PhD Programme in Biomedicine

Department of Medicine and Life Sciences (MELIS)

Universitat Pompeu Fabra

A new *in vitro* model to study the regulatory functions of β cells in early stages of Type 1 Diabetes

Beatriz Pérez-González

Thesis supervisor Dr. Lorenzo Pasquali

TESI DOCTORAL UPF / 2024

upf. Universitat Pompeu Fabra Barcelona

Acknowledgements

Completing this PhD thesis has been a difficult journey, filled with challenges and obstacles that are all too familiar to most PhD students. Despite the difficulties, the experience has been incredibly rewarding. I am both happy and proud to have reached this significant milestone, and I am grateful for the support and encouragement I have received along the way.

First of all, I would like to thank my thesis supervisor, Lorenzo, for his support and guidance throughout this project. I am really grateful for the confidence you put on me and all the collaboration opportunities, which have greatly enriched my academic journey. I also want to thank the other half of this project, my colleague Georgina, for all her amazing help with the bioinformatic analysis. Her technical skills and patience were a lifesaver, and I couldn't have done it without her. I'm really grateful for everything you've done and for always being there when I needed a hand. I have learnt a lot of how all the analysis are done thanks to you, I will always be grateful!

Our collaborators from Dr. Pere Santamaria's lab here in Barcelona for giving me the opportunity to expand my opportunities and learn key aspects of T1D immunology. Especially, to Nahir for the kindness and willingness to help me every time I was facing an obstacle with the lymphocytes. Your input was essential for the project!

Of course, I also want to thank all the other former and current members of the team: Mireia, Marc, Helena, Ángel, and Clara. Not only your professional help and technical expertise helped me along the way but also your personal advice and emotional support offered me comfort and encouragement. I will always remember Marc's adventures, the continuous laughs and endless talks about everything in life. I also want to thank Paula, our adopted lab member. You became the perfect addition to the lab as I gained a wonderful lab friend and someone to discuss problems and experiments with. Your insights, companionship, and motivation made this experience much more enjoyable.

També vull agrair I dedicar unes paraules a la Raquel. Crec que aquesta tesi no hagués sigut el mateix sense tota la teva ajuda, tant a nivell tècnic com a nivell personal. Moltes gràcies per totes les vegades que he necessitat alguna cosa i t'has mogut tot el que podies per ajudar, tant a mi com al laboratori. Per sempre estar allà quan necessitava parlar, la

teva comprensió i tots els consells que m'has donat al llarg d'aquests quatre anys.

I also want to thank Carla and Marta, who, despite not being part of the same research group, made it feel like we were all on the same team. Thank you for everything. Thanks to So-yeon, Jessica, Maria, Lorena, and all the UPF members I have met along the way for all the advice and long talks on the corridor or inside cell culture. I will never forget you!

I can't finish this thesis without mentioning Joan and Núria. I learnt so much from you at the IGTP and that definitely helped me through this thesis. You have been there whenever I needed it for any kind of advice and that is invaluable.

Finalmente, quiero dar las gracias a toda mi familia y amigos. Gracias a mis padres, por todo el apoyo y la confianza durante tantos años que me ha llevado a lograr este objetivo. Sé que no han sido unos años fáciles, pero ahora todo irá a mejor. A mi hermano y Estefanía, muchas gracias por estar ahí cuando lo he necesitado y, sobre todo, por el mejor regalo que podíais hacerme, Gael. Su alegría, energía y cariño son mágicos. A la Roser i en Joan, no tinc paraules per expressar l'agraïment pel que heu fet durant tots aquests anys. A en Jordi i la Núria, per l'altre increïble regal que ens heu fet, la Bianca. Poder passar temps amb ella sempre ens omple d'històries i amor. Anna i Nil, siempre habéis estado ahí, escuchando y ayudándome a pesar de no entender nada. Muchas gracias.

Y a ti Xavi, por la paciencia y todo lo que has hecho siempre, pero en especial estos últimos años. Sin ti no podría haberlo superado y todo lo que hemos construido no sería posible. Moltes gràcies, t'estimo molt.

Abstract

Type 1 Diabetes (T1D) is an autoimmune disease characterized by a selective destruction and a progressive loss of the pancreatic β cells. In the initial phase of inflammation, known as early insulitis, there is an interaction between immune cells and β cells. However, as the disease develops, the precise mechanisms leading to β cell death are not fully understood. On the other hand, there is an urgent need to develop in vitro human tissue models to capture the relevant regulatory responses of the cells involved in the early T1D stages. Investigating the gene regulatory networks that govern β cell responses during insulitis will enable the identification of key molecular pathways leading to β cell dysfunction and apoptosis, disentangling the role β cells play in their own destruction.

In the present thesis, I aimed to elucidate the β cell responses to an inflammatory environment by developing a new in vitro model, based on the co-culture of β cells with CD4+ T cells, to mimic early insulitis. After β cell exposure to activated CD4+ T cells, I characterized the cytokine profile of the culture milieu and the changes in chromatin landscape and gene expression. Cytokine profiling confirmed IFNγ presence and other inflammatory cytokines, indicating bidirectional communication between immune and β cells. Additionally, I observed significant chromatin remodeling coupled with extensive changes in gene expression. I identified a subset of new regulatory elements that become active or inactive only upon exposure, named Induced (IREs) and Lost (LoREs) Regulatory Elements. Comparative analysis revealed unique regulatory pathways activated in co-cultured β cells, containing IREs that directly overlap with some T1D-risk SNPs. Risk variants overlapping inflammatory responsive regulatory elements might disrupt the regular β cell response to inflammation, highlighting potential genes implicated in disease pathogenesis such as SOCS1. CRISPR activation experiments validated the functional impact of an IRE bearing a risk variant on SOCS1 gene expression.

In conclusion, the optimized co-culture model provides valuable insights into β cell responses to immune-mediated inflammation, emphasizing the importance of studying cell-cell interactions and chromatin dynamics in T1D pathogenesis. This approach offers a robust platform for identifying potential therapeutic targets for T1D treatment.

Resum

La diabetis tipus 1 (T1D) és una malaltia autoimmune caracteritzada per la destrucció selectiva de les cèl·lules β pancreàtiques. En la fase inicial d'inflamació, coneguda com a insulitis, hi ha una interacció entre les cèl·lules immunitàries i les cèl·lules β. No obstant això, els mecanismes exactes que condueixen a la mort de les cèl·lules β encara són desconeguts. A més, hi ha una necessitat urgent de desenvolupar models in vitro per captar les respostes moleculars rellevants més rellevants de les cèl·lules implicades en les primeres etapes de la T1D.

El principal objectiu d'aquesta tesi és aclarir la resposta de les cèl·lules β a un entorn inflamatori. Per aconseguir-ho, he desenvolupat un nou model in vitro basat en la co-cultivar cèl·lules β amb cèl·lules T CD4+ per reproduir la insulitis. Després de l'exposició a les cèl·lules T CD4+ activades, he caracteritzat la composició de citoquines del medi de cultiu i els canvis en la cromatina i l'expressió gènica. La presència d'IFNγ i altres citoquines inflamatòries confirma una comunicació bidireccional entre les cèl·lules immunitàries i les cèl·lules β. També hi ha una remodelació significativa de la cromatina i canvis en l'expressió gènica, caracteritzant elements reguladors induïts (IREs) i elements reguladors perduts (LoREs) després de l'exposició. La comparació amb altres models in vitro revela vies de regulació específiques activades en les cèl·lules β co-cultivades, amb IREs que se superposen amb variants genètiques de risc de T1D. Aquestes variants de risc que se superposen amb IREs podrien alterar la resposta normal de les cèl·lules β a la inflamació, destacant gens potencialment implicats en la patogènesi de la malaltia. Experiments amb CRISPR han validat l'impacte funcional d'un IRE amb una variant de risc en l'expressió del gen SOCS1.

En conclusió, el model de co-cultiu presentat proporciona valuoses perspectives sobre les respostes de les cèl·lules β a la inflamació, subratllant la importància d'estudiar les interaccions cel·lulars i la dinàmica de la cromatina en la patogènesi de la T1D. Aquest enfocament ofereix una plataforma robusta per identificar possibles dianes terapèutiques per al tractament de la T1D.

Contents

Abbreviations

AAb	Autoantibody
APC	Antigen Presenting Cell
ATAC-seq	Assay For Transposase-Accessible Chromatin Using Sequencing
bp	Base Pairs
ChIP	Chromatin Immunoprecipitation
CRISPRa	CRISPR Activation
CRISPRi	CRISPR Interference
CUT&RUN	Cleavage Under Targets and Release Using Nuclease
CUT&Tag	Cleavage Under Targets and Tagmentation
DEGs	Differentially Expressed Genes
EC	EndoC-βH1
ER	Endoplasmic Reticulum
FACS	Fluorescence Activated Cell Sorting
GADA	Glutamic Acid Decarboxylase Autoantibodies
gRNA	Guide RNAs
GSIS	Glucose-Stimulated Insulin Secretion
GWAS	Genome-Wide Association Study
hESCs	Human Embryonic Stem Cells
HLA	Human Leukocyte Antigen
HO	Hoechst 33342
IA-2A	Protein Phosphatase-Like Autoantibodies
IAA	Insulin Autoantibodies
IFN	Interferon
INS	Insulin
iPSCs	Induced Pluripotent Stem Cells
IREs	Induced Regulatory Elements
kb	Kilobase Pair
LoREs	Lost Regulatory Elements
Mb	Megabase Pair
MHC	Major Histocompatibility Complex

NK	Natural Killer Cells
NOD mouse	Nonobese Diabetic Mouse
PBMCs	Peripheral Blood Mononuclear Cells
PI	Propidium Iodide
qPCR	Quantitative PCR
REs	Regulatory Elements
SC	Stem Cells
scRNA-seq	Single-Cell RNA-Seq
SN	Supernatant
snATAC-seq	Single-Nucleus ATAC-Seq
SNPs	Single-Nucleotide Polymorphisms
SREs	Stable Regulatory Elements
T1D	Type 1 Diabetes
T2D	Type 2 Diabetes
TADs	Topologically Associating Domains
TCR	T Cell Receptor
TF	Transcription Factor
TFBSs	Transcription Factor Binding Sites
TSS	Transcription Start Site
UPR	Unfolded Protein Response

Preamble

This doctoral thesis was supervised by Dr. Lorenzo Pasquali, Principal Investigator of the Endocrine Regulatory Genomics group, at the Department of Medicine and Life Sciences of the Universitat Pompeu Fabra.

The results obtained by addressing the objectives of the present thesis have been organized in different sections: Introduction, Hypothesis and objectives, Materials and methods, Results, Discussion and Conclusions.

Despite advances in understanding the pathophysiological aspects of Type 1 diabetes (T1D), the exact mechanisms by which β cells become dysfunctional and die remain unclear. One of the main challenges in studying T1D is the limited availability of human pancreatic samples at the early stages of the disease to study the response of β cells in early stages of disease.

Regarding the content of this study, I have developed a novel co-culture model that mimics early insulitis in T1D by exposing human β cells to activated CD4+ T cells. This new model enabled a more detailed analysis of β cell-specific responses to immune-mediated inflammation as well as profiling the complex interactions between immune cells and β cells that occur during insulitis. This underscores the importance of studying β cell responses within a more physiologically relevant context, as it can provide a more comprehensive understanding of the mechanisms driving β cell dysfunction and death in T1D.

The results from the present study provide significant insights into the transcriptomic and epigenetic changes that occur in β cells under inflammatory stress. By analyzing gene expression and chromatin accessibility, I identified numerous regulatory elements and gene networks that are activated in response to inflammation. These findings highlight the dynamic changes in β cell gene regulation and suggest potential pathways that could be targeted to preserve β cell function and prevent cell death.

The broader implications of these findings extend to the uncovering of potentially new therapeutic targets. By integrating the data with genetic studies, I identified specific regulatory elements that may contribute to T1D susceptibility and progression. This knowledge can inform the

design of targeted therapies aimed at modulating the immune response or enhancing β cell resilience in the face of autoimmune attack.

In conclusion, my thesis project provides a comprehensive framework for studying β cell responses to immune-mediated inflammation, bridging the gap between *in vitro* models and the complex *in vivo* environment of T1D. The insights gained from this work pave the way for novel therapeutic approaches and enhance our understanding of the genetic and molecular mechanisms underlying T1D pathogenesis.

Part I

General Introduction

1

The immune system

The immune system has an essential role of protecting our organism from external pathogens to prevent infections and diseases. However, this is not the only role fulfilled by the immune system, being a key player in development, tissue homeostasis and repair[1]. The immune system achieves its functions and ensures system integrity through frequent interactions of two distinct compartments: the innate and adaptive responses. These are characterized by the speed and specificity of their responses, being the former an immediate unspecific defence response while the latter is antigen-specific which takes days to develop. Both compartments are composed of different cells including white blood cells and immune cells, which originate in the bone marrow and migrate to peripheral tissues to exert their function, through blood vessels and the lymphatic system. There are two main categories of white blood cells: the myeloid and lymphoid lineages[2, 3].

1.1 Innate immune response

The myeloid progenitor gives rise to granulocytes (neutrophils, eosinophils and basophils), monocytes/macrophages, mast cells and dendritic cells. These cells are key players in innate immunity, a system comprising different elements as immune cells and proteins that are highly conserved in evolution underscoring their crucial role in survival[2]. The innate immune response is a rapid response by the organism to defend against pathogens. However, due to its lack of specificity, it can sometime damage normal tissues. The main components of innate immune response are:

- **Neutrophils** are the main cell players of the innate response, with phagocytic activity. Their migration to the site of infection happens at early stages of infection or tissue damage, when activated macrophages release cytokines to recruit them, and they end up killing the organisms in their phagolysosome (membrane-bound vesicle with toxic granules).

- **Macrophages** are the mature form of monocytes in the tissues and acquire different phenotypes and functions upon activation. In general terms, they are in charge of cleaning the site of infection or the damage tissue by phagocyting microorganisms or dead cells, recruit other immune cell populations and favouring T cell activation at the site of inflammation.
- **Dendritic cells** are phagocytes specialized in taking up antigens and present them for recognition by the lymphocytes in the lymph nodes. They represent the direct communication between the innate and adaptive responses.
- **Complement** refers to a collection of 20 different serum glycoproteins that engage in an amplified cascade sequence upon activation. This cascade results in the formation of transmembrane pores, ultimately leading to the demise of microorganisms through osmotic lysis. Normal host cells are protected from this action due to expression of inhibitors of these proteins. Moreover, complement proteins play additional roles, including the ability to promote phagocytosis, enhancement of vascular permeability, and chemotactic activity, all of which contribute to bolstering the inflammatory response.
- **Natural Killer (NK) cells** are known as effector lymphocytes of innate immunity. Although they have similar morphology as lymphocytes, they do not require antigen-specific activation. In fact, these are the only cells, along with certain dendritic cells, from the innate immunity that are derived from the lymphoid progenitor. They recognize abnormal or stressed cells either when these cells are coated with antibodies produced by the adaptive response or when the cells lack the MHC I protein on the surface, which is normally present in healthy normal host cells. Once activated, the NK cells release perforins to make holes in the target cell's membrane and allowing granzymes to enter and induce apoptosis[4].
- **Cytokines** are soluble small molecules that function as messengers, transmitting signals from one cell to modulate the behaviour of another cell or even its own activity. Cytokines bind to membrane receptor and send intracellular signals. The biological function of cytokines varies depending on the cytokine and cell involved, acting as autocrine, paracrine or endocrine signals.
    - Chemokines are a subgroup of the cytokine family specialized in leucocyte migration. These molecules are produced by most

cells upon proinflammatory stimulus and act as chemotactic agents, attracting leucocytes to the site of inflammation.

- Interleukins are cytokines produced by leucocytes that are meant to act on other white blood cells.
- Interferons are a subfamily of cytokines secreted to interfere with viral replication in healthy cells and enhance recognition of infected cells (α and β). In contrast, IFNγ is produced only by immune cells and acts directly on the immune system to activate the inflammatory response.

In addition to the above mentioned, there are other cells types involved in the innate response such as eosinophils, that protect the host from parasitic infections and are involved in allergic reactions, or mast cells and/or basophils, which enhance the inflammatory response and are related to anaphylaxis. In general terms, the innate response is not antigen-specific but it is able to recognize patterns associated to pathogens to discriminate between self-molecules and external ones. However, the response is poorly targeted to only extracellular organisms and can eventually lead to collateral tissue damage[2, 3, 5].

1.2 Adaptive immune response

On the other hand, the lymphoid progenitor is the precursor for lymphocytes, mostly B and T cells but also some types of dendritic cells and NK cells. B and T lymphocytes belong to the adaptive response, which is highly specific but requires several days or even weeks to develop fully. Unlike the innate response, the adaptive response involves memory, meaning that upon subsequent exposure to the same antigen, the body mounts a swifter and more robust reaction[2, 6, 7]. This antigen-specific immunity is driven by T and B lymphocytes, although other cell populations are also involved:

- **CD4+ T cells** are the master regulators of the immune system and essential to achieve a regulated and effective immune response. They modulate the activation or inhibition of other immune populations through different subtypes of CD4+ T cells, including the T helper and regulatory CD4+ T cells. CD4+ T helper cells are a type of effector T cell that orchestrates the immune response, activating cell-mediated immune responses once activated by Antigen Presenting cells (APCs). In contrast, regulatory T cells

negatively regulate the immune response after infection resolution to preserve immunologic tolerance to both self and foreign antigens, thereby protecting against immunopathological consequences[8].

- **CD8+ T cells**, also known as cytotoxic T cells, are the other type of effector T cells. They detect and actively kill infected or abnormal cells, such as tumoral cells. They act similar to NK cells, inducing apoptosis in the targeted cell.
- **B lymphocytes** are the cells that detect free antigens and secrete antibodies. These antibodies not only help neutralising toxins or microorganisms but also help enhancing the innate immunity. These cells also help with antigen processing and presentation to activate T lymphocytes.

Although all immune cells need interaction at different levels with other immune cells to get fully activated, it is strictly necessary for T lymphocytes to physically interact with other cells that express MHC molecules to exert their function. These interactions involve different elements (Fig 1):

- **MHC molecules**, also known as HLA (Human Leukocyte Antigen) in humans, are a family of proteins that bind peptide fragments of abnormal proteins within the body, being self or from other microorganisms. These proteins display antigens for recognition by T cells, facilitating the immune response activation. MHC molecules are polygenic and polymorphic, with multiple gene variants, resulting in different combination of molecules varying between individuals. Additionally, the immune system is able to distinguish between self and nonself MHC molecules, activating an allogenic response in case of mismatch with the host MHC combination. There are two types of MHC molecules, class I and class II[3].
- **TCR** is the T cell receptor that recognizes the antigen. It is a complex of proteins that contains the TCR heterodimer subunit, which recognizes the antigen, and the CD3 molecules, the main responsible of transducing the activation signal to the T cell nucleus.
- **CD4/CD8** are co-receptors that stabilize the TCR/MHC-antigen interaction upon antigen recognition and help with the CD3 signal transduction.

- **Co-stimulatory molecules** are receptors on the T cell surface that cause activation of the cell only when the TCR is also engaged with the MHC. Co-activators are required to avoid anergy or cell death upon TCR-MHC interaction. CD28 is one of the main co-stimulatory receptors in T cells.

In the case of CD8+ T cells, they recognize abnormal endogenous antigens presented by host cells in MHC-I molecules and activate the cytotoxic attack targeted only to a specific cell. Whereas for CD4+ T cells, antigen recognition goes through MHC II antigen presentation only by APCs. This is due to the broad effect of CD4+ T cell activation, which leads to cytokine production to activate or inhibit a wide range of other cells, so the reaction needs to be controlled. Besides the need for processing and presenting the antigen, T cells also need the binding of coactivators, like CD28, to ensure that the adaptive response is not activated by innocuous antigens[2, 9]. In summary, when a CD4+ T cell encounters an APC with its antigen, it will first recognize the antigen in the MHC class II molecule through its TCR. Simultaneously, the CD4 molecule in the membrane will contact the MHC class II, identifying the APC as a cell from the host, and stabilizing the MHC-TCR interaction. This will activate the CD3 subunits leading to signal transduction and T cell activation. Additionally, the CD28 molecule on the lymphocyte will bind its ligand (B7 molecules) on the membrane of the APC, activating and enhancing the signals of the TCR. All these events promote T cell activation, leading to proliferation and differentiation into the different subtypes(Fig 1)[8].

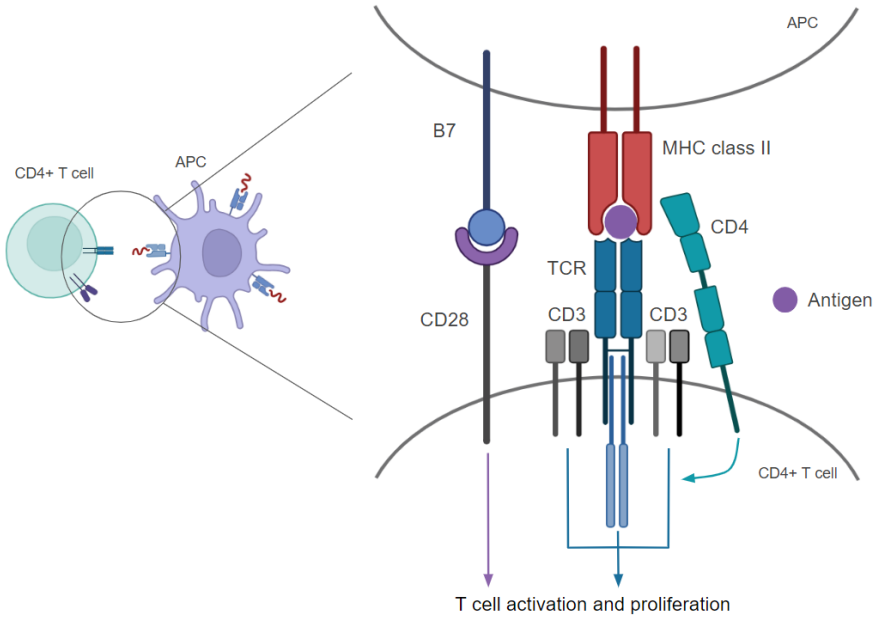


Figure 1. Immune synapse between CD4+ T cells and APCs to induce lymphocyte activation.

In brief, the immune system comprises a dynamic network of lymphoid organs, cells, humoral factors, and cytokines. Its crucial role becomes most evident when it malfunctions: underactivity can lead to severe infections or tumours, while excessive activity can cause allergies or autoimmune diseases, which are the focus of our interest. In this last case, the immune system has checkpoints to ensure that autoreactive receptors that recognize self-antigens are eliminated, a process called tolerance to reduce the risk of autoimmunity. However, due to genetic or environmental factors these checkpoints can fail, increasing the risk of developing an autoimmune disease like Type 1 Diabetes (T1D).

2

Human pancreatic islets

The pancreas is a gland composed of two distinct portions with different functions; the exocrine portion which secretes digestive enzymes and constitutes the majority of the organ, and the endocrine part which consists of pancreatic islets that produce and release different hormones, which represents only 1-2% of the entire pancreas. The pancreatic islets or islets of Langerhans are small organs with island-like structures embedded in the pancreas that play a key role in controlling the glucose homeostasis.

A human pancreas contains approximately one million islets, and although islets have shown great variability in size and cell abundance, each islet contains around 1,000 cells[10–12]. Regarding their composition, islets consist of different types of endocrine cells:

- **β cells** (65-80%) are the most common islet cells. β cells produce **insulin**, a hormone with an important role in the metabolism of carbohydrates, protein, and fat. Insulin promotes the uptake of glucose from peripheral tissues, inhibits glucose release by the liver, increases protein uptake in the muscles and prevents the release of fats. Beta cells release insulin in response to elevated blood glucose levels to counteract it.
- **α cells** (15-20%) produce a hormone called **glucagon**, countering insulin. Glucagon stimulates the release of glucose by the liver and fatty acids from adipose tissue. Its secretion is triggered in response to low blood glucose levels. Once blood glucose levels are normal, the release of insulin inhibits the secretion of further glucagon.
- **δ cells** (3-10%) produce **somatostatin**, which inhibits insulin and glucagon release although its actual role in metabolism is not clear, yet.
- **γ cells (PP cells or F cells)** (3-5%) produce the **pancreatic polypeptide** (PP), involved in regulating both exocrine and endocrine functions of the pancreas.

- **ε cells** (<1%) produce **ghrelin**, which has a role in appetite and adiposity by stimulating food intake, fat deposition and growth hormone release. Other functions have been associated to ghrelin such as glucose and energy homeostasis, cardioprotection and bone metabolism.

In humans, these cell types are highly organized, with a specific 3D tissue architecture, to ensure proper hormone secretion. Human islets are highly vascularized and quite heterogeneous in their composition, in contrast to rodent islets. While rodent islets show a core of β cells in the centre of the islet covered by a mantle of other endocrine cells (Fig 2), human islets show a more intermingled distribution of endocrine cells, forming trilaminar epithelial plates (Fig 2). Blood vessels surround these plates allowing the release of hormones into the blood. This distribution facilitates interactions between different β cells and between α and β cells[13–15].

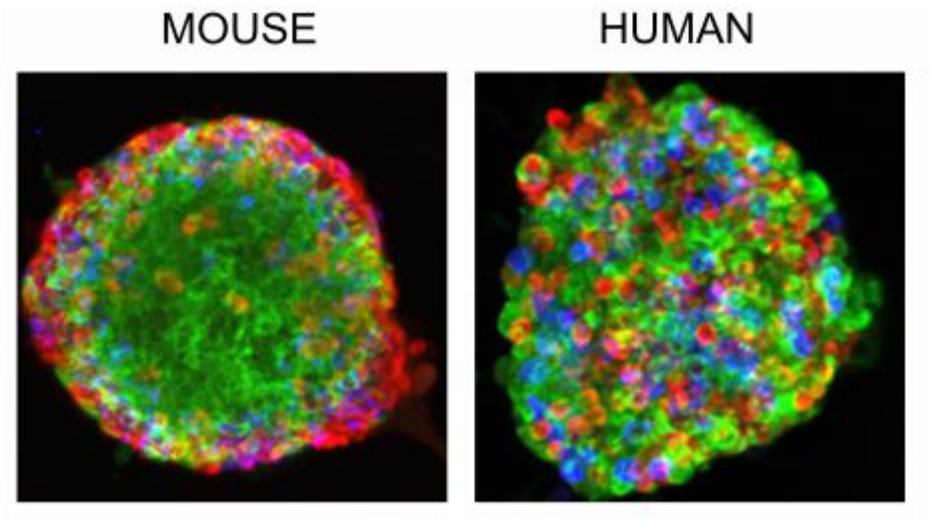


Figure 2. Pancreatic islet from mouse and human immunostained for different cell types. Immunofluorescent staining of insulin (green), glucagon (red), somatostatin (blue). Adapted from Powers and Brissova Research Group website (https://www.powersbrissovaresearch.org/projects).

The importance of the endocrine portion of the pancreas primarily resides in the central role of insulin and glucagon in energy metabolism. Body growth, development and maintenance require an adequate

glucose supply, which is especially important for the central nervous system. Therefore, the body maintains blood glucose levels within a narrow range, which underscore the significance of insulin and glucagon in metabolic regulation. The anabolic function of insulin in reducing free glucose in the bloodstream and incorporation of amino acids into proteins counteracts with the catabolic function of glucagon in a perfectly balanced regulation[16, 17].

Therefore, pancreatic islets are trivial in maintaining glucose homeostasis and dysregulation can cause different diseases. Indeed, one of the most known and common is diabetes mellitus, which is caused by a deficiency in insulin signalling because of a loss of the insulin producing beta cells (Type 1 Diabetes) or a relative reduction of insulin production/secretion and/or peripheral resistance to its effects (Type 2 Diabetes).

3

Type 1 diabetes

Type 1 Diabetes (T1D) is a chronic autoimmune disease that leads to the targeted destruction of insulin-producing β cells in the pancreas. The immune system recognizes the self-β cells as foreign and activates immune mechanisms to attack and eliminate them. As a consequence, patients face chronic insulin deficiency and must rely on insulin injections for life with a reduced quality of life and potential long-term complications[18]. The majority of T1D diagnoses occur when patients experience a life-threatening episode of ketoacidosis, which typically follows the common symptoms of T1D onset: increased urination, increased appetite and weight loss, persistent thirst, and fatigue. Besides the complications at the time of diagnosis, patients might also present long-term complications such as micro- and macrovascular defects along with hypoglycaemic episodes depending on long-term glycaemic management[19].

Historically, T1D has been also known as juvenile diabetes because it is usually diagnosed in children or young adults (age <20 years). Recent reports have shown that there is also an adult-onset form of T1D (age ≥20 years), often misdiagnosed as Type 2 Diabetes. Key differences exist between these two forms of T1D at genetic, metabolic and immune levels that make the adult form less aggressive, but many of these differences are still not well understood, yet[20, 21].

3.1 Epidemiology

Assessing the prevalence, incidence and associated mortality of T1D worldwide becomes difficult but there is an estimate prevalence of 8.5 million T1D patients worldwide in 2021. It represents about 10% of total diabetes cases and evidence suggest that the prevalence and incidence continue to rise globally. The top ten countries with highest T1D prevalence count for 60% of global cases, including USA, UK, Canada, China or Spain, among others. Only in Spain, the estimated prevalence of T1D in 2021 was about 200,000 patients all ages[22–24].

The incidence of T1D varies by geographic region and country (Fig 3). T1D represents the most common form of diabetes in patients ≤ 19 years, accounting for around 70% of new diagnosis of diabetes. The estimated incidence of T1D in children and young adults (<20 years) worldwide was > 350,000 new cases in 2021[25]. For USA and Spain, the incidence was estimated in >18,000 and 1,600 new patients per year, respectively[26].The incidence of adult-onset T1D is not that well characterized since historically T1D has been attributed to be a childhood-onset disease and the challenges to distinguish the adult-onset form from T2D. Recent reports point to a high incidence and prevalence of the adult-onset T1D, higher than the childhood-onset form, and a high percentage of diagnoses of total T1D occur during adulthood [21, 22, 27, 28]. Of note, Sardinian and Finnish populations present the highest T1D incidence rate with >36 cases/100,000 children per year [29].

In 2021, it was estimated that about 175,000 deaths worldwide were caused by T1D across all ages due to various complications. For individuals younger than 25 years, the primary cause of death was due to non-diagnosis. It is estimated that the prevalence of T1D will increase by 66% globally by 2040, also because of an improvement on the diagnosis both of young-onset and adult-onset forms[22].

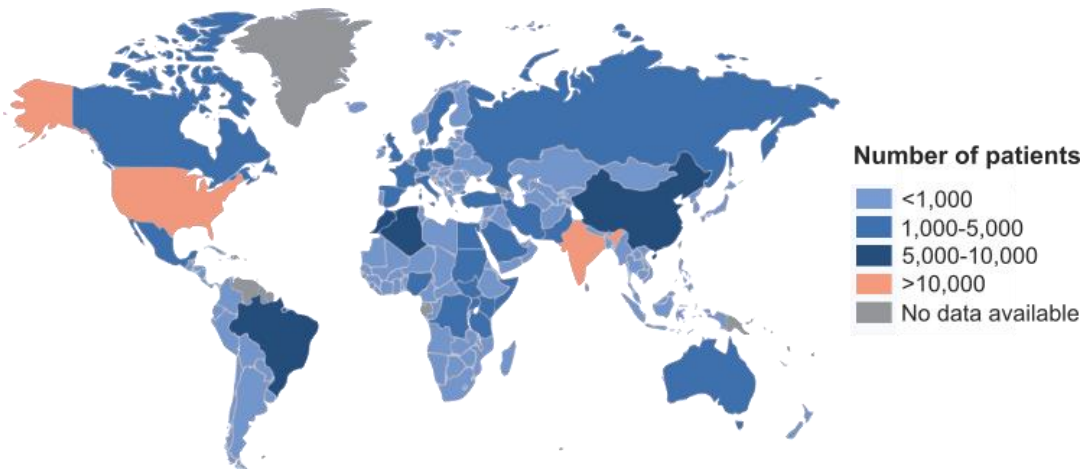


Figure 3. Estimated worldwide incidence of Type 1 Diabetes in 2021. Estimated new T1D cases per country in children (age <20 years) in 2021. Data and graph extracted from the International Diabetes Federation Diabetes Atlas (https://diabetesatlas.org/data/en/indicators/10/)

3.2 Disease pathogenesis

T1D is a chronic disease that follows a continuous progression since the trigger of the autoimmune attack to the appearance of the clinical symptoms. Thus, T1D pathogenic process starts long before, from months to even decades, the clinical onset and diagnosis. This asynchronicity between the trigger of the autoimmune attack on β cells and the clinical diagnosis makes the pathogenesis of T1D complex and difficult to study.

T1D disease pathogenesis has been divided in three stages (Fig 4):

- **Stage 1** is defined by the presence of β cell autoimmunity. Detection of more than one autoantibody (AAb) against insulin (IAA), glutamic acid decarboxylase GAD65 or GAD67 (GADA), IA-2/IA-2β (IA-2A), and/or ZnT8 (ZnT8A) means that the autoimmune attack has already started. However, the effect on insulin production is not critical yet and the patient preserves glycaemic control.
- **Stage 2** is characterized by the onset of dysglycaemia or glucose intolerance. At this stage, the blood glucose levels become abnormal due to the increase of β cell loss.
- **Stage 3** is defined by the clinical presentation of the disease. As a consequence of a significant β cell loss, the patient is no longer able to regulate glucose homeostasis. Symptoms of hyperglycaemia (polyuria, polydipsia, weight loss, blurred vision and fatigue) might be accompanied with diabetic ketoacidosis.

Stage 1 and 2 are still presymptomatic while stage 3 often coincides with the clinical diagnosis. This classification was adopted to standardize T1D diagnosis and facilitate the study of the pathogenesis at the different stages as well as for risk assessment, clinical trials and precision medicine[19, 30].

It is believed that one or more environmental factors such as viral infection, diet or toxins could trigger the autoimmune attack against β cells. The initial immune response to the trigger might lead to the activation of self-reactive immune cells leading to T1D pathogenesis.

Islet autoantibodies detection has been a really useful tool not only to stratify and diagnose T1D patients but also to develop predictive models to determine individuals at risk of T1D. Additionally they can be used as

markers of prognosis in islet transplantation or markers of treatment efficacy in clinical trials.

For children at risk, seroconversion can present two scenarios: a single autoantibody, often IAA or GADA, or, more commonly, multiple autoantibodies. Additionally, quantification of AAbs at time of seroconversion has been associated with risk of developing T1D. Those patients that show higher levels of IAA or IA-2A are at greater risk of finally developing T1D symptoms compared to patients with lower levels of the same AAbs. Finally, early age of seroconversion also correlates with increased T1D risk[30, 31].

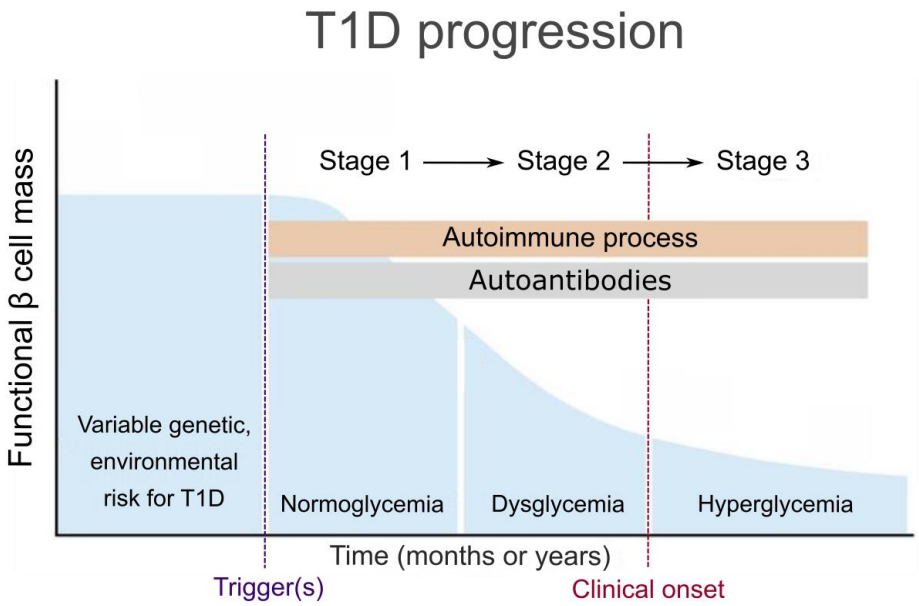


Figure 4. Schematic view of T1D progression. Adapted from Powers, AC. 2021[32].

Although AAbs represent a powerful biomarker tool to study the disease progression, their active role in T1D pathogenesis is not clear, yet. This is exemplified by the fact that having ≥ 2 AAbs is a determinant for T1D clinical progression, yet not all individuals at risk develop the disease[33]. Actually, T1D is considered a T-cell mediated disease while it is not yet clear whether the AAbs production has a functional role in the disease pathogenesis.

3.3 Inflammation in T1D

The inflammation within the pancreatic islets is called insulitis and it is one of the key features of T1D. T1D is heterogeneous, which translates into a very diverse infiltration process, both in the number of affected islets and the distribution of preserved β cell mass throughout the pancreas.

Several observations across different models of T1D, insulitis have been categorized into three distinct stages[34, 35]:

- **Induction** of the immune response. It is the first stage of inflammation corresponding to the activation of the innate immunity, likely in response to molecules from external microorganisms. This initial activation of the immune response is typically followed by an increase of the levels of interferons (IFNs).
- **Amplification** of the immune response in early stages of insulitis. The increase in IFNs causes cascade of autocrine and paracrine responses in the β cell that result in attraction of other immune cells. In this transition from innate immunity to an adaptive response, cytokines and chemokines are essential players that will determine the dialogue established between β cells and immune cells. As an example, the levels of certain chemokines (CXCL10, CCL3, CCL4, CCL2 or CCL7, among others) are increased in the serum of T1D patients. However, most of the events in early insulitis take place within the islet so examination of peripheral cells or blood samples usually is not entirely representative of disease pathogenesis.
- **Maintenance or resolution**. It consists on stabilization and maintenance of inflammation or its resolution. This stage is mainly dominated by the adaptive immune system. In the case of resolution, mild insulitis is disrupted and β cell function is restored. If the loop of chemokine production and immune cell activation does not stop, it will likely evolve to progressive accumulation of immune cells within the islet.

In healthy conditions the induction and amplification phases of insulitis are meant to eradicate an infection event. Each of these stages can be resolved, impeding T1D progression. However, in at risk T1D individuals, these mechanisms might malfunction. Inflammation may be enhanced, and there could be a reduced capacity to resolve each stage

of the inflammatory status, eventually leading to β cell loss and the induced exposure of autoantigens. Indeed, β cell death seems to happen simultaneously to strong inflammation in the islet.

3.4 Aetiology of T1D

The complexity and high heterogeneity of T1D limits our understanding of its aetiology. It is known that T1D development is influenced by many different known and unknown environmental factors and certain genetic predisposition. All these factors contribute to a multifactorial interplay that ultimately trigger an aggressive autoimmune attack against pancreatic β cells and clinical onset of the disease. Yet, our knowledge on the genetic and environmental mechanisms that lead to T1D and whether these factors serve as triggers, facilitate, or exacerbate the autoimmune reaction remains limited.

3.4.1 Environmental factors

In the past decades, the rapid increase in the incidence of T1D and the observation that immigrants tend to acquire a T1D risk similar to the native ethnicity[36], support the idea that different environmental factors are likely involved in T1D pathology. Many different factors have been shown to influence T1D pathogenesis, some of them have been strongly implicated with the risk of developing T1D.

Viral infections have been studied for decades since several evidence point to enteroviral, rubella and mumps infections as a key environmental factor. Most recent studies have shown that T1D patients show increased levels viral RNA and anti-enteroviral antibodies months before seroconversion and some studies reported detectable infections in the pancreas and islets of recently diagnosed patients[37, 38]. Many other circumstantial evidence supports the idea that a viral infection might trigger islet autoimmunity. For example, persistent infection of β cells with CVB4 affects insulin maturation, causing the release of antigenic forms of proinsulin and altering β cell functions[39]. Enteroviruses are also associated with β cell damage by cytolysis during viral replication and increased antigen presentation in response to infection[40]. Additionally, the genetic predisposition to T1D has been in part associated with β cell response to enteroviral infection[19, 41].

Nevertheless, a single viral infection is not likely to trigger autoimmunity alone and probably acts simultaneously with other disadvantageous factors[19, 40, 41].

Gut microbiome has also been the focus of many studies since the gut microbiota plays an important part in the development of innate and adaptive immunity and in maintaining the mucosal barrier. Their actual role in T1D pathogenesis or whether it is a risk or protective factor is not fully understood. It will likely depend on the microbiome composition and the type of mediators produced to facilitate the crosstalk between them and the host. In T1D patients, there is a dysbiosis (reduced diversity) of the microbiota, with an increase of Gram-negative bacteria that produce LPS, a proinflammatory compound. This dysbiosis can have other effects such as increased intestinal permeability and microbial components and exogenous antigens might reach the bloodstream promoting inflammation and enhancing autoimmunity[19, 42].

Nutritional factors such as diet are also very important. Diet is a key determinant of microbiome composition, and therefore stimulating an appropriate immune response. For example, cow milk, introduction of cereals and eggs or decreased vitamin D intake have been attributed to a higher risk of developing T1D. On the other hand, breastfeeding, vitamin D, nicotinamide and other vitamins have been linked to T1D risk protection. Indeed, exposure to maternal microbiome during birth and breastfeeding are of upmost importance during the establishment of mucosal immunity of newborns[19, 43, 44]. Obesity might also be related to T1D risk, especially in those children with reduced genetic risk[45].

3.4.2 Genetic susceptibility

Reflecting its multifactorial origin, T1D does not follow a Mendelian inheritance pattern. T1D has a significant hereditary component evidenced by up to 70% twin concordance[46] and an approximate 8% risk among siblings[47, 48]. Likewise, relatives of T1D patients show a 15 times higher risk of developing the disease[49].

