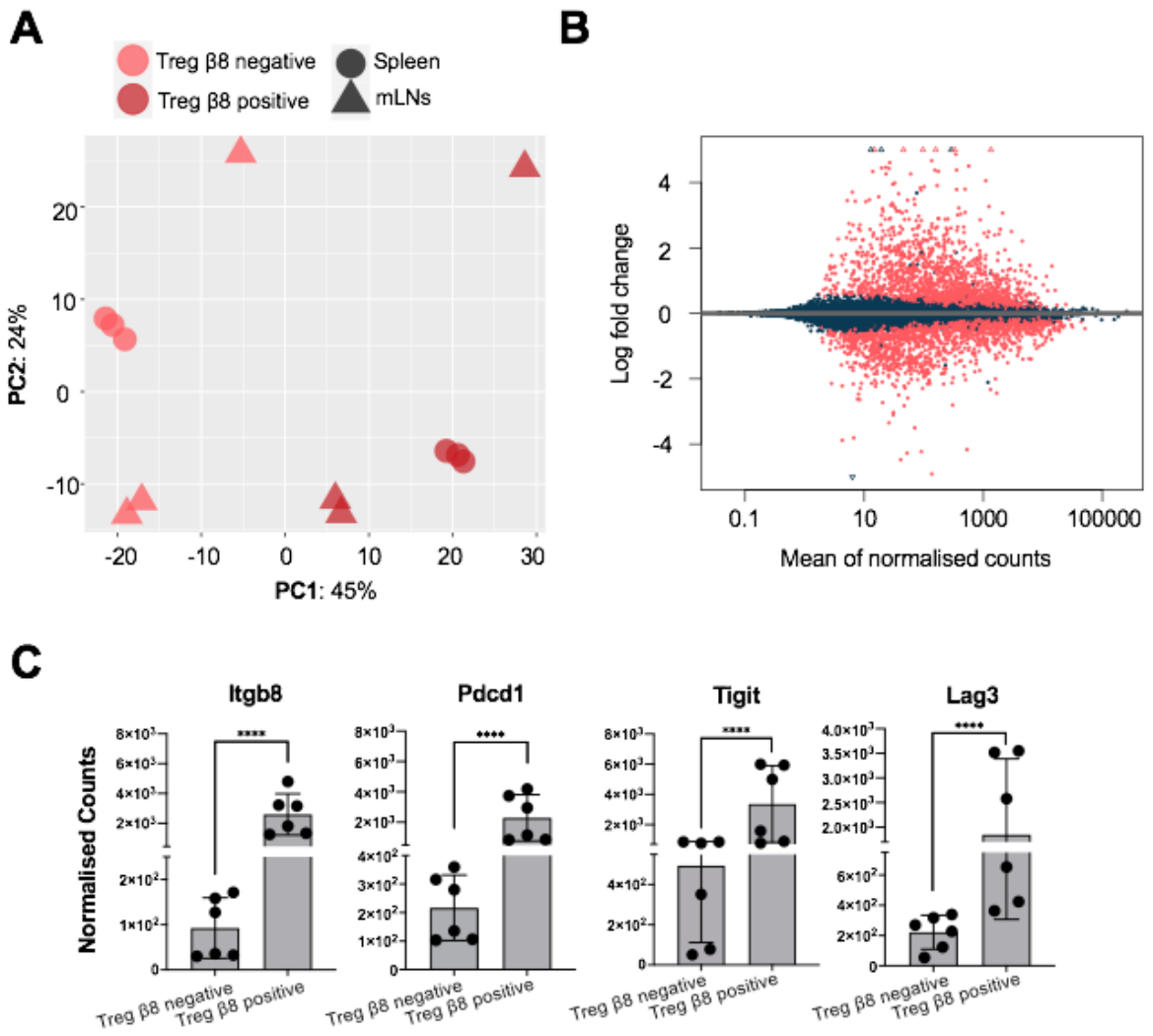


Figure 5.6 | Computational analysis performed on Treg, marked positive or negative for β8 expression, identifies potential pathways of interest. A) PCA plot of β8+ versus β8- Treg. On the X-axis is the experimental variance explained by the PC 1. On the Y-axis, is variance explained by PC 2. Circular and triangular shapes are Treg isolated from spleen and mLNs, respectively. B) MA plot, depicting gene distribution when β8+ Treg were compared to β8- counterparts. On the X-axis, is the mean of normalised counts; on the Y-axis is the log2 fold change generated with the DESeq2 package. Dots in red are genes differentially expressed. C) Normalised counts of selected genes, found to be differentially expressed with the DESeq2 package when the β8+ population of Treg was compared to β8- one. Asterisks mark a difference that reach significance with the Wald Test measured in DESeq2. (N=6; ****, p adjust ≤ 0.0001).

Furthermore, the MA plot indicated a tendency towards the upregulation for the medium expressed genes in the dataset when comparing Treg β8+ versus Treg β8- (Figure 5.6B). One of the most differentially expressed molecules in Treg β8+ compared to Treg β8- is the Itgb8 gene (Figure 5.6C). This suggests that, similarly to the Tem group (Figure 5.4B), transcription

of the latter gene is restricted to the β8+ subsets. Consequently, we also confirmed that regulatory targets genes, also seen in figure 5.4B, were also upregulated in the β8+ Treg compared to their β8- counterparts.

Next, to test transcriptional differences between αvβ8+ Tem and αvβ8- counterparts, we firstly measured concordance of differentially expressed genes between the two datasets generated in our lab (Figure 5.1 and 5.6). Differentially expressed genes, shared between the two subsets, showed a high degree of fold change correlation, especially for the upregulated ones (Figure 5.7A). Importantly, PCA plot of batch-controlled samples confirmed that Tem separated by their different expression of αvβ8 on the PC 1 – explaining the 59% experimental variation (Figure 5.7B, N=9).

Also, genes differentially expressed when comparing Tem β8+ versus Tem β8- were used to construct a matrix of regularised log transformed normalised counts represented as a heatmap (Figure 5.7C). Column dendrograms, confirmed results shown in figure 5.7B and defined two distinct groups identified by the presence or absence of the integrin αvβ8. Also, a heatmap showed three major clusters of gene expression.

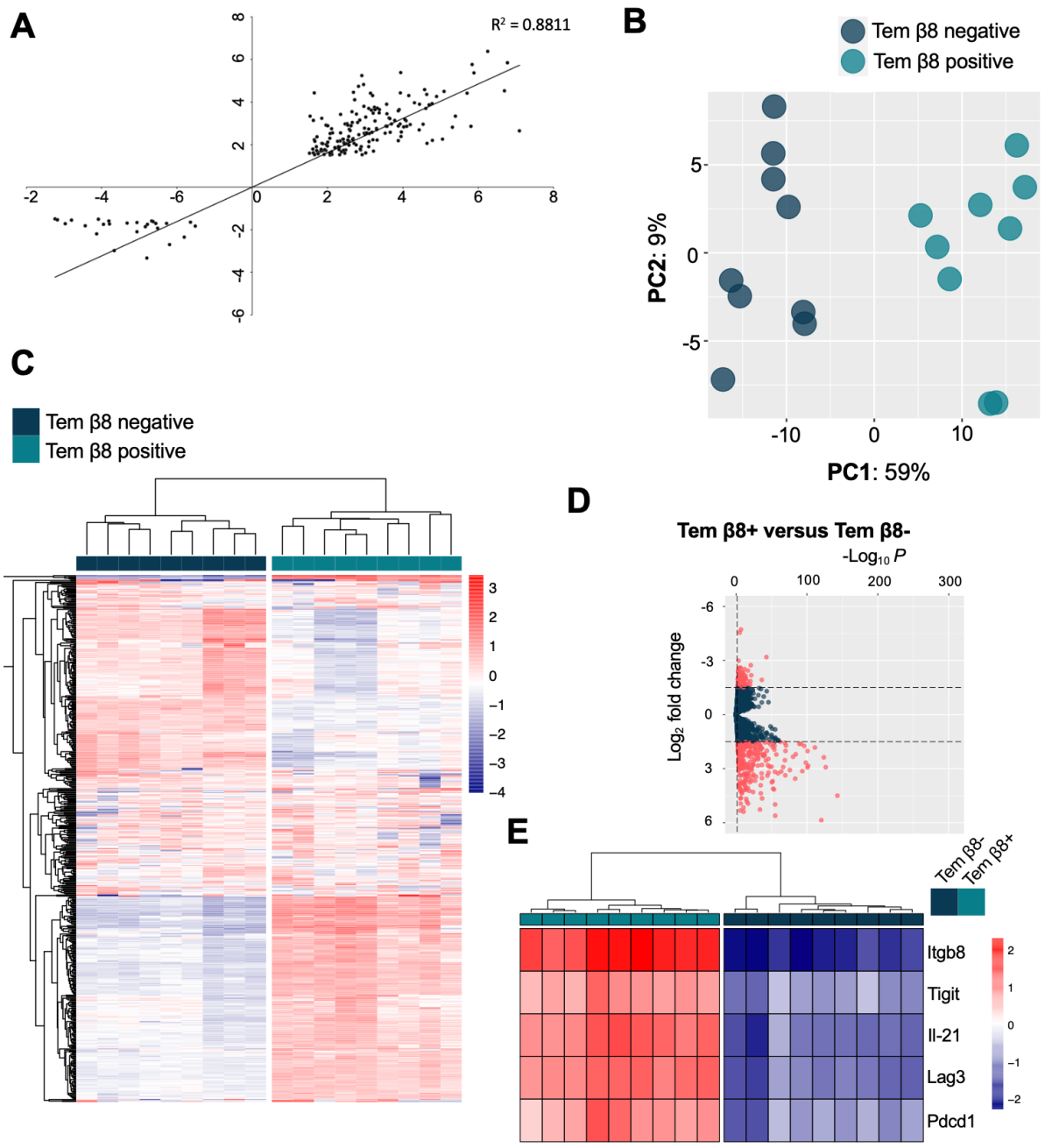


Figure 5.7 | Comparison between two separate experiments of Tem identifies core genes involved in important regulatory immunological activity. A) Log2 fold change correlation of common genes differentially expressed (log2 fold change cut off ≥1.5 and adjusted P value ≤0.05) in two separate experiments where the β8+ subpopulation of Tem is compared to the β8- counterpart. B) PCA plot of collated datasets, representing batch-corrected samples of β8 positive and negative Tem. C) Supervised hierarchical clustering of top 500 variable mean-centred, log-transformed normalised counts. Euclidean distance is calculated on normalised counts to draw column and row dendrograms. Red tiles are upregulated genes. D) Volcano plots representing genes differentially expressed in the merged dataset where transcriptional profile of batch effect corrected β8 positive and negative Tem is compared. On the X-axis, is the log2 fold change of the expression values. On the Y-axis, is the negative base 10 logarithm of the P value. Vertical and horizontal lines are used to visualise P value and log2 fold change threshold cut offs. E) Candidate genes of interest, found to be differentially regulated in β8+ Tem when compared to the β8- subpopulation. Results are represented as a

heatmap of regularised log transformed counts, on which Euclidean distance is calculated and used to draw column dendrograms (N=9).

The pattern of expression is in line with the Volcano plot, where the vast majority of genes changes measured in Tem β8+ were upregulated (Figure 5.7D). Also, we generated a heatmap of the top 5 differentially expressed genes in the above mentioned pairwise comparison and, in line with results showed in figure 5.4, we found that some of the greatest variations at the transcriptional level in this comparison involved genes encoding products involved in a regulatory function (Figure 5.7E). Success of this approach is that we integrated samples from different experimental batches to increase power of analysis while controlling for experimental variation, as indicated by segregation of samples in the PCA plot. Importantly, these results confirm previous observations according to which expression of αvβ8 marks transcriptionally distinct cell subsets, with an enrichment for expression of genes associated with an immune-suppressive function. Also, to gain better insight into biological processes that were significantly altered in β8+ Tem, we decided to perform further analytical techniques.

5.6 Pathway, Network and GO analysis of genes differentially expressed in αvβ8+ Tem identify important gene interactions

To attempt to identify potential biological processes that are associated to the expression of αvβ8 in CD4+ Tem, we sought to determine which networks and pathways were significantly enriched in this cell type. In brief, the overall scope of network and pathway analysis tool is to shrink data derived from hundreds of differentially expressed genes (Creixell *et al.,* 2015). In this way, sets of significantly altered pathways can be studied in a more interpretable way (Creixell *et al.,* 2015). To address which pathways, biological processes, and networks were significantly represented in β8+ Tem when compared to β8- ones, the outputs generated from the statistical analysis using the DESeq2 package were investigated further with the DAVID tool. This functional annotation tool returned a list of GO terms that were further analysed using the REViGO tool.

In brief, the REViGO tool aims to identify and remove redundant GO terms – i.e., annotations describing molecular function, cellular component, and biological processes of a given gene – based on the semantic similarity measure. Genes annotated with the same GO term will have similar function and semantic similarity. Structured as a graph, each GO term is conventionally, and logically related, to a parent and child term – i.e., reflecting a broader and a more specific annotation, specifically. Also, GO terms specify a domain of vocabulary terms that are computationally accessible. On the other hand, redundancy occurs when GO enrichment analysis is performed on differentially expressed genes, as many GO terms are related by inheritance. Therefore, a long list of redundant GO terms could be resolved by finding one GO term that best describe each gene cluster (Lord *et al.,* 2003; Supek *et al.,* 2011).

From the functional annotation analysis performed with the DAVID tool, using the list of genes differentially expressed when β8+ Tem are compared to β8- Tem as input, two types of categories of GO classes were chosen – referring to biological processes occurring at the molecular level (molecular function) and to their cellular role elicited in the cell (biological process). GO terms were processed with the REViGO tool and outputs analysed (Figure 5.8).

The most representative molecular function-related GO terms, which successfully passed the redundancy reduction step, indicated that gene products were involved mainly in protein binding (GO:000515), ion binding (GO:0043167), and signalling receptor activity (GO:0038023). Classes involved in cytokine activity (GO:0005125) and interleukin-17 receptor activity (GO:0030368) were also enriched (Figure 5.8, upper panel). On the other hand, the number of GO terms involved in biological processes, parsed by the REViGO tool, were higher. Generic GO terms involved in the inflammatory response (GO:0006954) and immune responses (GO:0006955) were found to be enriched in the REViGO tool, together with a more specific – yet less frequent – positive regulation of interleukin-10 production (GO:0032733). The latter class is in line with the positive regulation of MAPK cascade (GO:0043410) and upregulation

of Il-21 seen in DEA – as shown in figure 5.7E, might support the hypothesis of a shared regulation (Spolski & Leonard, 2014).