Genome-Wide Association Study is a method used to identify associations between single-nucleotide polymorphisms (SNPs) common in the population and particular traits or diseases. For GWAS, large cohorts of healthy controls and patients are compared to identify

genetic variants that are more common in people with a particular trait or disease. Thus, detection of genetic association signals by GWAS generates a set of common and low-impact risk variants associated to the disease. In contrast to monogenic diseases, the link between these associations and causal genes is not direct and the pathway to personalized medicine becomes more complex. However, GWAS brings the opportunity to study and understand pathogenic molecular mechanisms and highlight important genes in the development of T1D.

As for other multifactorial traits, collection of all these T1D-associated variants is useful to predict individuals at risk of the disease. To do so, polygenic risk scores (PRS) are used to determine the overall risk of certain individual. According to its contribution and direction of effect, each variant associated to disease has a specific value. Thus, PRS take into account the inherited combination of genetic variants associated to T1D in a given individual to calculate the predisposition to finally develop the disease. This strategy offers the possibility to develop surveillance protocols for at risk individuals and even selection of candidates for clinical trials focused on early stages of autoimmunity[50–52]. Importantly, most genetic studies are conducted in specific ethnic populations, usually with European ancestry, so the same PRS have limited transferability across different and underrepresented populations[47, 53].

GWAS studies have unveiled more than 60 different regions associated with T1D risk, which explain nearly 80% of disease heritability[54, 55]. These regions contain SNPs that can be classified into the HLA-associated variants and other non-HLA variants with lower impact.

3.4.2.1 HLA loci

The HLA-associated variants were the first ones associated to T1D risk and carry the highest impact on the risk of developing the disease[56, 57]. As already explained in the first section, HLA molecules are the Major Histocompatibility Complex (MHC) in humans. These molecules are highly polygenic and polymorphic, with different gene variants, so the combination of molecules varies a lot between individuals.

HLA proteins are key for antigen presentation to the immune system and can be classified in two main types of molecules:

- **MHC class I** are the proteins in charge of presenting endogenous antigens in host cells to be recognized by CD8+ T cells. Each individual presents three classical HLA I proteins: HLA-A, HLA-B and HLA-C. Non-classical ones are HLA-E, HLA-F and HLA-G.
- **MHC class II** are usually expressed in antigen-presenting cells to present exogenous phagocytosed proteins and activate the adaptive immune response through CD4+ T cells (section 1.2). Five different genes compose this subgroup: HLA-DP,HLA-DQ, HLA-DR, HLA-DM, and HLA-DO, although the last two ones are not usually profiled to determine the haplotype because they are not highly polymorphic

There is a third class, MHC class III, that is composed of proteins related to the complement system with important functions in the innate immune response (section 1.1).

The haplotype is the combination of HLA antigens inherited by an individual from one parent. As mentioned, there is a vast number of alleles for each HLA gene although this does not translate into a huge variety of different haplotypes. This is due to linkage disequilibrium events, a phenomenon in which two or more alleles tend to occur together more than expected rather than following independent segregating.

In T1D, HLA genes are strongly associated with risk and contributing to about 50% of the total heritability[48, 58]. However, not all HLA molecules show the same strength or direction in this association. Therefore, the patient risk is defined by the haplotype combination inherited from both parents.

The major contribution to the risk comes from **HLA class II** and almost all T1D cases, around 90%, exhibit HLA class II risk haplotypes. The variants with strongest association are located in the highly polymorphic pocket sites of the **DQ** and **DR** molecules. It is believed that the different polymorphisms can affect the affinity of the HLA for specific peptides or the generation of cross-reactive antigens, when two different peptides can be recognized by the same TCR. The most common associated haplotypes vary between different populations. As an example, in the Caucasian population, the DR4/DQ8 and DR3/DQ2 molecules are the most common haplotypes conferring susceptibility, specifically the DQ2/DQ8 show higher risk of T1D. In contrast, another association is

the DR1501-DQ6 haplotype, linked to a protective effect owing to a higher frequency of islet-specific T regulatory cells[19, 47].

In addition, certain HLA class II haplotypes have also been associated with the age of seroconversion or the appearance of specific autoantibodies. For instance, the DR4/DQ8 haplotype is linked to the IAA as first autoantibody at seroconversion, while the DR3/DQ2 is associated with the GADA[59, 60]. However, HLA II haplotypes have not been associated with other pathological traits such as age of onset or C-peptide loss[61].

Besides HLA class II, **HLA class I** genes, specifically HLA-A and HLA-B, have also been genetically associated with T1D risk, although with relatively modest effects. It is suggested that these HLA I risk haplotypes impact recognition and destruction of β cells. For example, HLA-A*24 allele is linked to lower residual β cell function, likely due to an increased immune-mediated destruction[57, 62]. Considering the nature and function of HLA class I proteins, these variants are likely influencing late stages of T1D pathogenesis, when an autoimmunity process is already in place.

3.4.2.2 Non-HLA loci

The other half heritability in T1D is explained by many other signals scattered across the genome that (with the relevant exception of two loci namely *INS* and *PTPN22*) present lower individual impact on the overall risk. This makes the detection of these loci more difficult although recent advances in DNA sequencing and powerful GWAS have help identifying many new associated signals.

The insulin locus (*INS*) was the first non-HLA region detected through small candidate gene association studies and represents the second region with strongest risk association, accounting for 10% of T1D susceptibility. The variants are clustered in the 5’ upstream section of the gene and most of the signal comes from the variable number of tandem repeats region (INS-VNTR). Alleles with shorter sequences are associated with increased risk, altering (pro)insulin expression in the thymus and affecting the self-tolerance immune mechanisms. In contrast, long alleles are associated with a protective effect due to an opposite role thus enhancing central immune tolerance[63, 64].

The other locus with strongest association signal, map to the *PTPN22* gene, which encodes a negative regulator of T cell receptor (TCR) signaling. Other relevant non-HLA T1D-associated loci include the *CTLA4* locus, a negative regulator of CD8+ T cell function, the interleukin 2 receptor alpha (*IL2RA*), ubiquitin-associated and SH3 domain-containing protein A (*UBASH3A*) and interferon-induced helicase c domain-containing protein 1 (*IFIH1*)[60, 65, 66]. Table 1 provides a list of potential genes that have been associated with T1D pathogenesis.

The use of GWAS is crucial for uncovering the genetic background of T1D and it has facilitated the discovery of new non-HLA regions linked to T1D risk. GWAS results offer valuable insights into the molecular mechanisms driving the disease pathogenesis. However, GWAS data is its interpretation is challenging. The detection of causative risk variants within the genetic signal of association is difficult due to linkage disequilibrium events. Additionally, assessing the individual impact of each the disease risk and understanding the underlying molecular mechanisms can be complicated by the fact that most GWAS-associated variants lay in non-coding parts of the genome, likely affecting gene regulatory networks rather than the gene code[67]. Finally, the ultimate goal is twofold: to translate the genetic signals associated with disease risk into precision medicine for T1D prediction and to apply the understanding of molecular mechanisms to target treatment for prevention or treatment[53].

Since most risk variants fall in the non-coding genome[67], additional efforts are needed to understand the gene regulatory pathways affected. Nonetheless, the gene regulatory landscape is highly dynamic, being specific for each cell type, and cell state. This implies knowledge of not only the regulatory networks active in each tissue or cell population implicated in disease but also how these regulatory networks may change upon disease development. Furthermore, associating non-coding variants to their respective gene targets requires a deeper understanding of tissue-specific and state-specific enhancer-promoter interactions. Addressing these challenges will improve our understanding of the mechanisms driving T1D pathogenesis.

Different studies have conducted co-localization analysis of T1D variants and regulatory elements active in distinct cell populations with the aim of uncovering the cell population and cell state capturing T1D

associated signals and their potential target genes [66, 68]. Recent association studies have focused on linking detected genetic signals with regulatory maps from immune and pancreatic cell types[69]. Early association studies [70–72] mapped T1D risk variants to enhancers active in different immune cell populations such as T cells, thymus and B cells. T1D risk variants have been primarily reported to be enriched in resting and stimulation-responsive T cell enhancers including the detection of a subtle enrichment in enhancers specifically active in regular pancreatic islet [71, 73]. Overall, these studies reported over 50 different non-HLA loci with a candidate gene assigned using co-localization and Quantitative trait loci (QTL) approaches.

The latest GWAS study [74], combined a large cohort of T1D patients with regulatory maps charted in pancreatic and immune tissues and obtained at the single-cell level. The increase in sample size and the use of single cell technology to map enabled the authors not only to confirm the enrichment of T1D-associated variants in T cell enhancers but also to identify specific enrichment in regulatory elements active in non-endocrine populations of acinar and ductal cells. Notably, the exocrine active enhancer bearing risk variants were found to be associated with genes specifically expressed in those populations, such as the *CFTR* gene. This was the first study linking T1D GWAS genetic risk to the exocrine portion of the pancreas.

Finally, the different genetic association between T1D and T2D has always been under study since β cells show a partially shared phenotype. Newer studies found several shared risk pathways or genes that interact to modulate islet functions and lead to β cell dysfunction in disease[19, 75, 76].

None of these studies linking T1D genetic risk with regulatory maps charted in disease-related tissues in basal conditions, captured enrichment of T1D-risk variants in pancreatic islet enhancers. These findings are supported by other studies showing that T2D only (and not T1D) risk variants are enriched in pancreatic islet enhancers[77]. Of note the studies above mentioned only explored unchallenged regulatory maps and did not include disease-relevant states. Increasing genetic evidence points to the implication of β cell in the predisposition to T1D [78, 79] which will be further discussed in section 6.

Candidate gene	Gene function
Immune-related	
CCR5	Involved in T(h) development and chemokine-induced signalling
CD226	Modulates thymic T cell selection (immune tolerance) Involved in effector CD8+ T cell activation and function Reduced functions of Tregs
CD69	Participates in lymphocyte activation and limits the inflammatory response Linked to the signalling of NK cells
CTLA4	Controls the proliferation of Tregs in the periphery and regulates pancreas autoimmunity
IFIH1	Mediates the response to virus by innate immune system, inducing interferon production Involved in the β cell response to viral dsRNA
IKZF1	Regulation of immune cell development
IL2/IL21	Implicated in T(h) cell differentiation and inflammatory response
IL27	Modulates T cell subsets, regulating inflammatory process
IL2RA	Receptor of IL-2, the main cytokine involved in T cell function. Variants cause different sensitivities to IL2 Potentially alters the balance between Tregs and Teffs
IL7R	Involvement in antigen binding, Ig production, and cytotoxicity
PRKCQ	Involved in T cell function, apoptosis and the innate immune response
PTPN22	Participates in TCR signalling pathway and central tolerance pathways
UBASH3A	Downregulates the NF-kB signalling pathway upon TCR stimulation, reducing the IL2 expression

Other genes

BACH2	Regulates proinflammatory cytokine-induced apoptosis in pancreatic β cells
C1QTNF6	Inflammation-related gene linked to the pathogenesis of tumors and arthritis, involved in proliferation and apoptosis regulation
CLEC16A	Regulates mitochondrial quality control with a putative involvement in β cell fragility
CTSH	Regulates cytokine-signalling inside β cells and proapoptotic signal transduction Protects from immune-mediated damage
ERBB3	Modulates antigen presentation and cytokine-induced β cell apoptosis
GLIS3	Involved in the generation of β cells, insulin expression, maintenance of β cell functions and mass It also shows antiapoptotic effects
HIP14	Regulates β cell apoptosis and insulin secretion
NRIR	Functional lncRNA likely acting as a negative regulator of interferon response
PTPN2	Induces β cell apoptosis due to increased local levels of interferon
SH2B3	Growth factor and negative regulator of cytokine signalling
STX4	Associated with insulin secretion Downregulates the expression of chemokine genes in inflammation Decreases apoptosis in pancreatic islets by reducing the translocation and activation of NF-kB
TASP1	Important for proper HOX gene expression, essential for MLL protein cleavage
TNFAIP3	Downregulates the intrinsic apoptotic pathway Regulates the expression of ZnT8 Essential for insulin production and secretion

TYK2	Regulates the proapoptotic effects of cytokines inside β cells Mediates interferon response in viral infections Mediates the immune reaction of different subtypes ofCD4+ T cells

Table 1. Selection of genes associated to T1D and their putative roles in immune or pancreatic pathways. *HLA* and *INS* loci are not included in the table. Adapted from Zajec, A. *et al* 2022[19].

3.5 Current therapies for T1D

The current gold standard treatment for T1D is the exogenous administration of insulin and lifestyle adjustments, which include regular exercise and a low-carbohydrate diet. Strict glycaemic control using insulin therapy can prevent ketoacidosis, delays the progression of hyperglycaemia-related complications and lower patient's mortality[80]. This treatment is focused on substituting the missing endogenous insulin and represents the only effective treatment for patients with no alternative that actually cures the disease. However, risk of severe hypoglycaemia episodes represents the major adverse effect of insulin treatment as maintaining good metabolic control requires continuous glucose monitoring[81, 82]. Many of the advances in T1D clinical care come from development of new devices to administer insulin or accurately measure blood glucose levels. Nevertheless, none of the current strategies on insulin administration avoid long-term complications so other alternatives have been explored to develop a treatment to stop or slow down the autoimmune attack. Nonetheless, the seek of effective treatments for T1D is extremely challenging especially due to the heterogeneity in both aetiology and disease pathogenesis.

Once the autoimmune nature of T1D was clarified, the focus for therapeutic efforts have been directed towards drugs designed to disrupt or modulate the autoimmune response with the aim of preserving the β cell mass. Numerous clinical trials involving the use of a wide range of immunosuppressive agents have been conducted over the years but with limited success. Most of the trials achieved a delay on onset or

progression of disease but only one was successful enough to be approved by the FDA for health care use. Teplizumab, a monoclonal antibody anti-CD3, delays the progression of type 1 diabetes and its onset in adults and children over 8 years old with stage 2 (pre-symptomatic) disease. It is the first drug approved for delaying the onset of any autoimmune disease before clinical symptoms appear[83, 84]. Although many reasons could account for the unfulfilled aim in immunotherapies, the main conclusion to highlight is the lack of complete understanding of T1D pathogenesis[85, 86]. Additionally, immunotherapies should be addressed for presymptomatic patients since most of the β cell mass is lost by the time of diagnosis.

Alternatively, transplantation of pancreas or islets from humans represents an alternative that has been proven successful in providing insulin independence in 70% of transplanted patients. However, a number of challenges limit the clinical implementation such asthe scarcity of donors and the significant side effects of immunosuppressors to protect the transplanted tissue. Xenotransplantation of islets from pigs has been explored and proven partially successful although with great limitations, too[87, 88].

The use of stem cells (SC), either human embryonic stem cells (hESCs) or induced pluripotent stem cells (iPSCs), has gained support in the last few years as it represents another alternative to human islet supply. Both sources of stem cells become a source to generate new β cells which can be transplanted into T1D patients. The islet-like organoids obtained contain functional glucose responsive insulin-producing cells. While the use of hESCs is ethically more controversial, the use of patient-specific iPSC-derived β cells is a promising tool for allogenic transplantation that can also prevent allograft rejection. When combined with genome editing techniques for gene correction is a promising strategy for T1D treatment.

Human islet isolation and post-transplantation protocols have been optimized very carefully in the past decades to improve effectiveness. However, there are still limitations, one of them being the survival of islet graft after transplantation. One promising technique is the encapsulation of islets before transplantation to physically protect them from the immune response of the host and, thus, avoid the need for immunosuppressors for life.

In 2021, Vertex Pharmaceuticals reported the results of first clinical trial of iPSC-derived islet transplantation (VX-880). The islets were transplanted into the liver of patients with no insulin production previous to the transplant and post-transplantation treatment included with immunosuppressant drugs. Transplanted patients showed improved levels of insulin production and glucose homeostasis, with even some patients able to be completely independent from exogenous insulin administration. Although still in early stages, research data suggests that iPSC-derived islets are a promising and potentially groundbreaking treatment for type 1 diabetes.

Vertex is also running another clinical trial using the same islets but encapsulated (VX-264). The company is also running preclinical studies combining the iPSC-derived islets with gene editing techniques to modify the islets and “hide” them from the immune system[89]. Both strategies focus on optimizing the protection of the implanted islet cells from the immune system as an alternative to immunosuppressant drugs.

In summary, many promising advances in T1D treatment have emerged over the past decades, a deeper understanding of the complex interplay between environment, genetic predisposition and immune dysregulation underlying T1D is essential for the successful clinical application of these strategies. A greater understanding of the precise mechanisms of disease progression will enable identification of new pathways relevant to the autoimmune attack and β cell death which can be modulated in SC-derived islets for cell replacement therapy.

4

Cells populations involved in T1D

4.1 T1D and the immune system

Being T1D an autoimmune disease, the immune system plays a crucial role in the pathophysiology of the disease. Decades ago, it was determined that the primary cause of beta cell demise is the immune attack by CD4+ and CD8+ T cells. T1D patients are believed to inherit an immune system that is genetically predisposed to recognize β cell antigens[90, 91]. The role of autoreactive T cells in this disease has been widely studied resulting in histological evidence of lymphocyte infiltration in the islets of T1D patients.

T1D has strong genetic association with HLA class II genes and HLA class I genes, along with other genes related to the maturation and control of the immune system. In fact, HLA associated variants show the highest impact on the risk of developing T1D with some haplotypes clearly associated with increased or reduced risk of disease. On the other hand, several pathways controlled by genes associated to T1D include those implicated in the establishment of central immune tolerance, in charge of reducing self-reactive T cells, or CD4+ T regulatory cell stability and function, which ensure peripheral tolerance. Diverse studies have found that peripheral Tregs for T1D antigens are found at similar frequency in both patients and controls but with reduced function in T1D[79, 92, 93].

Other immune cells have also been described to infiltrate the human islet in the context of T1D, such as B cells or NK cells, but their functions and interactions with other cell populations in disease progression are not well understood. Dendritic cells and macrophages are well defined as core components of human insulitis since they have an active role in connecting the innate and adaptive response, although their altered phenotypes and dysfunction in this context is still under study[91]. Additionally, macrophages have been the focus of several studies aimed at characterizing the early stages of T1D due to their role in releasing proinflammatory cytokines in response to β cell antigens that, in turn,

activate both the adaptive response and the β cell stress phenotype[79, 94].

In summary, it is clear that CD4+ and CD8+ T cells have a very active and essential role in the T1D. Self-reactive effector T cells infiltrate the pancreatic islets, with CD4+ lymphocytes orchestrating the immune attack and the CD8+ encompassing the direct killing of β cells. Experiments on mice demonstrated that β cell destruction in T1D requires the activation of both CD4+ and CD8+ effector T cell populations; importantly, independent activation of either population alone was insufficient to induce T1D[95]. On the other site, dysfunction of CD4+ T reg cells also seems to be a key factor as it was observed that in T1D, an effective suppression of self-reactive effector T cells cannot be achieved[96]. On the same line, cancer patients treated with immune checkpoint inhibitors, interfering with peripheral tolerance pathways, are at risk of developing an acute form of T1D[79, 96, 97]. Additionally, other immune cell populations, such as NK cells, B cells and dendritic/APC cells, are found to infiltrate islets of T1D patients although their roles and dysfunction in disease progression is not completely defined yet[96].

The immune system in T1D has been studied for decades, however the factors driving islet immune infiltration still remain poorly understood. In fact, it has been reported that pre-(pro)insulin reactive CD8+ T cells are present in the pancreas of healthy donors although at lower levels compared to T1D patients[98] and that autoreactive T cells are part of a normal T cell repertoire in healthy individuals[79, 91, 96]. These findings depict the complexity in understanding the role of immune cells in the pathogenesis of T1D and their interactions with the target β cells in the pancreatic islets. Thus, justifying the current efforts to investigate other cell populations besides those composing the immune system, potentially involved in the development of the disease.

4.2 Pancreatic cell populations

In last decades numerous pieces of evidence have supported the concept that the immune system is not solely responsible for T1D. Instead, **β cells** could also actively contribute to disease progression. The hypothesis that β cells contribute to their own demise was first proposed in the late 80s[99]. Studies analysing pancreas from T1D

donors show that β cells acquire a phenotype in response to stress characterized by increased endoplasmic reticulum (ER) stress and expression of unfolded protein response (UPR) markers while trying to restore cellular homeostasis[19, 79, 100, 101].

Additionally, it was shown that more than 80% of T1D-related candidate genes are expressed in β cells. Likewise, increasing genetic-based evidence also highlights the potential and important role of β cells in T1D development[19, 102]. For example, novel risk genes involved in the progression from islet autoimmunity to clinical onset are involved in interferon signalling in β cells in response to viral infections[19].

Activation of all the response pathways help the β cell cope with environmental stimuli but, due to their ability to constantly produce large amounts of insulin, they are inherently very sensitive to stress. If the proinflammatory stimulus becomes constant, it might eventually lead to β cell death[103–105]. A constant proinflammatory stimulus disrupts the activity of β cells, mainly based on insulin synthesis and release[106, 107]. Specifically, the prolonged inflammatory context can induce the loss of β cell identity, decreasing expression of key β cell marker genes, and loss of insulin production[34, 108]. Moreover, the coping mechanisms of β cells in insulitis might participate in the amplification of the immune attack and its own destruction. In fact, exposure to a proinflammatory environment induces expression of HLA class I and class II molecules, which in combination with presentation of aberrant antigens (neo-antigens) is now a current research focus of T1D pathogenesis[109–112]. Many altered processes contribute to the increased diversity in β cell neo-antigens such as alternative splicing events in gene transcription, higher error rate in mRNA translation into proteins or increased deposition of post-translational modifications to proteins[19, 79, 103, 113].

These neo-antigens can also be released into the environment as peptides or proteins which will be processed by APCs, exacerbating the immune activation. Over-expression of HLA class II molecules may also aggravate the immune response since β cells can themselves act as APCs. β cells also activate other pathways related to immune regulation such as overexpression and secretion of CXCL10, a chemokine with a potent role in T cell attraction, and other cytokines/chemokines.

Other pancreatic cell populations have also been studied, especially in primary islets from T1D donors using imaging and single cell techniques.

For example, α cells have also been reported to exhibit signs of immune-related stress such as altered gene expression, like overexpression of proinflammatory cytokines (e.g. CXCL10), or impairment of their endocrine function. However, β cells seem more sensitive to this process, suggesting a highly specific attack to insulin-presenting cells and/or an inherent fragility of the β cell[114, 115].

Finally, implication of other pancreatic populations in the development of T1D cannot be discarded either. A recent study found that some of the T1D risk variants overlap enhancers active in the ductal and acinar cells composing the exocrine pancreas. In addition, these enhancers were linked to genes only expressed in those specific populations[74].

In summary, extensive research suggests that β cells activate numerous pathways related to immune regulation, ER stress and apoptosis. These pathways likely result in increased antigen presentation, exacerbating the autoimmune assault. While the contribution of other pancreatic cell populations to disease pathogenesis still needs additional investigation, emerging evidence offers preliminary insights into their potential involvement.

Nowadays, it is becoming evident that T1D pathogenesis involves both immune cells and pancreatic β cells. Thus, understanding the interaction and crosstalk between them, especially at early stages of disease, is crucial for comprehending the molecular mechanisms leading to disease.

5

Different models to study T1D

The endocrine pancreas is difficult to access and isolate. In humans, pancreas biopsies can result in acute pancreatitis, as a result, biopsies in living donors are only performed under special circumstances, and most islets used for transplants or research are obtained from deceased donors. In addition, human pancreatic islets are embedded within the exocrine pancreas, comprising a very small portion of the total pancreatic mass. Therefore, obtaining a viable sample requires a complex and time-consuming isolation protocol developed more than 35 years ago[116].

Furthermore, since T1D pathogenesis is believed to develop over years before its clinical onset, early detection of individuals at risk is challenging. This is feasible to some extent in close relatives of diagnosed T1D patients but remains a significant challenge in the general population. Obtaining samples from T1D patients is already difficult, and collecting samples from patients in presymptomatic phases to study different disease stages is impractical.

Unlike other diseases, obtaining relevant blood biomarkers for personalized medicine or understanding mechanisms underlying disease pathogenesis has also been challenging. Since islets represent less than 2% of the pancreas, any local inflammatory process will inevitably be diluted in the peripheral circulation. Nevertheless, the presence of AAbs in the peripheral blood was shown to be a disease predictive marker. However, the role of AAbs in disease pathogenesis is unclear, and not all AAb+ individuals will eventually develop T1D.

For all these reasons, different *in vitro* and *in vivo* animal models have been developed over the years as an approach to study and comprehend the molecular mechanisms behind T1D[117].

5.1 Animal models

The **nonobese diabetic (NOD) mouse** model develops **T1D spontaneously** and is the most widely used T1D animal model. NOD mice develop T1D due to a high genetic risk associated with different gene polymorphisms found in genes that are also relevant in the human pathogenesis. The disease onset is around 12 weeks of age but, as for humans, the age of onset is highly heterogeneous. Regarding the initial stages of the disease, NOD mice show extensive immune infiltrates with similar characteristics as those observed in human insulitis. However, there are differences: NOD mice islets contain less CD8+ T cells and more CD4+ T cells compared to humans. Other immune populations have also been described to contribute to disease development in NOD including B cells producing AAbs against insulin[117].

This model has been extremely useful to understand the basic mechanisms of T1D pathogenesis because complex processes that involve many organs cannot be studied *in vitro*. The NOD mice have been particularly useful for studying the genetics of T1D and to understand how candidate genes expressed by immune cells or islet cells influence the autoimmune attack and to uncover the effect of environmental factor such as gut microbiome or infections. Additionally, the use of this model has also enabled the identification of many β cell antigens, which later have been confirmed in human patients.

Other models that spontaneously develop T1D upon an environmental perturbation include different **rat breeds**, with high and low incidence of the disease. Some **dog** breeds also develop T1D spontaneously at a prevalence similar to humans and following a pathology closer to humans compared to rodent models. However, due to the costs and difficulties in genetic manipulation, the canine models have been very valuable mainly for late preclinical studies[117].

Chemically induced T1D had also been used to model the disease pathogenesis. This approach was primarily applied to rodents, but in some cases to non-human primates. The animals were treated with agents to chemically induce T1D, mainly alloxan, streptozotocin (STZ) or cyclophosphamide. These diabetogenic agents generate pancreatic β cells toxicity resulting in hypoinsulinemia and hyperglycemia within a period of days. STZ causes DNA fragmentation on β cells and it is the most commonly used agent. The STZ rodent model reproduces both acute and chronic complications of human T1D. However, due to the

oncogenic effect of STZ, the animals show increased risk of insulinoma, kidney and liver tumors formation, amongst other complications. Thus, chemically induced models show some limitations which include the lack of disease heterogeneity observed in T1D patients in terms of disease variability and disease onset, the risk of toxicity and death of the animal. Moreover, the application of these models can be challenging in keeping consistency and reproducibility[117–119].

Other animal models including rodents and primates take advantage of the use of **viral infections to induce T1D**, both in genetically predisposed and non-predisposed animals. However, these models are not widely used since the infection is not only targeted to β cells and the complete mechanisms for getting T1D are not understood[119]. Finally, in order to study the behaviour of the human immune system in T1D *in vivo,* **humanized mice models have been used**. In this approach immunodeficient mice are transplanted with human immune cells to reproduce immune tolerance and induction of autoimmunity[117, 119].

Different non-human **animal cell lines** derived from pancreatic β cells have been developed over the years. Different strategies have been used to immortalize the β cells such as viral infections or radiation. Some of the mostly used β cell lines are: the **MIN6** mouse cell line; the **HIT** hamster cell line and the **INS-1** rat cell line. All these are derived from insulinomas, a functional β cell tumor. While all the cell lines available produce insulin, most of them are not responsive to glucose[120].

All the described animal models of T1D have been very useful in providing insights on the development and progression of the disease. However, none of these models are able to fully recapitulate the extensive spectrum of pathogenesis observed in human T1D patients. In addition, there are considerable differences in the ontology of each species, the phenotype presented, function of pathways and genes that limit the transferability of results in preclinical trials. Moreover, animal models are limited in recapitulating the genetic predisposition observed in human T1D. As already mentioned, most of the risk variants associated with T1D lie in the non-coding genome, likely affecting gene regulatory networks. While understanding variants that affect specific genes can be more straightforward in animal models, gene regulatory networks are highly specific to each species. Therefore, results from these studies are often not applicable to other species. Thus, modelling T1D in human cells or tissues is of upmost importance, not only for

studying the genetic contributions to the disease but also to fully understanding the complex human pathology and developing targeted treatments for T1D disease.

5.2 Human models

Mainly due to the accessibility, **peripheral blood** samples have been extensively collected in T1D patients to study human immune cells at different stages of disease. Different types of peripheral blood immune cells have been reported to be altered in T1D at the transcriptome and functional levels. The study of peripheral blood samples has also enabled the identification of AAbs as biomarkers of disease progression and characterization of different self-reactive T and B cells. These samples are of easy accessibility, however, biomarkers and changes observed in blood may not be comprehensively representative of the inflammatory process happening within the islet. Thus, to fully understand the mechanisms underlying disease pathogenesis and the effects on the cell types composing the pancreatic islet, more complex models including the combination of multiple disease relevant tissues may be required[117].

5.2.1 Primary human islets

Human pancreatic islet samples are scarce and difficult to obtain. The use of a Ricordi chamber has become the gold standard procedure to isolate pancreatic islets from the exocrine tissue in human and large animals. The procedure combines enzymatic and mechanical digestion of the exocrine pancreatic tissue to separate the different compartments of the pancreas. The protocol has been further optimized since its inception to increase the yield and purity of these samples, while maintaining the core principles of the original procedure [116]. This method has contributed to the success of islets transplantations and help with all research related to glucose metabolism.

Due to the limitations in obtaining human pancreatic islets samples from cadaveric donors, over the past two decades, different programs have been created to isolate and procure islets to cover needs for pancreatic islets transplantation and beta cell research. TrialNet[121, 122] or HPAP[123] are examples of consortia aiming in collecting vast number

of samples including clinical information on healthy and diabetic patients (T1D, T2D or prediabetic) donors. The main objective is to facilitate collaborative research to enhance the understanding of islet diseases and discovery of new treatment options. TrialNet is mainly focused on gathering information on T1D patients to provide insights into the pathology of T1D and developing clinical trials on prevention and intervention. HPAP is more focused on understanding the molecular function of healthy pancreatic islets and their dysfunction in both T1D and T2D.

Information gathering from these and other consortia is of great value to establish hallmarks of T1D in humans. For example, collection of human samples has been useful to characterize the T cell infiltration in insulitis or to uncover possible viral infections in T1D islet samples, evidences that was previously observed in animal models. However, human pancreatic samples are scarce, very sensitive to shipment and manipulation, and do not cover all disease stages. Thus, studying the contribution of specific genes or genetic variants in human islets is and has been challenging [117].

5.2.2 Models of insulitis

To further understand the process that leads to T1D progression and the contribution of genetic risk to these processes, *in vitro* models that mimic different stages of disease are used. To achieve this, these models must combine the appropriate cell types within the correct disease environment to accurately replicate interactions between β cells and immune cells.

Microfluidic devices, such as **microphysiological systems (MPS)**, create an environment with an artificial flow of nutrients, and, in some cases, immune cells, that allow the supply of desired conditions and recollection of secreted byproducts from the culture population. MPS systems can be used to study the reaction of the β or islet cells to autoreactive or activated immune cells. One of the major advances in T1D disease modelling has been the generation of SC-derived islets from hESC or iPSC, which are islet-like endocrine clusters, mostly composed of insulin positive and glucagon positive cells, derived from human pluripotent stem cells[124]. This system can be loaded with primary human islets, SC-derived islets or β cell lines and expose them

to different types of activated immune cells to study the effect on the target tissue. In the case of primary cells obtained from different donors, MPS can provide insights on the difference in response between T1D patients and healthy donors[125].

The development of protocols to produce of iPSC-derived islets has facilitated the creation of individualized models for different T1D genotypes and for control donors. Additionally, these cells can be genetically manipulated to study the specific contribution of different loci. In fact, the use of iPSC-derived islets facilitates the study of the interaction between β cells and immune cells[126]. Deriving T cells from iPSCs is still not feasible thus the use of autologous primary T cells in co-culture with iPSC-derived islets form the same donor is a promising approach to prevent alloreactivity when modelling insulitis in vitro. Moreover, this approach grants the possibility of maintaining the genetic background from the same individual. In one study, matched iPSC-derived β cells and α-cells with primary PBMCs from the same donor were co-cultured to characterize the physical interaction between CD8+ T cells and iPSC-β cells after induction of ER stress in β cells[127].Co-culture models taking advantage of iPSC-β cells can be adapted to study interactions with any type of immune cells purified from PBMCs.

SC-derived islets have transformed T1D research and treatment strategies. However, differentiation protocols from SCs do not generate fully mature islet populations and in any case pure β or α cell populations are obtained so, depending on the project hypothesis, further purification steps or single-cell techniques might be required. Additionally, obtaining iPSCs and differentiation protocols into islet-like organoids is expensive and requires highly specialized technical expertise[128, 129].

Other models of earlier stages of insulitis are mainly based on *in vitro* exposure to **proinflammatory cytokines**[34, 113, 130]. These models are focuses mainly on processes that precede CD8+ T cell-mediated killing and are had been extensively used to study the effect of insulitis on the β cell and human islets. Thus, different cytokines or cytokine cocktails have been used to mimic the inflammatory milieu present at early stages of disease. The specific cocktail of cytokines, their concentration and the different cycles of stimulation (unique stimulus ore multi-pulse) may vary depending on which immune cell type is likely acting at each stage. For example, IFN-α is secreted by innate immune

cells, mainly macrophages, and is involved in the response to viral infections. On the other hand, IFNγ is a cytokine mainly secreted by T cells, both CD4+ and CD8+, having a role in inducing an inflammatory response through the activation of other immune cells. IL1-β and TNFα are cytokines produced by activated macrophages and dendritic cells, and are key in regulating the immune response by connecting innate and adaptive responses. A list of the main cytokine cocktails used to mimic the insulitis preceding T1D can be found in Table 2.

These models have been really useful to study the β cell *in vitro* in a context that resembles insulitis. An advantage of cytokine cocktails is that they can be used with iPSCs, human islets and β cell lines, such as **EndoC-βH1**, an immortalized human pancreatic β cell line derived from fetal pancreas. EndoC-βH1 has been validated *in vitro* as a good model of human β cells since it expressed all key markers of β cell maturity and shows functional glucose-stimulated insulin secretion (GSIS). Additionally, this approach and the use of a cell line is compatible with genome editing techniques for gene candidate validation[131, 132]. Other human beta cell lines beyond EndoC-βH1 have also been developed, such as EndoC-βH3 that is a conditionally immortalized cell line in which tamoxifen-controlled de-immortalization increases insulin content and secretion, compared to EndoC-βH1.

Cytokine models have proven useful to understand specific disease mechanisms or how the β cell responds at different levels (epigenome, transcriptome, proteome, secretome, etc) to proinflammatory stimuli. However, these models have only partially elucidated the response of the β cells to inflammation and its connection to genetic predisposition.

Publication	Cytokines	Dose	Time	Tissue/ cell line	Readout techniques
Ramos-Rodriguez *et al* 2019[130]	IFNγ	1000 U/ml (50 ng/ml)	48 h	EC HI	ATAC-seq ChIP-seq RNA-seq DNA methylation 3D chromatin structure Proteomics
	IL-1β	50 U/ml (2.8 ng/ml)			
Benaglio *et al.* 2022[133]	IFNγ	HI: 10 ng/ml Low: 0.2 ng/ml	6 h 24 h 48 h 72 h	HI	Bulk: ATAC-seq Single-cell: snATAC-seq (only HI dose for 24 h)
	IL-1β	HI: 0.5 ng/ml Low: 0.01 ng/ml			
	+/- TNFα	HI: 1 ng/ml Low: 0.02 ng/ml			
Colli *et al.* 2020[134]	IFNα	2000 U/ml	2 h 8 h 24 h	EC HI	ATAC-seq RNA-seq Proteomics
Stancill *et al.* 2024[135]	IFNγ	500 U/mL	6 h 18 h	HI	scRNA-seq
	IL-1β	50 U/ml			
Eizirik *et al.* 1994[136]	IFNγ	1000 U/ml	6 h 30 h 48 h 6 d	HI	Nitric oxide (NO) formation GSIS
	TNFα	1000 U/ml			
	IL-1β	50 U/ml			
	IL-6	25 U/ml			
Dettmer *et al.* 2022[137]	IFNγ	14 U/ml	12 h 24 h 48 h	SC-β cells EC	Transcriptome microarray Cell apoptosis assays ROS formation Western Blot ELISA for cytokine quantification
	TNFα	185 U/ml			
	IL-1β	60 U/ml			
Oleson *et al.* 2015[138]	IL-1β	75 U/ml	72 h	EC	GSIS Metabolism tests Cell viability assays NO formation Western blot qPCR
	IFN-γ	750 U/ml			
	TNF-α	1,000 U/ml (10 ng/ml)			

Chen *et al.* 2005[139] **(Mouse cytokines)**	IL-1β	5 ng/ml	0-3 h	HI	GSIS Cell death Western blot Other functional assays
	IFN-γ	100 ng/ml			
	TNF-α	10 ng/ml			
Grunnet *et al.* 2009[140]	IL-1β	2 ng/ml	24 h	HI	Cell viability and death assays GSIS
	IFN-γ	100 ng/ml			
	TNF-α	100 ng/ml			

Table 2. Summary of the different cytokine cocktails used to mimic T1D insulitis in human models. Dose is sometimes stated as ng/ml or U/ml, depending on the information provided by the authors. However, U/ml or ng/ml between different studies are not always comparable unless the same product catalogue number was used. For standardized and comparable measure, IU should be stated. Of note, additional studies using similar models with same dose/cocktail were not included. Studies not reporting the dose of the cytokine cocktail used were not included either. Unless otherwise stated, cytokines are human recombinant. EC: EndoC-βH1; HI: Human Islets.

6

Chromatin remodelling and gene expression regulation

6.1 DNA structure

DNA, or deoxyribonucleic acid, is the molecule that contains all the genetic information required for the development and functioning of a human being. It consists of two intertwined strands that form a spiral-shaped structure known as the double helix. Each strand's backbone is composed of alternating sugar and phosphate molecules, to which one of four nitrogenous bases—adenine (A), thymine (T), guanine (G), and cytosine (C)—is attached. These bases pair specifically (A with T and G with C) through chemical bonds, creating the rungs of the ladder-like structure and stabilizing the double helix (Fig 5). Since the nitrogenous bases differentiate one nucleotide from another, they are often used to identify the entire monomer so a DNA molecule can be represented as a sequence of nucleotides (A, C, T, and G).

In humans, each somatic cell contains 46 molecules of genomic DNA, 23 pairs of chromosomes, along with a smaller DNA molecule located within the mitochondria, the mitochondrial DNA (mDNA). The genomic DNA is enclosed inside the nucleus and organized into molecular structures named chromosomes. When stretched out, the total amount of DNA in a human cell forms a very thin fiber of about 2 meters long[141]. Inside the cell the DNA is tightly compacted to suit the nucleus, which has a diameter measuring in average 6 μm. Thus, DNA is packaged into **chromatin**, a large nucleoprotein complex consisting of DNA and a variety of structural and regulatory proteins.

The high level of DNA organization is crucial in controlling chromatin-mediated cellular processes like transcription, DNA replication, and DNA repair. Genes are sequences within the genome that are transcribed into mRNA, a different type of nucleic acid, by RNA polymerase II (Pol II). These mRNA molecules are synthesized inside the nucleus and then exit the nucleus and undergo translation into proteins, which regulate cell metabolism and function. Usually, a single gene can be divided into different segments; **exons**, which are segments of nucleotides that are

transcribed into mRNA; interspersed with non-coding regions called **introns**. Often, gene exons are referred to as coding regions since they encode for proteins. The exons constitute about 2% of the human genome[142].

For many decades it was believed that only the coding parts of our genome held biological relevance. The remaining portion of the genome, comprising non-coding sequences, was largely dismissed and labelled as "junk DNA". Fortunately, numerous researchers persisted in studying the non-coding genome, aiming to comprehend why such a significant portion of the genome seemed to lack an apparent function. In 2012, the Encyclopedia of DNA Elements (ENCODE) project released findings from extensive studies spanning over a decade, challenging the perception of noncoding sequences as mere "junk". They proposed that approximately 80% of the human genome possesses a biochemically relevant function[143]. Thus, the non-coding genome has been further - and keeps being- characterized to unravel and describe these functions, such as non-coding genes and regulatory elements (REs), which play an essential role in gene transcription regulation.

6.2 Nucleosomes

While during mitosis chromosomes are compacted and easily visible, during interphase chromosomes are not easily distinguishable but still retain microscopic organization. Mitotic chromosomes represent the highest degree of condensation while during interphase chromosomes show different degrees of packaging, influenced by DNA-bound proteins. Thus, chromatin can be categorized into **euchromatin**, which is more loosely packed and accessible, and **heterochromatin**, which is more densely packed (Fig 5). Interestingly, the interphase organization has biological implications as euchromatin is enriched in actively transcribed regions, since facilitates access for the transcription machinery, while heterochromatin is predominant in inactive and repressed regions[144].

The chromatin fiber is fundamentally composed of **nucleosomes**, formed by a core body of proteins called histones to which the DNA wraps around every 150-200 base pairs (bp) (Fig 5).The nucleosome,

besides being an structural component of the chromatin, is implicated in the modulation of gene expression[144].

Nucleosomes are precisely positioned across DNA molecule to expose key transcription factor binding sites (TFBS) while maintaining a stable fiber conformation. Their positioning can be modulated due to specific interactions between the DNA sequence and nucleosome surface charges. Despite being discretely positioned, histones are mobile, which is important to ensure gene expression regulation.

6.2.1 Histone proteins

Different types of histones bind the DNA to facilitate organization and gene transcription regulation. Histones are main components of the chromatin which form the core complex of nucleosomes: H2A, H2B, H3 and H4. Core histones are positively charged proteins rich in lysine and arginine. They bind directly to the DNA fiber through non-covalent forces, primarily through interactions between the positively charged histone residues and negatively charged DNA phosphates. Core histone complexes are organized in octamers, and each complex is composed of two proteins of each core histone (Fig 5).

The terminal domains of core histones are phylogenetically conserved long tails that extend out of the nucleosome core. These terminal tails undergo various covalent post-translational modifications, which modulate the conformation and interactions of the chromatin fiber with adjacent DNA sequences[145].

Different variants of each histone gene also exist and are expressed at different developmental stages. Changes in the composition of core histone complexes affect the conformation and function of chromatin, affecting both physiological and pathogenic processes.

6.3 Chromatin organization

The chromatin fiber consists of a series of nucleosomes arranged in tandem array and folded into a compact, higher-order structure.

Genome organization is highly complex in order to maintain proper transcriptional regulation. Chromosomes are folded into various hierarchical domains, forming functional compartments, maintained by different architectural proteins including CTCF, cohesion or the mediator complex.

Recent advances in technologies to study the 3D genome organization in interphase chromatin allowed us to deepen our understanding of chromosome folding. The hierarchical layers of chromosome organization can be classified by decreasing size of interactions as following[146–148]:

- **Chromosome territories** refer to the localization of the chromosomes within the nucleus. This is the highest level of nuclear organization, with initial reports dating back to the early 1980s [144]. It promotes intra-chromosomal interactions of the chromatin.
- **A/B compartments** are large multi-Mb scale hubs with accessible chromatin and active gene transcription (A) or inactive and condensed chromatin (B)[149, 150]. The probability of interactions between genomic regions within the same compartment is higher than with regions in the other compartment. The composition and distribution of A/B compartments is cell-type specific, dynamic during development and seems to be altered in disease[151].
- **Topologically associating domains (TADs)** are genome regions marked by high degree of self-interaction with other regions within their boundaries, and limited interactions observed with regions outside the TAD, even when regions are at similar distances[152].First reported by Dixon *et al.*[149], TADs constitute the fundamental structural unit of the genome. They are conserved in different tissues and across species.
- **Chromatin loops** represent the ultimate structural unit controlling gene expression. The stability of chromatin loops is mediated by CTCF or enhancer-promoter contacts. These loops facilitate long-range interactions among distinct genomic regions (Fig 5). The establishment of a chromatin loop enhances the likelihood of interaction between two distant regions in the linear genomic space, which would otherwise occur at very low frequencies.

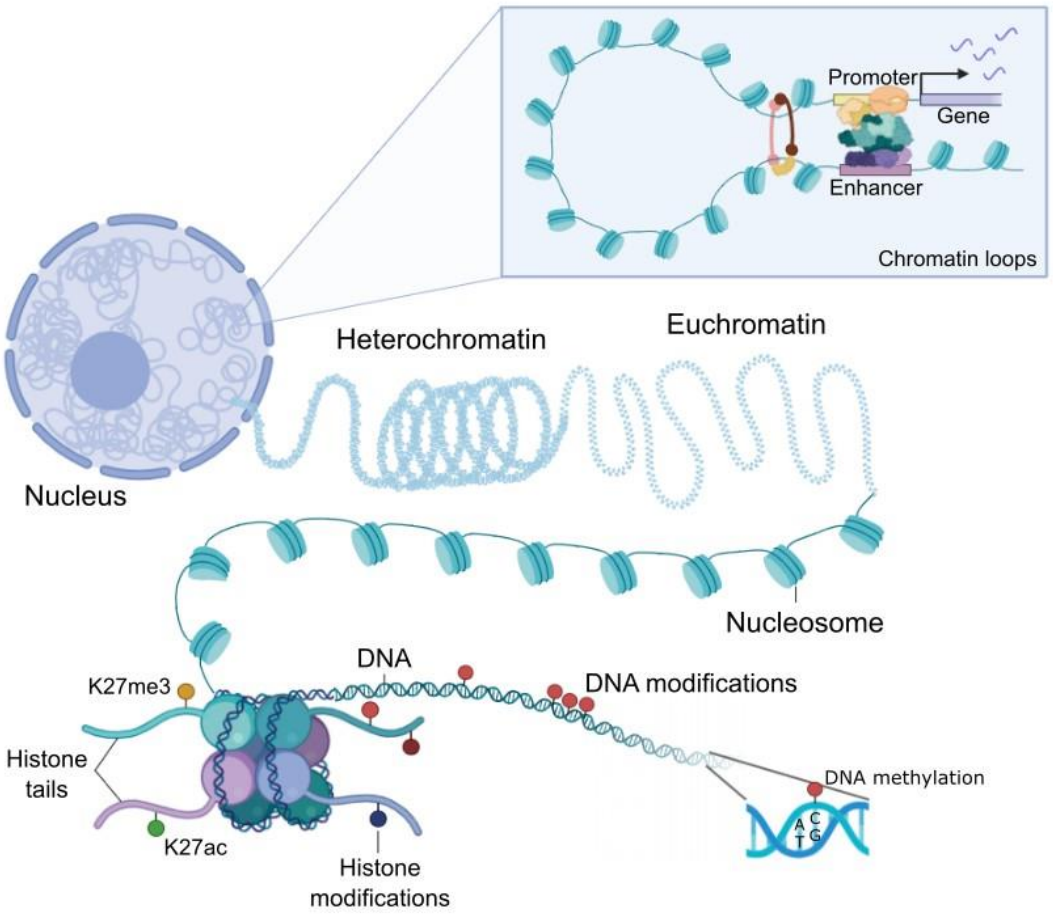


Figure 5. Chromatin structure and DNA composition. Summary of different levels of chromatin and DNA organization. From lower to higher degree: DNA consists of nucleotides (A, T, C or G) forming the double-helix. This DNA can be subjected to different modification, such as DNA methylation in guanine residues. DNA is wrapped around histone proteins (octamers) and compacted, forming nucleosomes. Histone proteins can present different posttranslational modifications and each modification will influence the condensation and accessibility of the chromatin. According to the degree of condensation, chromatin is divided into euchromatin and heterochromatin. Inside the nucleus, chromatin follows different layers of organization. Chromatin loops represent the smallest interaction domains where enhancer-promoter contacts occur, to regulate gene transcription.

6.4 Gene expression regulation

In general, all somatic cells in the body share the same DNA. However, each cell type varies in structure and function because somatic cells do not express all coding genes within the genome. Thus, different cell types express different sets of genes. The regulation of gene expression in each cell type involves is a highly regulated process with different levels of control to ensure that proteins are produced in the correct cells at the appropriate times. Interestingly, even in the adult organism, it remains essential to regulate the activation or suppression of gene expression to uphold cell function and respond to external stimuli.

The activation and repression of gene transcription is controlled by various mechanisms such as chromatin properties, transcription factor (TF) binding, non-coding RNAs, RNA processing and alternative splicing, among others. However, for the purpose of this thesis, I will focus my discussion on cis regulatory networks and non-coding REs.

Cis-regulatory elements, which are non-coding DNA sequences including promoters and enhancers. **Promoters** are located near the transcription start site (TSS) of genes and facilitate the binding of different proteins, such as TFs or RNA Pol II, to initiate transcription. On the other hand, **enhancers** can be positioned anywhere from a few bp to megabases (Mb) away in the linear space from their target genes. They are also bound by various proteins, such as TFs, and they likely physically interact with their target gene promoters to regulate gene expression.

TFs are proteins that bind short DNA sequences and regulate gene transcription. These DNA sequences, typically found at enhancers and promoters, include a consensus sequence, known as a motif, usually 6-10 base pairs in length. However, a TF motif alone is not sufficient to ensure TF binding, other cell type-specific factors also facilitate the binding. In fact, TFs bind to only a small fraction of the motifs in the genome as a consequence TFs that are expressed ubiquitously may show cell-type-specific binding and function.

Transcription factors are the ultimate regulators of gene transcription and help with the recruitment of RNA polymerase II (RNA pol II), the main enzyme in charge of facilitating gene transcription.

6.4.1 Gene promoters

A promoter is a DNA region where RNA polymerase initiates gene transcription, usually situated just upstream or at the 5' end of the transcription start site (TSS). RNA polymerase machinery and essential TFs bind to both promoters and transcription initiation sites to start the transcription process. Promoters, typically 100–1,000 bp long, have three elements in eukaryotic cells: core, proximal, and distal promoters. Each element plays a distinct role in DNA transcription and RNA polymerase function[153, 154]:

- **Core promoter** is a region approximately 40 bp upstream the TSS. It is the minimal unit for transcriptional initiation, where general TFs and Pol II bind. It contains the classical TATA-box that serves as the binding site for the TATA-box-Binding Protein (TBP), a subunit of the complex that mediates Pol II recruitment.
- **Proximal promoter** is located upstream to the core promoter, approximately 250 bp from TSS. It is characterized by the presence of TFBS that either activate or repress the expression of the target gene.
- **Distal promoter** are regions located upstream of the proximal promoter that also containing regulatory sequences with TFBS to regulate gene expression. These regions are not as well characterized as the other two but can expand up to 2 kb upstream the TSS.

The ability of a promoter to activate gene transcription depends on its sequence composition and the presence of specific transcription factors. While the core promoter can initiate basal transcription, the presence of a proximal promoter can boost transcriptional activity. Furthermore, enhancers can also interact with target promoters to further modulate gene transcription[155].

In broad terms, there are two types of promoters according to their basal activity. **Constitutive promoters** show general high expression levels independent of the cell type. **Inducible** or **regulated promoters** are those with cell-type and cell-state specific functions, only activated by environmental stimuli to adapt and respond accordingly.

6.4.2 Enhancers

The human genome is believed to contain hundreds of thousands to millions of enhancers. Enhancers are thought to play a crucial role in coordinating gene expression, essential for human development and homeostasis. They are also very relevant in common diseases as the heritable risk is largely associated with non-coding regions, particularly enhancers specific to disease-relevant cell types. This has heightened interest in annotating and understanding human enhancers. However, despite their clear importance in both basic and disease biology, much remains unknown about enhancers, including their locations, mechanisms of action, and the genes they influence.

As briefly explained above, enhancers are 100-700 bp non-coding DNA regulatory sequences that can be bound by transcription factors (TFs) and enhance the expression of genes. Unlike promoter, are located distally to the target gene TSS, from hundred bp to even 1M bp away. Enhancers are predominantly found in intergenic and intragenic regions, mainly introns and very few described within exons. Enhancers form complex loops to physically interact with their target gene promoter to boost transcription[156, 157].

Enhancer contain sequences that allow the binding of TFs (TFBS), co-regulators, chromatin modifiers and other enzymes to interact with gene promoters and induce the expression of target genes transcribed by the RNA Pol II. They act independent of orientation and location respect to the target gene. It has also been described that enhancers can form clusters of multiple enhancers forming super-enhancers[158].

In vertebrates, enhancers are usually conserved in sequence and function, although in some cases the sequences might diverge but their function may be conserved. The function of an enhancer is determined by its underlying sequence, since it primarily consists of dense clusters of TFBSs. Initially, inactive enhancers are likely bound by pioneer TFs that trigger the binding of other TFs, co-regulators, chromatin remodelling factors, among others. The recruitment of the TFs complexes may lead to enhancer priming and activation.

To impact gene transcription and regulation, enhancers must communicate with their target promoters. In higher eukaryotes,

enhancers are often physically distant from their target gene promoters along the genome. Different models have been proposed to understand the mechanisms by which enhancers meet their target promoters but the looping models is the most experimentally supported. The looping model suggests that those proteins assembled at the enhancer and promoter establish a direct physical contact between (Fig 5). Although these enhancer-promoter contacts do not ensure transcriptional activation, it seems to be one of the firsts steps needed to facilitate gene transcription. Furthermore, it is known that an enhancer can regulate multiple promoters, and a promoter can be influenced by several enhancers. Enhancers can be understood as modular units of gene expression, exhibiting additive and/or redundant effects on their target genes to provide robustness to gene expression.

The actual mechanisms of transcriptional activation by a single enhancer are not clearly described. Two main steps have been suggested[156]:

- Recruitment. Sequence-specific TFs are recruited to enhancers and promoters, these proteins will then recruit additional TFs and coactivators essential for the transcription process.
- Synergism. After their recruitment, TFs exhibit functional synergy, meaning that the combined transcriptional output of two TFs is greater than the sum of their individual outputs, indicating "functional amplification" of their action. This synergy has been extensively demonstrated in various contexts, though the precise mechanics remain unclear.

It seems that TFs create a necessary platform for coactivators and other enzymes to perform catalytic activities that result in transcriptional activation rather than directly activating transcription. Therefore, since enhancers are mainly TFBSs regions, they are known to increase the probability of transcription rather than its intensity or level. Therefore, a strong enhancer likely interacts with the promoter more often, with each interaction potentially initiating a transcription event. Thus, the strength of the enhancer determines the frequency of transcriptional bursts, which are rapid and repeated cycles of transcription. This scenario is only feasible if enhancer-promoter contacts are dynamic and transient, involving continual formation and disruption[156].

6.5 Methods to profile chromatin features that characterize regulatory elements

To completely understand the cis-regulatory networks governing gene expression, the first step is to characterize regulatory elements. Over the past decades, extensive efforts such as the ENCODE project have allowed the characterization of active regulatory elements in many human cell types, across different developmental stages and disease states. However, this library remains incomplete and many cell types or disease states haven't been characterized yet due to their tissue- and state-specific nature.

Many different features of regulatory elements have been studied and used to determine their function. However, regarding chromatin features, properly characterizing cis-regulatory elements involves considering three main aspects:

- **Genomic position**: Determining the genomic coordinates of a RE is crucial for studying its sequence features, like TF binding motifs, to characterize the networks involved in its activation.
- **Activity status**: Regulatory elements can exist in various activity states – active, poised or repressed – which may be specific to cell type and cell state. Defining the activity status contributes to decipher the role of the RE in gene regulatory programs of a cell in a specific state.
- **Gene targets**: Identifying the potential gene targets of a RE is probably the most challenging task, particularly because distal enhancers can be located Mb away from their target genes.

One of the most studied chromatin features of REs is **DNA methylation** (DNAme). It can be used as a proxy for active or repressed RE. It is a DNA modification that does not alter the nucleotide sequence. In this case, a methyl group is added to the cytosine residues within CpG dinucleotides, which may be dispersed throughout the genome or clustered in areas with a high density of CpG sites, known as CpG islands.

Promoters contain higher GC content than the rest of the genome. In fact, 70% of human proximal promoters contain CpG islands.

Methylation of these CpG islands is associated with reduced gene transcription, mechanistically it was suggested that this may result from the interference of the methylation residues with the binding of the transcription machinery[159]. Like promoters, enhancers may also contain CpG islands, which can repress enhancer activity if methylated[160]. DNAme solely does not provide information on the genomic coordinates of REs, as the majority of CpGs in mammal genomes are methylated[161]. However, examining DNA methylation at particular REs can provide information on their activity status. While DNAme is linked to inactive regulatory elements, the lack of DNA methylation does not always indicate active regulatory elements.

6.5.1 Chromatin accessibility

Chromatin accessibility refers to the extent to which nuclear proteins can physically interact with the chromatin. This event is influenced by the occupancy, spatial arrangement of nucleosomes, and other chromatin-binding factors that restrict access to DNA.

The arrangement of nucleosomes throughout the genome is not uniform. While they are densely packed in heterochromatin, they are depleted at active regulatory sites such as enhancers or promoters, and transcribed gene bodies. DNA between nucleosomes is often bound by transcription factors, RNA polymerases, or architectural proteins. The complexes recruited to an enhancer or promoter results in a large assembly of proteins that keeps the enhancer or promoter region nucleosome-free and hypersensitive to nucleases, a characteristic used to identify enhancers (Fig 6).

Nucleosome occupancy is dynamic and varies across the genome, a highly regulated feature during development, disease or in response to external stimuli. The landscape of chromatin accessibility reflects the regulatory capacity of the cell, crucial for determining chromatin organization and function[162].

While chromatin accessibility can be used to locate functional REs in the genome, accessibility alone does not necessarily mean that these regulatory elements are active and inducing expression of their target genes.

Different methodologies have been developed to profile chromatin accessibility. In general, it is typically assessed by measuring the susceptibility of nucleosome-free chromatin to enzymatic cleavage. The cleaved fragments of DNA can be used to generate NGS-ready libraries for high-throughput or quantitative PCR (qPCR) assays. Some of the most used technologies are micrococcal nuclease sequencing (MNase-seq), DNase I hypersensitive site sequencing (DNase-seq), Formaldehyde-Assisted Isolation of Regulatory Elements (FAIRE)[163] or Assay for transposase-accessible chromatin using sequencing (ATAC-seq)[162].

ATAC-seq, developed by Buenrostro *et al*.[164], employs a hyperactive Tn5 transposase to cut and integrate Illumina sequencing adaptors into regions of accessible chromatin. ATAC-seq is widely used due to its robust identification of accessible chromatin, straightforward and rapid implementation, and suitability for use with limited clinical and primary tissue samples. In fact, high quality ATAC-seq libraries have been successfully generated from a very reduced amount of starting material (500-1,000 cells) as a result of the high efficiency of Tn5-mediated adaptor ligation.

Slightly different modifications of the protocol have been developed to adapt the original protocol to cell-type specific requirements, like changing the detergents used or additional transposase enzyme, or to optimize the protocol for a wider range of samples, like Omni-ATAC[165]. Furthermore, ATAC-seq has been modified for single-cell applications to examine the regulatory networks of different cell populations within a heterogeneous tissue. The gold standard scATAC-seq technique was developed by 10x Genomics. The most recent advancement is the single-cell multiome technique, which enables simultaneous profiling of the transcriptome and chromatin accessibility from the same nucleus[166].

6.5.2 Histone modifications

The terminal tails of histones can be subjected to different post-translational modifications, such as acetylation, methylation, phosphorylation, SUMOylation and ubiquitination. These modifications can alter the structure of histone tails, changing the chromatin state and influencing gene expression. These post-transcriptional modifications

can alter the electronic charge and structure of histone tails that bind to DNA, thereby changing the chromatin state and influencing gene expression. Histone modifications have been identified as key players in various cellular processes.

Each modification has been associated with a specific effect on chromatin structure and gene regulation. However, in this brief overview of histone modifications, I will focus on some modifications of the lysine residues (K) of histone H3 (H3).

In lysine 4 (H3K4) different modifications have been described such as H3K4me2 (dimethylation), which marks the 5' end of transcribed genes, or **H3K4me3** (trimethylation), associated with **active promoters**. Interestingly, the deposition and association of these histone marks with activity status of the chromatin might vary in different stages of development[167].

On the other hand, **H3K4me1** (monomethylation) is enriched at **enhancer regions** but does not associate with the activity status of the RE. This histone mark is often profiled along with others, such as H3K27ac or H3K27me3, to determine if the RE is active or repressed. It has been suggested that the presence of this mark in distal REs may protect them from inhibitory complexes, such as DNA methylation enzymes, thereby keeping them in a poised state that can be activated by the appropriate stimulus[168].

Modifications in lysine 27 (H3K27) are also well studied due to their informative role in the activation status of REs.**H3K27me3** is usually linked to gene repression, found at **poised enhancers** and forming broad domains at **promoters of silenced genes**. In contrast, H3K27me1 is enriched at active transcribed promoters. **H3K27ac** is one of the most well studied acetylation modification. It is enriched at **active promoters and enhancers**. The combined profiling of H3K27ac, H3K27me3 and H3K4me1 helps classifying the enhancer regions as: **active** (H3K27ac + H3K4me1), **poised** (H3K27me3 + H3K4me1) and **primed** (H3K4me1 solely)[167](Fig 6).

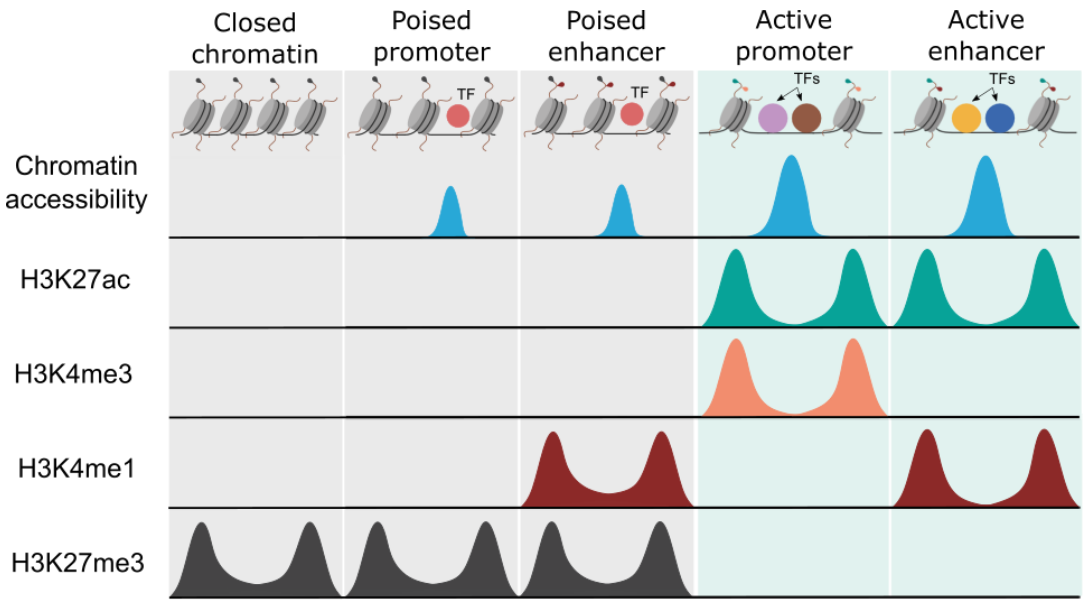


Figure 6. Chromatin accessibility and histone modifications as indicators of regulatory element activity. Overview of key epigenetic features used to identify regulatory elements and assess their activity status.

Many other modifications have also been described such as H3K9me2 or H3K9me3, linked to gene repression and heterochromatin, or H3K9me1, associated with active genes and found around their TSS[167]. In summary, histone modifications profiling not only provides information on the activity status of the RE but also indicates the location of RE, as histone modifications are found in the nucleosomes that border the accessible regulatory regions.

There are different methods to analyze histone modifications, although the gold standard to profile enrichment analysis on specific loci is chromatin immunoprecipitation (ChIP). From formaldehyde-fixed samples, it uses an antibody to detect the modification of interest and precipitate the DNA that is wrapped around the nucleosome containing that histone mark. After immunoprecipitation, the DNA is purified and subjected to qPCR (ChIP-qPCR) to interrogate known loci, microarray analysis or sequencing (ChIP-seq) to perform a genome-wide analysis. However, this technique requires a large number of cells (1-10 million cells) as starting material and it is directly dependent on the quality of the antibody used[169].

Over the years, different modifications to the original ChIP-seq protocol have been used, some to adapt the procedure to the requirements of specific samples and others to optimize the number of cells as starting material or the protocol time, like ChIPmentation[170]. Additionally, there have been some attempts to adapt the ChIP protocol to single cell applications[171, 172]. Single cell resolution of histone modifications did not fulfil the initial expectative although the authors managed to characterize different populations within a heterogeneous tissue and establish the basis for future improvements[171, 173].

Although ChIP-seq continues to be the technique of choice for histone profiling in most cases, in recent years, new techniques have been developed that improve some of the limitations of ChIP-seq. Two alternatives are Cleavage Under Targets and Release Using Nuclease (CUT&RUN)[174] and Cleavage Under Targets and Tagmentation (CUT&Tag)[175, 176]. Both are similar antibody-based techniques for profiling different chromatin features.

The increasing number of publications using the recently developed CUT&Tag protocol advocate for its advantages. This antibody-based technique profiles targeted proteins without relying on immunoprecipitation to capture genomic loci containing the protein of interest. Instead, after antibody binding, a recombinant Tn5 enzyme is used to directly shear the chromatin and prepare the library for sequencing.

CUT&Tag offers several benefits: 1) it requires only a small number of cells (5,000-100,000 cells); 2) it uses native samples, eliminating the fixation step and improving antibody performance; 3) it has a shorter protocol time; 4) it significantly reduces the signal-to-noise ratio; 5) it requires less sequencing depth; and it is more cost-effective.

Additionally, CUT&Tag shows great potential for single-cell applications. Various protocols have already adapted CUT&Tag to the 10x Genomics platform, as well as other platforms, to profile different histone modifications and transcription factors in heterogeneous tissues at the single-cell level[177–181].

6.5.3 3D chromatin organization

Once the distal RE is mapped in the genome, one remaining major challenge is that of determining the target gene promoter it regulates. Thus, charting three-dimensional chromatin contact profiles can help in

the identification of potential gene targets and in characterizing cell-type and cell-state specific regulatory programs.

Early studies of chromatin conformation relied on light microscopy techniques, such as fluorescence in situ hybridization (FISH), to determine the proximity between different loci. However, the low resolution of microscopy approaches only allowed for the characterization of higher chromatin architecture features. Nowadays, most assays are based on C-based methods, which have evolved in powerful technologies for profiling chromatin interactions at different levels of resolution.

Chromosome conformation capture techniques (3C and 3C-based) quantify the frequencies of contacts between different regions of the genome. These methods follow basic steps that start with the crosslinking of chromatin to preserve its 3D conformation. Next, the chromatin is isolated and digested with restriction enzymes, followed by re-ligation of DNA fragments that may be distant in linear distance but close in space at the time of cross-linking. The result is a library of chimeric molecules in which the two DNA fragments in physical contact but far from each other in the DNA linear sequence are stitched together. The libraries are then amplified by PCR or sequencing.

Various protocols have been developed based on these principles, differing in terms of genomic coverage and resolution. **3C** was the first assay used to confirm the existence of chromatin loops. It is used to study high resolution interactions between two known loci, “one vs one”, using PCR and gel electrophoresis. However, it is not suitable for high-throughput applications and requires a good PCR primer design on the regions of interest[182].

To address the limitations of 3C, other derived methods have been developed for high-throughput data. **HiC** is the technique with highest coverage along the genome, but with reduced resolution[150]. It is an "all vs all" method, interrogating contacts between all regions of the genome without relying on loci-specific primer design, as libraries generate genome-wide contact maps. Other techniques aimed in increasing the resolution of Hi-C by reducing the number of interrogated loci were developed and named "many vs all" approaches. Promoter Capture Hi-C is one example of such strategy, where only the contacts with the promoters are interrogated[183].

Another example with reduced coverage but higher resolution is **4C**, which analyses interactions between a viewpoint (region of interest) and the rest of the genome, following a "one vs all" strategy. It also relies on

locus-specific PCR primer design to amplify the library with contacts for sequencing. The advantage of this strategy is the ability to achieve high-resolution maps with low sequencing depth[146, 182]. A modification of 4C, called **UMI-4C**, allows for accurate quantification of chromatin interactions by reducing PCR amplification bias[184].

Many other variations of 3C-based techniques have been reported, each with specific applications and advantages. Interestingly, since linking distal regulatory elements with their target genes remains challenging, some studies have focused on using snATAC-seq data to predict chromatin interactions. Cicero is an algorithm designed to identify co-accessible pairs of DNA elements to link regulatory elements to their potential target genes[185].

6.6 Functional validation of enhancer-promoter contacts

Once an enhancer, or even an enhancer-promoter interaction, has been described, further experiments can be performed to confirm the functional effect of the distal RE.

One option as first strategy can be to evaluate the putative enhancer function of the detected sequence using reporter genes. There are different technologies available to measure enhancer activity, such as massively parallel reporter assays (**MPRA**). This approach uses a high-throughput assay that tests the functional activity of candidate regions using a reporter gene within a vector. The validation is independent of the position as the library of candidate sequences are cloned inside a reporter vector, to assess if the putative distal RE is able to enhance the expression of a reporter gene through a minimal promoter. Then, each reporter gene contains a barcode associated to a specific enhancer to quantify enhancer activity[157]. MPRAs have the advantage to simultaneously testlarge numbers of REs, although similar assays are also available to test the activity of only one or few candidate regions, like **luciferase reporter assays**. In this case, the conceptual basis is very similar as the sequence of interest is cloned into a vector with a reporter gene under a minimal promoter. Nonetheless, the readout this time is the activity of the luciferase enzyme, which synthesis is regulated by the putative enhancer sequence[186].

Nevertheless, MPRAs and other reporter assays only validate the potential enhancer function out of cellular context and this not always

translates into an active RE once the native context is restored. Thus, another approach to confirm the enhancer function over the potential target promoter whiteout interfering with the cell-specific chromatin context is to induce enhancer perturbations followed by transcriptional characterization of its potential/s target genes.

Nowadays, CRISPR-based technologies allow genetic or epigenetic perturbation in single assays to assess the function of selected numbers of enhancers or in high-throughput screening settings. CRISPR technology allows to study enhancers in their genomic context under the appropriate external conditions. Thus, a guide RNA (gRNA) or gRNAs are delivered with the CRISPR machinery to target the enhancer region, and the cells are tested to determine if the enhancer perturbation has an impact on the expression of a target gene or genes. Additionally, with single CRISPR perturbations, the modified cell lines can also be tested for other phenotypic traits, such as cell viability, proliferation or migration, among others, to further explore the phenotypic effects of such perturbation[157].

CRISPR screenings take advantage of large libraries of gRNAs which are delivered to the cells, resulting in the possibility of testing thousands of distal REs at the same time. Different variations of the CRISPR-based technology have been generated, with two main types: CRISPR machineries with active Cas9 for sequence disruption and CRISPR machineries with inactive Cas9 (dCas9) and fused to an epigenetic repressor or activator domain.

The original CRISPR tool contains a **gRNA-Cas9 active nuclease** complex that generates targeted DNA cuts that upon repair result in 1-10 bp deletions or 1 bp indels. In most cases, this technology has been used to target coding sequences of genes to generate mutations and study the function or dysfunction of such protein. Nevertheless, it has also been successfully applied to characterize regulatory elements although this presents additional challenges as compared to its application on coding sequences. Disruption of the enhancer function can be more challenging as the RE functions have to be perturbed, for example the binding site of a TF can be disrupted although, depending on the RE nature this may not be sufficient to alter its function. To address these limitations a common strategy that has been applied consists in using pairs of gRNAs to create a pair of cuts at both ends of the RE and generate larger deletion, although with very low efficiency.

An alternative methodology to study enhancer function is the use of CRISPR-dCas9 techniques which are designed to induce epigenetic

perturbations instead of genetic ones. In **CRISPR interference** (**CRISPRi**), a dCas9 with a KRAB repressor domain (dCas9-KRAB) is directed to silence gRNA defined sequences. On the other hand, **CRISPR activation** (**CRISPRa**) technology contains activator domains fused to the dCas9 able to induce the activation of poised regulatory elements. Other fused domains have also been described to disrupt enhancer activity[157].

Even though these perturbations have been successfully applied, the mechanisms of enhancer perturbation are not optimal. The epigenetic mechanisms mimicked with these techniques are partially accurate but they do not really recapitulate all layers of enhancer regulation. In contrast, CRISPR deletions of enhancer regions can result in a more drastic disruption of the enhancer function.

A limitation of CRISPR techniques are off-target effects that may affect complex regulatory interactions where multiple enhancers regulate a target gene or vice versa. Taking into account the limitations of CRISPR and other technologies, the most effective way to conclusively identify functional enhancer-gene pairs may be to integrate multiple methodologies. This include combining information from enhancer chromatin mapping, profiling of chromatin 3D contacts, and performing functional validation in a defined cell population.

7

Regulatory maps in pancreatic islets in health and disease

Understanding the molecular mechanisms underlying the genetic predisposition driving multifactorial diseases can shed light on the disease pathogenesis and eventually translate the findings towards a personalized therapy. Obtaining regulatory maps of T1D-relevant cell populations in the context of insulitis may be valuable approach to uncover molecular mechanisms linking T1D genetic risk to altered pancreatic islet gene regulatory functions (Fig 7A). However, this represents a novel hypothesis currently being tested by few research groups worldwide.

The application of next-generation sequencing techniques has allowed profiling, during the last decades, the transcriptome and the epigenome of hundreds of different cell populations and human tissues in detail. Different consortia such as Encyclopedia of DNA Elements (ENCODE) or Epigenome Roadmap produced tissue-specific regulatory maps and expression profiles from many different tissues. However, pancreatic islets, and thus islet cell populations, were not initially included in these databases as this tissue is less accessible than most other tissues.

As pancreatic islets are embedded in the exocrine tissue, their isolation is required to capture the islet-specific signal. In addition, islets are composed of different populations present at different proportions, which represents another challenge when studying a specific cell type population. To address the heterogeneity, different protocols have successfully applied to segregate the major cell populations, α and β cells, using FACS. However, sorting minor cell populations such as ε or γ cells requires further characterization of such population to select cell-specific markers[187].

Despite the difficulties, different laboratories took advantage of the optimization of the tissue isolation methods to profile the regulatory landscape of pancreatic islets. In an effort to dissect molecular mechanisms leading to different glucose metabolism diseases, the first studies focused on profiling adult pancreatic islets[77, 188–191] and

pancreatic progenitors[192]. These studies helped characterization of regulatory elements specifically active in pancreatic islets and link them to T2D genetic risk. Additional efforts have been made to deepen our understanding of pancreatic islet pathology. Another layer of complexity in linking genetic predisposition and regulatory maps is to understand the target gene of the regulatory element bearing a variant (Fig 7B-D). A more recent paper included 3D chromatin structure experiments in HI to characterize enhancer-promoter interactions, which are not necessarily close in the linear space[193].

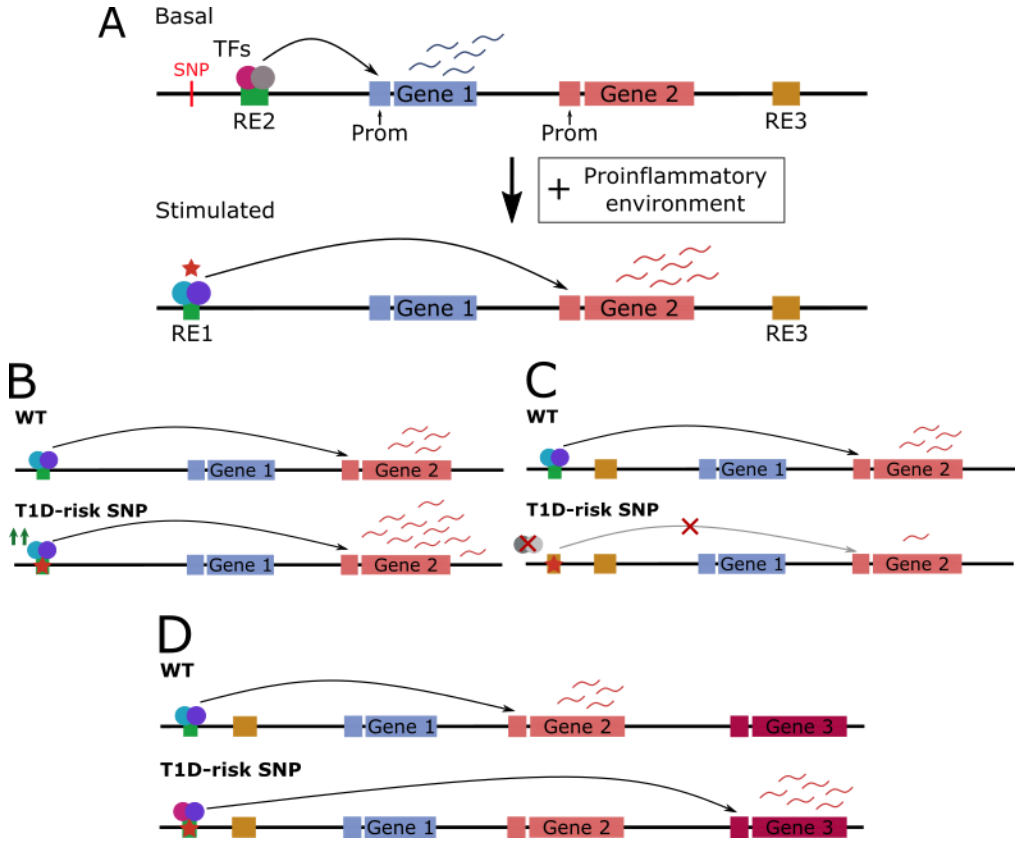


Figure 7. Schematic representation on how gene regulatory networks can change upon proinflammatory stimulus and how tissue-specific regulatory elements can capture the genetic risk signals. A. Proinflammatory stimuli can alter gene regulatory networks, by changing the activation status of different RE and 3D chromatin organization, leading to changes in gene expression. Genetic risk variants can be captured using disease-relevant chromatin maps, as key RE/genes might only be active in cell type and cell states important in disease pathogenesis. B. Relevant genetic risk variants can increase enhancer activity or TF binding in distal RE, leading

to amplify gene transcription. C. Risk variants can also disturb TFBSs in distal RE, causing a disruption of chromatin loops with target promoters and a decrease in gene expression. D. Risk variants may also modify TFBSs in distal RE, causing changes in TF machinery binding, creating new loops with different target promoters and leading to expression of new genes.

With the advent of single-cell technologies, an increasing number of studies are applying transcriptome and epigenomic analyses on the pancreatic islets which allow characterization of the different cell populations composing the tissue. Single-cell technologies have enabled researchers to study not only the different cell types but also their different potential cell states in health and disease state. These advancements in single-cell technologies have facilitate the collection of crucial information that may help understanding disease pathogenesis. For example, Chiou *et al*.[74], combined a powerful GWAS with regulatory maps from single cells of pancreatic and immune tissues from normal donors. This approach allowed them to confirm the enrichment of T1D risk variants in regulatory elements active in T cells while identifying an overlap of some T1D-associated variants with regulatory elements specific to the exocrine pancreas. This observation is relevant as it suggests a potential role of the exocrine pancreas in T1D pathogenesis.

In 2019, a study characterized islet regulatory maps in the context of early insulitis. Ramos-Rodriguez *et al*.[130] used a cytokine model to expose both β cells and HI to a proinflammatory stimulus mimicking early insulitis. The authors profiled the response at the chromatin, expression and protein levels. The β cell regulatory landscape showed marked plasticity upon exposure, with an induction of new and primed regulatory elements (**IREs**) already pre-bound by islet-specific transcription factors. Additionally, they identify for the first time T1D-associated variants that overlap and disrupt cytokine-responsive enhancer activity in human β cells.

Later, others laboratories explored how regulatory landscape of the β cell respond not only to a proinflammatory environment but also to viral islets infections [194, 195]. The combination of HI or β cell stimulation with single-cell profiling of chromatin regulatory networks has also been proven successful in deciphering the T1D genetic risk implication in β cell response. A main challenge now resides in the validation and interpretation of the variants potentially affecting the β cell, or other islet

cell populations, and assessing their impact on T1D pathogenesis. Different options have been used to address this limitation such as unbiased genome-wide CRISPR screening can help prioritize candidate genes[195], techniques to profile the 3D genome architecture can help linking regulatory elements to their target genes[130, 195] or creating SC-derived islets with specific knockouts to study the function of the regulatory region or target gene in β cell function[194].

In a big effort to fully characterize at the molecular level the different populations of islets cells in different contexts associated with diabetes, the Human Pancreas Analysis Program (HPAP) is dedicated to conducting in-depth phenotyping of the human endocrine pancreas. The goal is to gain a better understanding of the cellular and molecular events that precede and result in beta-cell loss and/or dysfunction in T1D and T2D. All the data gathered can be accessed in their open-source data repository PANC-DB, available for all the research community. It contains results from many types of experiments (histology, transcriptome and epigenome in bulk and in single-cell, DNA methylome, among others) in samples from T2D and T1D patients, AAb+ individuals and healthy controls.

In summary, many advances have been made in the past few years to understand the molecular mechanisms underlying T1D. It seems that the strategy of combining the power of single cell epigenomics in cell-types and cell-states relevant in T1D progression with GWAS results will help dissect how the genetic component of the disease modulates the start and progression of the disease.

Part II

Hypothesis and Objectives

8

Hypothesis and objectives

Given the substantial evidence implicating the β cell as an active player in the development of T1D, I strongly believe that investigating β cell behavior in the early stages of the disease will enhance our understanding of T1D pathogenesis before glucose homeostasis is disrupted. While *in vivo* animal models have been invaluable in identifying key hallmarks of T1D, they do not fully replicate human pathogenesis. Conversely, existing *in vitro* human models that mimic T1D primarily focus on the later stages of insulitis, characterized by the active destruction of β cells by CD8+ T cells, or are limited by the hypothesis that specific cytokines alone drive insulitis at the molecular level.

While the models using cytokines have significantly highlighted the crucial role of β cells in the early stages of T1D, moving beyond the notion that β cells are merely passive bystanders, they have only partially helped in understanding the response of β cells to inflammation and its relationship to genetic predisposition[130, 195]. This limitation may arise from the fact that the current models of early insulitis are confined to only studying β cell stress responses to abrupt and extreme inflammatory stimuli. Consequently, this limitation has hindered our ability to translate genetic risk into specific mechanisms or pathways affected in β cells in T1D, with few examples of potentially inaccurate β cell responses in the context of genetically susceptible individuals.

The main purpose of this and previous studies focused on early insulitis is to understand what happens to the β cell prior to the direct cytotoxic T cell-mediated destruction, with the aim of uncovering novel therapeutic targets to prevent, treat or cure T1D. Nevertheless, I believe that a critical gap persists in the field since current *in vitro* models fail to properly recapitulate the complex environment β cells are exposed to within the context of this disease.

Therefore, I **hypothesize** that the use of a novel model that better resembles the pathophysiology in early insulitis will help to unravel

unknown β cell-specific regulatory networks critical to T1D progression and β cell survival. To do so, I **aim to**:

1) Develop a new *in vitro* model to study the lymphocytes-beta cell interplay that may occur during early insulitis and precede T1D.
2) Characterize the regulatory genomics response of β cells to the inflammatory environment.
3) Study the T1D genetic risk modulating the ability of β cells to respond to the inflammatory stimulus and possibly contributing to the disease progression.

Part III

Materials and Methods

9

Wet lab techniques

9.1 Cell Culture

EndoC-βH1

The human insulin-producing EndoC-βH1 (EC) cells were purchased to Human Cell Design (formerly Univercell Biosolutions). EndoC-βH1 were cultured in DMEM medium low glucose (1 g/l), 2% BSA fraction V (#10775835001, Roche), 50 µM 2-mercaptoethanol (#31350010, Gibco), 10 mM nicotinamide, 5.5 µg/ml human transferrin (#T8158, Sigma), 6.7 ng/ml sodium selenite (#S9133, Sigma), 1% inactivated FBS, 100 units/ml penicillin and 100 µg/ml streptomycin. Cells were incubated at 37 ºC 5% CO2.