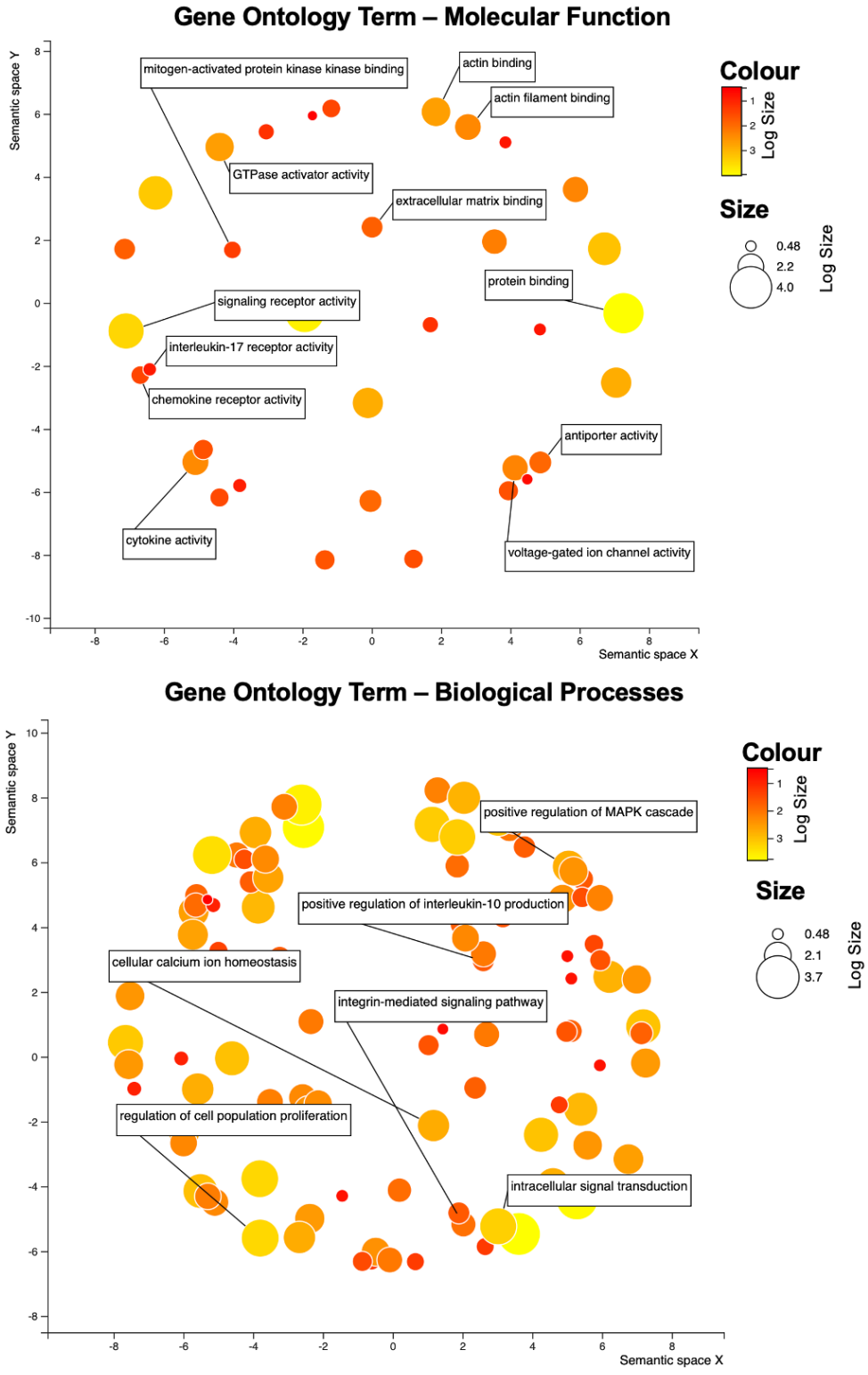


Figure 5.8 | Functional annotation analysis with the DAVID tool, after applying the REVIGO tool, in β8+ Tem compared to β8- counterparts. List of genes differentially expressed generated with the DESeq2 package were used as input for the DAVID tool, retuning list of enriched GO terms. After the analysis, list of entries from 'Molecular Function' (upper panel) and 'Biological Process' (lower panel) classes were chosen as inputs for the REViGO tool. In a bidirectional graph, the plots represent major clusters of the most representative GO terms, in which distances between each term is designated by the semantic similarity, on the X- and

Y-axes. Colour is the p-value generated from the statistical analysis performed with the DAVID tool. The size of the bubbles represent how generic is the GO term; the larger is the bubble the more general is the term.

Also, we wanted to test whether functional interaction of gene products found to be differentially expressed in β8+ Tem versus β8- Tem could be anticipated by analysing networks with an R implementation of the GiGA tool. Briefly, the GiGA network analysis tool, applying the iterative Group analysis method (Breitling *et al.,* 2004), identified 29 subnetworks (data not shown). As expected, one of the most relevant networks annotated was enriched for the Itgb8 gene – found to be the most significantly changed gene when β8+ Tem were compared to β8- Tem (Figure 5.9A). Of note, the latter subnetwork contained genes significantly overexpressed in Tem β8+ versus Tem β8- belonging to the collagen family of genes. Of note, Adamts3 and Itga genes – identified already when using the DAVID and the REViGO approaches (data not shown) – were also identified by the GiGA algorithm when comparing Tem β8+ versus Tem β8-. Adamts3 was significantly enriched, given its role in collagen processing (Fernandes *et al.,* 2001). Also, the Itga1 gene – encoding the α1 integrin subunit – has been described to heterodimerise with the β1 subunit to bind collagen, as a cell-surface receptor (Nunes *et al.,* 2018). In Tem, these findings might anticipate potential protein-protein interactions between integrin αvβ8 and molecules involved in collagen-related processes.

Likewise, we identified two other networks to be regulated in Tem β8+ versus Tem β8-, by grouping functional roles of Tigit, Lag3 and Pdcd1 – all upregulated in Tem β8+ (Figure 5.9B, upper panel) and IL17-related genes – in which expression of the gene encoding the cytokine was upregulated as opposed to their cognate receptors (Figure 5.9B, lower panel).

We also determined possible upstream regulators by using the IPA tool, given the list of genes generated with the DESeq2 package comparing β8+ versus β8- Tem. In brief, IPA, based on the list of differentially expressed genes, identifies clusters of genes which are potentially regulated by a common upstream regulator. It is worth mentioning that the upstream regulator is

not necessarily part of the list used as an input. Nevertheless, we found that – amongst different candidates found – the Pdcd1 and IL10 genes, transcription of which was positively upregulated in the β8+ Tem group, were enriched in this analysis (Figure 5.9C).

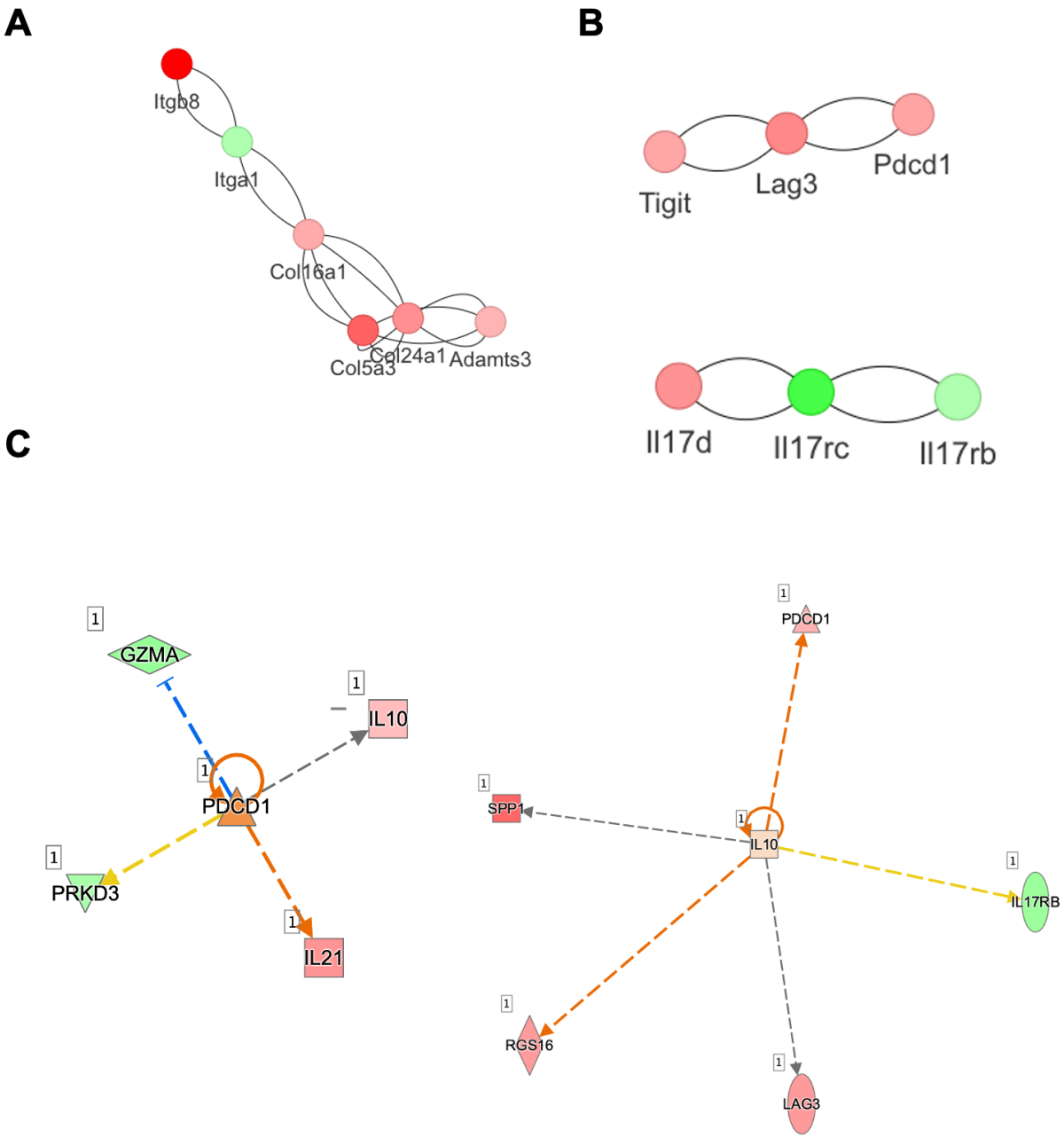


Figure 5.9 | GiGA analysis and upstream analysis with the IPA tool identifies potential gene interactions that are enriched when β8+ Tem are compared to β8- ones. A and B) Three representative subnetworks identified with the GiGA tool method where gene clusters are identified based on the mouse gene length, STRING interactome, and the input list generated with the DESeq2 package. Red edges are upregulated genes. Reciprocal interaction

is designated with a double-interaction node. C) Upstream analysis performed with the IPA tool. In the centre of the selected graphs are the upstream regulators. The colour of a predicted upregulation is darker for the PDCD1 gene (left panel) than IL-10 (right panel). Green denotes a gene found to be downregulated in the DESeq2 analysis. Square-shaped components are cytokines. Diamond-shaped genes are enzymes. Oval-shaped molecules are transmembrane receptors. Down- and up-pointing triangles are kinases and phosphatases, respectively. Dotted lines refer to an indirect interaction between the upstream regulator and a specific molecule. In IPA, a predicted relationship is coloured in orange or blue if the interaction leads to an activation or inhibition, respectively. Yellow connections refer to a discrepancy between the expected and measured state of the downstream molecule. If the effect cannot be predicted, based on the knowledge base IPA dataset, the connection is coloured in grey.

Interestingly, IPA results (Figure 5.9C) showed that selected upstream regulators – significantly enriched in the pathway analysis while being differentially regulated in β8+ versus β8- Tem, too – indirectly interacted with their downstream molecules identified. Therefore, IPA analysis suggested that other unidentified molecules could potentially be involved in these interactions.

In conclusion, this analysis generated pathways of interest that can be further interrogated experimentally. We found that genes involved in driving important regulatory functions in Tem β8+ were altered when compared to β8- counterparts, by using different pathways and network analysis tools. The importance of these observations confirmed that the αvβ8+ Tem group is a transcriptionally distinct subpopulation identified and enriched for specific regulatory pathways. These results complement experimental validation carried out in our lab.

5.7 Discussion

5.7.1 The expression of αvβ8 marks a transcriptionally distinct population of Tem

Our lab showed that the expression of αvβ8 plays a key immune-regulatory role in DC, Treg, and monocytes – in human and mouse (Fenton *et al.,* 2017; Kelly *et al.,* 2017; Travis *et al.,* 2007; Worthington *et al.,* 2015). Preliminary findings in our lab indicated that a subset of CD4+ Tem expressed αvβ8 and activated the latent TGF-β. Biological importance of this was found *in vivo* where mice lacking the expression of αvβ8 on CD4+ Tem, upon a secondary influenza infection, showed a vigorous CD8 response leading to a greater viral clearance. In such a

mouse model, transfer of the αvβ8+ CD4+ Tem population completely rescued the phenotype indicating that these cells have the potential to restrain anti-viral immunity. To better understand the effect elicited by the integrin αvβ8 in this cell population, transcriptome profiling was performed, in our lab. Bioinformatics analysis showed that expression of αvβ8 is correlated with a different transcriptional profile to when a Tem population does not express the αvβ8. RNA-seq indicated that β8+ versus β8- Tem were transcriptionally dissimilar, suggesting that expression of αvβ8 was crucial to segregate samples analysed in two distinct subsets. This confirms that expression of αvβ8 is not ubiquitous but rather compartmentalised to a specific subset of cells that confer uniquely distinct properties, as seen *in vivo*.