In CRISPR validation experiments, EndoC-βH1 were exposed to cytokines concentrations in the same medium without FBS for 48 hours. The cytokine concentrations used were as previously described[130]: recombinant human IL-1β (#201-LB-005, R&D Systems) at 50 U/ml; recombinant human IFN-γ (#AF-300-02, Peprotech) at 1000 U/ml.

Primary CD4+ T cells

CD4+ T cells were provided by Dr. Pere Santamaria's group in IDIBAPS, Barcelona. CD4+ T cells were isolated from PBMCs from 3 different human donors without a history of glucose intolerance and normal BMI using the Human CD4 T cell isolation kit (#130-096-533, Miltenyi Biotech). Frozen vials were thawed following standard procedures and centrifuged 400 g 10 min. Then, cells were resuspended at 1 million cells/ml and activated in CTL-Test Medium (CTLT-005, ImmunoSpot CTL) containing 1% Penicillin-streptomycin (Gibco), 1% Glutamine (#BE17-605E, Lonza), 10% inactivated Human Serum (#H4522, Sigma), 20 IU/ml Recombinant hIL2 (#202-IL-050, R&D Systems) and 25 µl/ml of ImmunoCult Human CD3/CD28/CD2 T Cell Activator (#10970, StemCell). Activated CD4+ T cells were cultured for 12-14

days and split at day 3, 7 and 10 using the same medium without CD3/CD28/CD2.

9.2 Co-culture of EndoC-βH1 with APC and CD4+ T cells

Both K562.DRB1*03:01 CD80 APC cell line and engineered CD4+ T cells expressing TCR specific of IGRP13-25 were kindly generated and provided by Dr. Santamaria's group.

CD4+ T cells were thawed following standard procedures and cultured at 1 M/ml in CTL-Test Medium containing 1% Penicillin-streptomycin, 1% Glutamine, 10% inactivated Human Serum, 20 IU/ml Recombinant hIL2, at 37 ºC 5% CO2 for a few hours. APC cells were thawed and cultured 1 M/ml in DMEM (#D6429, Sigma-Aldrich) supplemented with 10% FBS (#F7524, Sigma-Aldrich), 2 mM glutamine (#BE17-605E, Lonza), 1 mM sodium pyruvate (#S8636, Sigma-Aldrich), 100 U penicillin and 0'1 mg/mL streptomycin (Sigma-Aldrich), 50 mg/mL gentamycin sulphate (#17-518Z, Lonza) and 50 mg/mL Normocin (#NOL42-13, InvoGen) at 37 ºC 5% CO2. After a few hours, CD4+ T cells were washed with resting medium (CTL-Test Medium with 1% Penicillin-streptomycin and 1% Glutamine) and seeded 24 h for resting before co-culture. The same day, 100,000 (M1.1) or 500,000 (M1.2) EC cells were seeded in 96 or 48 well plates, for M1.1 and M1.2 respectively.

On day 0 of the co-culture, APC cells were collected and resuspended in EC medium. Then, EC medium from the wells was removed and 100,000 (M1.1) or 500,000 (M1.2) APC cells were seeded each well that contains EC. Then, CD4+ T cells were also collected, resuspended in EC medium and 100,000 (M1.1) or 500,000 (M1.2) CD4+ T cells were seeded each well containing both EC and APC cells. To activate the co-culture, 20 µl of IGRP13-25 peptide suspension (100 µg/ml) or 12.5 µl/ml of the CD3/CD28/CD2 T Cell Activator complex were added to the corresponding stimulated wells (IGRP and CD3, respectively). As a negative control, 20 µl of a DMSO control suspension was added to the corresponding wells.

After each time point, the supernatant (SN) from the whole well containing T and APC cells was collected and centrifuged 500 g 10 min.

The APC/T cell pellet was frozen and stored at -80 ºC and EC cells were detached using accutase treatment, collected in suspension, and processed for RNA extraction.

9.3 Co-culture of EndoC-βH1 with primary CD4+ T cells (CC-M2 and CC-M3)

After 12-14 days of growth, CD4+ T cells were washed and seeded in resting medium (CTL-Test Medium with 1% Penicillin-streptomycin and 1% Glutamine) at a concentration of 2 million cells/ml. They were incubated for 48 hours prior to starting the co-culture experiments. A day before the co-culture, 500,000 and 850,000 – 1 million EndoC-βH1 were seeded in 48 or 12 well plates, for CC-M2 and CC-M3 respectively.

On day 0 of the co-culture, EndoC-βH1 cells were washed once with 1X PBS, and fresh EC medium was added to each well. CD4+ T cells in resting medium were collected and resuspended in EndoC-βH1 medium. 500,000 or 700,000 CD4+ T cells were used in CC-M2 and CC-M3 respectively. In CC-M2 T cells were seeded directly into the well while in CC-M3 a cell culture insert (#3401, Corning) was placed into each well seeded with EC and T cells were added into each insert. Next, CD4+ T cells were activated using 12.5 µl/ml of the CD3/CD28/CD2 T Cell Activator complex and samples were left in co-culture for 2 h, 6 h, 24 h, 48 h or 72 h. Final culture volume was 500 µl for M2 and 1 ml for M3.

After each time point, the supernatant (SN) from the whole well containing T cells was collected in CC-M2 and centrifuged 500 g 10 min. In CC-M3, the SN from the upper chamber, containing medium and CD4+ T cells, was collected in a tube, and the SN in the bottom chamber, containing medium and dead EC, was collected in separate tube. Both SNs were centrifuged 500 g 10 min, and transferred to a new tube and mixed. In both cases, the CD4+ T cell pellet was frozen and stored at -80 ºC, while the dead EC pellet was discarded. EC cells were detached using accutase treatment, collected in suspension, and processed for ATAC-seq, Cut&Tag and RNA extraction.

9.4 ATAC-seq library preparation

ATAC-seq library preparations were done as previously described[130] with minor modifications. Briefly, 50,000 EC cells were centrifuged 5 min 500 g at 4 ºC and washed in ice-cold PBS for 15 min 500 g 4 °C. Then, I isolated the nuclei by incubating them in 300 µl cold lysis buffer (10 mM Tris-HCl pH 7.4, 10 mM NaCl, 3mM MgCl2, 0.1% Igepal CA-630) for 20 min and resuspending them after 10 min. The sample was centrifuged for 15 min at 500 g at 4 °C (with low acceleration and brake settings) and the pellet was next washed in 100 µl of cold lysis buffer and centrifuged again, same conditions. The transposition reaction was carried out in a 25 µl reaction mix containing 1.25 µl of Tn5 transposase (#C01070012, Diagenode), 12.5 µl of 2X tagmentation buffer (#C01019043, Diagenode) and 11.25 µl DEPC-treated water. The transposition reaction mix was incubated at 37 °C for 1 h following inactivation by incubating for 30 min at 40 °C after addition of 5 µl of clean up buffer (900 mM NaCl, 300 mM EDTA), 2 µl of 5% SDS and 2 µl of Proteinase K (#EO0491, ThermoScientific). Tagmented DNA was isolated with 2x SPRI beads cleanup (#A63880, Beckman Coulter) and was eluted in 21 µl 10mM Tris-HCl pH8.

Two sequential 9-cycle PCR were performed in order to enrich for small DNA fragments. The PCR mix consisted of 2 µl of 25 µM PCR Primer 1 (described in Buenrostro *et al.*[164]), 2 µl of 25 µM Barcoded PCR Primer 2 (described in Buenrostro *et al*.[196]), 25 µl of NeBNext High-Fidelity 2x PCR Master Mix (#M0541, NEB) and 21 µl of the eluted sample. The library was amplified in a thermocycler using the following program: 72 °C for 5 min; 98 °C for 30 s; 9 cycles of 98 °C for 10 s, 63 °C for 30 s; and 72 °C for 1 min; and at 4 °C hold. After the first PCR round, fragments smaller than 600 bp were selected using SPRI cleanup beads, 0.6x ratio with right side selection, and second round of PCR was performed with same conditions. The DNA library was finally purified using 1.8x SPRI cleanup beads, eluting in 21 µl 10mM Tris-HCl pH8.

Final library was quantified using Qubit dsDNA BR Assay (#Q32850, Invitrogen) and fragment analysis was performed using TapeStation, Agilent Bioanalyzer or similar to check library quality and nucleosomal pattern resulted from the tagmentation reaction. Once library quality was confirmed, samples were sequenced 100-150 bp paired-end on a NovaSeq X Plus system (Illumina) to obtain about 70 M reads per end.

9.5 CUT&Tag library preparation

CUT&Tag was performed as previously described[175, 176] with minor modifications. Cells were harvested, counted, and 100,000 cells were centrifuged in a LoBind tube (#30108051, Eppendorf) for 3 min at 600 g at room temperature. Cells were washed in Wash Buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 0.5 mM Spermidine (#S0266, Sigma-Aldrich), 10 mM Sodium butyrate (#B5887, Sigma), 1X Protease inhibitor cocktail (#11873580001, Sigma-Aldrich)) and resuspended in 300 µl of Wash Buffer. During washes, 10 µL per sample of Concanavalin A coated (Bangs Laboratories, #BP531) magnetic beads were mixed with 10 volumes of binding buffer (20 mM HEPES pH 7.9, 10 mM KCl, 1 mM CaCl2, 1 mM MnCl2, 10 mM Sodium butyrate), and washed using a magnet stand with 1.5 mL of binding buffer. Beads were resuspended in 1 volume of binding buffer, added to the cells, and mixture was incubated on an end-over-end rotator for 8 min.

After incubation, the unbound supernatant was removed, and bead-bound cells were resuspended in 50 µL ice-cold Antibody buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 0.5 mM Spermidine, 10 mM Sodium butyrate, 0.4 mM EDTA, 0.02% BSA, 0.05% Digitonin (#300410, Millipore), 1X Protease inhibitor cocktail). Primary antibody against H3K27ac (Abcam, #ab4729) was added (1:100) and incubated overnight on a rotating wheel at 4 ºC. Next day, tubes were placed on the magnet stand to clear and pull off the liquid. Secondary antibody (Antibodies Online, #ABIN101961) was diluted 1:100 in 100 µl of Dig-wash buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 0.5 mM Spermidine, 10 mM Sodium butyrate, 0.05% Digitonin, 1X Protease inhibitor cocktail) and incubated on a rotating wheel at room temperature for 1 h. Tubes were placed on the magnet stand to clear and withdraw the liquid and beads were washed three times with 1 mL of Dig-wash buffer. The pA-Tn5 adapter complex (#15-1017, Cutana) was diluted 1:20 in Dig-300 buffer (20 mM HEPES pH 7.5, 300 mM NaCl, 0.5 mM Spermidine, 10 mM Sodium butyrate, 0.01% Digitonin, 1X Protease inhibitor cocktail), and added to the beads. The tubes were mixed by soft vortexing and incubated on a rotating wheel at room temperature for 1 h. After incubation, beads were washed three times in 1 mL of Dig-300 buffer. Tubes were placed on the magnet stand to pull off the liquid, and beads were resuspended in 300 µL of Tagmentation buffer (Dig-300 buffer, 10 mM MgCl2) and incubated at 37 ºC for 1 h. To stop tagmentation and reverse cross-links, 10 µL 0.5M EDTA, 3 µL 10% SDS and 2.5 µL 20

mg/mL Proteinase K (#EO0491, Thermo Fisher Scientific) were added to each sample and incubated for 1 h at 55 ºC. DNA was purified by phenol-chloroform extraction and dissolved in 28 µL of TE (1 mM Tris-HCl pH 8, 0.1 mM EDTA).

For library amplification, the PCR mix consisted of: 21 µL of purified DNA, 2 µL of 10 µM universal i5, 2 µL of 10 µM a unique barcoded i7 primer (primers described in Buenrostro *et al.*[196]) and 25 µL NEBNext HiFi 2X PCR Master mix (#M0541, NEB). The following program was used for library amplification: 72 ºC for 5 min, 98 ºC for 30 s, 13 cycles of 98 ºC for 10 s and 63 ºC for 10 s, a final extension at 72 ºC for 1 min and hold at 8 ºC. Post-PCR clean-up was performed by adding 1.3X of Ampure XP beads (#A63880, Beckman-Coulter). Samples were eluted in 25 µL 10 mM Tris-HCl pH 8.

Final library was quantified using Qubit dsDNA BR Assay and fragment analysis was performed to check library quality. Once library quality was confirmed, samples were sequenced 100-150 bp paired-end on a NovaSeq X Plus system (Illumina) to obtain about 30 M reads per end.

9.6 RNA extraction

Total RNA was isolated from EC cells using the RNeasy Micro Kit (#74004, Qiagen). RNA quantity was measured using a NanoDrop Spectrophotometer and quality was assessed using gel electrophoresis. For sequencing samples, RNA integrity number values were evaluated using fragment analysis technologies. All the samples had RNA integrity number (RIN) values >8.5.

Stranded mRNA libraries were sequenced 100-150 bp paired-end on NovaSeq X Plus system (Illumina) to obtain about 60 M reads per end.

9.7 RT and qPCR

Reverse transcription (RT) was performed using 250 ng of total RNA in a final volume of 10 µl following the manufacturer's instructions (SuperScript IV Reverse Transcriptase, #18090, Invitrogen). Next, cDNA was diluted 1/10 and 1 µl was used to analyzed the expression of each gene by real time PCR. Experiments were performed in a LightCycler 480 platform using SYBR Green I Master (#4707516001,

Roche) and each sample was analyzed in triplicate. Samples were incubated for an initial denaturation at 95 °C for 10 min, then 45 PCR cycles were performed using the following conditions: 95 °C for 10 s, 64-66 °C for 20 s, and 72 °C for 15 s. An additional cycle to calculate the melting curve was added before cooling down. The expression levels were normalized according to the average of at least two independent reference genes (TBP, VAPA and/or RAB7A). Primer efficiency was calculated by extracting fluorescence raw data and using the Chainy tool (http://maplab.imppc.org/chainy/ [197]).

9.8 IFNγ quantification by ELISA

IFNγ quantification (at 2 h, 6 h, 24 h, 48 h and 72 h) was carried out using BioLegend's ELISA MAX™ Deluxe Set (#430104, BioLegend) following manufacturer's instructions. Briefly, SNs at each time point were collected and centrifuged 500 g 10 min to remove dead or floating cells, and stored at -20ºC.

The day before performing the ELISA, Capture Antibody was diluted in 1X Coating Buffer A to prepare the 96-well plate (#44-2404-21, Invitrogen), adding 100 µL of the Capture Antibody solution to all wells needed. Plate was sealed and incubated overnight between 2 °C and 8 °C. Next day, each well was washed four times with 300 µL Wash Buffer and non-specific binding to reduce background was blocked adding 200 µL 1X Assay Diluent A per well and incubating at RT for 1 hour with shaking. During this incubation, all standards and sample dilutions were prepared to be run in triplicate. Then, plate was washed four times with Wash Buffer and 100 µL/well of standards or samples was added to the appropriate wells. Plate was incubated at RT for 2 hours with shaking and washed four times with Wash Buffer. Next, 100 µL of diluted Detection Antibody solution were added to each well, and plate was incubated at RT for 1 hour with shaking. Plate was washed four times with Wash Buffer and 100 µL of diluted Avidin-HRP solution to each well was added and incubated at RT for 30 minutes with shaking. Afterwards, plate was washed five times with Wash Buffer and 100 µL of freshly mixed TMB Substrate Solution (#S5814, Sigma) was added and incubated in the dark for 20 minutes. Finally, reaction was stopped by adding 100 µL of Stop Solution to each well. Absorbance at 450 nm was read within 15 minutes and at 570 nm to subtract it from the absorbance

at 450 nm. ELISA data obtained was analyzed using GainData (Arigo Biolaboratories, https://www.arigobio.com/elisa-analysis).

9.9 Cytokine array

Multiplexing laser beads array was outsourced to Eve Technologies to profile 71 different human cytokines and chemokines related to inflammation and autoimmunity in the collected SNs at 2 h, 24 h and 72 h. Only those values above the limit of detection were reported in this work. Additionally, six quantiles were computed from all cytokine quantifications at 72 h, and the cytokines within the last quantile were excluded from the analysis.

9.10 FACS

Fluorescence activated cell sorting (FACS) to check CD4+ T cell contamination in co-culture samples was performed as follows. First, cells were counted and 500,000 cells were used per condition. Samples were transferred to a 1.5 ml tube and centrifuged 300 g 5 min. Cells were washed with 1X PBS to prevent inactivation of the blocking antibody. Samples were resuspended in 50 µl Blocking reagent (blocking antibody (#16-9161-73, eBioscience) diluted 1:20 in 1X PBS) and incubated 10 min covered at 4 ºC. Next, 1X PBS was added to each sample to dilute the blocking reagent and samples were centrifuged 300 g 5 min. Pellet was resuspended in 1X PBS with 1:16 anti-human CD4-PE (#12-0049-42, eBioscience) and incubated covered 25 min at 4 ºC. Then, the staining antibody was removed by diluting each sample with 1X PBS and centrifuged. Three more washes were done using 1X PBS. Finally, the pellet was resuspended in 200 µl of DAPI solution (1 µg/ml, #32670, Sigma) to discriminate live from dead cells. Samples were analyzed with a BD LSR II Flow Cytometer and FlowJo software was used for analyses.

9.11 Cell death assay

The percentage of viable and apoptotic cells was assessed by fluorescence microscopy as previously described[198]. Briefly, samples at 2 h, 6 h, and 24 h were stained with the DNA binding dyes Hoechst

33342 (HO, #382065, Sigma) and propidium iodide (PI, #P4170, Sigma). Cell death was determined in at least six different areas of each well. Results are expressed as percent of apoptotic cells, calculated as number of apoptotic cells/total number of cells × 100.

HO freely passes through the plasma membrane and enters both intact and damaged cells, staining DNA blue. In contrast, PI, a highly polar dye that cannot penetrate cells with intact membranes, stains DNA red.

9.12 CRISPRa experiments

The gene target of the RE containing the T1D-related SNP rs193778 was assessed using the CRISPR/Cas9 synergistic activation mediator (SAM) system. 6 different single guides RNAs (sgRNAs) were designed to target the RE but only 3 were finally selected for final experiments. The designed sgRNAs were cloned into the lentiSAMv2 vector (#75112, Addgene) and amplified by transformation into *Stbl3 E. coli*. EC cells were infected with lentiviruses containing the SAM vector with sgRNAs. The lentiviral particles were produced using HEK 293FT, filtered and concentrated before infection. EC cells were seeded in 12 well plates and infected with the volume of viral concentration determined by a previous titration test. After 24-72 h of infection, cells were selected using Blasticidine (6 µg/ml). After complete selection, cell lines were cultured for expansion. Finally, EC samples were collected for RNA extraction and qPCR analysis.

A control of EC cells without infection and antibiotic selection was included to ensure baseline cell viability. A selection control (no infection, with antibiotic) and a positive control of EC cells infected with GFP were also included. Additionally, three other negative controls for transcriptome changes were used, consisting of the same SAM backbone vector with sgRNAs targeting regions not present in the human genome. These vectors were also infected into EC cells, and changes in expression of potential target genes were normalized against these samples.

10

Bioinformatic analysis

The following bioinformatic processing and analysis were conducted by Georgina Fuentes, a PhD student from the laboratory.

10.1 ATAC-seq data processing

Raw fastq files were processed using the nf-core/atacseq pipeline version (**2.0**) with default parameters implemented in Nextflow v22.10.5. In short, reads were trimmed using trimGalore!, and aligned to the human reference genome GRCh38 release using Bowtie2, removing duplicate reads as well as reads mapping to non-canonical chromosomes or to ENCODE blacklisted regions using Picard tools and Samtools. Peak calling was performed using MACS2 with the argument "--narrow_peak".

10.2 CUT&Tag data processing

Raw fastq files were processed using the nf-core/atacseq pipeline version (2.0) with default parameters implemented in Nextflow v22.10.5. In short, reads were trimmed using trimGalore!, and aligned to the human reference genome GRCh38 release using Bowtie2, removing duplicate reads as well as reads mapping to non-canonical chromosomes or to ENCODE blacklisted regions using Picard tools and Samtools. Peak calling was performed using MACS2.

10.3 RNA-seq processing

Raw fastq files were processed using the nf-core/rnaseq pipeline version (3.11.1) with default parameters implemented in Nextflow v22.10.5. In short, reads were trimmed using trimGalore!, aligned to the human reference genome GRCh38 release using STAR and quantified by salmon.

10.4 Differential analysis of ATAC-seq, CUT&Tag and RNA-seq

For both ATAC-seq and CUT&Tag, aligned reads from all replicates were merged into a single BAM file to identify a comprehensive set of peaks. Next, the comprehensive peak set was used to compute read counts separately for each replicate and condition. In the case of the RNA-seq data, the output of salmon was used as the input matrix for downstream analysis. The generated matrices were normalized and differential analysis was performed using DESeq2 v.1.38.3 [199] using a paired sample design. Genes/regions were considered significantly gained when adjusted p-value < 0.05 and log2FC > 1 and significantly lost when adjusted p-value < 0.05 and log2FC < -1. All regions/genes that did not reach significance or did not pass the log2 fold change cutoff were classified as stable/equally regulated.

Table 1. List of islet-specific markers

ABCC8	G6PC2	INS	MYT1	PDX1	SNAP25
ADCYAP1	GAD2	INSM1	NBRE	PRKCA	SOX17
ARX	GATA4	ISL1	NCAM1	PRLR	SOX4
BHLHA15	GATA6	KCNB2	NEUROD1	PTF1A	SOX9
CACNA1C	GCK	KCNJ11	NEUROG3	PTPRN	ST18
CACNA1D	GIPR	KCNK10	NGN3	PTPRN2	STX1A
CACNA1E	GJD2	LMX1A	NKX2.2	RFX3	SULT4A1
CDKAL1	GLIS3	MAFA	NKX6.1	RFX6	SYT4
CHGA	GLP1R	MAFB	NR4A3	RGS4	SYT7
CPE	HB9	MEF2A	ONECUT1	RPBJ	TCF7L2
DACH1	HES1	MEF2C	PAX4	SCG2	TM4SF4
FFAR1	HNF1A	MEF2D	PAX6	SCG5	UCN3
FOXA2	HNF1B	MIST1	PCSK1	SCGN	WNT4
FOXO1	HNF6	MNX1	PCSK2	SLC2A2	
FXYD2	IAPP	MNX1	PDE3B	SLC30A8	

10.5 Defining classes of IREs

To characterize the dynamics of chromatin accessibility, the results obtained from the DESeq2 differential analysis were processed and the direct overlap between ATAC-seq peaks and H3K27ac-enriched sites was computed. Regions annotated as stable for both ATAC-seq and H3K27ac assays were classified as SREs. Regions classified as either stable or gained in ATAC-seq differential analysis and as gained in H3K27ac were classified as IREs. To remove redundant information and identify unique mechanisms over time, we further classified the set of IREs, ensuring that once a regulatory element was activated, it was not included in later time points. Thus, we only kept IREs that are unique for each specific time point. Regions classified as lost in ATAC-seq or H3K27ac differential analysis were classified as LoREs.

10.6 Assigning regulatory elements to target genes

To annotate regulatory elements as distal or proximal, we assigned each regulatory element to the nearest TSS of a coding gene (using GENCODE release 18 annotation 60). Those regions lying within 2 kilobases (kb) from the nearest TSS were annotated as promoters while the rest were considered as distal regulatory elements.

To analyze the effect of IREs and LoREs on gene expression changes, each IRE/LoRE was assigned to all DEGs whose TSS was closer than 40 kb. When an upregulated gene could not be found in <40 kb, the IRE was assigned to the closest, but <1 Mb far, induced.

10.7 T1D GWAS SNPs overlap with regulatory elements

T1D-associated SNPs were obtained from the National Human Genome Research Institute European Bioinformatics Institute (NHGRI-EBI) GWAS catalog. All variants in high linkage disequilibrium (1000 Genomes Project, phase 3 European population (EUR), R2 >0.8) with the leading SNP were obtained using the LDlinkR package (v1.3.0).

Part IV

Results

11

Co-culturing β cells with T lymphocytes and APCs

The first step in order to fulfil the aims of my project was to develop a new *in vitro* model mimicking early insulitis in T1D to study the β cell behaviour in this inflammatory environment. Current human models are mainly based on exposure of human β cells to specific cytokines. This can be modulated 1) in terms of composition, from just one cytokine to a cocktail of several cytokines, and 2) in terms of dose, by exposing the cells to cytokines at different concentrations. All these models have some limitations, as, to optimally mirror insulitis, they require a priori knowledge of all the concentration and type of chemokines and cytokines participating to the inflammatory process (Table 2, section 5.2.2, Introduction). Moreover, the stimulus is always abrupt, it goes from resting state to maximum stimulus directly while missing the effect of cell-to-cell communication that may occur during this process.

For models of later stages of insulitis, where direct CD8+ T cell killing is occurring, significant attention has been devoted to developing in vitro models that help unravel the interactions between β cells and other involved immune cell populations, primarily CD8+ T cells[127]. This need is apparent due to the physical contact required for direct β cell killing. However, these interactions may also be crucial in early insulitis, where the different populations involved “communicate” behaviour through the release and uptake of different cytokines and chemokines.

Therefore, in order to better recapitulate the complex scenario which β cells are exposed to, I decided to re-direct the focus on which cell types are important rather than which cytokines or molecules are present in the islets at this stage of the disease. As described in the introductory section, CD4+ T cells are one of the main players at early stages of T1D since they modulate the activation or suppression of downstream effectors of the immune response. Hence, in collaboration with Pere Santamaria, MD, PhD, team leader at IDIBAPS and professor at University of Calgary, I developed a new *in vitro* model of early insulitis based on the co-culture of β cells with CD4+ T cells to mirror the dynamic inflammatory changes at this stage and study the effects on the β cells regulatory networks.

11.1 Co-culture of β cells with APC and CD4+ T cells in 96-well plates

I initially set up a complex model by co-culture with EndoC-βH1, a primary human transgenic CD4+ T cell line modified to express a TCR that recognizes the IGRP13-25 T1D antigen in the context of a specific HLA class II T1D risk haplotype (DRB1*03:01), and an immortalized antigen presenting (APC) cell line that has been transduced to express the HLA class II haplotype that the CD4+ T cells recognize. I will from now on refer to this model as "**CC-M1.1**". The idea of combining these three cell populations is to study the response of the β cell in an immune context that resembles the likely scenario of a genetically predisposed individual within autoantigens-activated T lymphocytes. I co-cultured these three populations in 96-well plates and collected samples at different time points to capture the dynamics of molecular changes in β cells (Fig 1). The experiment design included 1) the co-culture samples treated with the IGRP13-25 antigen (IGRP), 2) a positive control in which CD4+ T cells were activated by CD3/CD28/CD2 antibody complex (CD3), 3) a negative control in which CD4+ T cells have not been activated (DMSO) and 4) negative control in which EndoC-βH1 (EC) were cultured alone. I analyzed the β cells transcriptome at different time points (24 h, 48 h) and collected the cell culture medium to confirm the activation of CD4+ T cells.

IFNγ abundance in the medium, measured by ELISA, confirmed the activation of the CD4+ T cells, both in the IGRP and the positive control samples. No activation was observed in the negative control (DMSO) sample (Fig 2A). Activation levels in the CD3 sample (artificially activated CD4+ T cells) were more than 10 folds higher as compared to the samples exposed to the IGRP antigen which were coupled with less marked changes in gene expression of key IFNγ-response genes. Overall, the activation of CD4+ T cells led to mild changes in transcription of EndoC-βH1 in the CD3 sample. Gene expression changes were also detected but to a less extend in the samples using the IGRP antigen (Fig 2B). The levels of IFNγ, used as a proxy for the CD4+ T cell activation, were quite low, especially in the IGRP samples, when compared to the IFNγ concentration used in other models of insulitis (Table 2, section 5.2.2 in introduction). Thus, driven by these results, I decided to further optimize the model in order to obtain a stronger activation of the lymphocytes.

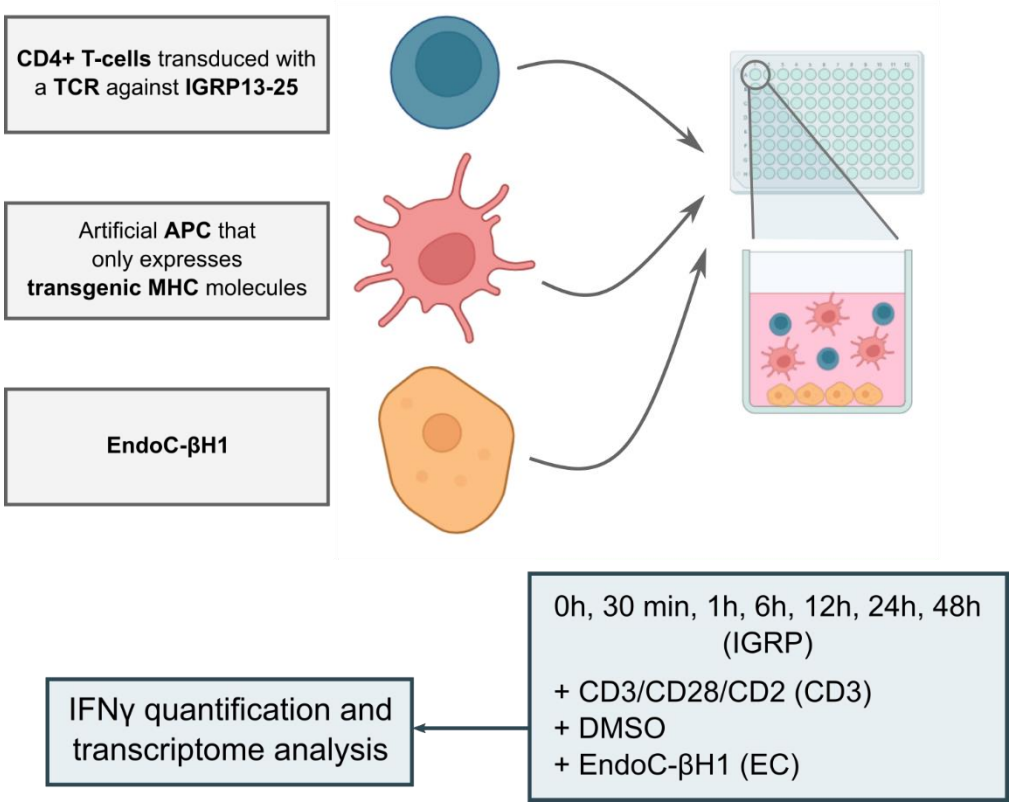


Figure 1. Schematic representation of Model 1.1 Co-culture (CC-M1.1).

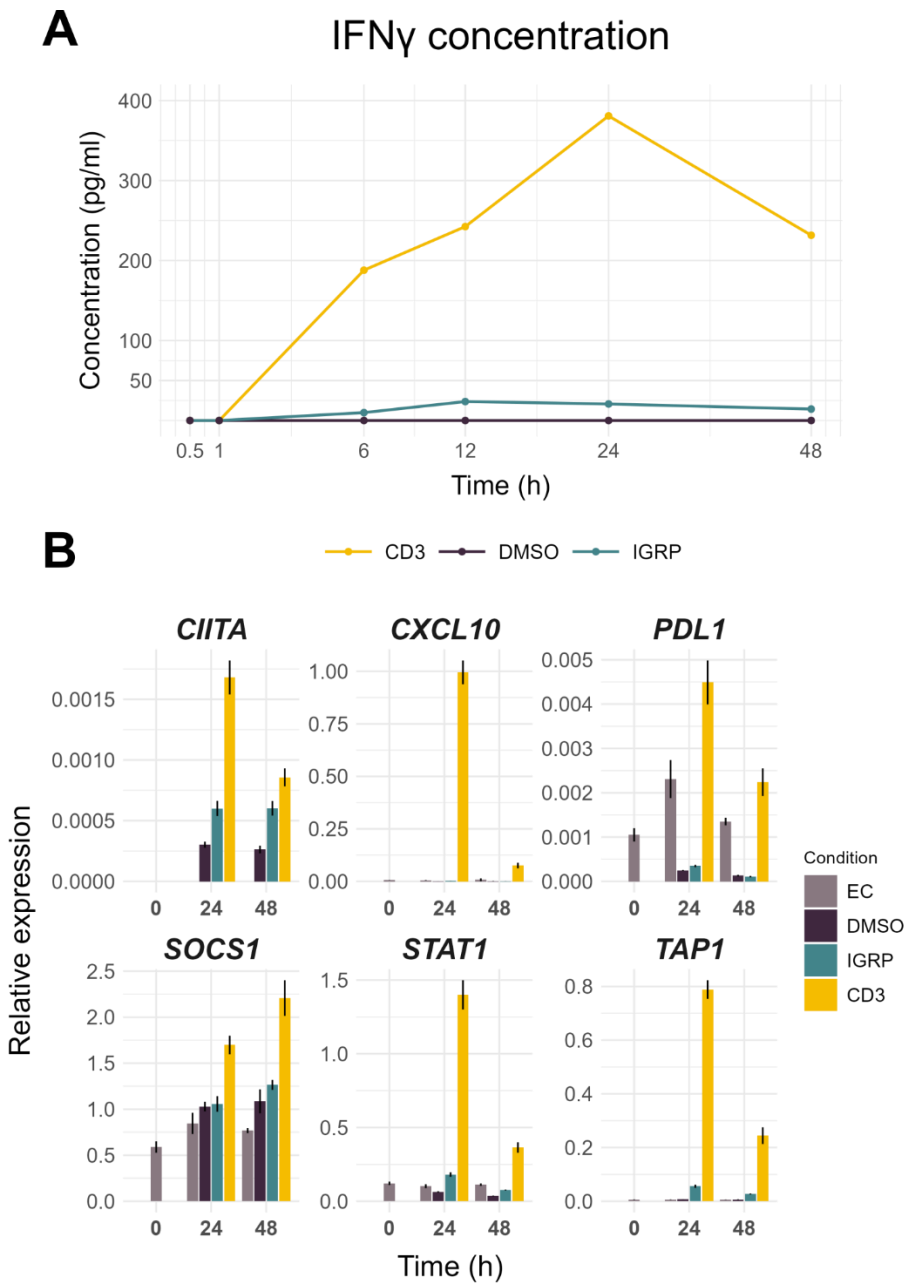


Figure 2. CC-M1.1 activation of CD4+ T cells led to a mild response in the transcriptome of the β cells. A. IFNγ levels present in the culture medium after each time point for each condition. Measured by ELISA and expressed as pg/ml. B. Relative expression of CXCL10, TAP1, STAT1 and SOCS1 at 24 h and 48 h of co-culture in EndoC-βH1. Gene expression was normalized to the mean expression of three reference genes. Bars represent mean +/- SD.

11.2 Co-culture of β cells with APC and CD4+ T cells in 48-well plates

In order to address the potential limitations revealed in the CC-M1.1, I increased the co-culture format and the density of every cell population to increase the stoichiometric probability of CD4+ T cell activation in the context of a new experimental setting (CC-M1.2) (Fig 3A). The IFNγ abundance at 24 h increased notably as compared to CC-M1.1 reflecting and improved induction of CD4+ T cell activation in both positive control and IGRP conditions (Fig 3B). However, the increase in IFNγ in both conditions did not translate much into increased changes in gene expression of marker genes measured by qPCR. Therefore, I performed RNA-seq to assess the breadth of transcriptional changes and evaluate the model's efficacy in capturing β cell responses to an inflammatory stimulus.

RNA-seq experiments revealed lymphocytes-induced changes in the transcriptome of the β cells. The transcriptome changes were in line with the abundance of INFγ detected in the media as a proxy of the CD4+ T cell activation (Fig 4A and 4B). Almost all the changes observed coincided with well characterized genes known to be part of the β cell response to a proinflammatory stimulus[113, 200, 201].However, taking into account the levels of CD4+ T cell activation and RNA-seq results, I hypothesized that these changes could reflect a mild response to a low inflammatory stimulus. Indeed, another work arrived to a similar conclusion when using a similar dose of IFNγ in a cytokine cocktail, but lacking power to profile transcriptome nor chromatin changes[195]. Therefore, although these results were promising, I decided to re-design the co-culture model to improve the activation of CD4+T cells to induce higher levels of cytokine production and increase the inflammatory stimulus to be able to better understand the β cell behaviour in early insulitis.

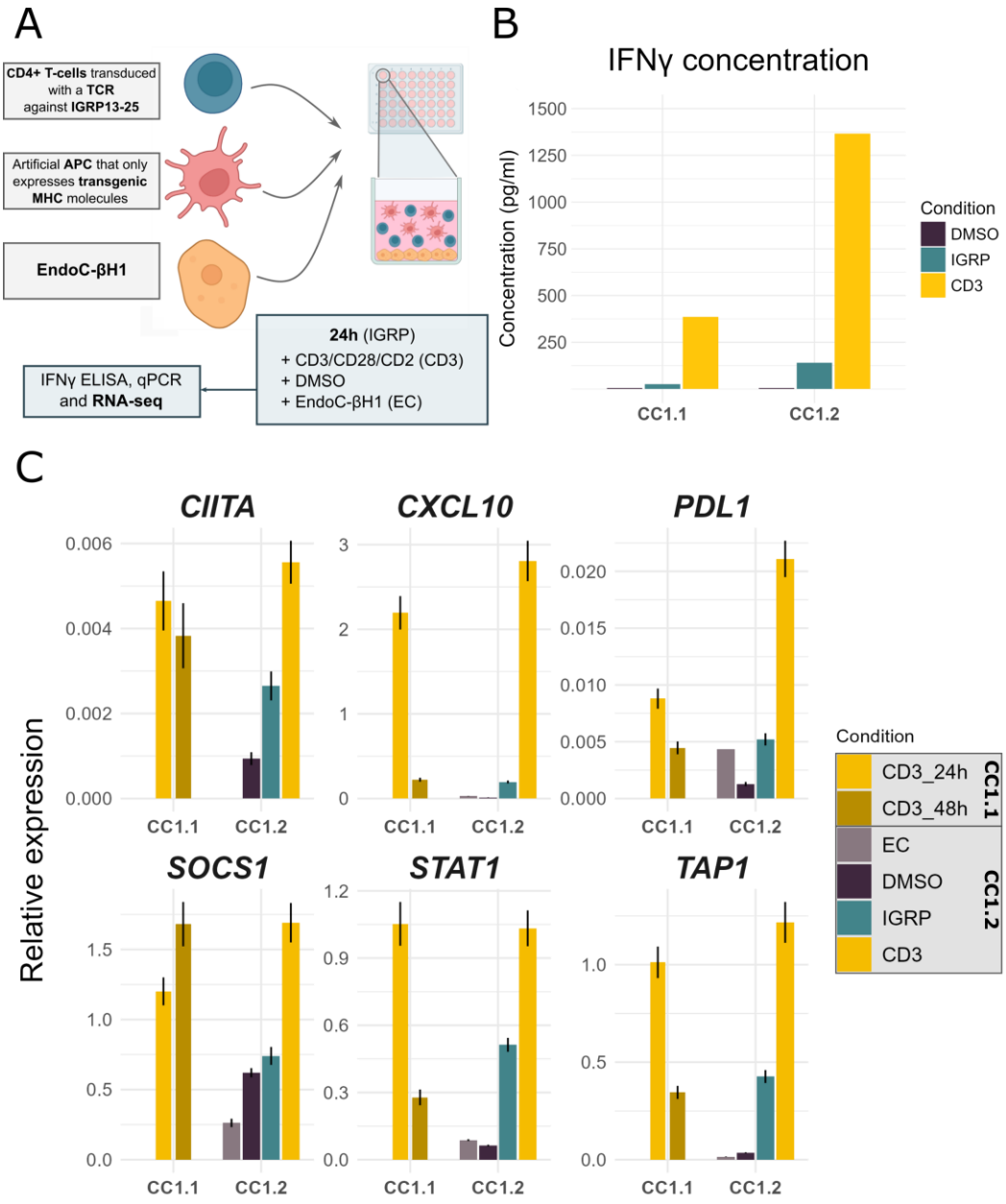


Figure 3. CC-M1.2 activation of CD4+ T cells led to an improved response in the transcriptome of the β cells in the IGRP condition. A. Schematic representation of Model 1.2 Co-culture (CC-M1.2). B. IFNγ levels present in the culture medium after 24 h compared to results of CC-M1.1 at same time point. Measured by ELISA and expressed as pg/ml. C. Relative expression of CXCL10, TAP1, STAT1 and SOCS1 at 24 h and 48 h of co-culture in EndoC-βH1 from CC-M1.1 and CC-1.2. Bars represent mean +/- SD.

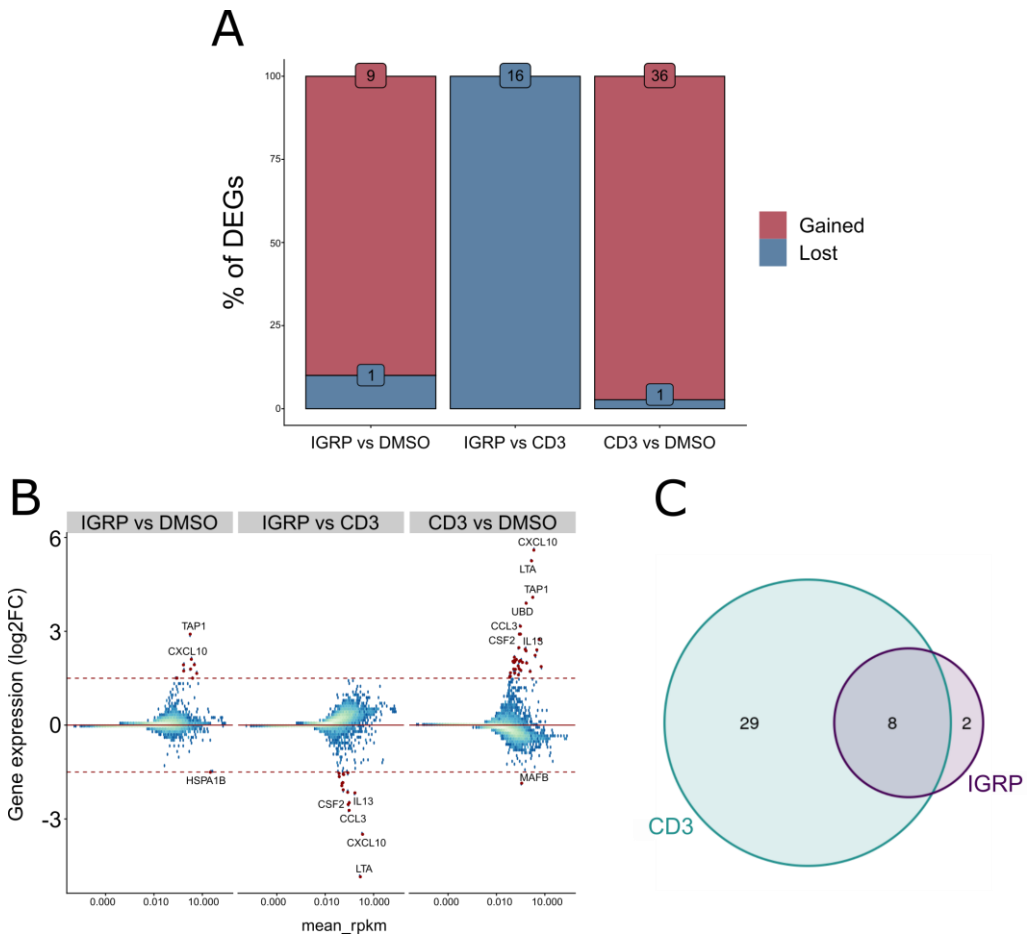


Figure 4. RNA-seq results from CC-M1.2 show limited transcriptional response in the IGRP sample. A. Barplot with the number of DEGs in each condition compared to the negative control (DMSO) and differential analysis between IGRP and CD3. B. MA plot comparing the gene expression changes of IGRP vs DMSO; IGRP vs CD3 and CD3 vs DMSO. C. Venn diagram overlapping the DEGs from CD3 vs DMSO and IGRP vs DMSO.

By observing these results, I reasoned a few important points to take into account when changing the model:

- The co-culturing model CCM1 is a complex setup which requires physical interaction of 3 distinct cell populations. The experimental design is thus particularly difficult to optimize.
- The β cell transcriptome changes observed by exposure of CD4+T cells CD3-activated (used in the experimental design as

a “positive control”) are qualitatively very similar to those obtained by IGRP CD4+ T cell activation. When directly comparing the response of both conditions, the only two DEGs in IGRP not present in CD3 are *MT2A* and *HSPA1B* (Fig 4C). However, when directly comparing the IGRP vs CD3 transcriptome, the only DEGs detected are downregulated, with no genes expressed in IGRP condition not found in CD3 (Fig 4B).

- Activation of the CD4+ T cells by CD3/CD28/CD2 complex can be further optimized in a much more flexible co-culture system.

Moreover, the process of engineering CD4+ T primary cells to recognize the IGRP13-25 antigen from primary cells is challenging and typically yields batches with a low cell count. These lymphocytes are thus often highly exhausted by the conclusion of the process. Based on these observations, I decided to optimize a co-culture experimental design based on CD3 complex CD4+ T lymphocytes activation in order to overcome some of the limitation encountered but maintaining the overall complexity of the experimental design as compared to exposure to specific cytokine cocktails.

12

Co-culture of β cells with T lymphocytes in 48 wells

In the previous setup, the CD3 condition involved co-culturing all cell populations similar to the IGRP condition but used an antibody complex (CD3/CD28/CD2) instead of the IGRP13-25 antigen for lymphocyte activation to ensure strong signal transduction to lymphocytes independently of APC cells. This method has proved more effective and RNA-seq results showed that the gene expression changes in the IGRP condition were also present in the CD3 condition (Fig 4A and 4B, IGRP vs CD3). Hence, I decided to focus on using the CD3 complex to activate lymphocytes. This approach reduces complexity ensuring an efficient response of CD4+ T cells consistent for each donor, as in this set up, the lymphocytes do not require engineering and selection.

I thus design a new experiment in which I combined in the same culture both human primary CD4+ T cells and EndoC-βH1. I included the 1) treated sample, named “CD3”, 2) a negative control in which the lymphocytes were not activated, named “DMSO” and 3) a negative control of bare EndoC-βH1. I collected samples at different time points for IFNγ quantification and transcriptome profiling (Fig 5A). Quantification of IFNγ revealed higher concentration of this cytokine present in the co-culture medium compared to the previous design (Fig 5B). After confirmation of high levels of lymphocyte activation, I analysed the transcriptome changes of the CD3 sample to profile the response of the β cells to the stimulus.

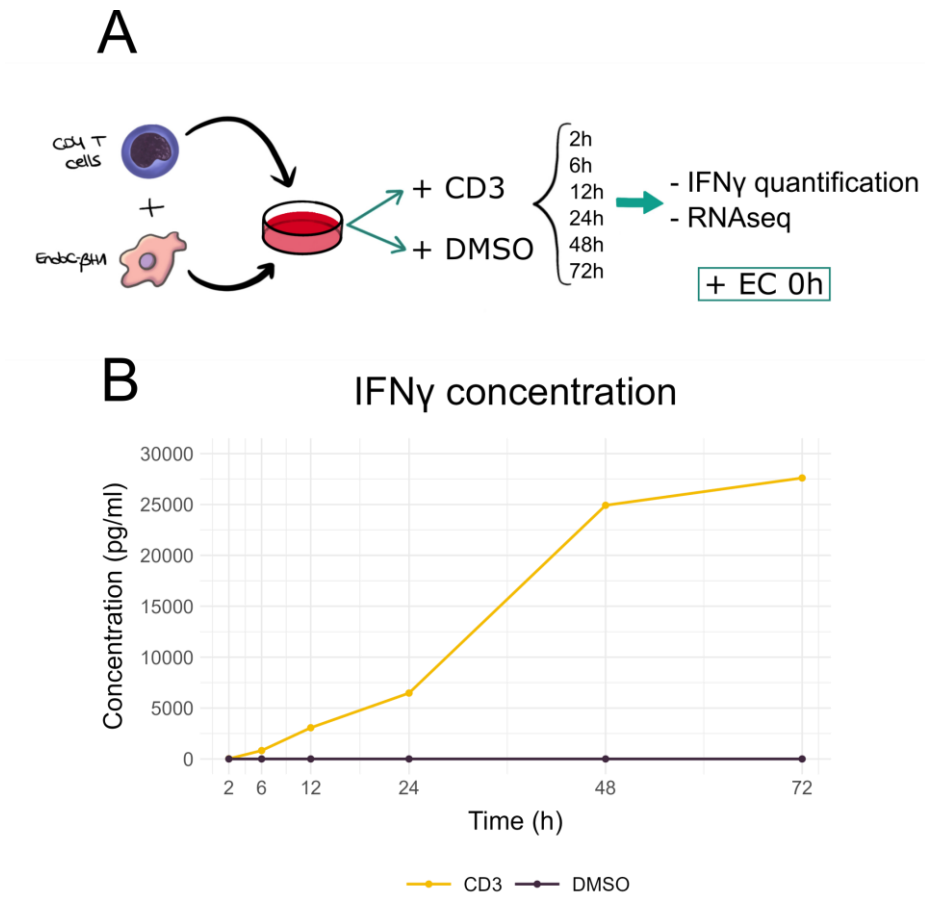


Figure 5. Re-design of co-culture led to a strong activation of CD4+ T cells. A. Schematic representation of Model 2 Co-culture (CC-M2). B. IFNγ levels present in the culture medium after each time point for each co-culture condition. Measured by ELISA and expressed as pg/ml.

RNA-seq experiments showed a relevant response of the β cells to the stimulus (Fig 6A) characterized, over time, by an increasing number of differentially expressed genes (DEGs), including both upregulated and downregulated genes. Gene upregulation seems to appear earlier in time than gene downregulation but both reaching the maximum number of DEGs at the later time point. Of note, most DEGs, with few exceptions, follow a similar trend in the time course as once a gene is upregulated or downregulated it maintains the expression levels over later time points (Fig 6B). Importantly, a relevant proportion of downregulated genes could be observed at later time points, around 40-

45% of DEGs, something that, to my knowledge, was not reported in other *in vitro* models of early insulitis (Fig 6A). I also checked gene expression changes for specific markers of inflammatory response, such as *SOCS1, IRF1, STAT1 and CXCL10*, and for β cell specific markers, *MAFB* and *PDX1*. Both β cell-specific markers follow a similar trend of downregulation after 48 h of co-culture, while inflammatory markers seem to follow different dynamics (Fig 6C).

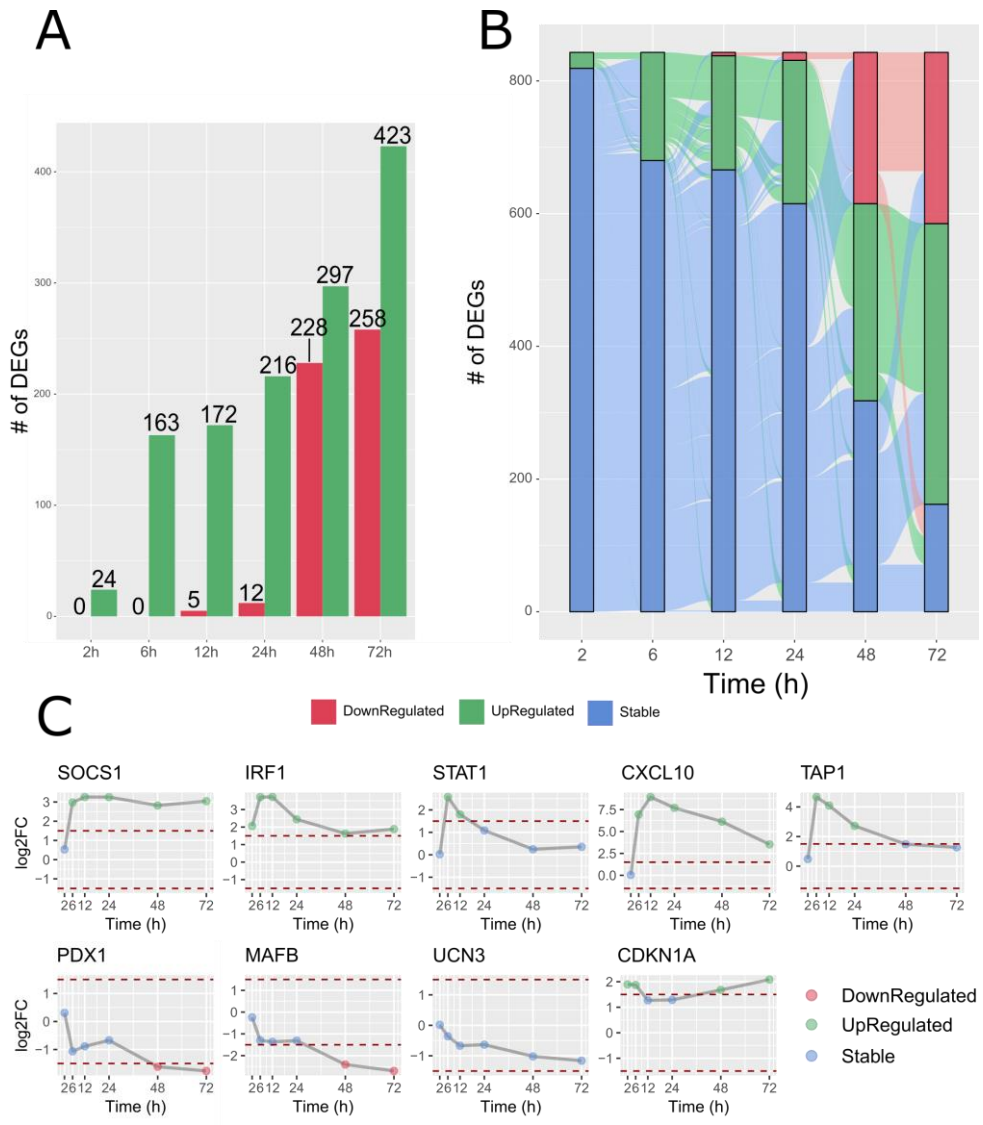


Figure 6. CC-M2 activation of CD4+ T cells induced a response in the β cells characterized by extensive changes in gene expression. A. Barplot depicting the number of differentially expressed genes (DEGs) in both directions, up or down, of CD3 compared to DMSO at the different time points. B. Alluvial plot showing the dynamic changes in gene expression in the CD3 sample through all the time course of the experiment. C. Gene expression changes of different marker genes of inflammation, β cell identity and senescence through time.

This model is based on culturing both populations in the same plate where EndoC-βH1 are expected to grow attached to the substrate and CD4+ T cells in suspension. Conceptually, separating both populations before sample collection should be straightforward. However, based on microscope observations indicating an incomplete separation of the two cell populations, I decided to check the amount of CD4+ T cells present in the β cell sample after co-culture before proceeding with further experiments. To do so, upon a 72 h co-culture I quantified the number of lymphocytes present in the cells detached from the plate and after removing those in suspension. Staining the sample with an anti-CD4 antibody and DAPI for FACS analysis revealed that over 70% of total cells were CD4+ T cells, while the DMSO sample only contained 14% of CD4+ T cells (Fig 7A and 7B). These results were consistent with previous concerns from microscope observations indicating that CD4+ T cells, when activated, tend to attach to the EndoC-βH1 cells rather than forming aggregates and growing in suspension. In fact, not only the levels of lymphocyte contamination represent an obvious problem but also the striking difference in cell composition between the CD3 and DMSO samples, that make impossible the comparison between samples.

In addition, the DAPI staining revealed that CD4+ T cells display good levels of cell viability (Fig 7C) while β cells displayed reduced viability with activation of CD4+ T cells, pointing to an increase in the levels of β cell death (Fig 7C).

Given the challenges encountered in separating CD4+ T cell from β cell after a co-culture time course, I decided to follow a new strategy aimed at achieving similar activation results but being able to discriminate the signal from the two cell populations.

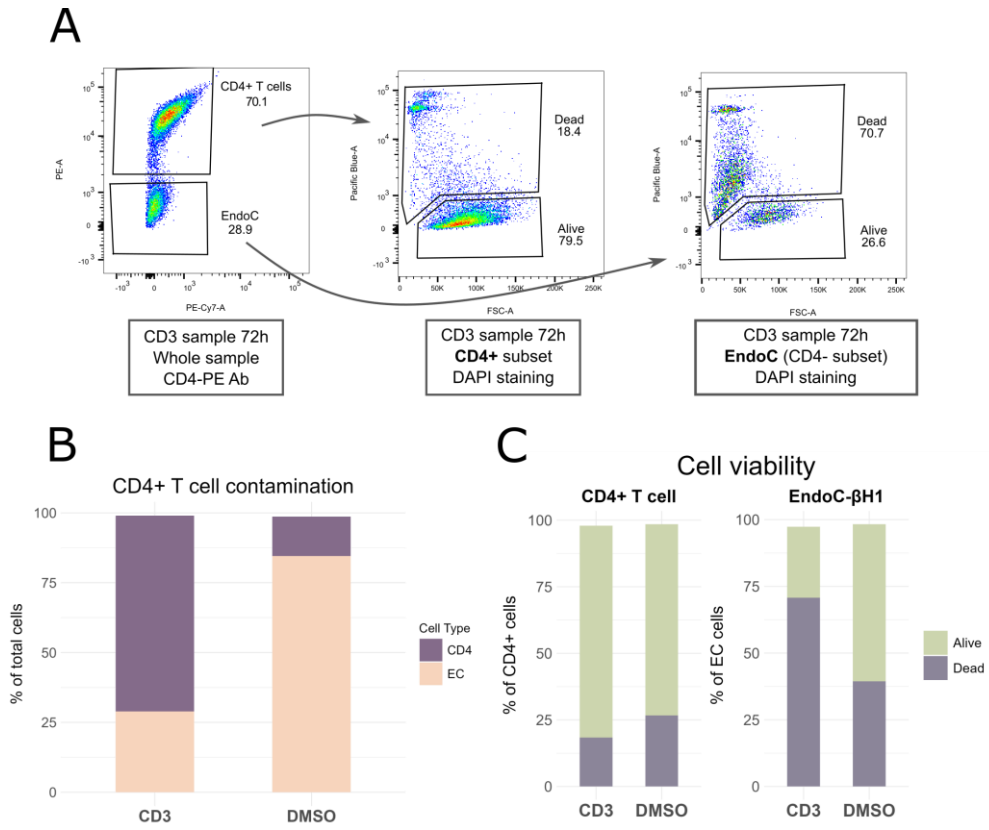


Figure 7. Sample of β cell from CC-M2 contains a high level of CD4+ T cell contamination. A. Gating strategy for FACS to distinguish between CD4+ and CD4- cells and viability of each subset. B. Percentage of CD4+ T cells and β cells in each condition after 72 h of co-culture. C. Percentage of CD4+ T cells or β cells alive and dead in each condition after 72 h of co-culture.

13

Co-culturing β cells with T lymphocytes using transwells

With the aim of achieving a proper discrimination of the co-cultured cells and being able to identify the β cell-specific molecular responses, I decided to use a transwell insert to physically separate both populations in culture. This system consists on a two-compartment culture vessel separated by a porous membrane (Fig 8A). Depending on the size of these pores, these cell culture inserts can have many applications such as co-culture or migration assays. In my case, I took advantage of this system to culture CD4+ T cells in the upper compartment with EndoC-βH1 in the bottom one, separated by a porous membrane that does not allow cell migration from one compartment to the other (Fig 8A). However, the small size of the pores does allow the cell culture medium to be shared between both compartments; in this way the molecules including chemokines and cytokines produced and secreted in one compartment will reach the other one. This approach allows both cell populations to share their media but at the same time permits to solely capture the β cells' molecular response signal.

13.1 Transwells co-culture set-up

In order to find the optimal conditions for the transwell co-culture set up, I performed an explorative experiment (CC-M3.1) using two different number of CD4+ T cells, 700,000 and 900,000 per well. As compared to the previous experiments described, the transwell co-culture set up requires increasing total culture volume while maintain the cell seeding density and nutrients used in previous experiments. IFNγ quantification of culture medium after different time points revealed high levels of CD4+ T cell activation, reaching concentrations of this cytokine very similar to those obtained in the CC-M2 (Fig 8B). Exposure to these concentrations of IFNγ in the medium also lead to a transcriptional response in the β cells, as it was evident from gene expression quantification by qPCR of specific inflammatory genes and β cell specific markers (Fig 8C). Inflammatory markers are highly upregulated in all

cases compared to both DMSO and EndoC-βH1 not exposed to lymphocytes, with almost no differences between experiments performed with different number of CD4+ T cells (Fig 8C). I noticed that exposure to activated lymphocytes seems to induce a downregulation of β cell markers although to a less extend when compared with CC-M2. Overall, the differences in the induced gene expression changes observed in the in CC-M3 as compared to CC-M2 are likely due to the lack of capacity, in the latter model, to discriminate the contribution of the two cell populations in the transcriptome data (Fig 8C and 7B).

Of note, in this preliminary test, I also included a 96 h time point in one condition (700,000 CD4+ T cells) to comprehend the dynamics of lymphocyte activation and β cell transcription changes after 72 h to better design the final experiment. In this 96 h time point, there is a slight increase in IFNγ concentration, although CD4+ T cell activation seems to be reaching a plateau (Fig 8B). In fact, the increase in IFNγ concentration and time of exposure do not translate into increased changes in gene expression of key markers, showing a response very similar to the 72 h time point.

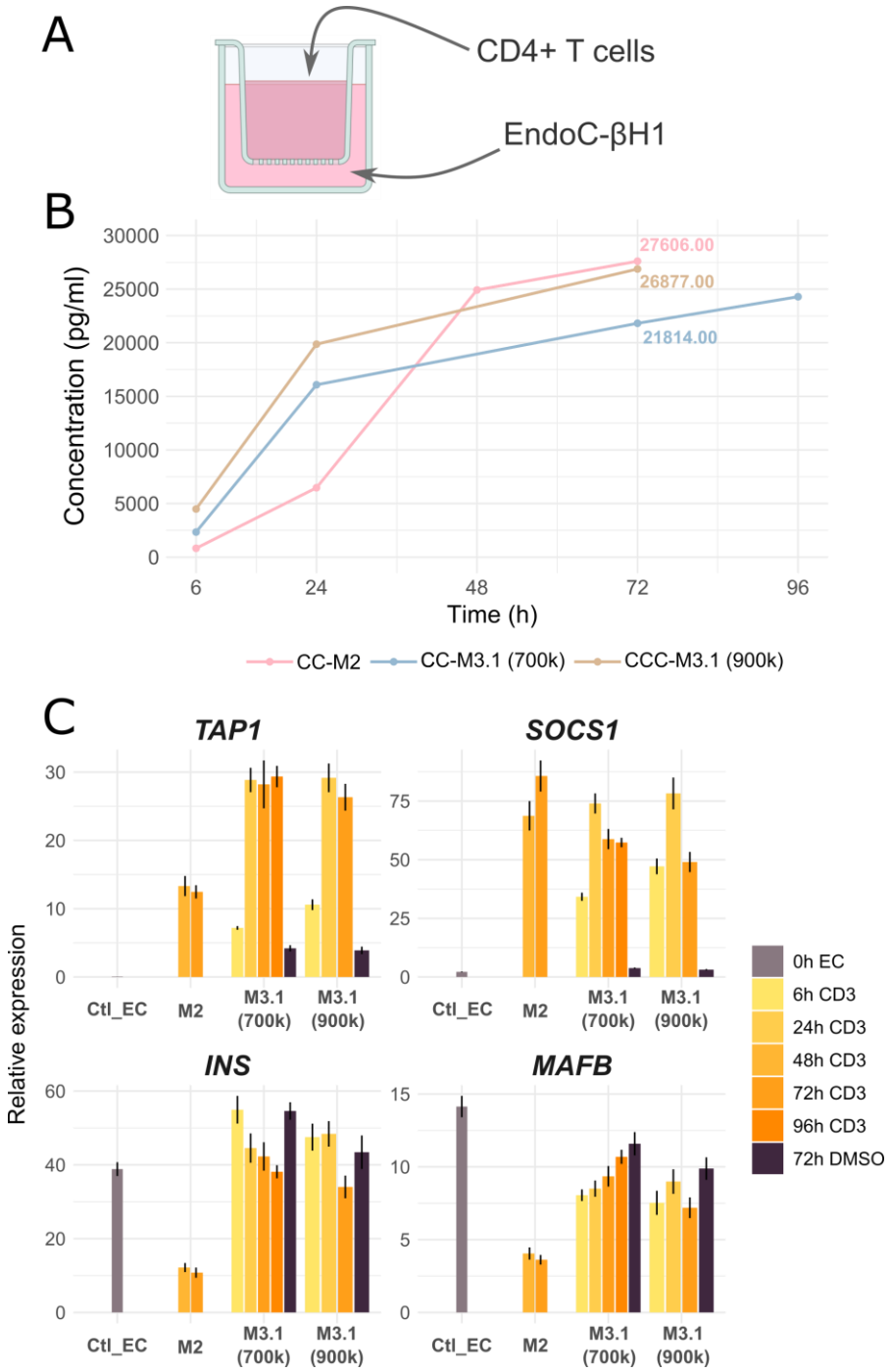


Figure 8. Activation of CD4+ T cells using a cell culture insert in CC-M3.1 led to the presence of IFNγ in the medium and changes in β cell's gene expression. A. Graphical representation of a transwell system, with the cell culture insert as top compartment and the regular well as bottom compartment.

B. IFNγ levels present in the culture medium of CC-M3.1 with 700,00 cells and 900,00 cells at the different time points compared to IFNγ levels of CC-M2. C. Relative expression of two inflammatory markers (TAP1 and SOCS1) and two β cell specific markers (INS and MAFB) in CC-M2 and CC-M3.1 at the different time points. Bars represent mean +/- SD.

In summary, the findings from CC-M3 indicate that the transwell system serves as a highly effective means to capture CD4+ T cell-induced alterations in a specific cell type such as the β cells. I observed that employing 700,000 CD4+ T cells for a 72-hour time course is adequate for comprehending the dynamic changes in β cells following this proinflammatory stimulus. Finally, after a 96-h exposure followed by a 48-hour resting period, in which the CD4+ T cells were removed from the co-culture, I measured gene expression in β cells (Fig 9A). Interestingly, expression levels of inflammatory markers were strongly reduced after 48 h of resting reaching similar levels of expression as in the negative controls while β cell-specific markers increased after resting, going back to normal levels. These results suggest that, even though β cells show a strong response when exposed to an inflammatory stimulus, they seem to be able to rapidly recover from it when the insult is ceased.

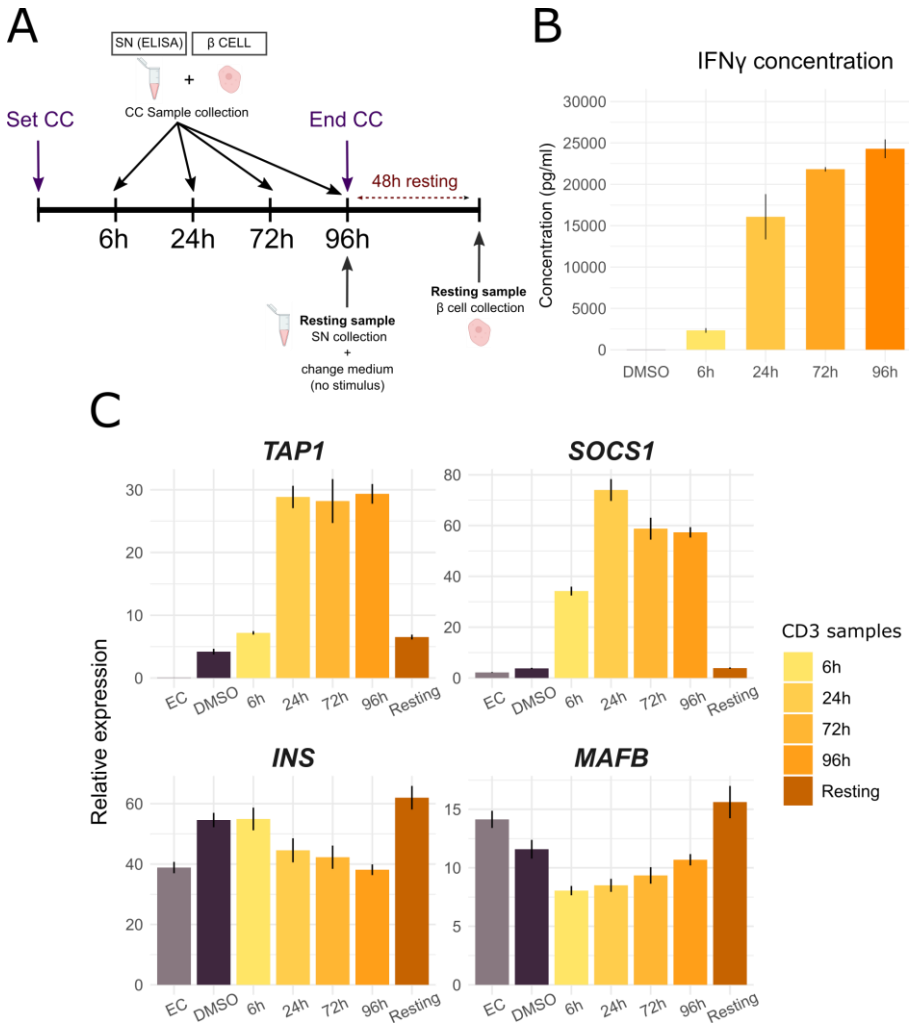


Figure 9. A resting period of 48 h after 96 h of co-culture seems to be enough for β cells to recover from the stimulus and progressively go back to normal gene expression levels. A. Schematic representation of CC-M3.1 using 700,000 CD4+ T cells. Resting sample was collected after 96 h exposure + 48 h resting period. SN of resting sample was collected at the end of exposure. B. IFNγ levels present in the medium of CC-M3.1 with 700,00 cells at the different time points. Of note, the levels of resting sample correspond to those present at 96 h of stimulus, right before removing the stimulus. C. Relative expression of two inflammatory markers (TAP1 and SOCS1) and two β cell specific markers (INS and MAFB) in CC-M3.1 with 700,000 CD4+ T cells at the different time points. Bars represent mean +/- SD.

14

Secretome profiling of β cells co-cultured with CD4+ T lymphocytes

Taking into consideration all the previous results, I designed an experiment to profile, at different time points, the inflammatory response of β cells at the transcriptional and gene regulation levels (Fig 10A). Of notice, for this experimental design I used regular EndoC-βH1 rather than DMSO treated co-cultures as negative controls. The rational of the use of this control was due to data recollected in previous experiments showing that resting lymphocytes in DMSO present residual activation resulting in the release of low levels of IFNγ in the co-culture medium. Although the activation is minimal it results in gene expression changes compared to non-co cultured EndoC-βH1 (the data is not shown but can be appreciated in Fig 8C).

The experiment was performed in replicates (n=3), each using a different human donor of CD4+ T cells. I studied the medium composition, the transcriptome and the chromatin profile in β cells at different time points.

Quantification of IFNγ was employed to ensure consistent lymphocyte activation across all replicates (Fig 10B). In an effort to examine the different proteins released in the medium after co-culture, I applied a commercial bead-based cytokine array for over 70 different cytokines related to autoimmunity and diabetes. Results underscore the predominant presence of IFNγ in the medium, alongside with numerous other cytokines present at varying concentrations (Fig 10C). Beyond IFNγ, the analysis uncovered several cytokines known to be secreted by different subtypes of CD4+ T cells, as well as others that could potentially be produced by β cells in response to the stimulus. Therefore, I decided to integrate the array results with RNA expression data obtained from the β cells at different co-culture time points (Fig 10C). I classified the medium components according to the likely cell of origin, considering the expression levels in β cells and their biological function. Interestingly, most of the components are attributed to CD4+ T cells (Fig 10C, green names), like IFNγ, IL13, CD40L, IL4, IL9, IL17A, or IL13, as

low or no expression of these genes was detected in β cells. In the case of VEGF-A and PDGF-A, both genes are expressed but with a tendency of downregulation in β cells. Taking into account their function and the opposite direction between expression and protein abundance, I assigned them to CD4+ T cell production. However, other regulatory mechanism of protein synthesis and secretion might be contributing to the presence of these molecules.

On the other hand, several cytokines are highly likely produced by β cells (Fig 10C, red names) or even both populations at the same time, confirming the bidirectional communication established between the immune cells and β cells within the inflammatory context (Fig 10C). It includes CXCL10, CXCL9, TRAIL and LIF, among others.

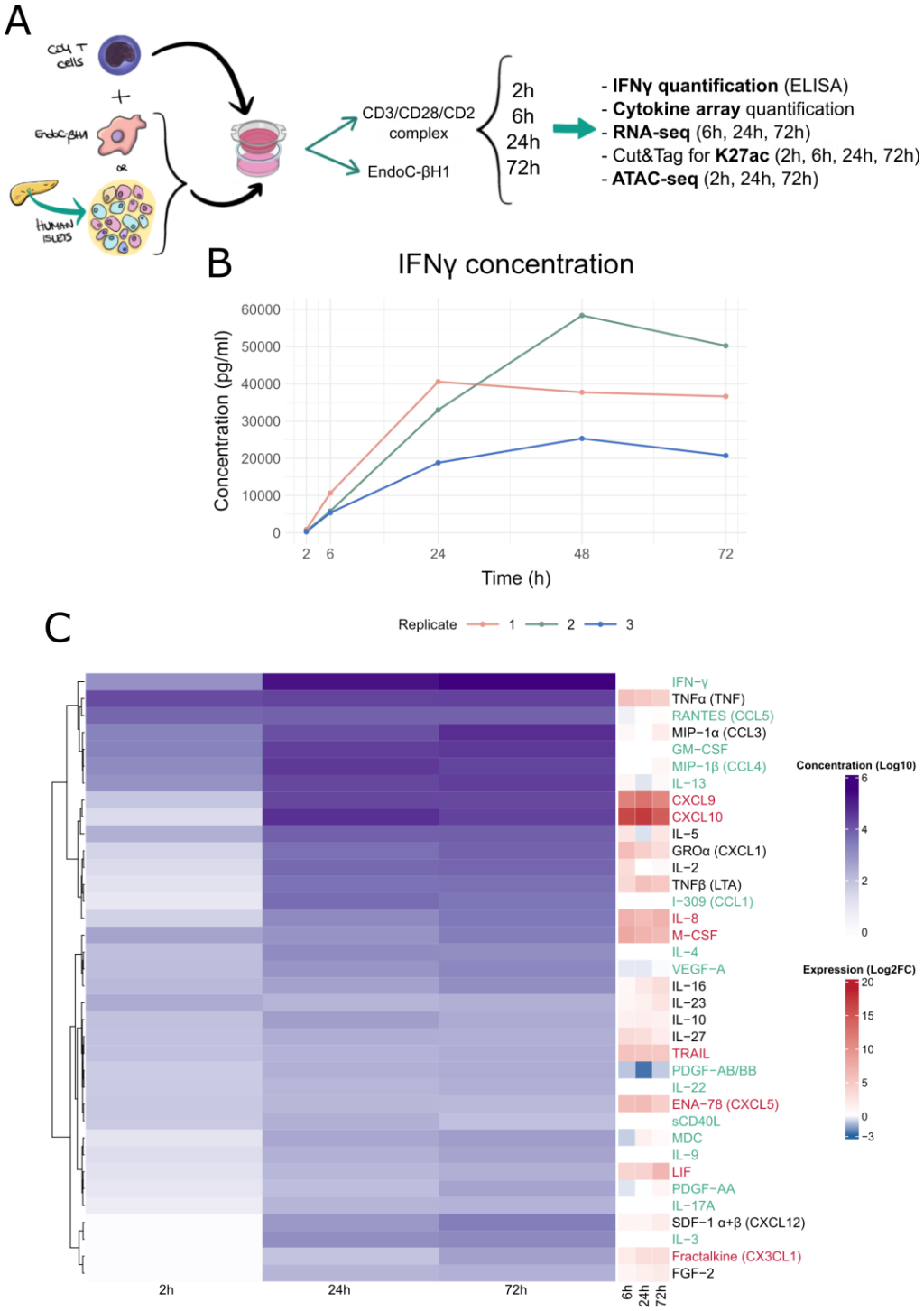

A
CD4 T cells
+
EndoC-βH1
or
HUMAN ISLETS
CD3/CD28/CD2 complex
EndoC-βH1
2h
6h
24h
72h
- **IFNγ quantification** (ELISA)
- **Cytokine array** quantification
- **RNA-seq** (6h, 24h, 72h)
- Cut&Tag for **K27ac** (2h, 6h, 24h, 72h)
- **ATAC-seq** (2h, 24h, 72h)
B
IFNγ concentration
Concentration (pg/ml)
60000
50000
40000
30000
20000
10000
0
2
6
24
48
72
Time (h)
Replicate 1 2 3
C
IFN−γ
TNFα (TNF)
RANTES (CCL5)
MIP−1α (CCL3)
GM−CSF
MIP−1β (CCL4)
IL−13
CXCL9
CXCL10
IL−5
GROα (CXCL1)
IL−2
TNFβ (LTA)
I−309 (CCL1)
IL−8
M−CSF
IL−4
VEGF−A
IL−16
IL−23
IL−10
IL−27
TRAIL
PDGF−AB/BB
IL−22
ENA−78 (CXCL5)
sCD40L
MDC
IL−9
LIF
PDGF−AA
IL−17A
SDF−1 α+β (CXCL12)
IL−3
Fractalkine (CX3CL1)
FGF−2
2h
24h
72h
6h
24h
72h
Concentration (Log10)
6
4
2
0
Expression (Log2FC)
20
15
10
5
0
−3

Figure 10. Activation of CD4+ T cells in CC-M3.2 leads to high levels of IFNγ in the medium. A. Schematic representation of CC-M3.2, done with EC.B. IFNγ levels present in the medium of CC-M3.2 at the different time points for each replicate quantified by ELISA. C. Co-culture medium composition at different time points (n=1) checked using a cytokine array. The heatmap on the left represents the protein concentration detected in the media at different timepoints. The heatmap on the right represents the changes in expression in β cells of the gene encoding the corresponding protein detected in the media at different time points. Green highlight means cytokines likely produced by CD4+ T cells with low or no expression in β cells, red one likely by EndoC-βH1 and no highlight likely by both populations.

15

EndoC-βH1 viability is disrupted when exposed to activated CD4+ T cells in co-culture

I found markers of apoptosis enriched in the upregulated genes in co-culture at all time points (Fig 14C), thus I decided to check the percentage of dead EC upon exposure to activated CD4+ T cells at the different time points. To do so, I used Hoechst (HO) and propidium iodide (PI) staining and quantified the number of PI positive cells (red) over the total number of cells (HO positive, blue) (Fig 11A). While treatment with a pro-apoptotic agent (POS) caused a significant increase in cell death, co-culture of EndoC-βH1 with activated CD4+ T cells induced β cell death upon 24 h (Fig 11B).

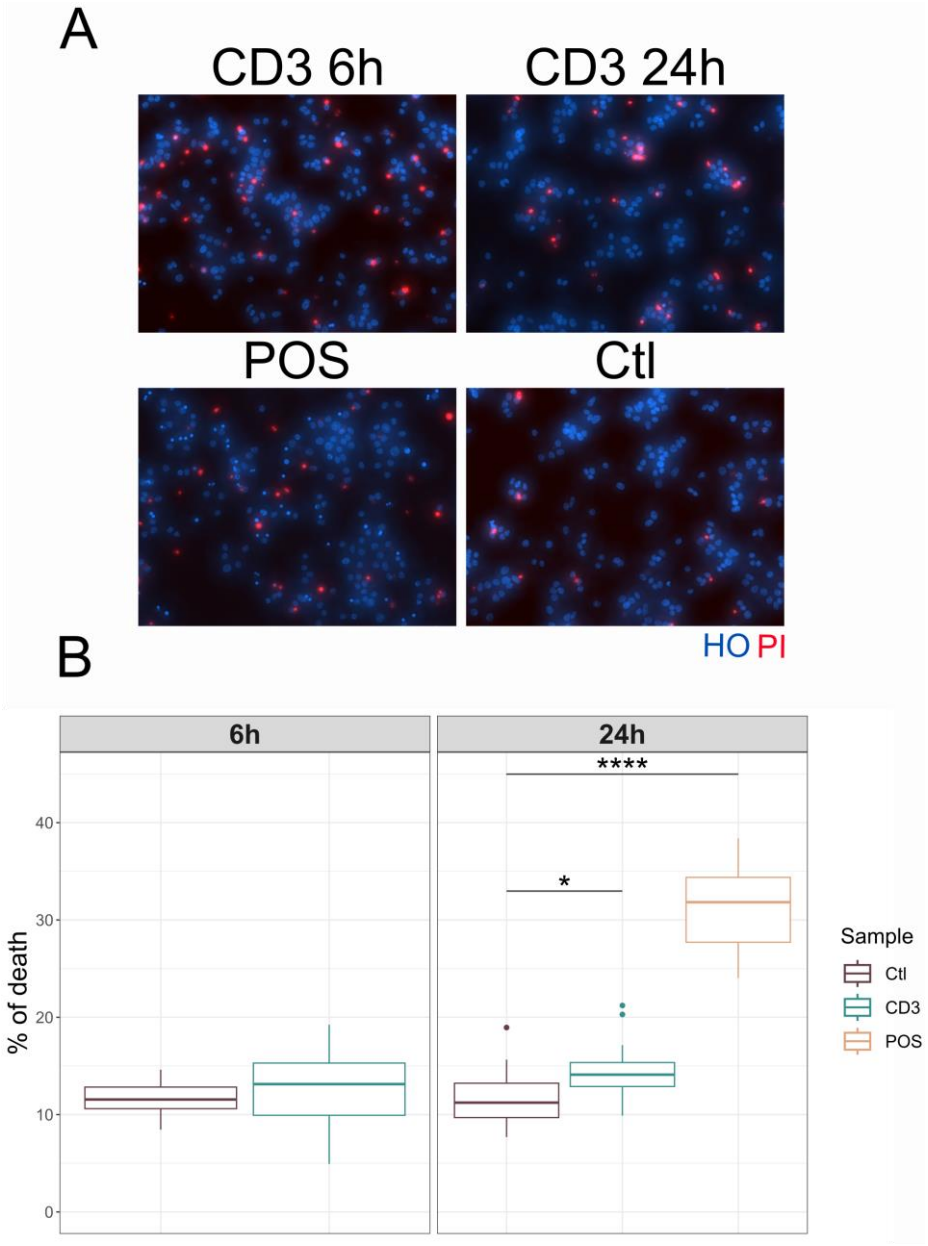


Figure 11. Co-culture exposure increases EndoC-βH1 cell death. A. Representative images of each condition at the different time points B. Percentage of dead cells of the different samples for each time point. Six different regions were quantified for each well, and each sample was run in triplicate. Ctl: untreated cells; CD3: co-cultured cells; POS: positive control. The lines/whiskers in the boxplot indicate the variability outside the upper and lower quartiles.

16

Transcriptomic Changes in β Cells after activated lymphocyte exposure

Next, I further characterize the impact of activated lymphocytes exposure on the β cells transcriptome. RNA-seq profiling of β cells unveiled a substantial number of gene expression changes, with 571 upregulated genes and 387 downregulated genes at 72 h (Fig 12A). Analysis of the dynamics of changes demonstrates that both gained and lost genes tend to maintain their respective categories across subsequent time points. However, the majority of upregulated genes are gained at 6 h and 24 h, while conversely, almost all downregulated genes are lost by 72 hours (Fig 12B). Curiously, the magnitude of changes is consistent over time, with the changes in gained genes exhibiting greater intensity compared to the intensity observed in the lost genes (Fig 12C).