5.7.2 αvβ8+ Tem are transcriptionally distinct from Treg

In our lab, *in vivo* studies indicated that expression of αvβ8 on Tem marked an immune suppressive phenotype. As previously mentioned, our results showed that β8+ Tem were able to restrain the anti-viral response upon secondary influenza infection (unpublished data, Credit: Craig McEntee). Importantly, this process occurred independently of Treg as we did not see the same phenotype in Foxp3-cre mice. These observations led to the hypothesis that β8+ Tem could play a role via activation of TGF-β, and therefore sharing similarities with Treg as showed by results published in our lab. Therefore, a first attempt to compare transcriptional profile was done through an intersection of our dataset with one generated elsewhere. Indeed, standardising different datasets with the same established pipeline – e.g., using the same adapter trimming and same reference transcriptome version during the pseudoalignment – we determined genes in common across the two lists. This suggested that, despite phenotypically distinct groups of cells, expression of αvβ8 correlated with a transcriptional pattern that was retained in Tem and Treg. The caveat of this observation was that samples were from different experiments and sequenced in two different ways (single-end versus paired-end) – which was not ideal for robust observations.

Therefore, we tried to use several bioinformatics tools and apply algorithms to identify which genes were uniquely expressed in each experimental group, in this experimental setting (data not shown). Given that the amount of technical bias could not be ruled out when samples from different groups are analysed in different batches, we decided was a better option to analyse samples from the same experiment thus generating a new RNA-seq library.

This strategy allowed us to directly measure cell clustering using the same library preparation. Noteworthy, preparation of another sequencing library, revealed that despite β8+ Treg and β8+ Tem shared regulation of a number of gene transcripts such as IL10, Tigit, and Lag3, these cells were markedly distinct. This led to the assumption that these cells might not share a common origin but rather that αvβ8 induce expression of the same transcripts in these different T cell subsets. Indeed, transcriptome profiling performed on the top variable genes in these experiments and PCA confirmed that the Tem group, expressing αvβ8, marks a completely distinct phenotype from αvβ8- Tem. This indicated that expression of αvβ8 commits cells to a pathway of differentiation that is potentially exclusive for the cell subsets analysed.

Thus, αvβ8 marks a transcriptionally different Tem subtype from β8- counterparts and Treg but do share some similarity to β8+ Treg, with some shared expression of genes that have a crucial regulatory role in modulating immune responses. It is worth noting that, apart from the immune suppressive phenotype induced by the enhanced expression of the anti-inflammatory cytokine IL10 and functional experiments carried out by colleagues in the lab, there is an evident T cell activation signature in the αvβ8+ Tem subset as indicated by the expression of co-inhibitory receptors (Joller & Kuchroo, 2017) and tumor necrosis factor superfamily members (Jones *et al.,* 1999).

5.7.3 αvβ8+ Tem have an enhanced expression of important genes involved in suppressor activity

We applied several bioinformatic analysis methods to determine which genes were differentially expressed when β8+ Tem were compared to β8- counterparts.

Standardisation of the two different datasets – i.e., using the same adapter trimming strategy, pseudoalignment to the same reference transcriptome, and batch effect correction approaches – allowed to increase the power of our experiment by merging data and perform statistically robust DEA. Supporting experimental results, thank to different pathway and network analysis tools used, we identified that segregation of the β8+ Tem as a novel subset was linked to regulatory pathways that need to be tested functionally.

Cross-comparing different computational approaches in our experiment, we were able to identify in the β8+ Tem population that several genes – regulated in the same direction within the Treg dataset when the β8+ group was compared to the β8- one – were involved in eliciting a suppressive function. Specifically, our analysis identified Lag3, Tigit and PD1 as significantly upregulated and functionally enriched.

Importantly, IL-21 was another gene significantly upregulated in Tem β8+ compared to β8- Tem, the expression of which was enhanced in Treg β8+ compared to the β8- counterparts. Also, GO enrichment analysis and the REViGO tool identified a positive regulation of the MAPK cascade, with cognate molecules significantly enriched. This, together with the increased expression of IL-10, could potentially indicate that in β8+ Tem, an intracellular pathway may link the upregulation of IL-21 with the increased expression of IL-10, as indicated by previous works (Spolski & Leonard, 2014). Indeed, as described by Spolski and colleagues, IL-21 elicited an immune-suppressive role by increasing the expression of IL-10, in TCR-stimulated naïve and activated murine T cells (Spolski *et al.,* 2009). Importance of these findings may help clarifying preliminary observations in our lab, correlating the ability of β8+ Tem in modulating an anti-viral immune response. Further work is needed to validate functional importance of this.

Therefore, the novelty of these findings could potentially pave to way to new validation experiments that could be carried out in the light of our results. For example, a Western blot analysis could be performed to study which MAPK signalling molecules are regulating such processes – studying the phosphorylation pattern of the latter ones. Furthermore, a cytokine assay for IL-21 and IL-10 could be carried out in β8+ Tem to verify if changes can be measured at protein level, too. Also, it would be interesting to investigate αvβ8-associated gene modules in different cell subsets, as the one studied in our project.

5.8 Conclusion

In this chapter, applying different computational methods, we provided evidence that the expression of αvβ8 on Tem is linked to a unique distinctive phenotype when compared to their β8- counterparts and Treg – regardless of their expression of the integrin αvβ8. Nevertheless, we found that genes linked to a regulatory phenotype were expressed in a similar way in both β8+ Treg and β8+ Tem.

These results indicate and corroborate previous experimental findings in our lab according to which the β8+ Tem subset is a markedly distinct subset of cell showing a regulatory function.

Chapter 6

General Discussion

The last 20 years has witnessed an increased understanding of how integrin αvβ8 regulates biological processes. Studies have clarified the role played by integrin αvβ8 in orchestrating vascular morphogenesis, with loss of β8 correlated to brain angiogenesis alterations in mice (Jiangwen Zhu *et al.,* 2002). The β8 cytoplasmic domain is involved in the regulation of the Rho small GTPases in glioblastoma cells (Reyes *et al.,* 2013). Likewise, developmental defects of the central nervous system associated with the αvβ8 mutant mice are linked with abnormal activation in Rho GTPase (Shin *et al.,* 2015). As reviewed by McCarty, whether an signalling pathway is propagated upon latent TGF-β binding is an appealing, but yet undescribed, research question (McCarty, 2020). Moreover, studies have confirmed the crucial role of αvβ8 in regulating immune homeostasis in mice and humans. Nonetheless, as previously mentioned, β8 is structurally and functionally different from other integrin molecules (McCarty, 2020). Therefore, its biological activity, following ligand engagement, was not described in immune cells, such as DC – on which β8 is highly expressed. Thus, in this thesis, we hypothesised that, through the use of high-throughput sequencing methods, we could describe a previously undefined role of prospective pathways regulated by αvβ8 following engagement with the latent TGF-β.

Previous study in our lab defined an existing correlation between loss of αvβ8 and the reduced ability of DC and Treg in regulating TGF-β activation – resulting in an increased inflammation in different settings (Fenton *et al.,* 2017; Travis *et al.,* 2007; Worthington *et al.,* 2015). The importance of these findings led to the hypothesis that, in parallel with TGFβ activation, the integrin αvβ8 may propagate cellular signaling cascades inside the DC. Before our studies, whether the integrin αvβ8 could elicit an intracellular signalling, upon ligand engagement, in immune cells was not determined. Yet, in other cell systems, the αvβ8 signalling pathway has been shown proposed to propagate important signals.

Recent work by Pollen *et al*, indicated that, in human, αvβ8 expression is activated in radial glia, indicating regulation of neural stem cell development in organising and structuring the human brain. Indeed, the gene encoding β8 subunit is highly transcribed in human brain organoids compared with organoids of non-human primates (Pollen *et al.,* 2019). In 2015, Lee *et al.,* identified that the tyrosine phosphatase PTP-PEST – described to elicit an important role in cell motility (Espejo *et al.,* 2010) – interacts with the cytoplasmic tail of the β8 subunit promoting the dephosphorylation of a dissociation inhibitor that modulates the activation state of small GTPases downstream (Lee *et al.,* 2015). Moreover, as described by Chen and colleagues, αvβ8 is important in promoting the glioblastoma cell invasion. Indeed, signalling via PTP-PEST and the valosin-containing protein, an ATP-dependent enzyme, has been linked to a much quicker turnover of focal adhesions at the protruding edge of the cell (Z. Chen *et al.,* 2018). In the light of these and other findings, we sought to determine biological pathways that were significantly altered upon αvβ8 engagement in DC. One way to address the biological question is the transcriptome profiling methodology. We demonstrated here that that LAP, at 5 µg/mL, induced only minor changes to murine DC. Importantly, abovementioned signalling pathways were not found in our experimental setting – possibly indicating that different signalling functions of αvβ8 exist in different cell types.

Our lab and others identified increased integrin αvβ8 expression on DC upon external stimuli in the intestinal environment (Boucard-Jourdin *et al.,* 2016; Fenton *et al.,* 2017; Païdassi *et al.,* 2011). Specifically, as indicated by Païdassi *et al.,* integrin expression is the highest in CD103+ DC. Therefore, the first attempt to address the overarching scope of this thesis, as described in Chapter 3, was studying a previously generated database in our lab containing CD103+ DC obtained from the gut-draining LNs, incubated with LAP or not, in cells expressing the β8 subunit or not (credit: Eleanor Sherwood). Given the high sample variability, possibly due to technical artefacts, we decided to generate a new dataset, as discussed in Chapter 4. This time, primary DC were incubated with LAP at two time points – with the hope to determine early or late gene expression changes within this timeframe. RNA-seq findings indicated that

a time-dependent segregation, regardless of any major effects induced by LAP were induced. Results identified in Chapter 3 and 4, that, in these conditions, the integrin αvβ8 do not propagate a biologically important signalling downstream when driving the activation of TGF-β, in primary DC.

To date, αvβ8 has been mainly identified for its crucial role in regulating TGF-β activation as discussed in different contexts, such as cancer, fibrosis, and immune homeostasis. For example, in glioblastoma – a brain tumour (Lathia *et al.,* 2015) – the integrin αvβ8, regulating the activation of TGF-β, induces vascular growth of tumour cells. Specifically, β8 expression was directly correlated with the invasiveness phenotype of glioblastoma cells (Reyes *et al.,* 2013; Tchaicha *et al.,* 2011). This and other studies confirmed that the upregulation of αvβ8 correlated to an enhanced TGF-β signalling associated with a higher degree of severity (McCarty, 2020). Markedly, Dodagatta-Marri and colleagues found that αvβ8, expressed on CD25+ CD4+ tumour cells, is correlated with tumour progression. In this context, inhibition of anti-tumour immunity was due to the αvβ8-mediated TGF-β activation (Dodagatta-Marri *et al.,* 2021). Expression of αvβ8 has also been associated with fibrosis. For example, in the lung, several pathways such as the transcription factor SP3 and AP1, have been described to stimulate expression of Itgb8 encoding the β8 subunit (Markovics *et al.,* 2010, 2011). Also, the integrin-mediated TGF-β activation, if upregulated, was associated with chronic obstructive pulmonary disease (Araya *et al.,* 2007). These findings indicated the pivotal importance elicited by the integrin αvβ8 in modulating, in different contexts, the activation of TGF-β. Moreover, our lab identified that maintenance of the immune homeostasis, in mice and humans, is also regulated by the integrin αvβ8 – through its role in promoting the activation of TGF-β (Fenton *et al.,* 2017; Kelly *et al.,* 2018; Travis *et al.,* 2007; Worthington *et al.,* 2015). Therefore, we wondered whether a subset of memory T cells – known to express high levels of integrin αvβ8 – might retain anti-inflammatory properties.

In this thesis, as shown in Chapter 5, we sought to determine whether the expression of β8 on Tem marked a uniquely distinct phenotype. Interestingly, our lab identified an expansion of influenza-specific CD8+ T cells, upon a secondary memory recall response, in mice with functional loss of αvβ8 on CD4+ T cells. These results indicated that targeting the integrin αvβ8, suppressing its role in promoting TGF-β activation, can enhance immune system in counteracting viral infections (Craig McEntee, unpublished data). Interestingly, our results are in line with recent work from Dodagatta-Marri and colleagues indicating that suppression of αvβ8 is linked to an enhanced expression of a pattern of genes enriched in anti-tumour activity. Importantly, results in our lab indicated that this expansion was not dependent on Treg as mice lacking β8 expression on these cells showed no phenotype (Craig McEntee, unpublished data). Markedly, we saw that a subpopulation of CD4+ Tem can activate TGF-β and this was dependent by high expression of αvβ8 and that β8 expression was on a subset of the CD4+ Tem population (Craig McEntee, unpublished data). Moreover, our bioinformatic analysis identified a transcriptionally distinct population marked by the expression of β8 on CD4+ Tem, associated with immune-suppressive genes differentially regulated. Indeed, transcriptome analysis was consistent with the ability of β8+ CD4+ Tem, but not the β8- counterpart, to suppress influenza-specific expansion of CD8+ T cells seen in mice lacking αvβ8+ on CD4+ T cells (Craig McEntee, unpublished data). Importance of these findings indicated a previously undetermined pathway that could be targeted therapeutically to strategically enhance immune responses against respiratory viruses.

Also, an elegant study conducted by Kitamura and colleagues indicated that conditional deletion of integrin αvβ8 prevented adenovirus-dependent airway inflammation. Specifically, integrin inhibition correlated with a markedly reduced recruitment of DC to the lung due to a TGF-β dependent expression of CCL2 and CCL20 (Kitamura *et al.,* 2011). This highlights the role of αvβ8 as a candidate target in the treatment of allergic asthma. All these works indicate

that the integrin αvβ8 has emerged as a more refined alternative to TGF-β blockade for the treatment of several ailments.

Although integrin αvβ8 has become a therapeutic target for the treatment of several ailments, such ambitious goals require further work. Indeed, refinement of leading-edge omics technologies, together with their cost reduction, will make it possible to characterise and analyse ECM protein ligands and αvβ8-related intracellular molecular partners. For example, harnessing RNA-seq techniques looking at the single cell level would allow the identification of αvβ8-dependent phenotypes seeded in different tissues and likely discover the impact of the tissue microenvironment in regulating such subtypes. Another important milestone in defining the role of αvβ8 in several biological settings would be determining which αvβ8-dependent signalling molecules are shared with other signalling pathways. Determination of cross-signalling interactions would certainly lead the way to new translational studies – in which altering such pathways could result in a clinical amelioration of αvβ8- or TGF-β-related diseases.

In this thesis, we attempted to address important immunological question relying on RNA-seq. Contrary to a mere static analysis performed by whole-genome sequencing techniques, transcriptome profiling has the great advantage of determining dynamic changes in the gene expression pattern (Whitley *et al.,* 2016). Microarray analysis would have been a viable alternative as a preliminary approach to address pathways regulated and regulated by the integrin αvβ8. Microarrays have been extensively used, especially in the past, to determine gene expression using DNA probes. Nevertheless, RNA-seq techniques are preferred over microarray analysis for different reasons. One major difference between RNA-seq and microarray analysis is that the former is genome- or transcriptome-wide, whereas microarrays have to be targeted against a subset of transcripts. On the other hand, microarrays are a much more affordable option compared to cost required to run an RNA-seq experiment – which is nowadays still prohibited for some labs.

Importantly, we could not anticipate which transcriptional differences may arise after treating DC with LAP or which transcripts were significantly altered in Tem versus Treg. Therefore, we decided it was better to perform bulk RNA-seq as a viable option to address these questions. Indeed, sequencing of bulk cell population has been extensively used in different biological contexts, such as the immune system (See *et al.,* 2018). As reviewed by See and colleagues, the major caveat of this approach is that bulk RNA-seq does not take into account existing variability in gene expression changes at the single-cell level. Therefore, relevant biological information could be averaged by such cell variability if analysis is performed with bulk RNA-seq (See *et al.,* 2018).

Indeed, previous studies have identified that αvβ8 expression is preferentially found on DC from mucosal sites (Païdassi *et al.,* 2011) and specifically on CD103+ CD11b- DC (Boucard-Jourdin *et al.,* 2016). Then, one could explore which subsets of DC have the highest response upon ligand engagement with LAP. Therefore, to best study cells expressing high level of integrin – after αvβ8 ligand engagement – one could rely on more advanced sequencing approaches looking at the transcriptional profile at the single-cell level (See *et al.,* 2018). In this way, one could study the effect of LAP on a single cell lineage – i.e., cDC1, described to increase β8 expression levels upon TGF-β stimulation, RA, and TLR stimulation (Boucard-Jourdin *et al.,* 2016). As discussed in this thesis and reviewed by Perkel in 2017, one of the major problems in sequencing and in single-cell RNA-sequencing (scRNA-seq) is the batch effect – in which identical cells might appear different simply because these are generated on different days. Another drawback of scRNA-seq techniques is the lack of a unified pipeline, which is commonly accepted by the scientific community to address big questions in science (Perkel, 2017). Moreover, the cost required to proceed with scRNA-seq are, to date, very expensive, making this approach still underutilised in many research environments.

Furthermore, chromatin accessibility assays are being increasingly used in biomedical research. One example of these approaches is the assay for transposase-accessible

chromatin with sequencing (ATAC-seq) (Buenrostro *et al.,* 2015). The ATAC-seq sequences accessible regions of open chromatin across the whole genome (Buenrostro *et al.,* 2015). In this way, ATAC-seq could determine which parts of chromatin impact gene expression changes (Buenrostro *et al.,* 2015). Therefore, to gain better insight into gene regulatory networks enriched significantly upon LAP-αvβ8 interaction in DC, gene expression changes determined by scRNA-seq may also be coupled to chromatin accessibility assays, such as the ATAC-seq. In this way, integrating these two approaches, would not only be possible to determine, in the appropriate cell responders, downstream signalling pathways propagated downstream of the integrin αvβ8, but also unveil crucial regulatory elements acting on a certain gene (Perkel, 2021). These combined approaches would also be useful to determine molecular information underpinning phenotypic differences seen in β8+ Tem versus β8- counterparts and Treg.

Another important limitation of transcriptome analysis is that transcript levels correlate with proteins encoded in a variable way (de Sousa Abreu *et al.,* 2009; Maier *et al.,* 2009). To study the biological flow of cell information, from DNA to proteins, RNA-seq may be coupled to translatome analysis – i.e., the analysis of the RNA fractions recruited to ribosomes and subjected to protein synthesis – and proteomics. Profiling the transcriptome and translatome would be crucial to understand how gene expression is regulated in selected cells (Smircich *et al.,* 2015). Therefore, analysing the fraction of RNA recruited to ribosomes may help elucidating key biological processes in a given experimental setting such as the intracellular pathways associated that are transcribed and translated upon ligand engagement with the integrin αvβ8.

This thesis aimed to unravel novel mechanisms by which the integrin αvβ8 regulates gene expression changes downstream and why it marks a phenotypically distinct phenotype on Tem, by using next-generation sequencing approaches. Future efforts are needed to clarify the functional importance of these findings and identify potential novel target molecules that

could be altered in disease conditions, such as chronic inflammation. This with the final aim to translate such findings into humans for future prospective therapeutic approaches.

References

Ahmadzadeh, M., & Rosenberg, S. A. (2005). TGF-β1 Attenuates the Acquisition and Expression of Effector Function by Tumor Antigen-Specific Human Memory CD8 T Cells. *The Journal of Immunology*, *174*(9), 5215 LP – 5223. https://doi.org/10.4049/jimmunol.174.9.5215

Andrews, S., Krueger, F., Seconds-Pichon, A., Biggins, F., & Wingett, S. (2015). *FastQC. A quality control tool for high throughput sequence data. Babraham Bioinformatics*. Babraham Institute.

Annes, J. P., Chen, Y., Munger, J. S., & Rifkin, D. B. (2004). Integrin alphaVbeta6-mediated activation of latent TGF-beta requires the latent TGF-beta binding protein-1. *The Journal of Cell Biology*, *165*(5), 723–734. https://doi.org/10.1083/jcb.200312172

Annes, J. P., Munger, J. S., & Rifkin, D. B. (2003). Making sense of latent TGFβ activation. *Journal of Cell Science*, *116*(2), 217–224. https://doi.org/10.1242/jcs.00229

Araya, J., Cambier, S., Markovics, J. A., Wolters, P., Jablons, D., Hill, A., Finkbeiner, W., Jones, K., Broaddus, V. C., Sheppard, D., Barzcak, A., Xiao, Y., Erle, D. J., & Nishimura, S. L. (2007). Squamous metaplasia amplifies pathologic epithelial-mesenchymal interactions in COPD patients. *The Journal of Clinical Investigation*, *117*(11), 3551–3562. https://doi.org/10.1172/JCI32526

Arnaout, M. A., Mahalingam, B., & Xiong, J.-P. (2005). INTEGRIN STRUCTURE, ALLOSTERY, AND BIDIRECTIONAL SIGNALING. *Annual Review of Cell and Developmental Biology*, *21*(1), 381–410. https://doi.org/10.1146/annurev.cellbio.21.090704.151217

Askari, J. A., Buckley, P. A., Mould, A. P., & Humphries, M. J. (2009). Linking integrin conformation to function. *Journal of Cell Science*, *122*(2), 165–170. https://doi.org/10.1242/jcs.018556

Assoian, R. K. (1997). Anchorage-dependent cell cycle progression. *The Journal of Cell Biology*, *136*(1), 1–4. https://doi.org/10.1083/jcb.136.1.1

Atherton, P., Stutchbury, B., Jethwa, D., & Ballestrem, C. (2016). Mechanosensitive components of integrin adhesions: Role of vinculin. *Experimental Cell Research*, *343*(1), 21–27. https://doi.org/10.1016/j.yexcr.2015.11.017

Banchereau, J., & Steinman, R. M. (1998). Dendritic cells and the control of immunity. *Nature*, *392*(6673), 245–252. https://doi.org/10.1038/32588

Barczyk, M., Carracedo, S., & Gullberg, D. (2009). Integrins. *Cell and Tissue Research*, *339*(1), 269. https://doi.org/10.1007/s00441-009-0834-6

Bianchi, M. E., & Mezzapelle, R. (2020). The Chemokine Receptor CXCR4 in Cell Proliferation and Tissue Regeneration. *Frontiers in Immunology*, *11*, 2109. https://doi.org/10.3389/fimmu.2020.02109

Boucard-Jourdin, M., Kugler, D., Endale Ahanda, M.-L., This, S., De Calisto, J., Zhang, A., Mora, J. R., Stuart, L. M., Savill, J., Lacy-Hulbert, A., & Paidassi, H. (2016). β8 Integrin Expression and Activation of TGF-β by Intestinal Dendritic Cells Are Determined by Both Tissue Microenvironment and Cell Lineage. *Journal of Immunology (Baltimore, Md. : 1950)*, *197*(5), 1968–1978. https://doi.org/10.4049/jimmunol.1600244

Bray, N. L., Pimentel, H., Melsted, P., & Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology*. https://doi.org/10.1038/nbt.3519

Breitling, R., Amtmann, A., & Herzyk, P. (2004). Graph-based iterative Group Analysis enhances microarray interpretation. *BMC Bioinformatics*, *5*, 100. https://doi.org/10.1186/1471-2105-5-100

Brunkow, M. E., Jeffery, E. W., Hjerrild, K. A., Paeper, B., Clark, L. B., Yasayko, S. A., Wilkinson, J. E., Galas, D., Ziegler, S. F., & Ramsdell, F. (2001). Disruption of a new forkhead/winged-helix protein, scurfin, results in the fatal lymphoproliferative disorder of the scurfy mouse. *Nature Genetics*, *27*(1), 68–73. https://doi.org/10.1038/83784

Buenrostro, J. D., Wu, B., Chang, H. Y., & Greenleaf, W. J. (2015). ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Current Protocols in Molecular Biology*, *109*, 21.29.1-21.29.9. https://doi.org/10.1002/0471142727.mb2129s109

Cabeza-Cabrerizo, M., Cardoso, A., Minutti, C. M., Pereira da Costa, M., & Reis e Sousa, C.

(2021). Dendritic Cells Revisited. *Annual Review of Immunology*, *39*(1), 131–166. https://doi.org/10.1146/annurev-immunol-061020-053707

Caligaris, C., Vázquez-Victorio, G., Sosa-Garrocho, M., Ríos-López, D. G., Marín-Hernández, A., & Macías-Silva, M. (2015). Actin-cytoskeleton polymerization differentially controls the stability of Ski and SnoN co-repressors in normal but not in transformed hepatocytes. *Biochimica et Biophysica Acta (BBA) - General Subjects*, *1850*(9), 1832–1841. https://doi.org/https://doi.org/10.1016/j.bbagen.2015.05.012

Campbell, I. D., & Humphries, M. J. (2011). Integrin structure, activation, and interactions. *Cold Spring Harbor Perspectives in Biology*, *3*(3), a004994. https://doi.org/10.1101/cshperspect.a004994

Campbell, M. G., Cormier, A., Ito, S., Seed, R. I., Bondesson, A. J., Lou, J., Marks, J. D., Baron, J. L., Cheng, Y., & Nishimura, S. L. (2020). Cryo-EM Reveals Integrin-Mediated TGF-β Activation without Release from Latent TGF-β. *Cell*, *180*(3), 490-501.e16. https://doi.org/10.1016/j.cell.2019.12.030

Carnero, A., & Paramio, J. M. (2014). The PTEN/PI3K/AKT Pathway in vivo, Cancer Mouse Models. *Frontiers in Oncology*, *4*, 252. https://doi.org/10.3389/fonc.2014.00252

Chen, C. S., Alonso, J. L., Ostuni, E., Whitesides, G. M., & Ingber, D. E. (2003). Cell shape provides global control of focal adhesion assembly. *Biochemical and Biophysical Research Communications*, *307*(2), 355–361. https://doi.org/https://doi.org/10.1016/S0006-291X(03)01165-3

Chen, Z., Morales, J. E., Guerrero, P. A., Sun, H., & McCarty, J. H. (2018). PTPN12/PTP-PEST Regulates Phosphorylation-Dependent Ubiquitination and Stability of Focal Adhesion Substrates in Invasive Glioblastoma Cells. *Cancer Research*, *78*(14), 3809–3822. https://doi.org/10.1158/0008-5472.CAN-18-0085

Chieppa, M., Rescigno, M., Huang, A. Y. C., & Germain, R. N. (2006). Dynamic imaging of dendritic cell extension into the small bowel lumen in response to epithelial cell TLR engagement. *The Journal of Experimental Medicine*, *203*(13), 2841–2852. https://doi.org/10.1084/jem.20061884

Cock, P. J. A., Fields, C. J., Goto, N., Heuer, M. L., & Rice, P. M. (2009). The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkp1137

Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., Szcześniak, M. W., Gaffney, D. J., Elo, L. L., Zhang, X., & Mortazavi, A. (2016). A survey of best practices for RNA-seq data analysis. *Genome Biology*, *17*(1), 13. https://doi.org/10.1186/s13059-016-0881-8

Corley, S. M., MacKenzie, K. L., Beverdam, A., Roddam, L. F., & Wilkins, M. R. (2017). Differentially expressed genes from RNA-Seq and functional enrichment results are affected by the choice of single-end versus paired-end reads and stranded versus non-stranded protocols. *BMC Genomics*, *18*(1), 399. https://doi.org/10.1186/s12864-017-3797-0

Cormier, A., Campbell, M. G., Ito, S., Wu, S., Lou, J., Marks, J., Baron, J. L., Nishimura, S. L., & Cheng, Y. (2018). Cryo-EM structure of the αvβ8 integrin reveals a mechanism for stabilizing integrin extension. *Nature Structural & Molecular Biology*, *25*(8), 698–704. https://doi.org/10.1038/s41594-018-0093-x

Costa-Silva, J., Domingues, D., & Lopes, F. M. (2017). RNA-Seq differential expression analysis: An extended review and a software tool. *PLOS ONE*, *12*(12), e0190152. https://doi.org/10.1371/journal.pone.0190152

Creixell, P., Reimand, J., Haider, S., Wu, G., Shibata, T., Vazquez, M., Mustonen, V., Gonzalez-Perez, A., Pearson, J., Sander, C., Raphael, B. J., Marks, D. S., Ouellette, B. F. F., Valencia, A., Bader, G. D., Boutros, P. C., Stuart, J. M., Linding, R., Lopez-Bigas, N., & Stein, L. D. (2015). Pathway and network analysis of cancer genomes. *Nature Methods*, *12*(7), 615–621. https://doi.org/10.1038/nmeth.3440