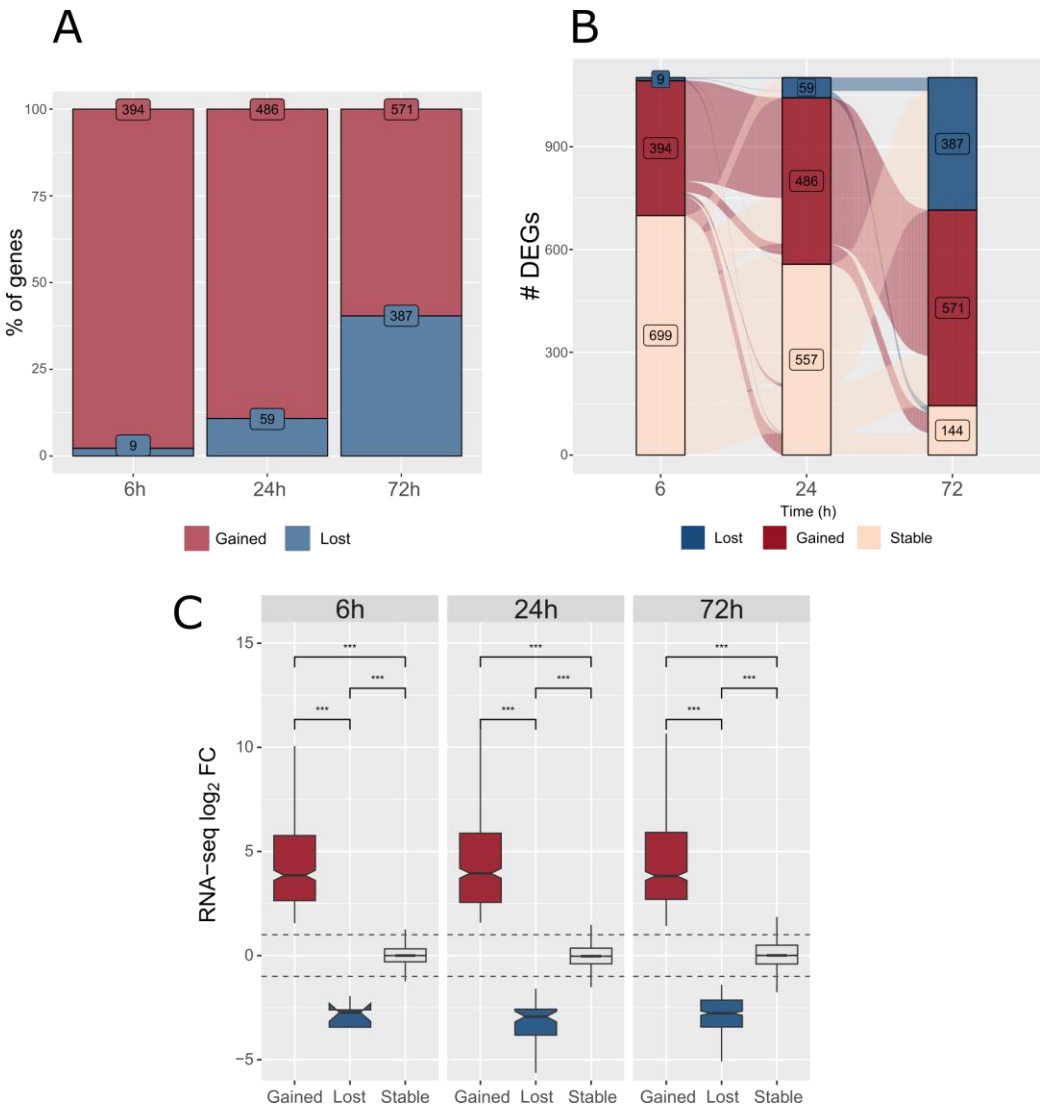


Figure 12. Exposure to activated lymphocytes leads to strong gene expression changes in β cells. A. Barplot with the number of changes in gene expression at the different time points B. Alluvial plot showing the dynamic changes of all DEGs over time. C. Boxplot with the magnitude of changes for each category at each time point. The lines/whiskers in the boxplot indicate the variability outside the upper and lower quartiles.

KEGG pathway enrichment analysis of gained and lost genes revealed that, already from an early time point, there is an induction of pathways related to the inflammatory response, such as cytokine and chemokine

signalling, as TNF or NF-κB signalling pathways, or antigen processing and presentation. Notably, a subset of upregulated genes linked to Type I Diabetes and other autoimmune diseases show enrichment from 24 hours, with their significance slightly increasing by 72 hours (Fig 13A).The subset of T1D-related genes includes *HLA* genes[96, 113], *FAS*[202], *LTA*[203], *IL12A*[204, 205] and *GAD1*[206]. As for the limited number of downregulated genes observed at 24 hours, the main enriched pathway is associated with both intracellular and extracellular signalling, which include hormones and neurotransmission receptors. Some of these genes are the nicotinic receptor subunits *CHRNA4* and *B4*, the GABA receptor *GABBR2* and glutamate receptor *GRIN1*. Conversely, genes downregulated at the later time point are mainly enriched in processes related to cell cycle and cellular division (Fig 13B).

I explored in more detail the differences between downregulated genes at 24 h and found that 19 genes were downregulated only at 24 h, including *CEBPA*, while 40 genes were downregulated at both 24 h and 72 h, including *SOCS2, G6PC2* and the previously mentioned *CHRNA4* and *B4*.

Gene expression levels of inflammatory markers related to IFNγ response significantly increase over time compared to the EC control, whereas HI-specific markers show a significant decrease, particularly at earlier time points (Fig 14A). Gene expression levels of EC treated with IFNγ and IL-1β for 48 h (CYT) from a previous publication of our lab[130] also show a significant increase of IFNγ response markers upon stimulus. However, in contrast to co-culture results, islet-specific markers do not exhibit the same decrease in expression, showing instead an increase in expression after 48 hours of stimulation (Fig 14B). Additionally, I also found an enrichment of genes related to apoptosis and senescence in the upregulated genes of co-culture at all time points (Fig 14C).

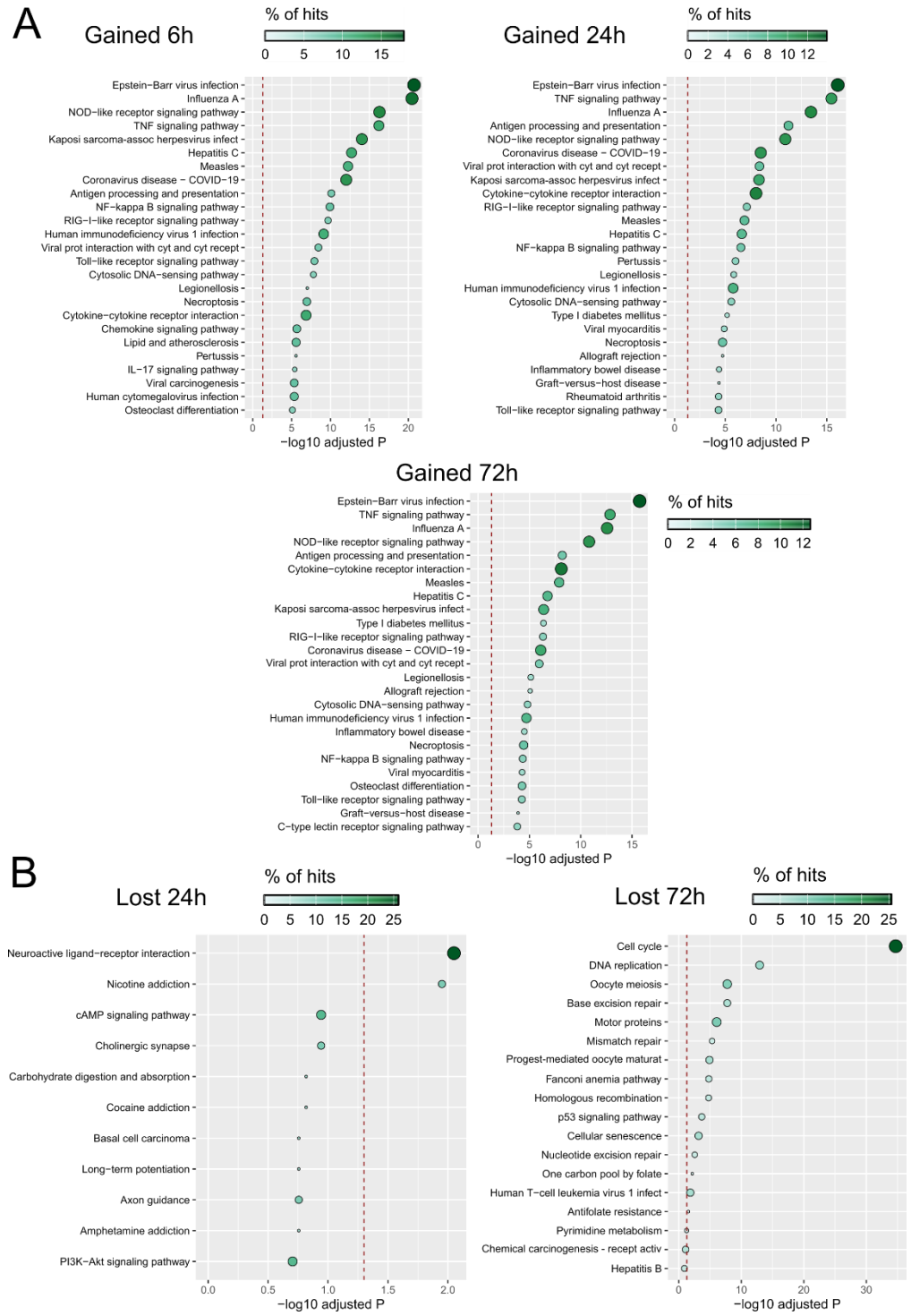


Figure 13. KEGG pathway enrichment analysis of gained and lost genes upon co-culture. A. Enriched terms in upregulated genes at the different time points B. Enriched terms in downregulated genes at 24 h and 72 h.

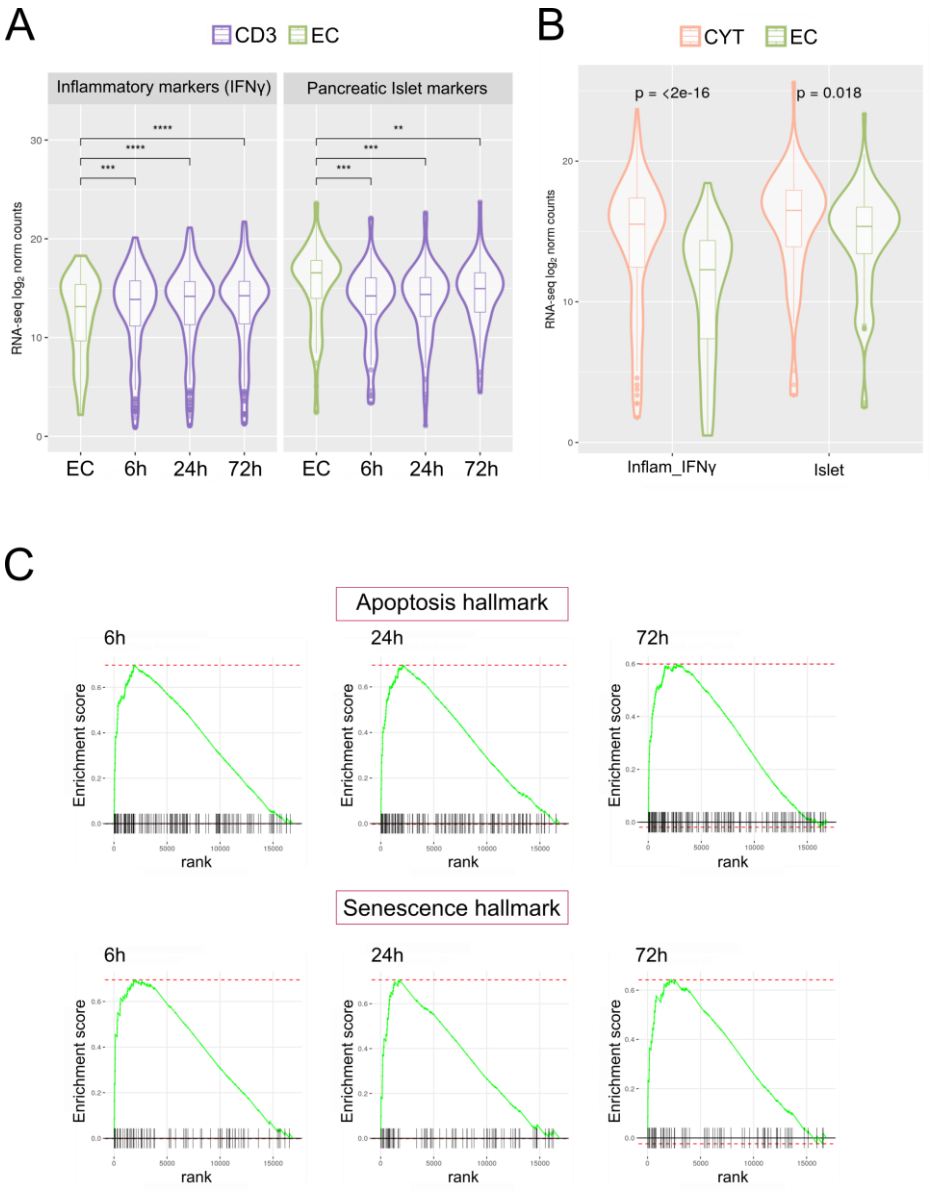


Figure 14. Gene expression changes reveal increased expression of inflammatory, apoptotic and senescence markers while a decrease in islet-specific markers. A. Gene expression of inflammatory and islet-specific markers in the different co-culture samples. B. Gene expression levels of same inflammatory and islet-specific markers (full list of islet-specific markers can be found in Table 1, Materials and Methods section) in EC treated with IFNγ + IL-1β. The lines/whiskers of each boxplot inside the violin plot indicate the variability outside the upper and lower quartiles. C. Gene Set Enrichment Analysis (GSEA) on all DEGs from co-culture samples for apoptosis and senescence markers.

17

Impact of activated lymphocytes on β cell regulatory landscape

Additionally, I investigated the impact of exposure to activated lymphocytes on the regulatory landscape of β cells. To accomplish this, I mapped all accessible chromatin sites and H3K27ac enriched regions in EndoC-βH1 upon the described co-culture setting. I used assay for transposase-accessible chromatin coupled with sequencing (ATAC-seq) and Cleavage Under Targets and Tagmentation (CUT&Tag), respectively, to map non-coding *cis*-regulatory elements responsive to the proinflammatory stimulus.

Consistent with the transcriptome findings, I observed substantial chromatin remodelling in response to activated lymphocytes, although, in general, the data suggests that most of the overall changes in the chromatin landscape organization required longer time of exposure compared to changes in transcription (Fig 15A). I identified, at different time points, over 11,000 differentially accessible regions (DAR). The majority of gained changes were detected at 24 hours and sustained at 72 hours, with minimal alterations at 6 h. Similarly, acetylation changes (DAcR) show no alterations at 2 h, but acetylation deposition required a longer exposure time than accessibility, with a progressive increase in DAcR over time. Similarly, almost no lost regions were detected before 24 h and the majority of all lost region were identified at 72 h (Fig 15A). As both layers of regulation provide insights into the activity of regulatory elements (RE), I integrated the signals and defined a set of regulatory elements based on the changes in activity observed upon the co-culture set up (Fig 15B):

- Lost Regulatory Elements (LoREs): these regions experience a decrease in activity upon co-culture. They are identified by the loss of either the ATAC signal, K27ac signal, or both, at least with a FC < -1.
- Stable Regulatory Elements (SREs), regions that do not change the activity status upon exposure.

- Induced Regulatory Elements (IREs): these regions exhibit increased activity when co-cultured with activated CD4+ T cells. They are subclassified into a) Neo-IREs, that gain both the ATAC and K27ac signal, and b) Primed-IREs, which are already accessible in stable state but increase the K27ac deposition upon activated lymphocytes exposure. Gained signal was considered when FC > 1.

Overall, I identified over 3,200 different IREs distributed along the different time points, with also a progressive increase over time, reaching the highest number at 72 hours. On the other hand, LoREs are rare, predominantly emerging at 72 hours with only 198 lost regulatory elements (Fig 15C). When analysing the dynamics of regulatory elements (Fig 15C), two key observations emerged: 1) half of the IREs at 72 h are also identified as IREs at both 24 h and 6 h, and 2) very few IREs activated at 6 h or 24 h transition to being classified as SREs at later time points (Fig 15C). Based on these observations, and with the aim of understanding whether specific pathways are activated earlier or later, providing a more nuanced understanding of the β cell response, we conducted a subclassification of the IREs based on their induction dynamics. To this end, IREs were categorized as **early, middle, or late responders**, based on their induction time point 6 h, 24 h, or 72 h, respectively. I found 650 early, 997 middle and 1613 late IREs (Fig 16A) and, overall, most IREs are neo while only 10-15% of them are primed (Fig 16B). Additionally, I found that the vast majority of IREs are located in distal regulatory elements, with solely 1-5% located in promoter regions, while almost all LoREs are found in distal regulatory elements (Fig 16C). The same strategy was not followed with LoREs since almost all LoREs are induced at 72 h and very few are specific of 6 h or 24 h (Fig 15C).

Moreover, changes in the activity of the different subgroups of REs correlate with gene expression changes of transcripts located in the same loci at the same time point, the effect being particularly evident for IREs (Fig 16D). Genes located near an IRE exhibit a significant increase in expression compared to those located near SREs at every time point analyzed. However, the correlation between LoREs and the expression changes of nearby genes become apparent at 72 hours, at which point most of the lost LoREs are detected and the nearby genes show significantly decreased expression (Fig 16D).

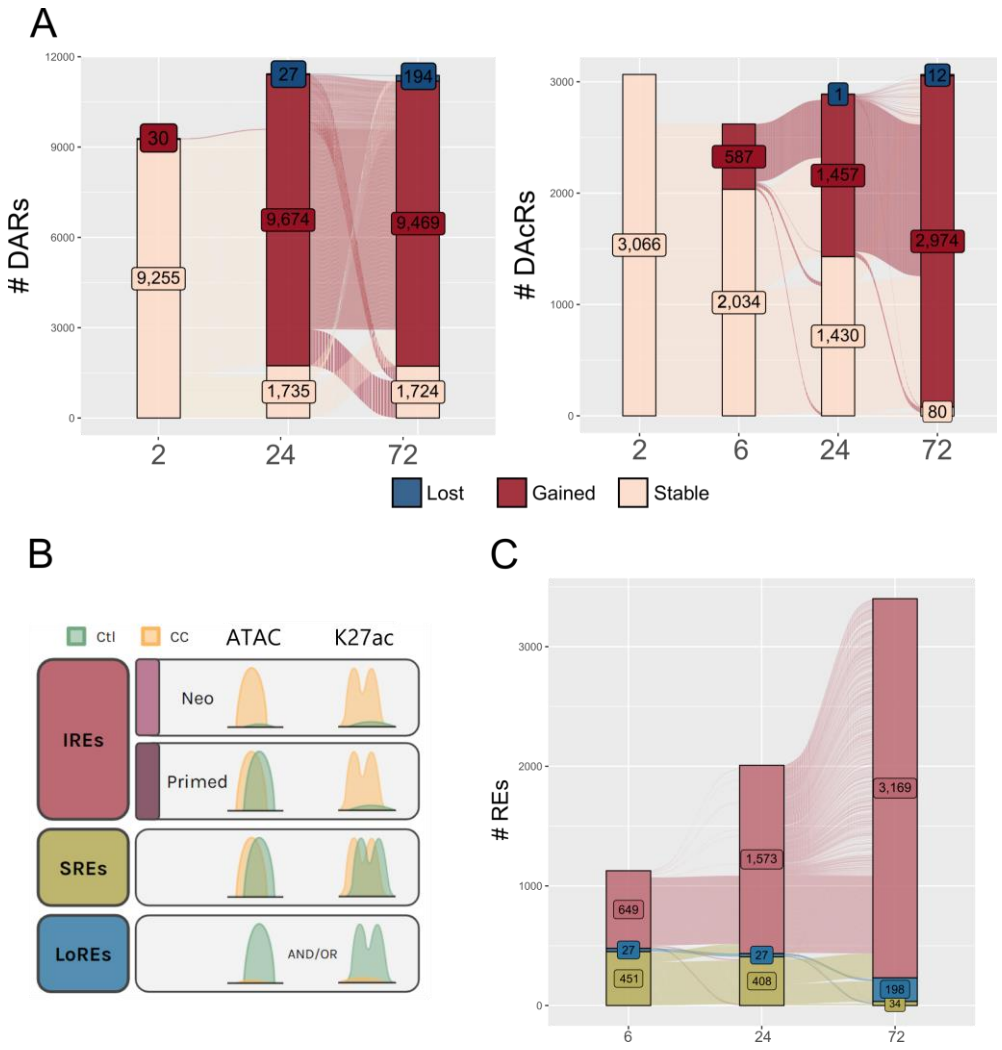


Figure 15. Classification of regulatory elements according to their changes in activity A. Dynamic changes of differentially accessible regions (DARs) and differentially acetylated regions (DAcRs) in time B. Scheme of classification of regulatory elements according to changes in accessibility and K27ac deposition. C. Alluvial plot showing classification dynamics across time.

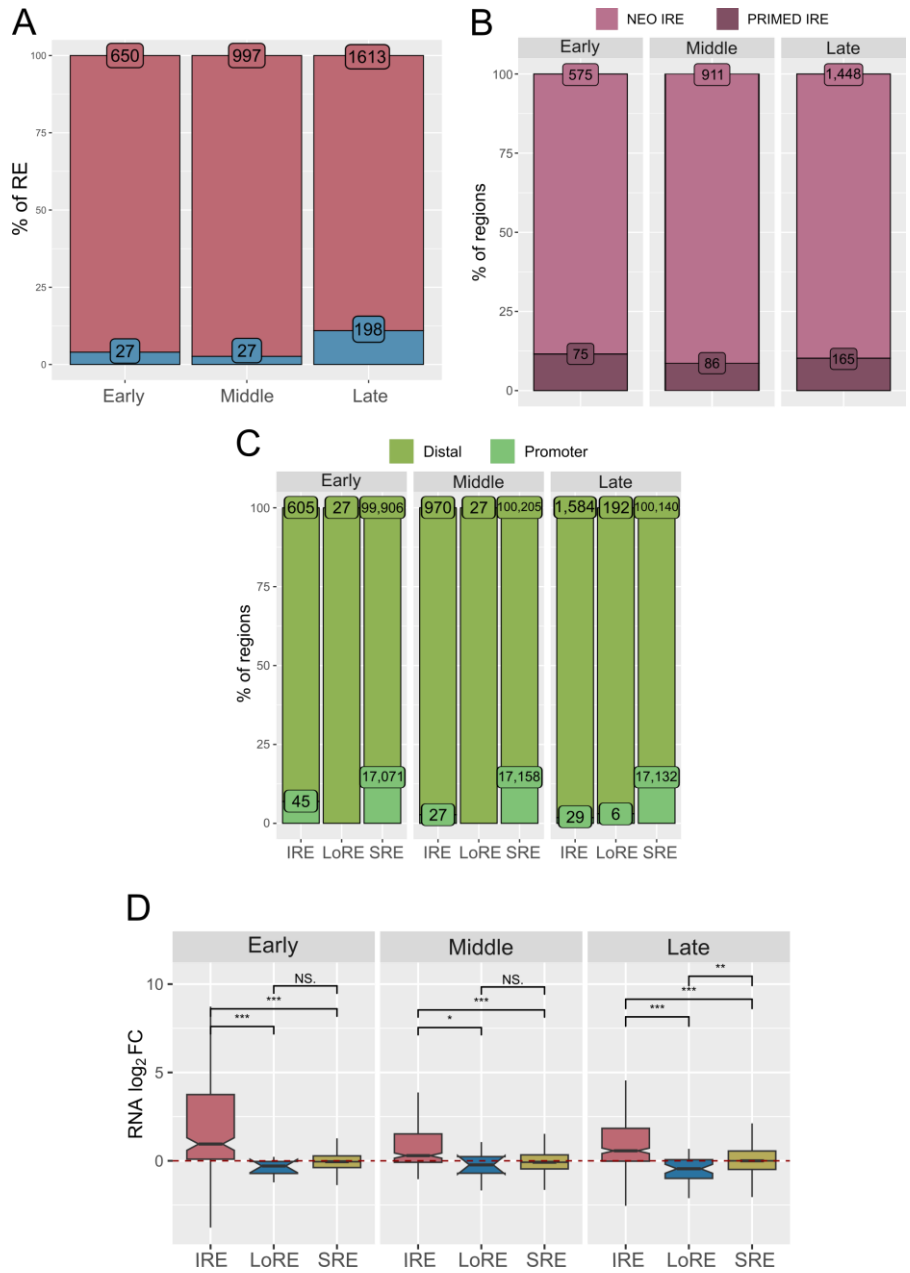


Figure 16. Sub-classification of regulatory elements according to time of induction is accompanied with changes in gene expression of nearby genes. A. Number of Induced Regulatory Elements (IREs) and Lost Regulatory Elements (LoREs) for each subclassification according to time of induction. B. Number of neo and primed IREs specifics of each induction time. C. Distribution of REs in distal RE or promoters. D. Gene expression changes broken down by exposure time points (6 h, 24 h, and 72 h, representing early, middle, and late stages respectively), for genes near IREs, SREs and LoREs. The

lines/whiskers in the boxplot indicate the variability outside the upper and lower quartiles.

Gene Ontology (GO) analysis of the DEGs genes around the different IREs highlighted a predominant response of stress and inflammation. As expected, early IREs are enriched in pathways related to the first inflammatory response and stress (Fig 17). Middle and late IREs are also enriched in early inflammatory pathways and more specific ones such as response to IFNs or TNF (Fig 17).

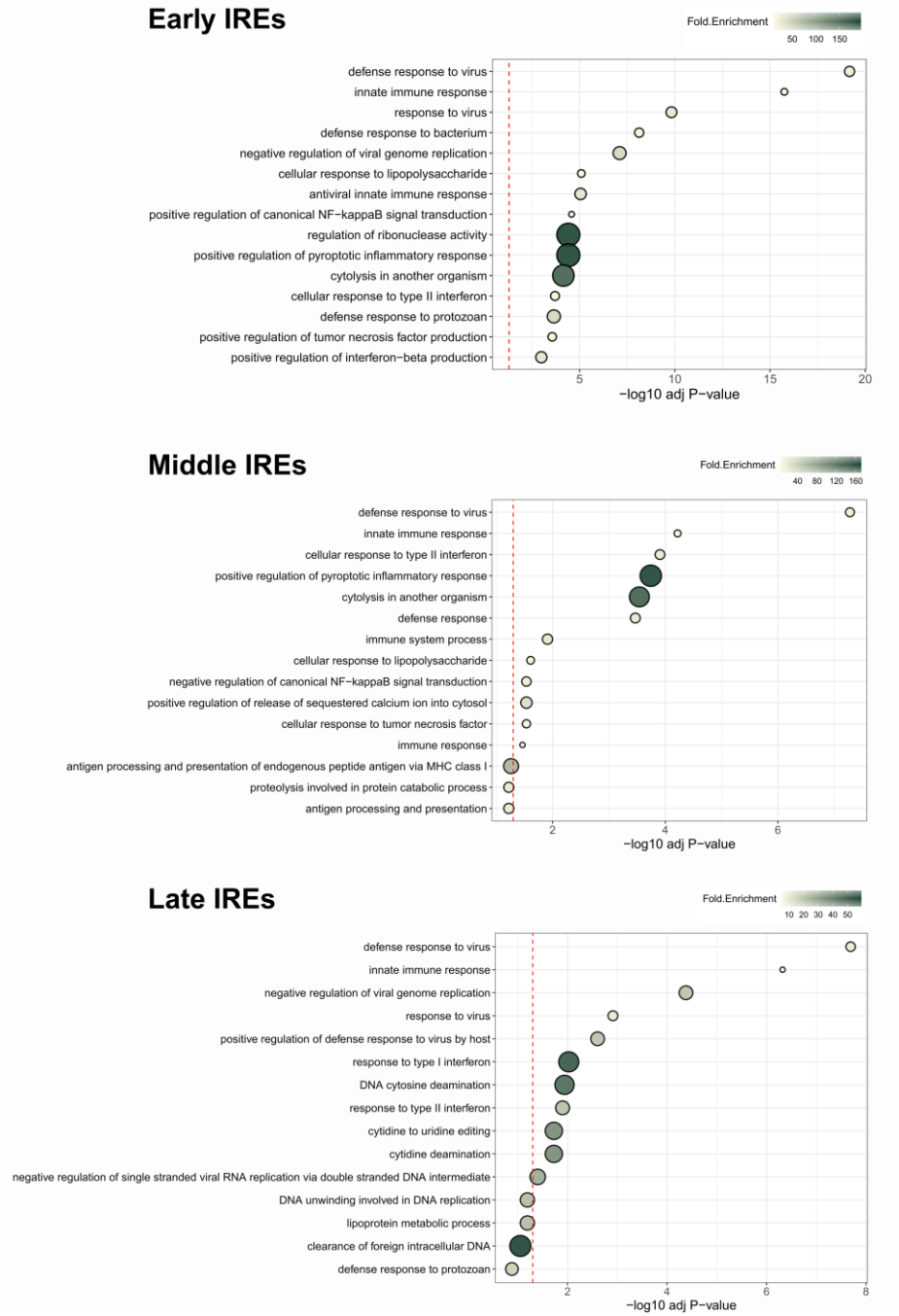


Figure 17. Classification of regulatory elements according to time of induction reveals inflammatory pathways activated upon exposure. GO analysis of early, middle and late IREs.

18

Co-culture model of EndoC-βH1 with activated CD4+ T cells provides novel chromatin regulatory maps

This new *in vitro* model based on the co-culture of β cells with activated CD4+ T cells to mimic early insulitis in T1D appears to replicate key factors of disease pathogenesis. However, it is imperative to understand the extent to which this model enhances our current understanding of early insulitis in T1D. To address this, I compared the findings from this model with existing datasets from other studies using diverse cytokine cocktails with EndoC-βH1.

Comparing the transcriptome results of the current model with two alternative cytokine exposure based experimental designs—one using IFNγ + IL1β[130] and the other employing IFNα[134]—, it became evident that the β cell response to IFNα is clearly distinct from the responses captured in the other two models (Fig 18A). As expected, the β cell response to activated CD4+ T lymphocytes aligns more closely with the model using IFNγ + IL1β (CYT), given that both models primarily involve IFNγ stimulation. Yet, samples from the co-culture model diverge, exhibiting separate clustering from those treated with IFNγ + IL1β (CYT), underscoring notable differences between the two approaches (Fig 18A). To explore this, I overlapped the overexpressed (Fig 18B) and downregulated genes (Fig 18C) from both models to assess the extent of similarities and differences in transcription.

In general terms, the transcriptomic response in the CYT model tends to be stronger than in the CC model. When comparing the upregulated genes, I found that while the majority of them are shared between both models, a subset of upregulated genes is specific of each group (Fig 18B). GO analysis on shared genes show enrichment in pathways related to inflammation and IFNγ signalling (data not shown). The CYT specific upregulated genes are enriched in pathways related to cell adhesion, regulation of cytokine signalling and production, and MHC-II production and assembly (data not shown). On the other hand, CC specific genes are enriched in viral response (data not shown). Comparison of downregulated genes revealed that most downregulated

genes in CC are shared with the CYT model whereas the majority of CYT-downregulated genes are specific to this group (Fig 18C). GO analysis showed that shared genes between both models are enriched in pathways related to synaptic transmission and cell division while CYT-specific genes include pathways of organ development, extracellular organization and cell cycle. In contrast, CC-specific genes are mainly enriched in cell cycle and division (not shown).

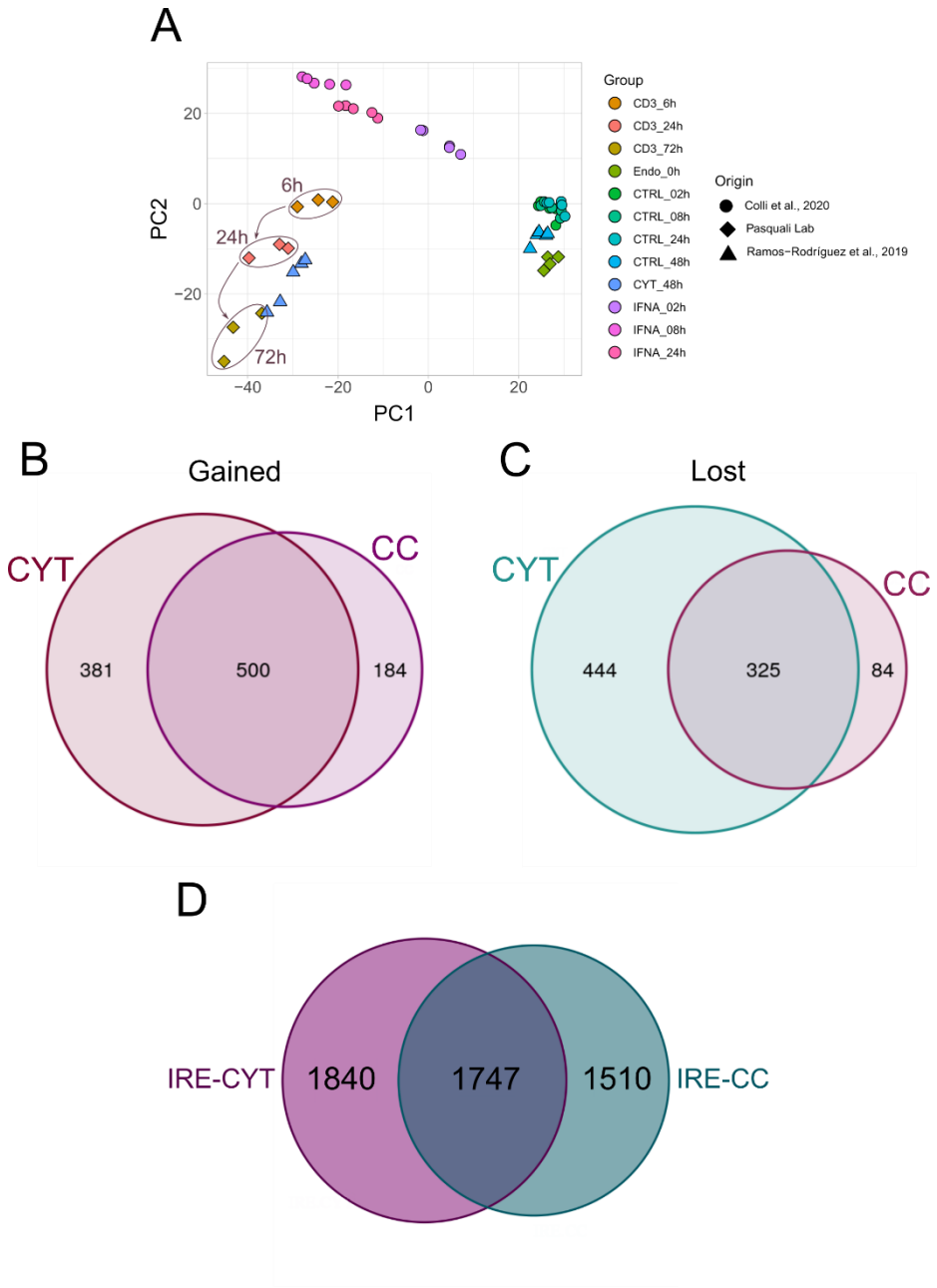


Figure 18. Comparison of RNA-seq data and IREs from the co-culture model and other cytokine models. A. PCA analysis of transcriptome data from available datasets with EndoC-βH1. B. Overlap of upregulated genes between co-culture model and IFNγ + IL1β cocktail. C. Overlap of downregulated genes between co-culture model and IFNγ + IL1β cocktail. D. Venn diagram overlapping IREs from CYT and co-culture. IFNA corresponds to

IFNα treatment, CYT corresponds to IFNγ + IL1β treatment and CD3/CC to co-culture.

Considering the transcriptome comparison, I decided to perform a similar analysis by comparing the chromatin changes induced by the different models. I compared the regulatory elements responsive to the IFNγ + IL1β cocktail (CYT) and co-culture (CC) approaches, as data of chromatin accessibility but not histone modification enrichments are available for the IFNα model. To do this, I overlapped the IREs obtained in the CC with those from CYT. Results showed only around a 50% overlap between the two models (Fig 18D), indicating that different regulatory pathways are activated by the different proinflammatory stimulus. These findings differ from the transcriptome comparison, where the majority of genes are common to both models. In this case, LoREs were not compared since there were no lost RE detected in the CYT model.

19

The response of β cells to the co-culture model recapitulates the genetic risk associated with T1D

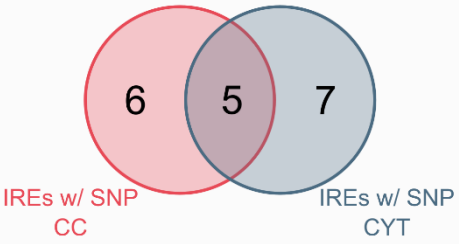
One of the main goals of current studies focused on the β cells role in T1D and its implication to the pathogenesis of the disease. To do so, I tested whether T1D genetic risk can be implicated with β cells inflammatory-responsive regulatory changes, by linking distal regulatory elements with their target genes, thereby highlighting affected pathways or mechanisms. In that regard, I overlapped all the REs defined at any time point in the co-culture model with available GWAS datasets of T1D SNPs. I also compared the results obtained with a similar analysis using the IREs defined in the IFNγ + IL1β model.

Overall, I observed that over 40% of T1D loci (n=182) encompass stress-induced β cell regulatory elements mapped in the CC model (Fig 19C). In search of potentially causative T1D variants that could mechanistically be implicated in the pathogenesis of the disease I computed a direct overlap of T1D risk variant and the recently mapped regulatory elements. I found a total of 11 IREs mapped in CC bearing 19 different T1D-SNPs (Fig 19A). As a comparison 12 IREs bearing 16 T1D SNPs overlapped CYT model IREs. Approximately half of the IREs bearing a T1D SNP are shared between the two models (Fig 19B). No overlap between SNPs and LoREs was found. A list of shared and CC-specific IREs can be found in Table 1.

A

	IRE	SRE
Co-culture		
SNPs	19	437
RE	**11**	235
IFNγ + IL1β		
SNPs	16	197
RE	**12**	114

B



C

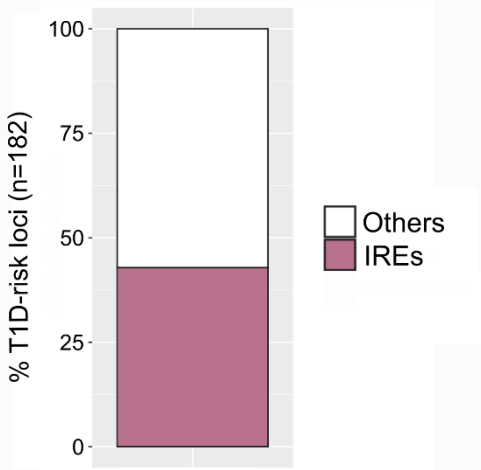


Figure 19. Overlap of T1D-risk SNPs with IREs from CC and CYT models. A. Number of IREs and SREs that bear a T1D-risk SNP for CC and CYT dataset. B. Overlap between the IREs from CC and CYT that bear at least one T1D-SNP.

Shared SNPs

SNP	Chrom location (hg38)	Nearest DEGs	Distance to TSS	Potential candidate gene
rs61674550	chr1:113540310	*DCLRE1B* (down) *OLFML3* (up)	360 kb 440 kb	*OLFML3*
rs78037977	chr1:172746562	-	-	-
rs12083773	chr1:197374633	*ASPM* (down) *CFH* (up)	220 kb 720 kb	-
rs4065274	chr17:39918763	*GSDMB* (up)	Near promoter	*GSDMB*
rs3024495	chr1:206769068	*RASSF5* (up) *IKBKE* (up) *C4BPB* (up)	260 kb 295 kb 315 kb	*RASSF5; IKBKE*

CC-specific SNPs

SNP	Chrom location (hg38)	Nearest DEGs	Distance to TSS	Potential candidate gene
rs11741255	chr5:132475490	*IRF1* (up)	15 kb	*IRF1*
rs10863989	chr1:212721403	*BATF3* (up)	20 kb	*BATF3*
rs2427749	chr6:30111111	*TRIM31* (up)	2 kb	*TRIM31*
rs2523988	chr6:30111352			
rs118026715	chr6:30111530			
rs17187931	chr6:30113179	*TRIM31* (up)	Prom	*TRIM31*
rs117994940	chr6:30113412			
rs17455103	chr6:30113469			
rs17187945	chr6:30113557			
rs17389539	chr6:30114226			
rs3822914	chr6:167020379	-	-	-
rs117306933	chr6:167020640			
rs117702482	chr6:167020646			
rs413809	chr16:11259447	*SOCS1* (up) *CIITA* (up)	3 kb 390 kb	*SOCS1*/*CIITA*

Table 1. List of SNPs overlapping IREs from either CC and CYT models (shared) or only from CC (CC-specific). List of SNPs with their genomic location, the nearest DEGs to the RE and the distance to it.

The T1D SNPs overlapping CC defined IREs were found at *BATF3, IRF1, OLFML3, RASSF5* or *SOCS1*, loci amongst others (Table 1). Interestingly I found rs4065274 T1D associated variant overlapping and IRE responsive to both CYT and CC located near the TSS of *GSDMB* (Fig 20). This gene is the only upregulated gene in the locus in both models. This observation makes *GSDMB* being very likely the target gene. Among the T1D-risk SNPs overlapping IREs specifically induced in CC, I found 8 different variants (rs2427749, rs2523988, rs118026715, rs17187931, rs117994940, rs17455103, rs17187945 and rs17389539) overlapping 2 different IREs located in the promoter or near the TSS of *TRIM31* (Fig 20). In this case, two genes are upregulated in the locus, *TRIM31* and *TRIM40* which makes both transcripts candidate T1D targets.