Datto, M. B., Frederick, J. P., Pan, L., Borton, A. J., Zhuang, Y., & Wang, X.-F. (1999). Targeted Disruption of Smad3 Reveals an Essential Role in Transforming Growth Factor β-Mediated Signal Transduction. *Molecular and Cellular Biology*. https://doi.org/10.1128/mcb.19.4.2495

David, C. J., & Massagué, J. (2018). Contextual determinants of TGFβ action in development, immunity and cancer. *Nature Reviews Molecular Cell Biology*, *19*(7), 419–435. https://doi.org/10.1038/s41580-018-0007-0

de Sousa Abreu, R., Penalva, L. O., Marcotte, E. M., & Vogel, C. (2009). Global signatures of protein and mRNA expression levels. *Molecular BioSystems*, *5*(12), 1512–1526. https://doi.org/10.1039/b908315d

Dodagatta-Marri, E., Ma, H.-Y., Liang, B., Li, J., Meyer, D. S., Chen, S.-Y., Sun, K.-H., Ren, X., Zivak, B., Rosenblum, M. D., Headley, M. B., Pinzas, L., Reed, N. I., Del Cid, J. S., Hann, B. C., Yang, S., Giddabasappa, A., Noorbehesht, K., Yang, B., … Sheppard, D. (2021). Integrin αvβ8 on T cells suppresses anti-tumor immunity in multiple models and is a promising target for tumor immunotherapy. *Cell Reports*, *36*(1), 109309. https://doi.org/https://doi.org/10.1016/j.celrep.2021.109309

Dong, X., Zhao, B., Iacob, R. E., Zhu, J., Koksal, A. C., Lu, C., Engen, J. R., & Springer, T. A. (2017). Force interacts with macromolecular structure in activation of TGF-β. *Nature*, *542*(7639), 55–59. https://doi.org/10.1038/nature21035

Espejo, R., Rengifo-Cam, W., Schaller, M. D., Evers, B. M., & Sastry, S. K. (2010). PTP-PEST controls motility, adherens junction assembly, and Rho GTPase activity in colon cancer cells. *American Journal of Physiology. Cell Physiology*, *299*(2), C454-63. https://doi.org/10.1152/ajpcell.00148.2010

Ewels, P., Magnusson, M., Lundin, S., & Käller, M. (2016). MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*. https://doi.org/10.1093/bioinformatics/btw354

Fahlén, L., Read, S., Gorelik, L., Hurst, S. D., Coffman, R. L., Flavell, R. A., & Powrie, F. (2005). T cells that cannot respond to TGF-β escape control by CD4+CD25+ regulatory T cells . *Journal of Experimental Medicine*, *201*(5), 737–746. https://doi.org/10.1084/jem.20040685

Feng, X.-H., & Derynck, R. (2005). Specificity and versatility in tgf-beta signaling through Smads. *Annual Review of Cell and Developmental Biology*, *21*, 659–693.

https://doi.org/10.1146/annurev.cellbio.21.022404.142018

Fenton, T. M., Kelly, A., Shuttleworth, E. E., Smedley, C., Atakilit, A., Powrie, F., Campbell, S., Nishimura, S. L., Sheppard, D., Levison, S., Worthington, J. J., Lehtinen, M. J., & Travis, M. A. (2017). Inflammatory cues enhance TGFβ activation by distinct subsets of human intestinal dendritic cells via integrin αvβ8. *Mucosal Immunology*, *10*(3), 624–634. https://doi.org/10.1038/mi.2016.94

Fernandes, R. J., Hirohata, S., Engle, J. M., Colige, A., Cohn, D. H., Eyre, D. R., & Apte, S. S. (2001). Procollagen II amino propeptide processing by ADAMTS-3. Insights on dermatosparaxis. *The Journal of Biological Chemistry*, *276*(34), 31502–31509. https://doi.org/10.1074/jbc.M103466200

Förster, R., Schubel, A., Breitfeld, D., Kremmer, E., Renner-Müller, I., Wolf, E., & Lipp, M. (1999). CCR7 coordinates the primary immune response by establishing functional microenvironments in secondary lymphoid organs. *Cell*, *99*(1), 23–33. https://doi.org/10.1016/s0092-8674(00)80059-8

Frisch, S. M., & Francis, H. (1994). Disruption of epithelial cell-matrix interactions induces apoptosis. *The Journal of Cell Biology*, *124*(4), 619–626. https://doi.org/10.1083/jcb.124.4.619

Gaspal, F. M. C., Kim, M.-Y., McConnell, F. M., Raykundalia, C., Bekiaris, V., & Lane, P. J. L. (2005). Mice deficient in OX40 and CD30 signals lack memory antibody responses because of deficient CD4 T cell memory. *The Journal of Immunology*, *174*(7), 3891–3896.

Ghoreschi, K., Laurence, A., Yang, X.-P., Tato, C. M., McGeachy, M. J., Konkel, J. E., Ramos, H. L., Wei, L., Davidson, T. S., Bouladoux, N., Grainger, J. R., Chen, Q., Kanno, Y., Watford, W. T., Sun, H.-W., Eberl, G., Shevach, E. M., Belkaid, Y., Cua, D. J., … O'Shea, J. J. (2010). Generation of pathogenic TH17 cells in the absence of TGF-β signalling. *Nature*, *467*(7318), 967–971. https://doi.org/10.1038/nature09447

Giancotti, F. G., & Ruoslahti, E. (1999). Integrin Signaling. *Science*, *285*(5430), 1028 LP – 1033. https://doi.org/10.1126/science.285.5430.1028

Gilmore, A. P., & Burridge, K. (1996). Regulation of vinculin binding to talin and actin by phosphatidyl-inositol-4-5-bisphosphate. *Nature*, *381*(6582), 531–535. https://doi.org/10.1038/381531a0

Ginsberg, M. H. (2014). Integrin activation. *BMB Reports*, *47*(12), 655–659. https://doi.org/10.5483/bmbrep.2014.47.12.241

Gorelik, L., Constant, S., & Flavell, R. A. (2002). Mechanism of transforming growth factor beta-induced inhibition of T helper type 1 differentiation. *The Journal of Experimental Medicine*, *195*(11), 1499–1505. https://doi.org/10.1084/jem.20012076

Gorelik, L., Fields, P. E., & Flavell, R. A. (2000). Cutting Edge: TGF-β Inhibits Th Type 2 Development Through Inhibition of GATA-3 Expression. *The Journal of Immunology*, *165*(9), 4773 LP – 4777. https://doi.org/10.4049/jimmunol.165.9.4773

Greenwald, R. J., Freeman, G. J., & Sharpe, A. H. (2005). The B7 family revisited. *Annual Review of Immunology*, *23*, 515–548. https://doi.org/10.1146/annurev.immunol.23.021704.115611

Guilliams, M., Lambrecht, B. N., & Hammad, H. (2013). Division of labor between lung dendritic cells and macrophages in the defense against pulmonary infections. *Mucosal Immunology*, *6*(3), 464–473. https://doi.org/10.1038/mi.2013.14

Heldin, C.-H., & Moustakas, A. (2016). Signaling Receptors for TGF-β Family Members. *Cold Spring Harbor Perspectives in Biology*, *8*(8). https://doi.org/10.1101/cshperspect.a022053

Hemmings, L., Rees, D. J., Ohanian, V., Bolton, S. J., Gilmore, A. P., Patel, B., Priddle, H., Trevithick, J. E., Hynes, R. O., & Critchley, D. R. (1996). Talin contains three actin-binding sites each of which is adjacent to a vinculin-binding site. *Journal of Cell Science*, *109 (Pt 1*, 2715–2726.

Horwitz, D. A., Zheng, S. G., Wang, J., & Gray, J. D. (2008). Critical role of IL-2 and TGF-beta in generation, function and stabilization of Foxp3+CD4+ Treg. *European Journal of Immunology*, *38*(4), 912–915. https://doi.org/10.1002/eji.200738109

Huang, C.-T., Workman, C. J., Flies, D., Pan, X., Marson, A. L., Zhou, G., Hipkiss, E. L.,

Ravi, S., Kowalski, J., Levitsky, H. I., Powell, J. D., Pardoll, D. M., Drake, C. G., & Vignali, D. A. A. (2004). Role of LAG-3 in regulatory T cells. *Immunity*, *21*(4), 503–513. https://doi.org/10.1016/j.immuni.2004.08.010

Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, *4*(1), 44–57. https://doi.org/10.1038/nprot.2008.211

Humphries, M. J., Symonds, E. J. H., & Mould, A. P. (2003). Mapping functional residues onto integrin crystal structures. *Current Opinion in Structural Biology*, *13*(2), 236–243. https://doi.org/https://doi.org/10.1016/S0959-440X(03)00035-6

Huttenlocher, A., & Horwitz, A. R. (2011). Integrins in cell migration. *Cold Spring Harbor Perspectives in Biology*, *3*(9), a005074–a005074. https://doi.org/10.1101/cshperspect.a005074

Hynes, R. O. (2002). Integrins: Bidirectional, Allosteric Signaling Machines. *Cell*, *110*(6), 673–687. https://doi.org/https://doi.org/10.1016/S0092-8674(02)00971-6

Iborra, S., Izquierdo, H. M., Martínez-López, M., Blanco-Menéndez, N., Reis e Sousa, C., & Sancho, D. (2012). The DC receptor DNGR-1 mediates cross-priming of CTLs during vaccinia virus infection in mice. *The Journal of Clinical Investigation*, *122*(5), 1628–1643. https://doi.org/10.1172/JCI60660

Iborra, S., Martínez-López, M., Khouili, S. C., Enamorado, M., Cueto, F. J., Conde-Garrosa, R., Del Fresno, C., & Sancho, D. (2016). Optimal Generation of Tissue-Resident but Not Circulating Memory T Cells during Viral Infection Requires Crosspriming by DNGR-1(+) Dendritic Cells. *Immunity*, *45*(4), 847–860. https://doi.org/10.1016/j.immuni.2016.08.019

Iliev, I. D., Mileti, E., Matteoli, G., Chieppa, M., & Rescigno, M. (2009). Intestinal epithelial cells promote colitis-protective regulatory T-cell differentiation through dendritic cell conditioning. *Mucosal Immunology*, *2*(4), 340–350. https://doi.org/10.1038/mi.2009.13

Iliev, I. D., Spadoni, I., Mileti, E., Matteoli, G., Sonzogni, A., Sampietro, G. M., Foschi, D., Caprioli, F., Viale, G., & Rescigno, M. (2009). Human intestinal epithelial cells promote

the differentiation of tolerogenic dendritic cells. *Gut*, *58*(11), 1481 LP – 1489. https://doi.org/10.1136/gut.2008.175166

Jameson, S. C., & Masopust, D. (2018). Understanding Subset Diversity in T Cell Memory. *Immunity*, *48*(2), 214–226. https://doi.org/10.1016/j.immuni.2018.02.010

Joller, N., & Kuchroo, V. K. (2017). Tim-3, Lag-3, and TIGIT. *Current Topics in Microbiology and Immunology*, *410*, 127–156. https://doi.org/10.1007/82_2017_62

Joller, N., Lozano, E., Burkett, P. R., Patel, B., Xiao, S., Zhu, C., Xia, J., Tan, T. G., Sefik, E., Yajnik, V., Sharpe, A. H., Quintana, F. J., Mathis, D., Benoist, C., Hafler, D. A., & Kuchroo, V. K. (2014). Treg cells expressing the coinhibitory molecule TIGIT selectively inhibit proinflammatory Th1 and Th17 cell responses. *Immunity*, *40*(4), 569–581. https://doi.org/10.1016/j.immuni.2014.02.012

Jones, D., Fletcher, C. D. M., Pulford, K., Shahsafaei, A., & Dorfman, D. M. (1999). The T-Cell Activation Markers CD30 and OX40/CD134 Are Expressed in Nonoverlapping Subsets of Peripheral T-Cell Lymphoma. *Blood*, *93*(10), 3487–3493. https://doi.org/https://doi.org/10.1182/blood.V93.10.3487.410k39_3487_3493

Josefowicz, S. Z., Niec, R. E., Kim, H. Y., Treuting, P., Chinen, T., Zheng, Y., Umetsu, D. T., & Rudensky, A. Y. (2012). Extrathymically generated regulatory T cells control mucosal TH2 inflammation. *Nature*, *482*(7385), 395–399. https://doi.org/10.1038/nature10772

Kanehisa, M., & Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, *28*(1), 27–30. https://doi.org/10.1093/nar/28.1.27

Kapsenberg, M. L. (2003). Dendritic-cell control of pathogen-driven T-cell polarization. *Nature Reviews. Immunology*, *3*(12), 984–993. https://doi.org/10.1038/nri1246

Kavsak, P., Rasmussen, R. K., Causing, C. G., Bonni, S., Zhu, H., Thomsen, G. H., & Wrana, J. L. (2000). Smad7 binds to Smurf2 to form an E3 ubiquitin ligase that targets the TGF beta receptor for degradation. *Molecular Cell*, *6*(6), 1365–1375. https://doi.org/10.1016/s1097-2765(00)00134-9

Kechagia, J. Z., Ivaska, J., & Roca-Cusachs, P. (2019). Integrins as biomechanical sensors of the microenvironment. In *Nature Reviews Molecular Cell Biology*.

https://doi.org/10.1038/s41580-019-0134-2

Kelly, A., Gunaltay, S., McEntee, C. P., Shuttleworth, E. E., Smedley, C., Houston, S. A., Fenton, T. M., Levison, S., Mann, E. R., & Travis, M. A. (2018). Human monocytes and macrophages regulate immune tolerance via integrin αvβ8–mediated TGFβ activation. *Journal of Experimental Medicine*, *215*(11), 2725–2736. https://doi.org/10.1084/jem.20171491

Kelly, A., Houston, S. A., Sherwood, E., Casulli, J., & Travis, M. A. (2017). Regulation of Innate and Adaptive Immunity by TGFβ. *Advances in Immunology*, *134*, 137–233. https://doi.org/10.1016/bs.ai.2017.01.001

Kim, C., Ye, F., & Ginsberg, M. H. (2011). Regulation of Integrin Activation. *Annual Review of Cell and Developmental Biology*, *27*(1), 321–345. https://doi.org/10.1146/annurev-cellbio-100109-104104

Kitamura, H., Cambier, S., Somanath, S., Barker, T., Minagawa, S., Markovics, J., Goodsell, A., Publicover, J., Reichardt, L., Jablons, D., Wolters, P., Hill, A., Marks, J. D., Lou, J., Pittet, J.-F., Gauldie, J., Baron, J. L., & Nishimura, S. L. (2011). Mouse and human lung fibroblasts regulate dendritic cell trafficking, airway inflammation, and fibrosis through integrin αvβ8-mediated activation of TGF-β. *The Journal of Clinical Investigation*, *121*(7), 2863–2875. https://doi.org/10.1172/JCI45589

Korn, T., Bettelli, E., Gao, W., Awasthi, A., Jäger, A., Strom, T. B., Oukka, M., & Kuchroo, V. K. (2007). IL-21 initiates an alternative pathway to induce proinflammatory TH17 cells. *Nature*, *448*(7152), 484–487. https://doi.org/10.1038/nature05970

Korn, T., Bettelli, E., Oukka, M., & Kuchroo, V. K. (2009). IL-17 and Th17 Cells. *Annual Review of Immunology*, *27*, 485–517. https://doi.org/10.1146/annurev.immunol.021908.132710

Kramer, Ij. M. (2016). *Chapter 11 - Signal Transduction to and from Adhesion Molecules* (Ij. M. B. T.-S. T. (Third E. Kramer (ed.); pp. 655–702). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-394803-8.00011-5

Krause, M., Dent, E. W., Bear, J. E., Loureiro, J. J., & Gertler, F. B. (2003). Ena/VASP

proteins: regulators of the actin cytoskeleton and cell migration. *Annual Review of Cell and Developmental Biology*, *19*, 541–564. https://doi.org/10.1146/annurev.cellbio.19.050103.103356

Kuwahara, M., Yamashita, M., Shinoda, K., Tofukuji, S., Onodera, A., Shinnakasu, R., Motohashi, S., Hosokawa, H., Tumes, D., Iwamura, C., Lefebvre, V., & Nakayama, T. (2012). The transcription factor Sox4 is a downstream target of signaling by the cytokine TGF-β and suppresses TH2 differentiation. *Nature Immunology*, *13*(8), 778–786. https://doi.org/10.1038/ni.2362

Lacy-Hulbert, A., Smith, A. M., Tissire, H., Barry, M., Crowley, D., Bronson, R. T., Roes, J. T., Savill, J. S., & Hynes, R. O. (2007). Ulcerative colitis and autoimmunity induced by loss of myeloid alphav integrins. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(40), 15823–15828. https://doi.org/10.1073/pnas.0707421104

Lanzavecchia, A., & Sallusto, F. (2005). Understanding the generation and function of memory T cell subsets. *Current Opinion in Immunology*, *17*(3), 326–332. https://doi.org/https://doi.org/10.1016/j.coi.2005.04.010

Larson, R. S., Corbi, A. L., Berman, L., & Springer, T. (1989). Primary structure of the leukocyte function-associated molecule-1 alpha subunit: an integrin with an embedded domain defining a protein superfamily. *Journal of Cell Biology*, *108*(2), 703–712. https://doi.org/10.1083/jcb.108.2.703

Lathia, J. D., Mack, S. C., Mulkearns-Hubert, E. E., Valentim, C. L. L., & Rich, J. N. (2015). Cancer stem cells in glioblastoma. *Genes & Development*, *29*(12), 1203–1217. https://doi.org/10.1101/gad.261982.115

Lau, T.-L., Kim, C., Ginsberg, M. H., & Ulmer, T. S. (2009). The structure of the integrin αIIbβ3 transmembrane complex explains integrin transmembrane signalling. *The EMBO Journal*, *28*(9), 1351–1361. https://doi.org/https://doi.org/10.1038/emboj.2009.63

Lee, H. S., Cheerathodi, M., Chaki, S. P., Reyes, S. B., Zheng, Y., Lu, Z., Paidassi, H., DerMardirossian, C., Lacy-Hulbert, A., Rivera, G. M., & McCarty, J. H. (2015). Protein

tyrosine phosphatase-PEST and β8 integrin regulate spatiotemporal patterns of RhoGDI1 activation in migrating cells. *Molecular and Cellular Biology*, *35*(8), 1401–1413. https://doi.org/10.1128/MCB.00112-15

Leek, J. T., & Storey, J. D. (2007). Capturing heterogeneity in gene expression studies by surrogate variable analysis. *PLoS Genetics*. https://doi.org/10.1371/journal.pgen.0030161

Li, H., Deng, Y., Sun, K., Yang, H., Liu, J., Wang, M., Zhang, Z., Lin, J., Wu, C., Wei, Z., & Yu, C. (2017). Structural basis of kindlin-mediated integrin recognition and activation. *Proceedings of the National Academy of Sciences*, *114*(35), 9349 LP – 9354. https://doi.org/10.1073/pnas.1703064114

Li, M. O., Sanjabi, S., & Flavell, R. A. (2006). Transforming growth factor-beta controls development, homeostasis, and tolerance of T cells by regulatory T cell-dependent and -independent mechanisms. *Immunity*, *25*(3), 455–471. https://doi.org/10.1016/j.immuni.2006.07.011

Lim, Y., Han, I., Jeon, J., Park, H., Bahk, Y.-Y., & Oh, E.-S. (2004). Phosphorylation of focal adhesion kinase at tyrosine 861 is crucial for Ras transformation of fibroblasts. *The Journal of Biological Chemistry*, *279*(28), 29060–29065. https://doi.org/10.1074/jbc.M401183200

Liu, Y., Zhang, P., Li, J., Kulkarni, A. B., Perruche, S., & Chen, W. (2008). A critical function for TGF-beta signaling in the development of natural CD4+CD25+Foxp3+ regulatory T cells. *Nature Immunology*, *9*(6), 632–640. https://doi.org/10.1038/ni.1607

Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2-ΔΔCT method. *Methods*. https://doi.org/10.1006/meth.2001.1262

Lord, P. W., Stevens, R. D., Brass, A., & Goble, C. A. (2003). Investigating semantic similarity measures across the Gene Ontology: the relationship between sequence and annotation. *Bioinformatics*, *19*(10), 1275–1283. https://doi.org/10.1093/bioinformatics/btg153

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*. https://doi.org/10.1186/s13059-014-0550-8

Lucas, P. J., Kim, S. J., Melby, S. J., & Gress, R. E. (2000). Disruption of T cell homeostasis in mice expressing a T cell-specific dominant negative transforming growth factor beta II receptor. *The Journal of Experimental Medicine*, *191*(7), 1187–1196. https://doi.org/10.1084/jem.191.7.1187

Ma, C., & Zhang, N. (2015). Transforming growth factor-β signaling is constantly shaping memory T-cell population. *Proceedings of the National Academy of Sciences*, *112*(35), 11013 LP – 11017. https://doi.org/10.1073/pnas.1510119112

Mackay, L. K., Rahimpour, A., Ma, J. Z., Collins, N., Stock, A. T., Hafon, M.-L., Vega-Ramos, J., Lauzurica, P., Mueller, S. N., Stefanovic, T., Tscharke, D. C., Heath, W. R., Inouye, M., Carbone, F. R., & Gebhardt, T. (2013). The developmental pathway for CD103(+)CD8+ tissue-resident memory T cells of skin. *Nature Immunology*, *14*(12), 1294–1301. https://doi.org/10.1038/ni.2744

MacLeod, M. K. L., Clambey, E. T., Kappler, J. W., & Marrack, P. (2009). CD4 memory T cells: what are they and what can they do? *Seminars in Immunology*, *21*(2), 53–61. https://doi.org/10.1016/j.smim.2009.02.006

Maier, T., Güell, M., & Serrano, L. (2009). Correlation of mRNA and protein in complex biological samples. *FEBS Letters*, *583*(24), 3966–3973. https://doi.org/10.1016/j.febslet.2009.10.036

Mani, V., Bromley, S. K., Äijö, T., Mora-Buch, R., Carrizosa, E., Warner, R. D., Hamze, M., Sen, D. R., Chasse, A. Y., Lorant, A., Griffith, J. W., Rahimi, R. A., McEntee, C. P., Jeffrey, K. L., Marangoni, F., Travis, M. A., Lacy-Hulbert, A., Luster, A. D., & Mempel, T. R. (2019). Migratory DCs activate TGF-β to precondition naïve CD8(+) T cells for tissue-resident memory fate. *Science (New York, N.Y.)*, *366*(6462). https://doi.org/10.1126/science.aav5728

Marie, J. C., Liggitt, D., & Rudensky, A. Y. (2006). Cellular mechanisms of fatal early-onset

autoimmunity in mice with the T cell-specific targeting of transforming growth factor-beta receptor. *Immunity*, *25*(3), 441–454. https://doi.org/10.1016/j.immuni.2006.07.012

Markovics, J. A., Araya, J., Cambier, S., Jablons, D., Hill, A., Wolters, P. J., & Nishimura, S. L. (2010). Transcription of the transforming growth factor beta activating integrin beta8 subunit is regulated by SP3, AP-1, and the p38 pathway. *The Journal of Biological Chemistry*, *285*(32), 24695–24706. https://doi.org/10.1074/jbc.M110.113977

Markovics, J. A., Araya, J., Cambier, S., Somanath, S., Gline, S., Jablons, D., Hill, A., Wolters, P. J., & Nishimura, S. L. (2011). Interleukin-1beta induces increased transcriptional activation of the transforming growth factor-beta-activating integrin subunit beta8 through altering chromatin architecture. *The Journal of Biological Chemistry*, *286*(42), 36864–36874. https://doi.org/10.1074/jbc.M111.276790

Masopust, D., Choo, D., Vezys, V., Wherry, E. J., Duraiswamy, J., Akondy, R., Wang, J., Casey, K. A., Barber, D. L., Kawamura, K. S., Fraser, K. A., Webby, R. J., Brinkmann, V., Butcher, E. C., Newell, K. A., & Ahmed, R. (2010). Dynamic T cell migration program provides resident memory within intestinal epithelium. *The Journal of Experimental Medicine*, *207*(3), 553–564. https://doi.org/10.1084/jem.20090858

McCarty, J. H. (2020). αvβ8 integrin adhesion and signaling pathways in development, physiology and disease. *Journal of Cell Science*, *133*(12). https://doi.org/10.1242/jcs.239434

McCarty, J. H., Cook, A. A., & Hynes, R. O. (2005). An interaction between αvβ8 integrin and Band 4.1B via a highly conserved region of the Band 4.1 C-terminal domain. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(38), 13479 LP – 13483. https://doi.org/10.1073/pnas.0506068102

McEntee, C. P., Gunaltay, S., & Travis, M. A. (2020). Regulation of barrier immunity and homeostasis by integrin-mediated transforming growth factor β activation. *Immunology*, *160*(2), 139–148. https://doi.org/https://doi.org/10.1111/imm.13162

McKenna, H. J., Stocking, K. L., Miller, R. E., Brasel, K., De Smedt, T., Maraskovsky, E., Maliszewski, C. R., Lynch, D. H., Smith, J., Pulendran, B., Roux, E. R., Teepe, M.,

Lyman, S. D., & Peschon, J. J. (2000). Mice lacking flt3 ligand have deficient hematopoiesis affecting hematopoietic progenitor cells, dendritic cells, and natural killer cells. *Blood*, *95*(11), 3489–3497. https://doi.org/https://doi.org/10.1182/blood.V95.11.3489

Merad, M., Sathe, P., Helft, J., Miller, J., & Mortha, A. (2013). The dendritic cell lineage: ontogeny and function of dendritic cells and their subsets in the steady state and the inflamed setting. *Annual Review of Immunology*, *31*, 563–604. https://doi.org/10.1146/annurev-immunol-020711-074950

Meredith, M. M., Liu, K., Darrasse-Jeze, G., Kamphorst, A. O., Schreiber, H. A., Guermonprez, P., Idoyaga, J., Cheong, C., Yao, K.-H., Niec, R. E., & Nussenzweig, M. C. (2012). Expression of the zinc finger transcription factor zDC (Zbtb46, Btbd4) defines the classical dendritic cell lineage. *Journal of Experimental Medicine*, *209*(6), 1153–1165. https://doi.org/10.1084/jem.20112675

Miranti, C. K., & Brugge, J. S. (2002). Sensing the environment: a historical perspective on integrin signal transduction. *Nature Cell Biology*, *4*(4), E83–E90. https://doi.org/10.1038/ncb0402-e83

Mitra, S. K., Mikolon, D., Molina, J. E., Hsia, D. A., Hanson, D. A., Chi, A., Lim, S.-T., Bernard-Trifilo, J. A., Ilic, D., Stupack, D. G., Cheresh, D. A., & Schlaepfer, D. D. (2006). Intrinsic FAK activity and Y925 phosphorylation facilitate an angiogenic switch in tumors. *Oncogene*, *25*(44), 5969–5984. https://doi.org/10.1038/sj.onc.1209588

Mohammed, J., Beura, L. K., Bobr, A., Astry, B., Chicoine, B., Kashem, S. W., Welty, N. E., Igyártó, B. Z., Wijeyesinghe, S., Thompson, E. A., Matte, C., Bartholin, L., Kaplan, A., Sheppard, D., Bridges, A. G., Shlomchik, W. D., Masopust, D., & Kaplan, D. H. (2016). Stromal cells control the epithelial residence of DCs and memory T cells by regulated activation of TGF-β. *Nature Immunology*, *17*(4), 414–421. https://doi.org/10.1038/ni.3396

Montanez, E., Ussar, S., Schifferer, M., Bösl, M., Zent, R., Moser, M., & Fässler, R. (2008). Kindlin-2 controls bidirectional signaling of integrins. *Genes & Development*, *22*(10),

1325–1330. https://doi.org/10.1101/gad.469408

Moyle, M., Napier, M. A., & McLean, J. W. (1991). Cloning and expression of a divergent integrin subunit β8. *Journal of Biological Chemistry*. https://doi.org/10.1016/S0021-9258(18)55042-0