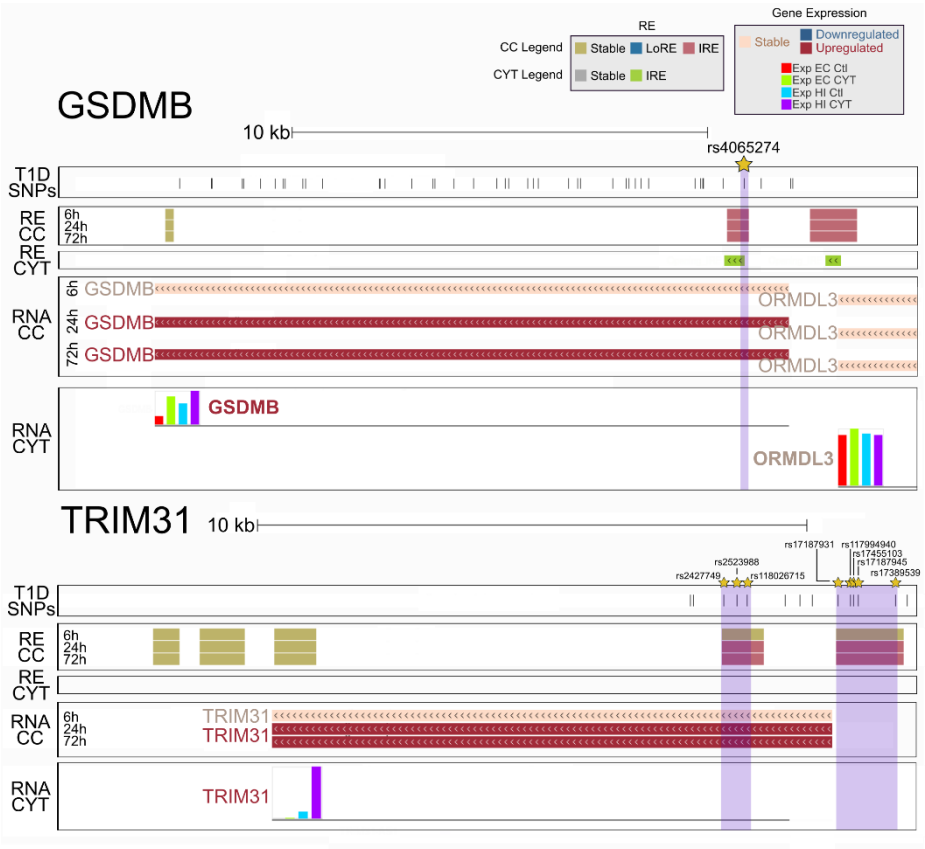


Figure 20. Schematic representation of chromatin regulatory maps and gene expression data of interesting loci containing IREs overlapping T1D associated variants. Schematic representation of the loci containing *GSDMB*

and *TRIM31* genes with the RE and gene expression data obtained in both models, co-culture (CC) and IFNγ + IL1β (CYT).

In addition to T1D-SNPs overlapping IREs, I also found a significant overlap between variants and SREs in both models. The locus containing the *DCLRE1B* gene has a region with several T1D-risk variants, one of which, rs11552449, overlaps with an SRE located in the promoter of *DCLRE1B* (Fig 21). This gene is downregulated and is the only DEG in the locus. Similarly, in the *ZFP36L1* locus, there is another region with different T1D-risk variants, and two of them overlap with an SRE located in the promoter of *ZFP36L1* (Fig 21). This gene is upregulated and the only DEG in the locus.

Furthermore, I also examined whether any genes encoding components present in the co-culture medium detected in the cytokine array could be modulated by T1D genetic risk. I focused on components likely secreted by EndoC-βH1 or both populations for easier interpretation. I found that both TNF-β (*LTA* gene) and LIF (*LIF* gene) have SNPs in the loci. In the case of *LTA*, there is just one SNP, and it does not overlap with any RE. In contrast, the *LIF* locus contains a region with several T1D-risk SNPs downstream of the gene. Although none of the SNPs directly overlap a RE, two of them fall really close to two different SREs, rs57043769 and rs1807711 (Fig 21). Interestingly, *LIF* is the only DEGs within a 5 Mb window from the gene.

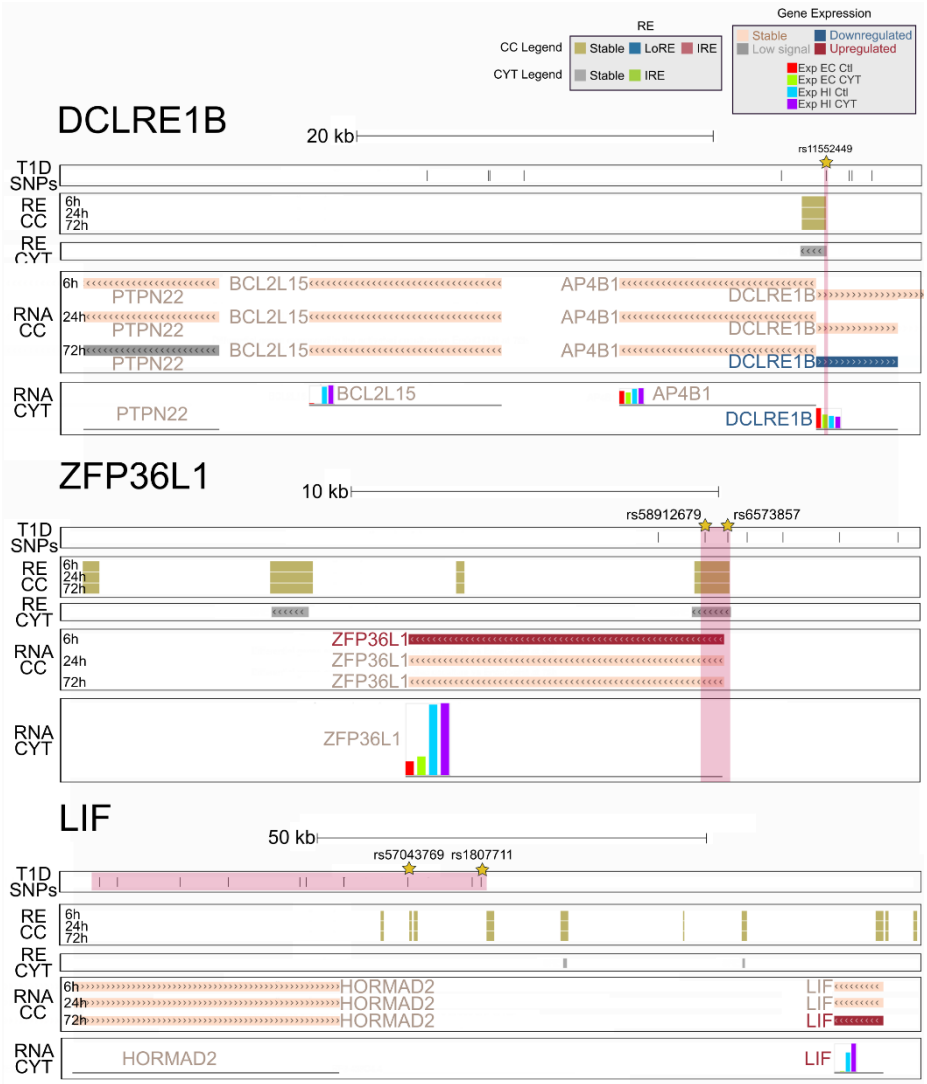


Figure 21. Schematic representation of chromatin regulatory maps and gene expression data of interesting loci. Schematic representation of the loci containing *DCLRE1B, ZFP36L1* and *LIF* genes with the RE and gene expression data obtained in both models, co-culture (CC) and IFNγ + IL1β (CYT).

20

CRISPR validation of cytokine-induced regulatory element reveals potential target gene

In parallel to co-culture experiments, I also started a validation experiment for one of the IREs identified using the IFNγ + IL1β model. As reported in a previous publication by the group[130], the *SOCS1* locus contains two distinct regions enriched in T1D-risk variants. Among the regulatory elements detected in this locus, rs193778 overlaps with an IRE near the *SOCS1* promoter (Fig 22A). To investigate the impact of this variant on β cell response, two approaches were employed. The first approach was to assess whether the variant affects enhancer activity using a luciferase reporter assays (Fig 4g from Ramos-Rodriguez *et al*[130]). The second approach involved profiling 3D contact maps of the locus to explore potential interactions between the regulatory element and gene promoters. The results demonstrated that the risk SNP increases the enhancer activity of the RE only upon cytokine stimulation. Additionally, the 3D contact maps revealed a highly complex organization of elements within the locus, identifying potential gene targets for this regulatory element (Fig 4h from Ramos-Rodriguez *et al*[130]).

Therefore, I decided to validate the gene target of the RE using CRISPR activation (CRISPRa) experiments. Six different sgRNAs were designed to target the CYT IRE containing rs193778 (Fig 22A, red star), and three of them (gRNA1, gRNA3 and gRNA4) were selected for validation in a preliminary experiment. For this experiment, in addition to the cloned gRNAs, I used four different control gRNAs (non-targeting ctls) that target non-human sequences. Cells were infected in triplicate, and gene expression was profiled by qPCR after antibiotic selection (Fig 22B).

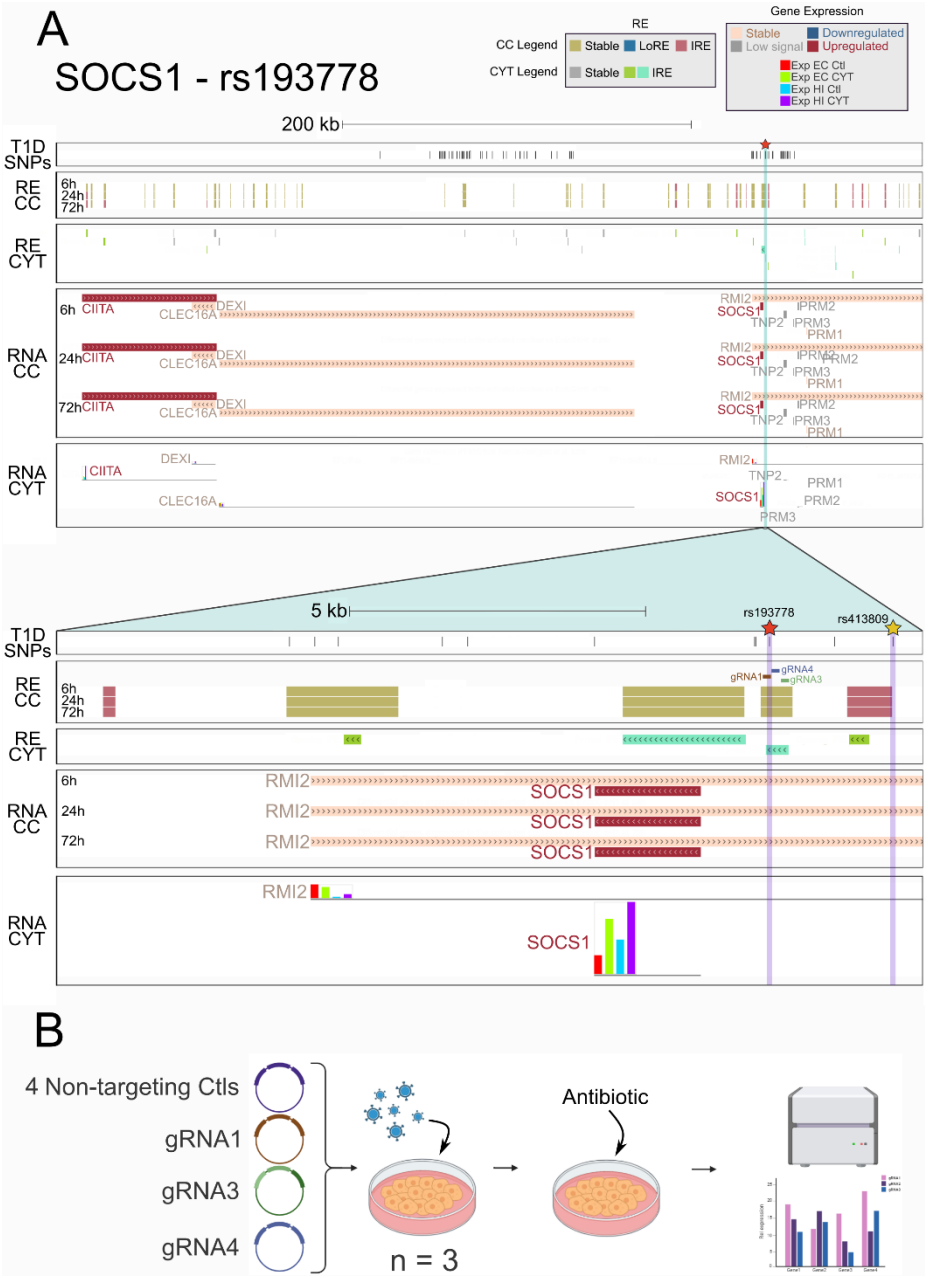


Figure 22. Schematic representation of the locus containing rs193778 (red star) and rs413809 (yellow star). A. Locus of *SOCS1* containing both T1D-risk SNPs overlapping two different RE, one IRE in CYT and another IRE in CC. B. Simplified scheme of CRISPR workflow. 3 gRNAs targeting the RE and 4 non-targeting controls for normalization were used. Each infection was done in triplicate.

After harvesting the cells, the expression levels of *SOCS1, DEXI, RMI2* and *CIITA* were quantified by qPCR. In both CYT and CC models, *SOCS1* and *CIITA* were the only genes within the locus that were upregulated in EC (Fig 22A). Additionally, transcriptome data from HI exposed to IFNγ + IL1β showed that *RMI2* and *DEXI* were also upregulated in HI. Therefore, I examined the expression levels of all four genes and found that SOCS1 was the only gene to show a significant increase in expression following the activation of the regulatory element by CRISPRa (Fig 23), when compared to the non-targeting controls. quantification by qPCR, which was performed in unstimulated EC, did not detected *CIITA* expression in the controls or in any of the gRNA samples.

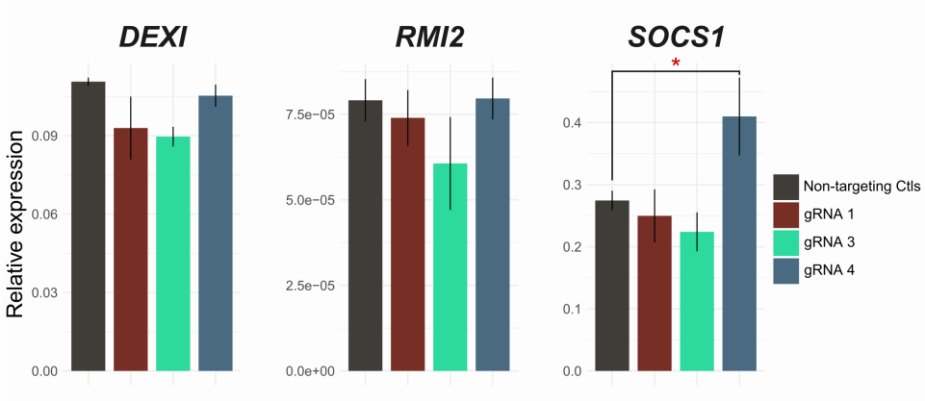


Figure 23. Gene expression quantification after CRISPRa experiments revealed an increase in *SOCS1* expression. Relative expression of three potential target genes normalized to three reference genes (*GAPDH*, *RPLP0, TBP*). The expression value corresponds to the mean of all the replicates (n=3) and the bars represent mean +/- SD.

Part V

Discussion

21

Optimization of a co-culture model to characterize β Cell Inflammatory Responses in T1D

Insulitis, a key process in the development of T1D, is marked by the infiltration of immune cells into the pancreatic islets. This infiltration leads to the secretion of different cytokines and chemokines by both immune and pancreatic islet cells, including β cells. These signalling molecules facilitate communication among the different cell populations and induce changes in nearby cells, altering their gene expression patterns as they adapt and respond to the inflammatory environment. Understanding the regulatory responses in β cells within this proinflammatory context is crucial for uncovering new pathological mechanisms of T1D and identifying potential targets for therapeutic interventions.

However, human primary samples at early stages of disease are of difficult access as most patients have not been diagnosed yet. Thus, in order to study the response of the β cell in these early stages of disease, different *in vitro* models have been developed over the years, mainly based on exposure to different cocktails of cytokines (Table 2, section 5.2.2, Introduction). All these models have been very useful in characterizing some of the key pathways activated in β cells within a proinflammatory environment, but the exact mechanisms leading to β cell dysfunction and death remain unclear. Here, I present a new *in vitro* model based on the co-culture of β cells with activated CD4+ T cells to explore new regulatory pathways that can be affected in early stages of T1D. Importantly, instead of using a fixed cytokine cocktail to mimic the exact mixture of molecules produced by T lymphocytes, this innovative experimental design allow reproducing, at least in part, the inflammatory environment experienced by β cells during early insulitis

In the first approach used, CC-M1, the activation of the T cells was limited, in particular when relying on the IGRP antigen, and induced mild changes in β cell's transcription. Transcriptome profiling showed only 9 upregulated genes and 1 downregulated gene in the IGRP sample. Thus, I decided to redesign the set up as it did not seem optimal for the

aims of this project. As a reference this experimental design resulted in the induction in the coculture media of IFNγ levels proximal to 200 pg/ml, a dose that was recently tested in another study. Consistent with our findings, this study explored the β cell regulatory landscape in insulitis and reached similar conclusions[195]. They observed few changes in chromatin accessibility and subsequently decided to work with higher doses of the cytokine.

In contrast, the non-specific activated “CD3” sample resulted in 36 upregulated and 1 downregulated genes. When comparing the DEGs from the two approaches, only two genes (*MT2A* and *HSPA1B*) were found as DEG in the antigen-specific activated “IGRP” sample but not in “CD3”. Nevertheless, *MT2A* and *HSPA1B,* although not identified as DEG, were close to the threshold and these results were based only on one replicate. In fact, when comparing directly “IGRP” against “CD3”, no IGRP-specific DEG gene could be found.

Re-design of the model and focus on co-culturing only β cells with activated CD4+ T cells facilitated the optimization, in terms of cell availability and activation efficiency. Preliminary results of transcriptome profiling in CC-M2 were promising although segregation of the signal from β cells and CD4+ T cells was difficult. Transwell systems or cell culture inserts, have been widely used to prevent cellular interactions or contaminations[207]. The application of a transwell system allowed to unequivocally distinguish the β cell from the CD4+ T cells for downstream analyses.

Interestingly, β cells showed a tendency of recovery when left in resting for 48 h after co-culture as markers of inflammation went back to pre-co-culturing levels and β cell-specific markers returned to unchallenged β cells levels. Although these results were obtained only from one replicate, the observation is supported by other studies that have shown a similar behaviour of β cell plasticity. Chen*et al*. [208] demonstrated that chronic ER stress causes a transcriptional reprogramming and impaired β cell function that is restored after recovery. Nonetheless, according to this study, β cell plasticity is gradually lost after repeated episodes of ER stress. The CC-M3 co-culture model could be utilized in follow-up studies to further investigate and characterize the gene regulatory landscape associated with recurrent inflammation.

22

Co-culture of β cells with activated CD4+ T cells as a novel *in vitro* approach to study early insulitis in T1D

22.1 Exposure to activated CD4+ T cells generates a bidirectional cytokine communication and reduced β cell viability

Exposure of β cells to activated CD4+ T cells induces a response not only at the gene regulation and transcriptional levels but also in the secretome observed after 72 h of co-culture.

The variability observed in terms of lymphocytes activation was expected as each batch of primary CD4+ T cells came from a different human donor. Besides activation variability, the use of primary cells offers the advantage of avoiding clone or lymphocyte subtype selection, thereby providing a wide spectrum of CD4+ T cell responses. Thus, it will be of great value to obtain the cytokine profile of all samples used to capture the potential variability of CD4+ T cell subtypes amongst different donors. For example, I found a response with IFNγ, IL2, TNFα, and TNFβ, commonly associated to Th1 CD4+ lymphocyte, a subtype related to cell-mediated immune response[2, 209]. While IL-4, IL-5, IL-9, IL10, and IL13 is typically associated with Th2 CD4+ T cells, related to B cell activation and humoral response, the presence of IL-17A, and IL-22 is associated with another subtype of CD4+ T cells, Th17. This subtype is involved in the immune response against bacteria and fungi, and it has been linked to the generation of autoimmune disease[8, 96, 209]. In the context of T1D, many of these cytokines have been associated with a proinflammatory effect, such as TNF-α, and IL-17, while others have been linked to an anti-inflammatory effect, like IL-10, IL-5, and IL-4. However, many of the cytokines detected, like IL-2, have shown a pleiotropic effect in immune cell activation[210].

The cytokine profiling of co-culture medium at the different time points confirms the bidirectional communication between CD4+ T cells and β cells. Some of the cytokines and chemokines are clearly secreted by the lymphocytes upon activation, but I also found molecular mediators that

are very likely secreted by β cells, and others by both populations, in response to the stimulus. Although cytokine function can be very broad and should be carefully assessed for each specific condition, upon co-culture, there is a clear induction of CXCL9 and CXCL10, markers of β cell response in T1D and key chemokines in the regulation of immune cell migration, activation and differentiation[91, 96, 211]. Other cytokines produced by β cells probably with effects on the immune system are IL-8, M-CSF, CXCL5, Fractalkine, TRAIL, and LIF. For example, fractalkine, also known as CX3CL1, is a special chemokine that can be found as a membrane bound, as an adhesion molecule, or soluble protein, acting as chemoattractant for monocytes, NK and T cells[212, 213]. Increased circulatory levels of fractalkine have been associated with T2D pathogenesis[214, 215]. On the other hand, while TRAIL is known as a death ligand involved in the regulation of innate immunity[216], LIF plays a crucial role in numerous biological processes from tissue homeostasis to immune response regulation[217]. Interestingly, LIF is found in one T1D-risk loci (HORMAD2 locus)[218, 219]. TRAIL deficient mice exhibit increased autoimmunity, indicating a function as a negative immune regulator causing anergy in diabetogenic T cells rather than an apoptotic effector[220–222]. However, these last two mediators show a broad spectrum of activity and their function in co-culture context should be carefully assessed as well as the affected cell types[223].

In addition to secretome changes, I also assessed cell viability upon co-culture exposure as I found the apoptosis pathway enriched in upregulated genes from early time points. Quantification of mortality confirmed that upon 24 h, β cells show increased cell death which is in line with reported results of cell viability after exposure to IFNγ and IL-1β[224, 225]. Limitations of these analysis include the lack of viability assessment at 72 h due to technical issues and the use of an additional complementary method to check cell viability, such as quantification of caspases enzymatic activity.

22.2 Co-culture of β cells with activated CD4+ T cells induces relevant transcriptome changes

In regards to transcriptome changes, I found over 680 upregulated genes and over 400 downregulated genes throughout the co-culture

time course. Upregulated genes are enriched in T1D relevant pathways that have been previously reported to be upregulated in β cells upon inflammation, such as inflammatory response, cytokine signalling andantigen processing and presentation. Moreover, gene ontology terms included a specific enrichment for T1D-related genes [113, 226].

Remarkably, while inflammatory markers show a significant increase in gene expression upon co-culture, islet-specific markers show a significant decrease in expression. Importantly, the downregulation of islet-specific markers is not observed in β cells exposed to IFNγ+IL-1β (data from Ramos-Rodriguez *et al*.[130]), indicating that this effect is driven by different combinations of cytokine signals. Moreover, senescence markers were also upregulated in co-culture, particularly at early and middle time points. A recent work demonstrated the presence of a subset of β cells that acquires a senescence phenotype, in both human and mice, and its elimination prevented T1D development in mice[227]. Another work, showed that triggering an early senescence phenotype in β cells can promote an anti-inflammatory response and prevent T1D development[228]. The molecular mechanisms driving the senescence phenotype in β cells and its implications in disease pathogenesis are not yet clear. However the co-culture model could serve as a tool to disentangle the molecular mechanisms driving these specific pathways and their implication in the disease pathogenesis[229].

Most downregulated genes were detected after 72 h of co-culture. While general characterization of this subset of genes provided minimal information on mechanisms of response, detailed inspection highlighted some interesting genes related to β cell differentiation and insulin secretion. I here highlight 3 genes consistently downregulated in β cells upon 24-72 h of co-culture whose loss function could be implicated in the disease pathogenesis.

The nicotinic receptor subunits *CHRNA4* and *B4* have been associated to T2D genetic risk as the receptor has a function in insulin secretion and its function is impaired in T2D patients. Its promoter, which bears some T1D-risk variants, is regulated by *MAFA*, one of the main β cell-specific TFs[230]. Additionally, *MAFB* regulation of this receptor has been associated with pancreatic islet formation and development in mice[231].

GABBR2 encodes a GABA receptor, and its expression is downregulated at 24 h. It has been reported that activation of this receptor in β cells reduces expression of β cell-specific markers and insulin secretion[232]. Whether the downregulation of *GABBR2* is a consequence of the increased GABA production due to upregulation of *GAD1* in β cells is not known.

G6PC2 is an important enzyme implicated in glucose metabolism. The non-coding T1Drisk variantsupstream of*G6PC2* have been linked to a more rapid loss of insulin secretion in T1D[233]. In this study, T1D-associated variants rs59681820 and rs16855123 were found to overlap a SREs upstream of *G6PC2*.

22.3 Co-culture with activated CD4+ T cells reveals novel chromatin regulatory pathways in β cells

In terms of gene regulatory networks, I found that co-culture with activated CD4+ T cells induced an extensive chromatin remodelling in β cells. I found >11,000 opening chromatin sites and >3,000 newly acetylated regions upon co-culture. Curiously, very few regions lose their accessibility or acetylation upon exposure compared to gained regions, and almost all are detected at later time points. This observation contrast with gene expression changes, as approximately 40% of the DEGs are downregulated genes. Thus, it is likely that this decrease in gene expression is primarily mediated by mechanisms other than chromatin accessibility, such as loss of TF binding/replication machinery or changes in 3D chromatin structure.

In order to better characterize the cis-regulatory networks that drive the response of the β cell to activated CD4+ T cells, both chromatin features (chromatin accessibility and enrichment of H3K27ac) were used to define responsive non-coding regulatory elements. I found >3,200 IREs and > 200 LoREs. IREs correlated positively with changes in expression of nearby genes at the same time of induction at all time points while late LoREs were associated with a decreased expression of nearby genes.

In general terms, the magnitude of changes for gene expression, chromatin accessibility and histone acetylation are comparable to other published studies using cytokine cocktails[130, 134, 195]. Specific

comparison with other works is limited by the use in many studies of human islets instead of a human β cell line resulting in higher variability. Future experiments of co-culture with human islets will allow correct comparison with the different cytokine cocktail used in other studies.

When comparing the results obtained in this study with exposure of EndoC-βH1 to IFNγ + IL-1β[130], I observed relevant differences in inflammatory-responsive non-coding regulatory functions. Only half of the induced regulatory elements (IREs) are shared between both models, uncovering novel regulatory pathways in β cells. As genetic susceptibility mainly acts in the non-coding genome, obtaining new regulatory maps of β cells in response to inflammation is of upmost importance to dissect potential mechanisms of disease development.

Thus, the results in the present work show that the co-culture model is a valid approach to generate datasets of transcriptome and chromatin regulatory maps of β cells in response to a physiological proinflammatory stimulus. It captures well known features of T1D pathogenesis providing novel regulatory maps that could help disentangle the genetic susceptibility of the disease.

Additionally, the co-culture model described can be adapted to other sources of β cells, such as primary human islets or SC-derived islets. Future efforts should focus on profiling the response of primary human islets to co-culture both in bulk and at the single cell level to dissect response of the different islet populations to inflammation. Current efforts are focusing on profiling the transcriptomic and gene regulatory networks at single cell level to uncover the specific molecular pathways active in each cell population in the context of T1D and identifying the cell populations affected by genetic predisposition [74, 194, 195, 226, 234].

23

Linking the β cell response to inflammatory clues with T1D genetic susceptibility

GWAS data provides valuable information to understand the genetic risk of complex traits, such as T1D. However, most of the variants associated to T1D are located in the non-coding genome, potentially affecting gene regulatory networks. Thus, intersecting GWAS datasets with detailed regulatory maps of cell types and cell states relevant in disease pathogenesis can be a useful strategy to shed light onto potential disease mechanisms. However, this approach relies on two factors: 1) the power of the GWAS which determines the number of leading SNPs, and 2) the quality of the regulatory maps used to elucidate new disease mechanisms.

Nearly 40% of T1D-associated loci contain stress-induced RE in β cells from the co-culture model, where genetic susceptibility can potentially disrupt the β cell response to inflammation. From the direct overlap between IREs and T1D-variants, only half of the SNP-bearing IREs are shared between the co-culture and IFNγ + IL-1β model. This supports the conclusion that the novel regulatory maps obtained with the co-culture model can help dissecting new disease mechanisms that were not previously captured.

Examples of loci in which T1D risk variants overlap inflammatory-responsive β cell regulatory elements included the GSDMB and the TRIM31 loci.

GSDMB belongs to the gasdermin-domain containing protein family, and it is involved in pyroptosis, an inflammatory form of cell death. This mechanisms have been suggested to be implicated in T1D pathogenesis[235]. Variants found near the promoter of *GSDMB* might increase the expression of the gene, increasing β cell death and potentially exacerbating the immune response. *TRIM31* is an E3 ubiquitin ligase that has been described to regulate the innate immune response and deficiency has been associated with impaired glucose metabolism and gut microbiota disruption in mice[236, 237]. Variants found in the promoter of the gene or its vicinity can potentially alter the

innate immune response in the pancreatic islets of susceptible individuals.

Although the main analysis focused on T1D-risk variants overlapping IREs or LoREs, I uncovered a large number of variants overlapping SRE (437 T1D associated SNPs overlapping 235 SREs). While gene expression changes can derive from regulatory variants disrupting TF binding sites in stress-induced REs, other mechanisms could explain gene expression changes without affecting the RE status directly, such as changes in the binding of co-activators or co-repressors or changes in the chromatin organization. The *ZFP36L1* promoter contains an SRE bearing two T1D-risk variants. *ZFP36L1* is an early-responder gene to stress that has been implicated in a number of different diseases and linked to protein synthesis attenuation[238–241]. Upregulation of this gene is only captured at 6 h, and not detected in the cytokine model. T1D-risk variants affecting the expression of this gene might modulate the earliest response of the β cell to inflammation, increasing the risk of exacerbating the immune assault. Similarly, *DCLRE1B* is a gene that blocks gene transcription and replication to facilitate DNA repair[242]. Additionally, it has an essential role in telomere maintenance[243]. Altered regulation of this gene upon stimulus might have detrimental effects on β cells. Nonetheless, the mechanisms by which each of these variants disrupts the β cell response to a proinflammatory stimulus should be studied in detail and addressed with other experimental approaches.

The *LIF* locus contains a nearby region enriched in T1D-risk variants. Although none overlaps with any of the RE directly, some of them are located in close proximity to two SRE that are not classified as IREs due to a lack of statistical power. Both SRE show FC > 1 in chromatin openness and acetylation that are not statistically significant. However, LIF is one of the components found in the co-culture medium, and its expression is upregulated in β cells. The *LIF* pathway has been linked to β cell replication in a subset of β cells that express the LIF receptor and have increased proliferative capacity[244]. Upregulation of LIF could stimulate β cell proliferation during inflammation and, its altered expression could contribute to T1D progression.

24

SOCS1 as the likely target gene of T1D-risk variants

Elucidating the molecular mechanisms by which candidate risk variants influence enhancer function requires significant effort. One of the key challenges is identifying the potential gene targets of candidate enhancers carrying T1D-risk variants. To this end various strategies can be employed including mapping 3D chromatin contacts and profiling physical contacts between different RE in the genome. However, it is crucial to functionally validate the effect of the regulatory element on the target gene under disease-relevant conditions to uncover the gene regulatory impact of the variant.

Previous work from the group demonstrated that rs193778 overlapped a β cell cytokine-responsive RE. Allele-specific luciferase reporter assay showed significant differences in enhancer activation for constructs bearing or not the T1D-associated variant. Thus, I performed CRISPRa experiments on the enhancer to uncover and functionally validate the target gene. Activation of the RE in basal conditions led to increased expression of the*SOCS1* transcript while other genes in the locus were not perturbed.

While this experiment clearly implicates the *SOCS1* inflammatory-dependant gene expression regulation to the susceptibility of developing T1D, I acknowledge 2 potential experimental limitations.

First, the expression of β cells was evaluated under basal conditions, while a more complete characterization of enhancer function would require assessing expression following a proinflammatory exposure. Nevertheless, it is unlikely that a proinflammatory stimulus would alter the regulatory landscape of the locus by switching the enhancer gene target. Secondly, the proximity between the CRISPRa-target enhancer and the TSS of *SOCS1* experiments could affect the experiment's outcome. However, it has been reported that optimal distance of the gRNA tothe promoter of a gene to activate its expression is around 100 bp from the TSS[245]. The regulatory element targeted in this experiment is located 1.6 kb from the TSS of *SOCS1*, so it should not interfere with the promoter of the gene.

Interestingly, an independent, recently published study, found *SOCS1* as a T1D risk gene regulating cytokine-induced β cell death by CRISPR-KO screening[195]. *SOCS1* is a negative regulator of cytokine signalling and protects β cells from death in response to an inflammatory stimulus[195]. The regulation of the locus is complex including several regulatory elements potentially regulating the expression of the gene, for example, by co-culturing beta cells and activated CD4+ T cells, I found a previously unappreciated IRE, proximal to the enhancer tested in this experiment that overlaps an additional T1D-risk SNP, rs413809. A similar experimental approach as the one described above could be employed to functionally test whether rs413809 has an impact on the regulation of the *SOCS1* gene in inflammatory conditions thus further implicating the protein encoded by this gene in the pathogenesis of T1D.

Part VI

Conclusions

25

Conclusions

1. An optimized co-culture model featuring **β cells** and activated **CD4+ T cells**, using a transwell system, can be employed to investigate T-cell-mediated inflammatory effects on β cells, which are crucial for understanding T1D pathogenesis.

2. Co-culturing CD4+ T cells and β cells leads to bidirectional communication, characterized by a diverse array of cytokines likely secreted by both cell populations.

3. Exposure to activated CD4+ T cells causes **extensive** changes in β cells gene expression. Transcriptomic changes are characterized by induction of **inflammatory** pathways and **T1D-related** genes over time and decreased expression of pancreatic **islet-specific** markers.

4. Exposure of β cells to activated lymphocytes induces substantial **chromatin remodeling**. This epigenetic rearrangement results in activation of over 3,200 and loss of ~200 regulatory elements across the β cell genome.

5. Co-culturing CD4+ T cells and β cells leads results in **unique regulatory pathways** compared to other models using specific combinations of cytokines cocktail including IFNα, IFNγ, and/or IL-1β.

6. Inflammatory-responsive β cell regulatory elements map to ~40% of T1D-associated genetic loci with specific **T1D-risk SNPs** overlapping these **IREs**. These observations implicate β cell regulatory functions to the susceptibility of developing T1D.

7. Regulatory variants overlapping inflammatory-responsive β cell regulatory elements implicate gene expression regulation of *SOCS1* to the risk of developing T1D.

Part VII

References

26

References

1. Sattler S (2017) Chapter 1: The role of the immune system beyond the fight against infection. In: Immunol. Cardiovasc. Homeost. Pathol. Adv. Exp. Med. Biol. pp 3–14
2. Parkin J, Cohen B (2001) Overview of the Immune System. Lancet 357:1777–89
3. Janeway CA Jr, Travers P, Walport M et al. (2001) Immunobiology: The Immune System in Health and Disease., 5th editio. New York: Garland Science
4. Vivier E, Tomasello E, Baratin M, Walzer T, Ugolini S (2008) Functions of natural killer cells. Nat Immunol 9:503–510
5. Lendeckel U, Venz S, Wolke C (2022) Macrophages: shapes and functions. Chemtexts 8:12
6. DELVES PJ, ROITT IM (2000) The immune system. First of Two Parts. N Engl J Med 343:37–49
7. DELVES PJ, ROITT IM (2000) The immune system. Second of Two Parts. N Engl J Med 343:108–117
8. Luckheeram RV, Zhou R, Verma AD, Xia B (2012) $CD4^+$T cells: differentiation and functions. Clin Dev Immunol 2012:925135
9. Shah K, Al-Haidari A, Sun J, Kazi JU (2021) T cell receptor (TCR) signaling in health and disease. Signal Transduct Target Ther. https://doi.org/10.1038/s41392-021-00823-w
10. Pisania A, Weir GC, O'Neil JJ, Omer A, Tchipashvili V, Lei J, Colton CK, Bonner-Weir S (2010) Quantitative analysis of cell composition and purity of human pancreatic islet preparations. Lab Invest 90:1661–1675
11. Dybala MP, Hara M (2019) Heterogeneity of the Human Pancreatic Islet. Diabetes 68:1230–1239
12. Ramachandran K, Huang H-H, Stehno-Bittel L (2015) A Simple Method to Replace Islet Equivalents for Volume Quantification of Human Islets. Cell Transplant 24:1183–1194

13. Brissova M, Powers AC (2008) Architecture of Pancreatic Islets. In: Seino S, Bell GI (eds) Pancreat. Beta Cell Heal. Dis. Springer Japan, Tokyo, pp 3–11

14. Bosco D, Armanet M, Morel P, Niclauss N, Sgroi A, Muller YD, Giovannoni L, Parnaud G, Berney T (2010) Unique Arrangement of α- and β-Cells in Human Islets of Langerhans. Diabetes 59:1202–1210

15. Cabrera O, Berman DM, Kenyon NS, Ricordi C, Berggren P-O, Caicedo A (2006) The unique cytoarchitecture of human pancreatic islets has implications for islet cell function. Proc Natl Acad Sci U S A 103:2334–2339

16. Röder P V, Wu B, Liu Y, Han W (2016) Pancreatic regulation of glucose homeostasis. Exp Mol Med 48:e219–e219

17. Pradhan G, Samson SL, Sun Y (2013) Ghrelin: much more than a hunger hormone. Curr Opin Clin Nutr Metab Care 16:619–624

18. Chan JCN, Lim L-L, Wareham NJ, et al (2021) The Lancet Commission on diabetes: using data to transform diabetes care and patient lives. Lancet (London, England) 396:2019–2082

19. Zajec A, Trebušak Podkrajšek K, Tesovnik T, Šket R, Čugalj Kern B, Jenko Bizjan B, Šmigoc Schweiger D, Battelino T, Kovač J (2022) Pathogenesis of Type 1 Diabetes: Established Facts and New Insights. Genes (Basel). https://doi.org/10.3390/genes13040706

20. Buzzetti R, Zampetti S, Maddaloni E (2017) Adult-onset autoimmune diabetes: current knowledge and implications for management. Nat Rev Endocrinol 13:674–686

21. Leslie RD, Evans-Molina C, Freund-Brown J, et al (2021) Adult-Onset Type 1 Diabetes: Current Understanding and Challenges. Diabetes Care 44:2449–2456

22. Gregory GA, Robinson TIG, Linklater SE, et al (2022) Global incidence, prevalence, and mortality of type 1 diabetes in 2021 with projection to 2040: a modelling study. lancet Diabetes Endocrinol 10:741–760

23. Mobasseri M, Shirmohammadi M, Amiri T, Vahed N, Hosseini Fard H, Ghojazadeh M (2020) Prevalence and incidence of type 1 diabetes in the world: a systematic review and meta-analysis. Heal Promot Perspect 10:98–115

24. Gomez-Lopera N, Pineda-Trujillo N, Diaz-Valencia PA (2019) Correlating the global increase in type 1 diabetes incidence

across age groups with national economic prosperity: A systematic review. World J Diabetes 10:560–580

25. Ward ZJ, Yeh JM, Reddy CL, et al (2022) Estimating the total incidence of type 1 diabetes in children and adolescents aged 0-19 years from 1990 to 2050: a global simulation-based analysis. lancet Diabetes Endocrinol 10:848–858

26. International Diabetes Federation IDF Diabetes Atlas, 10th edition. In: IDF Diabetes Atlas 2021. https://diabetesatlas.org/data/en/. Accessed 3 May 2024

27. Harding JL, Wander PL, Zhang X, et al (2022) The Incidence of Adult-Onset Type 1 Diabetes: A Systematic Review From 32 Countries and Regions. Diabetes Care 45:994–1006

28. Rodríguez Escobedo R, Delgado Álvarez E, Menéndez Torre EL (2023) Incidence of type 1 diabetes mellitus in Asturias (Spain) between 2011 and 2020. Endocrinol Diabetes y Nutr (English ed) 70:189–195

29. Karvonen M, Viik-Kajander M, Moltchanova E, Libman I, LaPorte R, Tuomilehto J (2000) Incidence of childhood type 1 diabetes worldwide. Diabetes Mondiale (DiaMond) Project Group. Diabetes Care 23:1516–1526

30. Insel RA, Dunne JL, Atkinson MA, et al (2015) Staging presymptomatic type 1 diabetes: a scientific statement of JDRF, the Endocrine Society, and the American Diabetes Association. Diabetes Care 38:1964–1974

31. Lampasona V, Liberati D (2016) Islet Autoantibodies. Curr Diab Rep 16:53

32. Powers AC (2021) Type 1 diabetes mellitus: much progress, many opportunities. J Clin Invest. https://doi.org/10.1172/JCI142242

33. Jacobsen LM, Bocchino L, Evans-Molina C, et al (2020) The risk of progression to type 1 diabetes is highly variable in individuals with multiple autoantibodies following screening. Diabetologia 63:588–596

34. Eizirik DL, Colli ML, Ortis F (2009) The role of inflammation in insulitis and beta-cell loss in type 1 diabetes. Nat Rev Endocrinol 5:219–226

35. Eizirik DL, Colli ML (2020) Revisiting the role of inflammation in the loss of pancreatic β-cells in T1DM. Nat Rev Endocrinol 16:611–612

36. Hussen HI, Moradi T, Persson M (2015) The Risk of Type 1 Diabetes Among Offspring of Immigrant Mothers in Relation to the Duration of Residency in Sweden. Diabetes Care 38:934–936

37. Geravandi S, Richardson S, Pugliese A, Maedler K (2021) Localization of enteroviral RNA within the pancreas in donors with T1D and T1D-associated autoantibodies. Cell reports Med 2:100371

38. Vehik K, Lynch KF, Wong MC, et al (2019) Prospective virome analyses in young children at increased genetic risk for type 1 diabetes. Nat Med 25:1865–1872

39. Nekoua MP, Bertin A, Sane F, Gimeno J-P, Fournier I, Salzet M, Engelmann I, Alidjinou EK, Hober D (2021) Persistence of Coxsackievirus B4 in Pancreatic β Cells Disturbs Insulin Maturation, Pattern of Cellular Proteins, and DNA Methylation. Microorganisms. https://doi.org/10.3390/microorganisms9061125

40. Härkönen T, Paananen A, Lankinen H, Hovi T, Vaarala O, Roivainen M (2003) Enterovirus infection may induce humoral immune response reacting with islet cell autoantigens in humans. J Med Virol 69:426–440