Mu, D., Cambier, S., Fjellbirkeland, L., Baron, J. L., Munger, J. S., Kawakatsu, H., Sheppard, D., Courtney Broaddus, V., & Nishimura, S. L. (2002). The integrin αvβ8 mediates epithelial homeostasis through MT1-MMP-dependent activation of TGF-β1. *Journal of Cell Biology*. https://doi.org/10.1083/jcb.200109100

Mueller, S. N., & Mackay, L. K. (2016). Tissue-resident memory T cells: local specialists in immune defence. *Nature Reviews. Immunology*, *16*(2), 79–89. https://doi.org/10.1038/nri.2015.3

Munger, J. S., Huang, X., Kawakatsu, H., Griffiths, M. J., Dalton, S. L., Wu, J., Pittet, J. F., Kaminski, N., Garat, C., Matthay, M. A., Rifkin, D. B., & Sheppard, D. (1999). The integrin alpha v beta 6 binds and activates latent TGF beta 1: a mechanism for regulating pulmonary inflammation and fibrosis. *Cell*, *96*(3), 319–328. https://doi.org/10.1016/s0092-8674(00)80545-0

Murphy, K., Janeway Jr., 1943-2003 (viaf)100308245, C. A., Travers, P. (viaf)84844575, & Walport, M. J. S. (2012). *Janeway's immunobiology* (8th ed.). New York : Garland Science. http://lib.ugent.be/catalog/rug01:002038139

Nakamura, K., Kitani, A., Fuss, I., Pedersen, A., Harada, N., Nawata, H., & Strober, W. (2004). TGF-beta 1 plays an important role in the mechanism of CD4+CD25+ regulatory T cell activity in both humans and mice. *Journal of Immunology (Baltimore, Md. : 1950)*, *172*(2), 834–842. https://doi.org/10.4049/jimmunol.172.2.834

Nakawesi, J., This, S., Hütter, J., Boucard-Jourdin, M., Barateau, V., Muleta, K. G., Gooday, L. J., Thomsen, K. F., López, A. G., Ulmert, I., Poncet, D., Malissen, B., Greenberg, H., Thaunat, O., Defrance, T., Paidassi, H., & Lahl, K. (2021). αvβ8 integrin-expression by BATF3-dependent dendritic cells facilitates early IgA responses to Rotavirus. *Mucosal Immunology*, *14*(1), 53–67. https://doi.org/10.1038/s41385-020-0276-8

Nishimura, S L, Sheppard, D., & Pytela, R. (1994). Integrin alpha v beta 8. Interaction with vitronectin and functional divergence of the beta 8 cytoplasmic domain. *Journal of Biological Chemistry*, *269*(46), 28708–28715. https://doi.org/https://doi.org/10.1016/S0021-9258(19)61963-0

Nishimura, Stephen L. (2009). Integrin-mediated transforming growth factor-βactivation, a potential therapeutic target in fibrogenic disorders. In *American Journal of Pathology*. https://doi.org/10.2353/ajpath.2009.090393

Notley, C. A., Brown, M. A., McGovern, J. L., Jordan, C. K., & Ehrenstein, M. R. (2015). Engulfment of Activated Apoptotic Cells Abolishes TGF-β–Mediated Immunoregulation via the Induction of IL-6. *The Journal of Immunology*, *194*(4), 1621 LP – 1627. https://doi.org/10.4049/jimmunol.1401256

Nuckolls, G. H., Romer, L. H., & Burridge, K. (1992). Microinjection of antibodies against talin inhibits the spreading and migration of fibroblasts. *Journal of Cell Science*, *102*(4), 753–762. https://doi.org/10.1242/jcs.102.4.753

Nunes, A. M., Minetti, C. A. S. A., Remeta, D. P., & Baum, J. (2018). Magnesium Activates Microsecond Dynamics to Regulate Integrin-Collagen Recognition. *Structure (London, England : 1993)*, *26*(8), 1080-1090.e5. https://doi.org/10.1016/j.str.2018.05.010

Ozawa, A., Sato, Y., Imabayashi, T., Uemura, T., Takagi, J., & Sekiguchi, K. (2017). Molecular basis of the ligand binding specificity of αvβ8 integrin. *Journal of Biological Chemistry*, *291*(22), 11551–11565.

Pagani, G., & Gohlke, H. (2018). On the contributing role of the transmembrane domain for subunit-specific sensitivity of integrin activation. *Scientific Reports*, *8*(1), 5733. https://doi.org/10.1038/s41598-018-23778-5

Païdassi, H., Acharya, M., Zhang, A., Mukhopadhyay, S., Kwon, M., Chow, C., Stuart, L. M., Savill, J., & Lacy-Hulbert, A. (2011). Preferential expression of integrin αvβ8 promotes generation of regulatory T cells by mouse CD103+ dendritic cells. *Gastroenterology*. https://doi.org/10.1053/j.gastro.2011.06.076

Pan, Y., Zhang, K., Qi, J., Yue, J., Springer, T. A., & Chen, J. (2010). Cation-pi interaction

regulates ligand-binding affinity and signaling of integrin alpha4beta7. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(50), 21388–21393. https://doi.org/10.1073/pnas.1015487107

Perkel, J. M. (2017). Single-cell sequencing made simple. *Nature*, *547*(7661), 125–126. https://doi.org/10.1038/547125a

Perkel, J. M. (2021). Single-cell analysis enters the multiomics age. In *Nature* (Vol. 595, Issue 7868, pp. 614–616). https://doi.org/10.1038/d41586-021-01994-w

Perruche, S., Zhang, P., Liu, Y., Saas, P., Bluestone, J. A., & Chen, W. (2008). CD3-specific antibody–induced immune tolerance involves transforming growth factor-β from phagocytes digesting apoptotic T cells. *Nature Medicine*, *14*(5), 528–535. https://doi.org/10.1038/nm1749

Pollen, A. A., Bhaduri, A., Andrews, M. G., Nowakowski, T. J., Meyerson, O. S., Mostajo-Radji, M. A., Di Lullo, E., Alvarado, B., Bedolli, M., Dougherty, M. L., Fiddes, I. T., Kronenberg, Z. N., Shuga, J., Leyrat, A. A., West, J. A., Bershteyn, M., Lowe, C. B., Pavlovic, B. J., Salama, S. R., … Kriegstein, A. R. (2019). Establishing Cerebral Organoids as Models of Human-Specific Brain Evolution. *Cell*, *176*(4), 743-756.e17. https://doi.org/10.1016/j.cell.2019.01.017

Powrie, F., Carlino, J., Leach, M. W., Mauze, S., & Coffman, R. L. (1996). A critical role for transforming growth factor-beta but not interleukin 4 in the suppression of T helper type 1-mediated colitis by CD45RB(low) CD4+ T cells. *The Journal of Experimental Medicine*, *183*(6), 2669–2674. https://doi.org/10.1084/jem.183.6.2669

Ramalingam, R., Larmonier, C. B., Thurston, R. D., Midura-Kiela, M. T., Zheng, S. G., Ghishan, F. K., & Kiela, P. R. (2012). Dendritic Cell-Specific Disruption of TGF-β Receptor II Leads to Altered Regulatory T Cell Phenotype and Spontaneous Multiorgan Autoimmunity. *The Journal of Immunology*, *189*(8), 3878 LP – 3893. https://doi.org/10.4049/jimmunol.1201029

Rebhun, J. F., Castro, A. F., & Quilliam, L. A. (2000). Identification of guanine nucleotide exchange factors (GEFs) for the Rap1 GTPase. Regulation of MR-GEF by M-Ras-GTP

interaction. *The Journal of Biological Chemistry*, *275*(45), 34901–34908. https://doi.org/10.1074/jbc.M005327200

Reis e Sousa, C. (2004). Activation of dendritic cells: translating innate into adaptive immunity. *Current Opinion in Immunology*, *16*(1), 21–25. https://doi.org/https://doi.org/10.1016/j.coi.2003.11.007

Reyes, S. B., Narayanan, A. S., Lee, H. S., Tchaicha, J. H., Aldape, K. D., Lang, F. F., Tolias, K. F., & McCarty, J. H. (2013). αvβ8 integrin interacts with RhoGDI1 to regulate Rac1 and Cdc42 activation and drive glioblastoma cell invasion. *Molecular Biology of the Cell*, *24*(4), 474–482. https://doi.org/10.1091/mbc.e12-07-0521

Reynolds, G., & Haniffa, M. (2015). Human and Mouse Mononuclear Phagocyte Networks: A Tale of Two Species? *Frontiers in Immunology*, *6*, 330. https://doi.org/10.3389/fimmu.2015.00330

Ricart, B. G., John, B., Lee, D., Hunter, C. A., & Hammer, D. A. (2011). Dendritic Cells Distinguish Individual Chemokine Signals through CCR7 and CXCR4. *The Journal of Immunology*. https://doi.org/10.4049/jimmunol.1002358

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., & Smyth, G. K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*, *43*(7), e47–e47. https://doi.org/10.1093/nar/gkv007

Rivero, F., Illenberger, D., Somesh, B. P., Dislich, H., Adam, N., & Meyer, A.-K. (2002). Defects in cytokinesis, actin reorganization and the contractile vacuole in cells deficient in RhoGDI. *The EMBO Journal*, *21*(17), 4539–4549. https://doi.org/10.1093/emboj/cdf449

Romer, L. H., Birukov, K. G., & Garcia, J. G. N. (2006). Focal Adhesions. *Circulation Research*, *98*(5), 606–616. https://doi.org/10.1161/01.RES.0000207408.31270.db

Sallusto, F., Geginat, J., & Lanzavecchia, A. (2004). Central memory and effector memory T cell subsets: function, generation, and maintenance. *Annual Review of Immunology*, *22*, 745–763. https://doi.org/10.1146/annurev.immunol.22.012703.104702

Sallusto, F., Lenig, D., Förster, R., Lipp, M., & Lanzavecchia, A. (1999). Two subsets of

memory T lymphocytes with distinct homing potentials and effector functions. *Nature*, *401*(6754), 708–712. https://doi.org/10.1038/44385

Sanjabi, S., Oh, S. A., & Li, M. O. (2017). Regulation of the Immune Response by TGF-β: From Conception to Autoimmunity and Infection. *Cold Spring Harbor Perspectives in Biology*, *9*(6). https://doi.org/10.1101/cshperspect.a022236

Schenkel, J. M., & Masopust, D. (2014). Tissue-resident memory T cells. *Immunity*, *41*(6), 886–897. https://doi.org/10.1016/j.immuni.2014.12.007

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J. Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., & Cardona, A. (2012). Fiji: An open-source platform for biological-image analysis. In *Nature Methods*. https://doi.org/10.1038/nmeth.2019

Schulz, O., Diebold, S. S., Chen, M., Näslund, T. I., Nolte, M. A., Alexopoulou, L., Azuma, Y.-T., Flavell, R. A., Liljeström, P., & Reis e Sousa, C. (2005). Toll-like receptor 3 promotes cross-priming to virus-infected cells. *Nature*, *433*(7028), 887–892. https://doi.org/10.1038/nature03326

See, P., Lum, J., Chen, J., & Ginhoux, F. (2018). A Single-Cell Sequencing Guide for Immunologists. *Frontiers in Immunology*, *9*, 2425. https://doi.org/10.3389/fimmu.2018.02425

Sergeant, S., Rahbar, E., & Chilton, F. H. (2016). Gamma-linolenic acid, Dihommo-gamma linolenic, Eicosanoids and Inflammatory Processes. *European Journal of Pharmacology*, *785*, 77–86. https://doi.org/10.1016/j.ejphar.2016.04.020

Shen, B., Delaney, M. K., & Du, X. (2012). Inside-out, outside-in, and inside-outside-in: G protein signaling in integrin-mediated cell adhesion, spreading, and retraction. *Current Opinion in Cell Biology*, *24*(5), 600–606. https://doi.org/10.1016/j.ceb.2012.08.011

Shen, B., Zhao, X., O'Brien, K. A., Stojanovic-Terpo, A., Delaney, M. K., Kim, K., Cho, J., Lam, S. C. T., & Du, X. (2013). A directional switch of integrin signalling and a new anti-thrombotic strategy. *Nature*. https://doi.org/10.1038/nature12613

Shi, M., Lin, T. H., Appell, K. C., & Berg, L. J. (2009). Cell cycle progression following naive T cell activation is independent of Jak3/common gamma-chain cytokine signals. *Journal of Immunology (Baltimore, Md. : 1950)*, *183*(7), 4493–4501. https://doi.org/10.4049/jimmunol.0804339

Shi, Y., & Massagué, J. (2003). Mechanisms of TGF-beta signaling from cell membrane to the nucleus. *Cell*, *113*(6), 685–700. https://doi.org/10.1016/s0092-8674(03)00432-x

Shin, L. H., Mujeeburahiman, C., P., C. S., B., R. S., Yanhua, Z., Zhimin, L., Helena, P., Celine, D., Adam, L.-H., M., R. G., & H., M. J. (2015). Protein Tyrosine Phosphatase-PEST and β8 Integrin Regulate Spatiotemporal Patterns of RhoGDI1 Activation in Migrating Cells. *Molecular and Cellular Biology*, *35*(8), 1401–1413. https://doi.org/10.1128/MCB.00112-15

Siegel, D. H., Ashton, G. H. S., Penagos, H. G., Lee, J. V, Feiler, H. S., Wilhelmsen, K. C., South, A. P., Smith, F. J. D., Prescott, A. R., Wessagowit, V., Oyama, N., Akiyama, M., Al Aboud, D., Al Aboud, K., Al Githami, A., Al Hawsawi, K., Al Ismaily, A., Al-Suwaid, R., Atherton, D. J., … Epstein, E. H. (2003). Loss of kindlin-1, a human homolog of the Caenorhabditis elegans actin-extracellular-matrix linker protein UNC-112, causes Kindler syndrome. *American Journal of Human Genetics*, *73*(1), 174–187. https://doi.org/10.1086/376609

Singh, R., Kapur, N., Mir, H., Singh, N., Lillard Jr., J. W., & Singh, S. (2016). CXCR6-CXCL16 axis promotes prostate cancer by mediating cytoskeleton rearrangement via Ezrin activation and α v β 3 integrin clustering. *Oncotarget; Vol 7, No 6*. https://www.oncotarget.com/article/6944/text/

Smircich, P., Eastman, G., Bispo, S., Duhagon, M. A., Guerra-Slompo, E. P., Garat, B., Goldenberg, S., Munroe, D. J., Dallagiovanna, B., Holetz, F., & Sotelo-Silveira, J. R. (2015). Ribosome profiling reveals translation control as a key mechanism generating differential gene expression in Trypanosoma cruzi. *BMC Genomics*, *16*(1), 443. https://doi.org/10.1186/s12864-015-1563-8

Smyth, M. J., Strobl, S. L., Young, H. A., Ortaldo, J. R., & Ochoa, A. C. (1991). Regulation of

lymphokine-activated killer activity and pore-forming protein gene expression in human peripheral blood CD8+ T lymphocytes. Inhibition by transforming growth factor-beta. *The Journal of Immunology*, *146*(10), 3289 LP – 3297. http://www.jimmunol.org/content/146/10/3289.abstract

Soto, J. A., Gálvez, N. M. S., Andrade, C. A., Pacheco, G. A., Bohmwald, K., Berrios, R. V, Bueno, S. M., & Kalergis, A. M. (2020). The Role of Dendritic Cells During Infections Caused by Highly Prevalent Viruses . In *Frontiers in Immunology* (Vol. 11, p. 1513). https://www.frontiersin.org/article/10.3389/fimmu.2020.01513

Spolski, R., Kim, H.-P., Zhu, W., Levy, D. E., & Leonard, W. J. (2009). IL-21 mediates suppressive effects via its induction of IL-10. *Journal of Immunology (Baltimore, Md. : 1950)*, *182*(5), 2859–2867. https://doi.org/10.4049/jimmunol.0802978

Spolski, R., & Leonard, W. J. (2014). Interleukin-21: a double-edged sword with therapeutic potential. *Nature Reviews Drug Discovery*, *13*(5), 379–395. https://doi.org/10.1038/nrd4296

Stark, R., Grzelak, M., & Hadfield, J. (2019). RNA sequencing: the teenage years. *Nature Reviews Genetics*, *20*(11), 631–656. https://doi.org/10.1038/s41576-019-0150-2

Steel, N., Faniyi, A. A., Rahman, S., Swietlik, S., Czajkowska, B. I., Chan, B. T., Hardgrave, A., Steel, A., Sparwasser, T. D., Assas, M. B., Grencis, R. K., Travis, M. A., & Worthington, J. J. (2019). TGFβ-activation by dendritic cells drives Th17 induction and intestinal contractility and augments the expulsion of the parasite Trichinella spiralis in mice. *PLOS Pathogens*, *15*(4), e1007657. https://doi.org/10.1371/journal.ppat.1007657

Steinert, E. M., Schenkel, J. M., Fraser, K. A., Beura, L. K., Manlove, L. S., Igyártó, B. Z., Southern, P. J., & Masopust, D. (2015). Quantifying Memory CD8 T Cells Reveals Regionalization of Immunosurveillance. *Cell*, *161*(4), 737–749. https://doi.org/10.1016/j.cell.2015.03.031

Stockis, J., Liénart, S., Colau, D., Collignon, A., Nishimura, S. L., Sheppard, D., Coulie, P. G., & Lucas, S. (2017). Blocking immunosuppression by human Tregs in vivo with antibodies targeting integrin αVβ8. *Proceedings of the National Academy of Sciences*,

114(47), E10161–E10168.

Sun, G., Sun, X., Li, W., Liu, K., Tian, D., Dong, Y., Sun, X., Xu, H., & Zhang, D. (2018). Critical role of OX40 in the expansion and survival of CD4 T-cell-derived double-negative T cells. *Cell Death & Disease*, *9*(6), 616. https://doi.org/10.1038/s41419-018-0659-x

Supek, F., Bošnjak, M., Škunca, N., & Šmuc, T. (2011). REVIGO summarizes and visualizes long lists of gene ontology terms. *PloS One*, *6*(7), e21800–e21800. https://doi.org/10.1371/journal.pone.0021800

Tailor, P., Tamura, T., Morse, H. C. 3rd, & Ozato, K. (2008). The BXH2 mutation in IRF8 differentially impairs dendritic cell subset development in the mouse. *Blood*, *111*(4), 1942–1945. https://doi.org/10.1182/blood-2007-07-100750

Takai, S., Schlom, J., Tucker, J., Tsang, K. Y., & Greiner, J. W. (2013). Inhibition of TGF-β1 signaling promotes central memory T cell differentiation. *Journal of Immunology (Baltimore, Md. : 1950)*, *191*(5), 2299–2307. https://doi.org/10.4049/jimmunol.1300472

Takimoto, T., Wakabayashi, Y., Sekiya, T., Inoue, N., Morita, R., Ichiyama, K., Takahashi, R., Asakawa, M., Muto, G., Mori, T., Hasegawa, E., Shizuya, S., Hara, T., Nomura, M., & Yoshimura, A. (2010). Smad2 and Smad3 Are Redundantly Essential for the TGF-β–Mediated Regulation of Regulatory T Plasticity and Th1 Development. *The Journal of Immunology*, *185*(2), 842 LP – 855. https://doi.org/10.4049/jimmunol.0904100

Tamkun, J. W., DeSimone, D. W., Fonda, D., Patel, R. S., Buck, C., Horwitz, A. F., & Hynes, R. O. (1986). Structure of integrin, a glycoprotein involved in the transmembrane linkage between fibronectin and actin. *Cell*, *46*(2), 271–282. https://doi.org/https://doi.org/10.1016/0092-8674(86)90744-0

Tanaka, S., Morishita, T., Hashimoto, Y., Hattori, S., Nakamura, S., Shibuya, M., Matuoka, K., Takenawa, T., Kurata, T., & Nagashima, K. (1994). C3G, a guanine nucleotide-releasing protein expressed ubiquitously, binds to the Src homology 3 domains of CRK and GRB2/ASH proteins. *Proceedings of the National Academy of Sciences of the United States of America*, *91*(8), 3443–3447. https://doi.org/10.1073/pnas.91.8.3443

Tchaicha, J. H., Reyes, S. B., Shin, J., Hossain, M. G., Lang, F. F., & McCarty, J. H. (2011). Glioblastoma angiogenesis and tumor cell invasiveness are differentially regulated by β8 integrin. *Cancer Research*, *71*(20), 6371–6381. https://doi.org/10.1158/0008-5472.CAN-11-0991

Thompson, P. M., Ramachandran, S., Case, L. B., Tolbert, C. E., Tandon, A., Pershad, M., Dokholyan, N. V, Waterman, C. M., & Campbell, S. L. (2017). A Structural Model for Vinculin Insertion into PIP(2)-Containing Membranes and the Effect of Insertion on Vinculin Activation and Localization. *Structure (London, England : 1993)*, *25*(2), 264–275. https://doi.org/10.1016/j.str.2016.12.002

Torchinsky, M. B., Garaude, J., Martin, A. P., & Blander, J. M. (2009). Innate immune recognition of infected apoptotic cells directs TH17 cell differentiation. *Nature*, *458*(7234), 78–82. https://doi.org/10.1038/nature07781

Torres-Gomez, A., Sanchez-Trincado, J. L., Toribio, V., Torres-Ruiz, R., Rodríguez-Perales, S., Yáñez-Mó, M., Reche, P. A., Cabañas, C., & Lafuente, E. M. (2020). RIAM-VASP Module Relays Integrin Complement Receptors in Outside-In Signaling Driving Particle Engulfment. In *Cells* (Vol. 9, Issue 5). https://doi.org/10.3390/cells9051166

Travis, M. A., Reizis, B., Melton, A. C., Masteller, E., Tang, Q., Proctor, J. M., Wang, Y., Bernstein, X., Huang, X., Reichardt, L. F., Bluestone, J. A., & Sheppard, D. (2007). Loss of integrin αvβ8 on dendritic cells causes autoimmunity and colitis in mice. *Nature*. https://doi.org/10.1038/nature06110

Veldhoen, M., Hocking, R. J., Flavell, R. A., & Stockinger, B. (2006). Signals mediated by transforming growth factor-β initiate autoimmune encephalomyelitis, but chronic inflammation is needed to sustain disease. *Nature Immunology*, *7*(11), 1151–1156. https://doi.org/10.1038/ni1391

Villadangos, J. A., & Schnorrer, P. (2007). Intrinsic and cooperative antigen-presenting functions of dendritic-cell subsets in vivo. *Nature Reviews. Immunology*, *7*(7), 543–555. https://doi.org/10.1038/nri2103

Wallace, D. M. E., Lindsay, A. J., Hendrick, A. G., & McCaffrey, M. W. (2002). Rab11-FIP4

interacts with Rab11 in a GTP-dependent manner and its overexpression condenses the Rab11 positive compartment in HeLa cells. *Biochemical and Biophysical Research Communications*, *299*(5), 770–779. https://doi.org/10.1016/s0006-291x(02)02720-1

Wang, J., Dong, X., Zhao, B., Li, J., Lu, C., & Springer, T. A. (2017). Atypical interactions of integrin α_Vβ₈ with pro-TGF-β1. *Proceedings of the National Academy of Sciences*, *114*(21), E4168 LP-E4174. https://doi.org/10.1073/pnas.1705129114

Wang, Z., Gerstein, M., & Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews. Genetics*, *10*(1), 57–63. https://doi.org/10.1038/nrg2484

Whitley, S. K., Horne, W. T., & Kolls, J. K. (2016). Research Techniques Made Simple: Methodology and Clinical Applications of RNA Sequencing. *Journal of Investigative Dermatology*, *136*(8), e77–e82. https://doi.org/https://doi.org/10.1016/j.jid.2016.06.003

Wipff, P.-J., Rifkin, D. B., Meister, J.-J., & Hinz, B. (2007). Myofibroblast contraction activates latent TGF-beta1 from the extracellular matrix. *The Journal of Cell Biology*, *179*(6), 1311–1323. https://doi.org/10.1083/jcb.200704042

Worthington, J. J., Fenton, T. M., Czajkowska, B. I., Klementowicz, J. E., & Travis, M. A. (2012). Regulation of TGFβ in the immune system: An emerging role for integrins and dendritic cells. *Immunobiology*. https://doi.org/10.1016/j.imbio.2012.06.009

Worthington, J. J., Kelly, A., Smedley, C., Bauché, D., Campbell, S., Marie, J. C., & Travis, M. A. (2015). Integrin αvβ8-Mediated TGF-β Activation by Effector Regulatory T Cells Is Essential for Suppression of T-Cell-Mediated Inflammation. *Immunity*, *42*(5), 903–915. https://doi.org/10.1016/j.immuni.2015.04.012

Worthington, J. J., Klementowicz, J. E., Rahman, S., Czajkowska, B. I., Smedley, C., Waldmann, H., Sparwasser, T., Grencis, R. K., & Travis, M. A. (2013). Loss of the TGFβ-Activating Integrin αvβ8 on Dendritic Cells Protects Mice from Chronic Intestinal Parasitic Infection via Control of Type 2 Immunity. *PLOS Pathogens*, *9*(10), e1003675. https://doi.org/10.1371/journal.ppat.1003675

Worthington, J. J., Klementowicz, J. E., & Travis, M. A. (2011). TGFβ: a sleeping giant awoken by integrins. *Trends in Biochemical Sciences*, *36*(1), 47–54. https://doi.org/10.1016/j.tibs.2010.08.002

Wu, J., Wu, H., An, J., Ballantyne, C. M., & Cyster, J. G. (2018). Critical role of integrin CD11c in splenic dendritic cell capture of missing-self CD47 cells to induce adaptive immunity. *Proceedings of the National Academy of Sciences*, *115*(26), 6786 LP – 6791. https://doi.org/10.1073/pnas.1805542115

Xiao, T., Takagi, J., Coller, B. S., Wang, J.-H., & Springer, T. A. (2004). Structural basis for allostery in integrins and binding to fibrinogen-mimetic therapeutics. *Nature*, *432*(7013), 59–67. https://doi.org/10.1038/nature02976

Yakymovych, I., Yakymovych, M., & Heldin, C.-H. (2018). Intracellular trafficking of transforming growth factor β receptors. *Acta Biochimica et Biophysica Sinica*, *50*(1), 3–11. https://doi.org/10.1093/abbs/gmx119

Yang, Z., Mu, Z., Dabovic, B., Jurukovski, V., Yu, D., Sung, J., Xiong, X., & Munger, J. S. (2007). Absence of integrin-mediated TGFβ1 activation in vivo recapitulates the phenotype of TGFβ1-null mice. *The Journal of Cell Biology*, *176*(6), 787–793.

Zhang, X., Tang, N., Hadden, T. J., & Rishi, A. K. (2011). Akt, FoxO and regulation of apoptosis. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, *1813*(11), 1978–1986. https://doi.org/https://doi.org/10.1016/j.bbamcr.2011.03.010

Zhou, A. C., Snell, L. M., Wortzman, M. E., & Watts, T. H. (2017). CD30 Is Dispensable for T-Cell Responses to Influenza Virus and Lymphocytic Choriomeningitis Virus Clone 13 but Contributes to Age-Associated T-Cell Expansion in Mice. *Frontiers in Immunology*, *8*, 1156. https://doi.org/10.3389/fimmu.2017.01156

Zhou, L., Lopes, J. E., Chong, M. M. W., Ivanov, I. I., Min, R., Victora, G. D., Shen, Y., Du, J., Rubtsov, Y. P., Rudensky, A. Y., Ziegler, S. F., & Littman, D. R. (2008). TGF-beta-induced Foxp3 inhibits T(H)17 cell differentiation by antagonizing RORgammat function. *Nature*, *453*(7192), 236–240. https://doi.org/10.1038/nature06878

Zhu, Jiangwen, Motejlek, K., Wang, D., Zang, K., Schmidt, A., & Reichardt, L. F. (2002).

beta8 integrins are required for vascular morphogenesis in mouse embryos. *Development (Cambridge, England)*, *129*(12), 2891–2903.

Zhu, Jinfang. (2018). T Helper Cell Differentiation, Heterogeneity, and Plasticity. *Cold Spring Harbor Perspectives in Biology*, *10*(10), a030338. https://doi.org/10.1101/cshperspect.a030338