41. Akhbari P, Richardson SJ, Morgan NG (2020) Type 1 Diabetes: Interferons and the Aftermath of Pancreatic Beta-Cell Enteroviral Infection. Microorganisms. https://doi.org/10.3390/microorganisms8091419

42. Dedrick S, Sundaresh B, Huang Q, et al (2020) The Role of Gut Microbiota and Environmental Factors in Type 1 Diabetes Pathogenesis. Front Endocrinol (Lausanne) 11:78

43. Tremblay J, Hamet P (2019) Environmental and genetic contributions to diabetes. Metabolism 100S:153952

44. Khamis T, Abdelalim AF, Saeed AA, Edress NM, Nafea A, Ebian HF, Algendy R, Hendawy DM, Arisha AH, Abdallah SH (2021) Breast milk MSCs upregulated β-cells PDX1, Ngn3, and PCNA expression via remodeling ER stress /inflammatory /apoptotic signaling pathways in type 1 diabetic rats. Eur J Pharmacol 905:174188

45. Corbin KD, Driscoll KA, Pratley RE, Smith SR, Maahs DM, Mayer-Davis EJ (2018) Obesity in Type 1 Diabetes: Pathophysiology, Clinical Impact, and Mechanisms. Endocr Rev 39:629–663

46. Redondo MJ, Jeffrey J, Fain PR, Eisenbarth GS, Orban T (2008)

Concordance for islet autoimmunity among monozygotic twins. N Engl J Med 359:2849–2850

47. Pociot F (2017) Type 1 diabetes genome-wide association studies: not to be lost in translation. Clin Transl Immunol 6:e162

48. Pociot F, Lernmark Å (2016) Genetic risk factors for type 1 diabetes. Lancet (London, England) 387:2331–2339

49. Redondo MJ, Steck AK, Pugliese A (2018) Genetics of type 1 diabetes. Pediatr Diabetes 19:346–353

50. Eizirik DL, Pasquali L, Cnop M (2020) Pancreatic β-cells in type 1 and type 2 diabetes mellitus: different pathways to failure. Nat Rev Endocrinol. https://doi.org/10.1038/s41574-020-0355-7

51. Ramos-Rodríguez M, Pérez-González B, Pasquali L (2021) The β-Cell Genomic Landscape in T1D: Implications for Disease Pathogenesis. Curr Diab Rep 21:1–9

52. Chatterjee N, Shi J, García-Closas M (2016) Developing and evaluating polygenic risk prediction models for stratified disease prevention. Nat Rev Genet 17:392–406

53. Cano-Gamez E, Trynka G (2020) From GWAS to Function: Using Functional Genomics to Identify the Mechanisms Underlying Complex Diseases. Front Genet 11:1–21

54. Størling J, Pociot F (2017) Type 1 diabetes candidate genes linked to pancreatic islet cell inflammation and beta-cell apoptosis. Genes (Basel) 8:1–12

55. Inshaw JRJJ, Cutler AJ, Burren OS, Stefana MI, Todd JA (2018) Approaches and advances in the genetic causes of autoimmune disease and their implications. Nat Immunol 19:674–684

56. Cudworth AG, Woodrow JC (1974) Letter: HL-A antigens and diabetes mellitus. Lancet (London, England) 2:1153

57. Polychronakos C, Li Q (2011) Understanding type 1 diabetes through genetics: advances and prospects. Nat Rev Genet 12:781–792

58. Noble JA, Valdes AM, Cook M, Klitz W, Thomson G, Erlich HA (1996) The role of HLA class II genes in insulin-dependent diabetes mellitus: molecular analysis of 180 Caucasian, multiplex families. Am J Hum Genet 59:1134–1148

59. Ziegler AG, Rewers M, Simell O, et al (2013) Seroconversion to Multiple Islet Autoantibodies and Risk of Progression to Diabetes in Children. JAMA 309:2473–2479

60. Ilonen J, Lempainen J, Veijola R (2019) The heterogeneous pathogenesis of type 1 diabetes mellitus. Nat Rev Endocrinol 15:635–650

61. Carré A, Richardson SJ, Larger E, Mallone R (2021) Presumption of guilt for T cells in type 1 diabetes: lead culprits or partners in crime depending on age of onset? Diabetologia 64:15–25

62. Nyaga DM, Vickers MH, Jefferies C, Perry JK, O'Sullivan JM (2018) The genetic architecture of type 1 diabetes mellitus. Mol Cell Endocrinol 477:70–80

63. Pugliese A, Zeller M, Fernandez AJ, Zalcberg LJ, Bartlett RJ, Ricordi C, Pietropaolo M, Eisenbarth GS, Bennett ST, Patel DD (1997) The insulin gene is transcribed in the human thymus and transcription levels correlated with allelic variation at the INS VNTR-IDDM2 susceptibility locus for type 1 diabetes. Nat Genet 15:293–297

64. Bell GI, Horita S, Karam JH (1984) A Polymorphic Locus Near the Human Insulin Gene Is Associated with Insulin-dependent Diabetes Melliitus. Diabetes 33:176–183

65. Grant SFA, Wells AD, Rich SS (2020) Next steps in the identification of gene targets for type 1 diabetes. Diabetologia. https://doi.org/10.1007/s00125-020-05248-8

66. Ramos-Rodríguez M, Pérez-González B, Pasquali L (2021) The β-Cell Genomic Landscape in T1D: Implications for Disease Pathogenesis. Curr Diab Rep 21:1

67. Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, Manolio TA (2009) Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proc Natl Acad Sci U S A 106:9362–9367

68. Nyaga DM, Vickers MH, Jefferies C, Fadason T, O'Sullivan JM (2021) Untangling the genetic link between type 1 and type 2 diabetes using functional genomics. Sci Rep 11:13871

69. Nyaga DM, Vickers MH, Jefferies C, Perry JK, O'Sullivan JM (2018) Type 1 Diabetes Mellitus-Associated Genetic Variants Contribute to Overlapping Immune Regulatory Networks. Front Genet 9:535

70. Onengut-Gumuscu S, Chen W-M, Burren O, et al (2015) Fine mapping of type 1 diabetes susceptibility loci and evidence for colocalization of causal variants with lymphoid gene enhancers. Nat Genet 47:381–386

71. Farh KK-H, Marson A, Zhu J, et al (2015) Genetic and epigenetic fine mapping of causal autoimmune disease variants. Nature 518:337–343

72. Fasolino M, Goldman N, Wang W, et al (2020) Genetic Variation in Type 1 Diabetes Reconfigures the 3D Chromatin Organization of T Cells and Alters Gene Expression. Immunity 52:257-274.e11

73. Robertson CC, Inshaw JRJ, Onengut-Gumuscu S, et al (2021) Fine-mapping, trans-ancestral and genomic analyses identify causal variants, cells, genes and drug targets for type 1 diabetes. Nat Genet 53:962–971

74. Chiou J, Geusz RJ, Okino M-L, et al (2021) Interpreting type 1 diabetes risk with genetics and single-cell epigenomics. Nature 594:398–402

75. Kaur S, Mirza AH, Overgaard AJ, Pociot F, Størling J (2021) A Dual Systems Genetics Approach Identifies Common Genes, Networks, and Pathways for Type 1 and 2 Diabetes in Human Islets. Front Genet 12:630109

76. Inshaw JRJ, Sidore C, Cucca F, Stefana MI, Crouch DJM, McCarthy MI, Mahajan A, Todd JA (2021) Analysis of overlapping genetic association in type 1 and type 2 diabetes. Diabetologia 64:1342–1347

77. Pasquali L, Gaulton KJ, Rodríguez-Seguí SA, et al (2014) Pancreatic islet enhancer clusters enriched in type 2 diabetes risk-associated variants. Nat Genet 46:136–143

78. Sims EK, Mirmira RG, Evans-Molina C (2020) The role of beta-cell dysfunction in early type 1 diabetes. Curr Opin Endocrinol Diabetes Obes 27:215–224

79. Roep BO, Thomaidou S, van Tienhoven R, Zaldumbide A (2021) Type 1 diabetes mellitus as a disease of the β-cell (do not blame the immune system?). Nat Rev Endocrinol 17:150–161

80. Nathan DM, Genuth S, Lachin J, Cleary P, Crofford O, Davis M, Rand L, Siebert C (1993) The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. N Engl J Med 329:977–986

81. Nathan DM (2014) The diabetes control and complications trial/epidemiology of diabetes interventions and complications study at 30 years: overview. Diabetes Care 37:9–16

82. Alonso Martín DE, Roldán Martín MB, Álvarez Gómez MÁ, Yelmo

Valverde R, Martín-Frías M, Alonso Blanco M, Barrio Castellanos R (2016) Impact of diabetes education on type 1 diabetes mellitus control in children. Endocrinol y Nutr organo la Soc Esp Endocrinol y Nutr 63:536–542

83. Ramos EL, Dayan CM, Chatenoud L, et al (2023) Teplizumab and β-Cell Function in Newly Diagnosed Type 1 Diabetes. N Engl J Med 389:2151–2161

84. Herold KC, Gitelman SE, Gottlieb PA, Knecht LA, Raymond R, Ramos EL (2023) Teplizumab: A Disease-Modifying Therapy for Type 1 Diabetes That Preserves β-Cell Function. Diabetes Care 46:1848–1856

85. Atkinson MA, Roep BO, Posgai A, Wheeler DCS, Peakman M (2019) The challenge of modulating β-cell autoimmunity in type 1 diabetes. lancet Diabetes Endocrinol 7:52–64

86. Nagy G, Szekely TE, Somogyi A, Herold M, Herold Z (2022) New therapeutic approaches for type 1 diabetes: Disease-modifying therapies. World J Diabetes 13:835–850

87. Pathak V, Pathak NM, O'Neill CL, Guduric-Fuchs J, Medina RJ (2019) Therapies for Type 1 Diabetes: Current Scenario and Future Perspectives. Clin Med Insights Endocrinol Diabetes 12:1179551419844521

88. Matsumoto S (2010) Islet cell transplantation for Type 1 diabetes. J Diabetes 2:16–22

89. Vertex Type 1 diabetes. https://www.vrtx.com/our-science/pipeline/type-1-diabetes/. Accessed 2 Sep 2024

90. Zhong J, Xu J-F, Yang P, Liang Y, Wang C-Y (2011) Innate Immunity in the Recognition of β-Cell Antigens in Type 1 Diabetes. Type 1 Diabetes. https://doi.org/10.5772/22264

91. James EA, Joglekar A V., Linnemann AK, Russ HA, Kent SC (2023) The beta cell-immune cell interface in type 1 diabetes (T1D). Mol Metab 78:101809

92. Lindley S, Dayan CM, Bishop A, Roep BO, Peakman M, Tree TIM (2005) Defective suppressor function in CD4(+)CD25(+) T-cells from patients with type 1 diabetes. Diabetes 54:92–99

93. Tree TIM, Lawson J, Edwards H, Skowera A, Arif S, Roep BO, Peakman M (2010) Naturally arising human CD4 T-cells that recognize islet autoantigens and secrete interleukin-10 regulate proinflammatory T-cell responses via linked suppression. Diabetes 59:1451–1460

94. Kolb-Bachofen V, Kolb H (1989) A role for macrophages in the pathogenesis of type 1 diabetes. Autoimmunity 3:145–154

95. Phillips JM, Parish NM, Raine T, Bland C, Sawyer Y, De La Peña H, Cooke A (2009) Type 1 diabetes development requires both CD4+ and CD8+ T cells and can be reversed by non-depleting antibodies targeting both T cell populations. Rev Diabet Stud 6:97–103

96. Scherm MG, Wyatt RC, Serr I, Anz D, Richardson SJ, Daniel C (2022) Beta cell and immune cell interactions in autoimmune type 1 diabetes: How they meet and talk to each other. Mol Metab 64:1–18

97. de Filette JMK, Pen JJ, Decoster L, et al (2019) Immune checkpoint inhibitors and type 1 diabetes mellitus: a case report and systematic review. Eur J Endocrinol 181:363–374

98. Bender C, Rodriguez-Calvo T, Amirian N, Coppieters KT, von Herrath MG (2020) The healthy exocrine pancreas contains preproinsulin-specific CD8 T cells that attack islets in type 1 diabetes. Sci Adv. https://doi.org/10.1126/sciadv.abc5586

99. Bottazzo GF (1986) Death of a- Beta Cell: Homicide or Suicide? Diabet Med 3:119–130

100. Meyerovich K, Ortis F, Allagnat F, Cardozo AK (2016) Endoplasmic reticulum stress and the unfolded protein response in pancreatic islet inflammation. J Mol Endocrinol 57:R1–R17

101. Eizirik DL, Pasquali L, Cnop M (2020) Pancreatic β-cells in type 1 and type 2 diabetes mellitus: different pathways to failure. Nat Rev Endocrinol 16:349–362

102. Pang H, Luo S, Huang G, Xia Y, Xie Z, Zhou Z (2020) Advances in Knowledge of Candidate Genes Acting at the Beta-Cell Level in the Pathogenesis of T1DM. Front Endocrinol (Lausanne) 11:119

103. Marré ML, James EA, Piganelli JD (2015) β cell ER stress and the implications for immunogenicity in type 1 diabetes. Front cell Dev Biol 3:67

104. Miani M, Barthson J, Colli ML, Brozzi F, Cnop M, Eizirik DL (2013) Endoplasmic reticulum stress sensitizes pancreatic beta cells to interleukin-1β-induced apoptosis via Bim/A1 imbalance. Cell Death Dis 4:e701

105. Gurzov EN, Ortis F, Cunha DA, Gosset G, Li M, Cardozo AK, Eizirik DL (2009) Signaling by IL-1beta+IFN-gamma and ER

stress converge on DP5/Hrk activation: a novel mechanism for pancreatic beta-cell apoptosis. Cell Death Differ 16:1539–1550

106. Toren E, Burnette KS, Banerjee RR, Hunter CS, Tse HM (2021) Partners in Crime: Beta-Cells and Autoimmune Responses Complicit in Type 1 Diabetes Pathogenesis. Front Immunol 12:756548

107. Dula SB, Jecmenica M, Wu R, Jahanshahi P, Verrilli GM, Carter JD, Brayman KL, Nunemaker CS (2010) Evidence that low-grade systemic inflammation can induce islet dysfunction as measured by impaired calcium handling. Cell Calcium 48:133–142

108. Darville MI, Eizirik DL (2006) Notch signaling: a mediator of beta-cell de-differentiation in diabetes? Biochem Biophys Res Commun 339:1063–1068

109. Firdessa Fite R, Bechi Genzano C, Mallone R, Creusot RJ (2023) Epitope-based precision immunotherapy of Type 1 diabetes. Hum Vaccin Immunother 19:2154098

110. Gonzalez-Duque S, Azoury ME, Colli ML, et al (2018) Conventional and Neo-antigenic Peptides Presented by β Cells Are Targeted by Circulating Naïve CD8+ T Cells in Type 1 Diabetic and Healthy Donors. Cell Metab 28:946-960.e6

111. Rodriguez-Calvo T, Johnson JD, Overbergh L, Dunne JL (2021) Neoepitopes in Type 1 Diabetes: Etiological Insights, Biomarkers and Therapeutic Targets. Front Immunol 12:667989

112. Roep BO, Kracht MJ, van Lummel M, Zaldumbide A (2016) A roadmap of the generation of neoantigens as targets of the immune system in type 1 diabetes. Curr Opin Immunol 43:67–73

113. Eizirik DL, Sammeth M, Bouckenooghe T, et al (2012) The human pancreatic islet transcriptome: expression of candidate genes for type 1 diabetes and the impact of pro-inflammatory cytokines. PLoS Genet 8:e1002552

114. Brissova M, Haliyur R, Saunders D, et al (2018) α Cell Function and Gene Expression Are Compromised in Type 1 Diabetes. Cell Rep 22:2667–2676

115. Eizirik DL, Szymczak F, Mallone R (2023) Why does the immune system destroy pancreatic β-cells but not α-cells in type 1 diabetes? Nat Rev Endocrinol 19:425–434

116. Piemonti L, Pileggi A (2013) 25 YEARS OF THE RICORDI AUTOMATED METHOD FOR ISLET ISOLATION. CellR4--repair, Replace. Regen. reprogramming 1:

117. Khosravi-maharlooei M, Madley R, Borsotti C, Ferreira LMR, Sharp RC, Brehm MA, Greiner DL, Parent A V, Anderson MS, Sykes M (2022) Modeling human T1D-associated autoimmune processes. Mol Metab 56:101417

118. Chatzigeorgiou A, Halapas A, Kalafatakis K, Kamper E (2009) The use of animal models in the study of diabetes mellitus. In Vivo 23:245–258

119. Singh R, Gholipourmalekabadi M, Shafikhani SH (2024) Animal models for type 1 and type 2 diabetes: advantages and limitations. Front Endocrinol (Lausanne) 15:1359685

120. Skelin M, Rupnik M, Cencic A (2010) Pancreatic beta cell lines and their applications in diabetes mellitus research. ALTEX 27:105–113

121. TrialNet TrialNet, Type 1 Diabetes. https://www.trialnet.org/our-research.

122. Bingley PJ, Wherrett DK, Shultz A, Rafkin LE, Atkinson MA, Greenbaum CJ (2018) Type 1 Diabetes TrialNet: A Multifaceted Approach to Bringing Disease-Modifying Therapy to Clinical Use in Type 1 Diabetes. Diabetes Care 41:653–661

123. Kaestner KH, Powers AC, Naji A, Atkinson MA (2019) NIH Initiative to Improve Understanding of the Pancreas, Islet, and Autoimmunity in Type 1 Diabetes: The Human Pancreas Analysis Program (HPAP). Diabetes 68:1394–1402

124. Hogrebe NJ, Ishahak M, Millman JR (2023) Developments in stem cell-derived islet replacement therapy for treating type 1 diabetes. Cell Stem Cell 30:530–548

125. Peters L, Posgai A, Brusko TM (2019) Islet–immune interactions in type 1 diabetes: the nexus of beta cell destruction. Clin Exp Immunol 198:326–340

126. Maxwell KG, Millman JR (2021) Applications of iPSC-derived beta cells from patients with diabetes. Cell Reports Med 2:100238

127. Leite NC, Sintov E, Meissner TB, Brehm MA, Greiner DL, Harlan DM, Melton DA (2020) Modeling Type 1 Diabetes In Vitro Using Human Pluripotent Stem Cells. Cell Rep 32:107894

128. Jin W, Jiang W (2022) Stepwise differentiation of functional pancreatic β cells from human pluripotent stem cells. Cell Regen (London, England) 11:24

129. Al-Khawaga S, Memon B, Butler AE, Taheri S, Abou-Samra AB,

Abdelalim EM (2018) Pathways governing development of stem cell-derived pancreatic β cells: lessons from embryogenesis. Biol Rev Camb Philos Soc 93:364–389

130. Ramos-Rodríguez M, Raurell-Vila H, Colli ML, et al (2019) The impact of proinflammatory cytokines on the β-cell regulatory landscape provides insights into the genetics of type 1 diabetes. Nat Genet 51:1588–1595

131. Tsonkova VG, Sand FW, Wolf XA, et al (2018) The EndoC-βH1 cell line is a valid model of human beta cells and applicable for screenings to identify novel drug target candidates. Mol Metab 8:144–157

132. Human Cell Design EndoC-βH1® The human beta cell line of reference. https://www.humancelldesign.com/human-beta-cells-endoc-bh1/. Accessed 23 May 2024

133. Benaglio P, Zhu H, Okino M, Frazer KA, Sander M, Gaulton KJ (2022) Type 1 diabetes risk genes mediate pancreatic beta cell survival in response to proinflammatory cytokines ll Type 1 diabetes risk genes mediate pancreatic beta cell survival in response to proinflammatory cytokines. Cell Genomics 2:100214

134. Colli ML, Ramos-Rodríguez M, Nakayasu ES, et al (2020) An integrated multi-omics approach identifies the landscape of interferon-α-mediated responses of human pancreatic beta cells. Nat Commun 11:2584

135. Stancill JS, Kasmani MY, Cui W, Corbett JA (2024) Single Cell RNAseq analysis of cytokine-treated human islets: Association of cellular stress with impaired cytokine responsiveness. Function zqae015

136. Eizirik DL, Sandler S, Welsh N, Cetkovic-Cvrlje M, Nieman A, Geller DA, Pipeleers DG, Bendtzen K, Hellerström C (1994) Cytokines suppress human islet function irrespective of their effects on nitric oxide generation. J Clin Invest 93:1968–1974

137. Dettmer R, Niwolik I, Cirksena K, Yoshimoto T, Tang Y, Mehmeti I, Gurgul-Convey E, Naujok O (2022) Proinflammatory cytokines induce rapid, NO-independent apoptosis, expression of chemotactic mediators and interleukin-32 secretion in human pluripotent stem cell-derived beta cells. Diabetologia 65:829–843

138. Oleson BJ, McGraw JA, Broniowska KA, Annamalai M, Chen J, Bushkofsky JR, Davis DB, Corbett JA, Mathews CE (2015) Distinct differences in the responses of the human pancreatic β-cell line EndoC-βH1 and human islets to proinflammatory

cytokines. Am J Physiol Regul Integr Comp Physiol 309:R525-34

139. Chen M, Yang ZD, Smith KM, Carter JD, Nadler JL (2005) Activation of 12-lipoxygenase in proinflammatory cytokine-mediated beta cell toxicity. Diabetologia 48:486–495

140. Grunnet LG, Aikin R, Tonnesen MF, Paraskevas S, Blaabjerg L, Størling J, Rosenberg L, Billestrup N, Maysinger D, Mandrup-Poulsen T (2009) Proinflammatory cytokines activate the intrinsic apoptotic pathway in beta-cells. Diabetes 58:1807–1815

141. Piovesan A, Pelleri MC, Antonaros F, Strippoli P, Caracausi M, Vitale L (2019) On the length, weight and GC content of the human genome. BMC Res Notes 12:106

142. Carninci P, Kasukawa T, Katayama S, et al (2005) Molecular biology: The transcriptional landscape of the mammalian genome. Science (80-) 309:1559–1563

143. Pennisi E (2012) Genomics. ENCODE project writes eulogy for junk DNA. Science 337:1159,1161

144. Gilbert N, Gilchrist S, Bickmore WA (2004) Chromatin organization in the mammalian nucleus. Int Rev Cytol 242:283–336

145. Buschbeck M, Hake SB (2017) Variants of core histones and their roles in cell fate decisions, development and cancer. Nat Rev Mol Cell Biol 18:299–314

146. Kempfer R, Pombo A (2020) Methods for mapping 3D chromosome architecture. Nat Rev Genet 21:207–226

147. Ruiz-Velasco M, Zaugg JB (2017) Structure meets function: How chromatin organisation conveys functionality. Curr Opin Syst Biol 1:129–136

148. Anania C, Lupiáñez DG (2020) Order and disorder: abnormal 3D chromatin organization in human disease. Brief Funct Genomics 19:128–138

149. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature 485:376–380

150. Lieberman-Aiden E, van Berkum NL, Williams L, et al (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 326:289–293

151. Schmitt AD, Hu M, Jung I, et al (2016) A Compendium of Chromatin Contact Maps Reveals Spatially Active Regions in the Human Genome. Cell Rep 17:2042–2059

152. Long HS, Greenaway S, Powell G, Mallon A-M, Lindgren CM, Simon MM (2022) Making sense of the linear genome, gene function and TADs. Epigenetics Chromatin 15:4

153. Le NQK, Yapp EKY, Nagasundaram N, Yeh H-Y (2019) Classifying Promoters by Interpreting the Hidden Information of DNA Sequences via Deep Learning and Combination of Continuous FastText N-Grams. Front Bioeng Biotechnol 7:305

154. Hernandez-Garcia CM, Finer JJ (2014) Identification and validation of promoters and cis-acting regulatory elements. Plant Sci 217–218:109–119

155. Haberle V, Lenhard B (2016) Promoter architectures and developmental gene regulation. Semin Cell Dev Biol 57:11–23

156. Panigrahi A, O'Malley BW (2021) Mechanisms of enhancer action: the known and the unknown. Genome Biol 22:108

157. Gasperini M, Tome JM, Shendure J (2020) Towards a comprehensive catalogue of validated and target-linked human enhancers. Nat Rev Genet 21:292–310

158. Pott S, Lieb JD (2015) What are super-enhancers? Nat Genet 47:8–12

159. Saxonov S, Berg P, Brutlag DL (2006) A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. Proc Natl Acad Sci U S A 103:1412–1417

160. Bell JSK, Vertino PM (2017) Orphan CpG islands define a novel class of highly active enhancers. Epigenetics 12:449–464

161. Jabbari K, Bernardi G (2004) Cytosine methylation and CpG, TpG (CpA) and TpA frequencies. Gene 333:143–149

162. Klemm SL, Shipony Z, Greenleaf WJ (2019) Chromatin accessibility and the regulatory epigenome. Nat Rev Genet 20:207–220

163. Giresi PG, Kim J, McDaniell RM, Iyer VR, Lieb JD (2007) FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin. Genome Res 17:877–885

164. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ

(2013) Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods 10:1213–1218

165. Grandi FC, Modi H, Kampman L, Corces MR (2022) Chromatin accessibility profiling by ATAC-seq. Nat Protoc 17:1518–1552

166. Baek S, Lee I (2020) Single-cell ATAC sequencing analysis: From data preprocessing to hypothesis generation. Comput Struct Biotechnol J 18:1429–1439

167. Zhang Y, Sun Z, Jia J, Du T, Zhang N, Tang Y, Fang Y, Fang D (2021) Overview of Histone Modification. In: Fang D, Han J (eds) Histone Mutat. Cancer. Springer Singapore, Singapore, pp 1–16

168. Ooi SKT, Qiu C, Bernstein E, et al (2007) DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. Nature 448:714–717

169. Kimura H (2013) Histone modifications for human epigenome analysis. J Hum Genet 58:439–445

170. Schmidl C, Rendeiro AF, Sheffield NC, Bock C (2015) ChIPmentation: fast, robust, low-input ChIP-seq for histones and transcription factors. Nat Methods 12:963–965

171. Rotem A, Ram O, Shoresh N, Sperling RA, Goren A, Weitz DA, Bernstein BE (2015) Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state. Nat Biotechnol 33:1165–1172

172. Grosselin K, Durand A, Marsolier J, et al (2019) High-throughput single-cell ChIP-seq identifies heterogeneity of chromatin states in breast cancer. Nat Genet 51:1060–1066

173. Grosselin K, Durand A, Marsolier J, et al (2019) High-throughput single-cell ChIP-seq identifies heterogeneity of chromatin states in breast cancer. Nat Genet 51:1060–1066

174. Skene PJ, Henikoff S (2017) An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. Elife 6:1–35

175. Kaya-Okur HS, Wu SJ, Codomo CA, Pledger ES, Bryson TD, Henikoff JG, Ahmad K, Henikoff S (2019) CUT&Tag for efficient epigenomic profiling of small samples and single cells. Nat Commun 10:1–10

176. Kaya-Okur HS, Janssens DH, Henikoff JG, Ahmad K, Henikoff S (2020) Efficient low-cost chromatin profiling with CUT&Tag. Nat

Protoc 15:3264–3283

177. Bartosovic M, Kabbe M, Castelo-Branco G (2021) Single-cell CUT&Tag profiles histone modifications and transcription factors in complex tissues. Nat Biotechnol 39:825–835

178. Wu SJ, Furlan SN, Mihalas AB, et al (2021) Single-cell CUT&Tag analysis of chromatin modifications in differentiation and tumor progression. Nat Biotechnol 39:819–824

179. Janssens DH, Otto DJ, Meers MP, Setty M, Ahmad K, Henikoff S (2022) CUT&Tag2for1: a modified method for simultaneous profiling of the accessible and silenced regulome in single cells. Genome Biol 23:81

180. Janssens DH, Greene JE, Wu SJ, Codomo CA, Minot SS, Furlan SN, Ahmad K, Henikoff S (2024) Scalable single-cell profiling of chromatin modifications with sciCUT&Tag. Nat Protoc 19:83–112

181. Zhijun Fu Sanjie Jiang YSSZLZ, Li P (2024) Cut\&tag: a powerful epigenetic tool for chromatin profiling. Epigenetics 19:2293411

182. Han J, Zhang Z, Wang K (2018) 3C and 3C-based techniques: the powerful tools for spatial genome organization deciphering. Mol Cytogenet 11:21

183. Javierre BM, Burren OS, Wilder SP, et al (2016) Lineage-Specific Genome Architecture Links Enhancers and Non-coding Disease Variants to Target Gene Promoters. Cell 167:1369-1384.e19

184. Schwartzman O, Mukamel Z, Oded-Elkayam N, Olivares-Chauvet P, Lubling Y, Landan G, Izraeli S, Tanay A (2016) UMI-4C for quantitative and targeted chromosomal contact profiling. Nat Methods 13:685–691

185. Pliner HA, Packer JS, McFaline-Figueroa JL, et al (2018) Cicero Predicts cis-Regulatory DNA Interactions from Single-Cell Chromatin Accessibility Data. Mol Cell 71:858-871.e8

186. Smale ST (2010) Luciferase assay. Cold Spring Harb Protoc 2010:pdb.prot5421

187. Arda HE, Tsai J, Rosli YR, Giresi P, Bottino R, Greenleaf WJ, Chang HY, Kim SK (2018) A Chromatin Basis for Cell Lineage and Disease Risk in the Human Pancreas. Cell Syst 7:310-322.e4

188. Parker SCJ, Stitzel ML, Taylor DL, et al (2013) Chromatin stretch enhancer states drive cell-specific gene regulation and harbor human disease risk variants. Proc Natl Acad Sci U S A

110:17921–17926

189. Bhandare R, Schug J, Le Lay J, Fox A, Smirnova O, Liu C, Naji A, Kaestner KH (2010) Genome-wide analysis of histone modifications in human pancreatic islets. Genome Res 20:428–433

190. Gaulton KJ, Nammo T, Pasquali L, et al (2010) A map of open chromatin in human pancreatic islets. Nat Genet 42:255–259

191. Stitzel ML, Sethupathy P, Pearson DS, et al (2010) Global epigenomic analysis of primary human pancreatic islets provides insights into type 2 diabetes susceptibility loci. Cell Metab 12:443–455

192. Cebola I, Rodríguez-Seguí SA, Cho CHH, et al (2015) TEAD and YAP regulate the enhancer network of human embryonic pancreatic progenitors. Nat Cell Biol 17:615–626

193. Miguel-Escalada I, Bonàs-Guarch S, Cebola I, et al (2019) Human pancreatic islet three-dimensional chromatin architecture provides insights into the genetics of type 2 diabetes. Nat Genet 51:1137–1148

194. Zhao Z, D'Oliveira Albanus R, Taylor H, et al (2023) An integrative single-cell multi-omics profiling of human pancreatic islets identifies T1D associated genes and regulatory signals. Res Sq. https://doi.org/10.21203/rs.3.rs-3343318/v1

195. Benaglio P, Zhu H, Okino ML, et al (2022) Type 1 diabetes risk genes mediate pancreatic beta cell survival in response to proinflammatory cytokines. Cell genomics 2:100214

196. Buenrostro JD, Wu B, Chang HY, Greenleaf WJ (2015) ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. Curr Protoc Mol Biol 109:21.29.1-21.29.9

197. Mallona I, Díez-Villanueva A, Martín B, Peinado MA (2017) Chainy: an universal tool for standardized relative quantification in real-time PCR. Bioinformatics 33:1411–1413

198. Roca-Rivada A, Marín-Cañas S, Colli ML, Vinci C, Sawatani T, Marselli L, Cnop M, Marchetti P, Eizirik DL (2023) Inhibition of the type 1 diabetes candidate gene PTPN2 aggravates TNF-α-induced human beta cell dysfunction and death. Diabetologia. https://doi.org/10.1007/S00125-023-05908-5

199. Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15:550

200. Coomans de Brachène A, Alvelos MI, Szymczak F, et al (2024) Interferons are key cytokines acting on pancreatic islets in type 1 diabetes. Diabetologia 67:908–927

201. Schulthess FT, Paroni F, Sauter NS, Shu L, Ribaux P, Haataja L, Strieter RM, Oberholzer J, King CC, Maedler K (2009) CXCL10 impairs beta cell function and viability in diabetes through TLR4 signaling. Cell Metab 9:125–139

202. Chervonsky A V, Wang Y, Wong FS, Visintin I, Flavell RA, Janeway CAJ, Matis LA (1997) The role of Fas in autoimmune diabetes. Cell 89:17–24

203. Liu S-W, Sun F, Rong S-J, Wang T, Wang C-Y (2022) Lymphotoxins Serve as a Novel Orchestrator in T1D Pathogenesis. Front Immunol 13:917577

204. Boraska V, Zeggini E, Groves CJ, Rayner NW, Škrabić V, Diakite M, Rockett KA, Kwiatkowski D, McCarthy MI, Zemunik T (2009) Family-based analysis of tumor necrosis factor and lymphotoxin-α tag polymorphisms with type 1 diabetes in the population of South Croatia. Hum Immunol 70:195–199

205. Timasheva YR, Balkhiyarova ZR, Nasibullin TR, Avzaletdinova DS, Morugova T V, Mustafina OE, Prokopenko I (2019) Multilocus associations of inflammatory genes with the risk of type 1 diabetes. Gene 707:1–8

206. Hagan DW, Ferreira SM, Santos GJ, Phelps EA (2022) The role of GABA in islet function. Front Endocrinol (Lausanne) 13:972115

207. Vis MAM, Ito K, Hofmann S (2020) Impact of Culture Medium on Cellular Interactions in in vitro Co-culture Systems. Front Bioeng Biotechnol 8:1–8

208. Chen C-W, Guan B-J, Alzahrani MR, et al (2022) Adaptation to chronic ER stress enforces pancreatic β-cell plasticity. Nat Commun 13:4621

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

210. Lu J, Liu J, Li L, Lan Y, Liang Y (2020) Cytokines in type 1 diabetes: mechanisms of action and immunotherapeutic targets. Clin Transl Immunol 9:e1122

211. Tokunaga R, Zhang W, Naseem M, Puccini A, Berger MD, Soni S, McSkane M, Baba H, Lenz H-J (2018) CXCL9, CXCL10, CXCL11/CXCR3 axis for immune activation - A target for novel cancer therapy. Cancer Treat Rev 63:40–47

212. Jones BA, Beamer M, Ahmed S (2010) Fractalkine/CX3CL1: a potential new target for inflammatory diseases. Mol Interv 10:263–270

213. Cormican S, Griffin MD (2021) Fractalkine (CX3CL1) and Its Receptor CX3CR1: A Promising Therapeutic Target in Chronic Kidney Disease? Front Immunol 12:664202

214. Sindhu S, Akhter N, Arefanian H, Al-Roub AA, Ali S, Wilson A, Al-Hubail A, Al-Beloushi S, Al-Zanki S, Ahmad R (2017) Increased circulatory levels of fractalkine (CX3CL1) are associated with inflammatory chemokines and cytokines in individuals with type-2 diabetes. J Diabetes Metab Disord 16:15

215. Shah R, Hinkle CC, Ferguson JF, Mehta NN, Li M, Qu L, Lu Y, Putt ME, Ahima RS, Reilly MP (2011) Fractalkine is a novel human adipochemokine associated with type 2 diabetes. Diabetes 60:1512–1518

216. Shepard BD, Badley AD (2009) The Biology of TRAIL and the Role of TRAIL-Based Therapeutics in Infectious Diseases. Antiinfect Agents Med Chem 8:87–101

217. Yue X, Wu L, Hu W (2015) The regulation of leukemia inhibitory factor. Cancer cell Microenviron. https://doi.org/10.14800/ccm.877

218. Klak M, Gomółka M, Kowalska P, Cichoń J, Ambrożkiewicz F, Serwańska-Świętek M, Berman A, Wszoła M (2020) Type 1 diabetes: genes associated with disease development. Cent J Immunol 45:439–453

219. Rich SS (2017) Genetics and its potential to improve type 1 diabetes care. Curr Opin Endocrinol Diabetes Obes 24:279–284

220. Vaccarezza M, Delbello G, Zauli G (2007) A role of the TRAIL-TRAIL receptor system in the pathogenesis of diabetes. Acta Biomed 78 Suppl 1:262–267

221. Lamhamedi-Cherradi S-E, Zheng S-J, Maguschak KA, Peschon J, Chen YH (2003) Defective thymocyte apoptosis and accelerated autoimmune diseases in TRAIL-/- mice. Nat Immunol 4:255–260

222. Lamhamedi-Cherradi S-E, Zheng S, Tisch RM, Chen YH (2003) Critical roles of tumor necrosis factor-related apoptosis-inducing ligand in type 1 diabetes. Diabetes 52:2274–2278

223. Cui A, Huang T, Li S, Ma A, Pérez JL, Sander C, Keskin DB, Wu CJ, Fraenkel E, Hacohen N (2024) Dictionary of immune

responses to cytokines at single-cell resolution. Nature 625:377–384

224. Kutlu B, Cardozo AK, Darville MI, Kruhøffer M, Magnusson N, Ørntoft T, Eizirik DL (2003) Discovery of Gene Networks Regulating Cytokine-Induced Dysfunction and Apoptosis in Insulin-Producing INS-1 Cells. Diabetes 52:2701–2719

225. Moore F, Colli ML, Cnop M, Esteve MI, Cardozo AK, Cunha DA, Bugliani M, Marchetti P, Eizirik DL (2009) PTPN2, a Candidate Gene for Type 1 Diabetes, Modulates Interferon-γ–Induced Pancreatic β-Cell Apoptosis. Diabetes 58:1283–1291

226. Elgamal RM, Kudtarkar P, Melton RL, Mummey HM, Benaglio P, Okino M-L, Gaulton KJ (2023) An Integrated Map of Cell Type-Specific Gene Expression in Pancreatic Islets. Diabetes 72:1719–1728

227. Thompson PJ, Shah A, Ntranos V, Van Gool F, Atkinson M, Bhushan A (2019) Targeted Elimination of Senescent Beta Cells Prevents Type 1 Diabetes. Cell Metab 29:1045-1060.e10

228. Lee H, Sahin GS, Chen C-W, et al (2023) Stress-induced β cell early senescence confers protection against type 1 diabetes. Cell Metab 35:2200-2215.e9

229. Cha J, Aguayo-Mazzucato C, Thompson PJ (2023) Pancreatic β-cell senescence in diabetes: mechanisms, markers and therapies. Front Endocrinol (Lausanne) 14:1212716

230. Ganic E, Singh T, Luan C, et al (2016) MafA-Controlled Nicotinic Receptor Expression Is Essential for Insulin Secretion and Is Impaired in Patients with Type 2 Diabetes. Cell Rep 14:1991–2002

231. Bsharat S, Monni E, Singh T, et al (2023) MafB-dependent neurotransmitter signaling promotes β cell migration in the developing pancreas. Development. https://doi.org/10.1242/dev.201009

232. Rachdi L, Maugein A, Pechberty S, Armanet M, Hamroune J, Ravassard P, Marullo S, Albagli O, Scharfmann R (2020) Regulated expression and function of the GABAB receptor in human pancreatic beta cell line and islets. Sci Rep 10:13469

233. Sanda S, Wei S, Rue T, Shilling H, Greenbaum C (2013) A SNP in G6PC2 predicts insulin secretion in type 1 diabetes. Acta Diabetol 50:459–462

234. Fasolino M, Schwartz GW, Patil AR, et al (2022) Single-cell multi-

omics analysis of human pancreatic islets reveals novel cellular states in type 1 diabetes. Nat Metab 4:284–299

235. Frørup C, Svane CAS, Henriksen K, Kaur S, Størling J (2024) Beta-Cell Pyroptosis – A Burning Flame in Type 1 Diabetes? bioRxiv 2024.01.05.574294

236. Liu B, Zhang M, Chu H, et al (2017) The ubiquitin E3 ligase TRIM31 promotes aggregation and activation of the signaling adaptor MAVS through Lys63-linked polyubiquitination. Nat Immunol 18:214–224

237. Cheng J, Xue F, Zhang M, et al (2018) TRIM31 Deficiency Is Associated with Impaired Glucose Metabolism and Disrupted Gut Microbiota in Mice. Front Physiol 9:24

238. Schmidlin M, Lu M, Leuenberger SA, Stoecklin G, Mallaun M, Gross B, Gherzi R, Hess D, Hemmings BA, Moroni C (2004) The ARE-dependent mRNA-destabilizing activity of BRF1 is regulated by protein kinase B. EMBO J 23:4760–4769

239. Ciais D, Cherradi N, Bailly S, Grenier E, Berra E, Pouyssegur J, Lamarre J, Feige J-J (2004) Destabilization of vascular endothelial growth factor mRNA by the zinc-finger protein TIS11b. Oncogene 23:8673–8680

240. Duan H, Cherradi N, Feige J-J, Jefcoate C (2009) cAMP-dependent posttranscriptional regulation of steroidogenic acute regulatory (STAR) protein by the zinc finger protein ZFP36L1/TIS11b. Mol Endocrinol 23:497–509

241. Lykke-Andersen J, Wagner E (2005) Recruitment and activation of mRNA decay enzymes by two ARE-mediated decay activation domains in the proteins TTP and BRF-1. Genes Dev 19:351–361

242. Dronkert ML, de Wit J, Boeve M, Vasconcelos ML, van Steeg H, Tan TL, Hoeijmakers JH, Kanaar R (2000) Disruption of mouse SNM1 causes increased sensitivity to the DNA interstrand cross-linking agent mitomycin C. Mol Cell Biol 20:4553–4561

243. Touzot F, Callebaut I, Soulier J, Gaillard L, Azerrad C, Durandy A, Fischer A, de Villartay J-P, Revy P (2010) Function of Apollo (SNM1B) at telomere highlighted by a splice variant identified in a patient with Hoyeraal–Hreidarsson syndrome. Proc Natl Acad Sci 107:10097–10102

244. Rosado-Olivieri EA, Aigha II, Kenty JH, Melton DA (2020) Identification of a LIF-Responsive, Replication-Competent Subpopulation of Human β Cells. Cell Metab 31:327-338.e6

245. Hoberecht L, Perampalam P, Lun A, Fortin J-P (2022) A comprehensive Bioconductor ecosystem for the design of CRISPR guide RNAs across nucleases and technologies. Nat Commun 13:6568