eman ta zabal zazu

Universidad del País Vasco
Euskal Herriko Unibertsitatea

DOCTORAL THESIS

IDENTIFICATION AND FUNCTIONAL CHARACTERIZATION OF LNCRNAS ASSOCIATED WITH TYPE 1 DIABETES

Itziar González Moro

Leioa, 2023

Supervised by:

Izortze Santin Gómez

Luis Antonio Castaño González

Mucha gente pequeña,
en lugares pequeños,
haciendo cosas pequeñas,
PODEMOS cambiar el mundo.

AKNOWLEDGEMENT

I would like to thank the University of the Basque Country for supporting me during these four years with the fellowship grant from the University of the Basque Country UPV/EHU to Itziar González Moro (PIF2018/150).

Biochemistry and Molecular Biology department (UPV-EHU), the Biotecnology Center María Goyri; and the members from both domains.

Finally, I would like to thank Mariana Igoillo-Esteve and her group for hosting me during my stay in the ULB Center for Diabetes Research (UCDR) in Brussels.

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following publications:

Gonzalez-Moro, I., Olazagoitia-Garmendia, A., Colli, M. L., Cobo-Vuilleumier, N., Postler, T. S., Marselli, L., Marchetti, P., Ghosh, S., Gauthier, B. R., Eizirik, D. L., Castellanos-Rubio, A., & Santin, I. (2020). The T1D-associated lncRNA Lnc13 modulates human pancreatic β cell inflammation by allele-specific stabilization of STAT1 mRNA. Proceedings of the National Academy of Sciences of the United States of America, 117(16), 9022–9031. https://doi.org/10.1073/pnas.1914353117

González-Moro I., Garcia-Etxebarria K., Mendoza L.M., Fernández-Jimenez N., Mentxaka-Salgado J., Olazagoitia-Garmendia A, Arroyo M.N., Sawatani T., de Beek A.O., Cnop M., Igoillo-Esteve M. & Santin I. (2022) The type 1 diabetes-associated lncRNA *ARGI* participates in virus-induced pancreatic β cell inflammation. bioRxiv; DOI: 10.1101/2022.12.01.518685.

During this thesis I have also participated in the following publications as first author:

González-Moro, I., & Santin, I. (2021). Long non-coding RNA-regulated pathways in pancreatic β cells: Their role in diabetes. International review of cell and molecular biology, 359, 325–355. https://doi.org/10.1016/bs.ircmb.2021.02.007

González-Moro, I., Rojas-Márquez, H., Sebastian-delaCruz, M., Mentxaka-Salgado, J., Olazagoitia-Garmendia, A., Mendoza, L. M., Lluch, A., Fantuzzi, F., Lambert, C., Ares Blanco, J., Marselli, L., Marchetti, P., Cnop, M., Delgado, E., Fernández-Real, J. M., Ortega, F. J., Castellanos-Rubio, A., & Santin, I. (2023). A long non-coding RNA that harbors a SNP associated with type 2 diabetes regulates the expression of TGM2 gene in pancreatic beta cells. Frontiers in endocrinology, 14, 1101934. https://doi.org/10.3389/fendo.2023.1101934

And as co-author in the following publications:

Sebastian-delaCruz, M., **Gonzalez-Moro, I**., Olazagoitia-Garmendia, A., Castellanos-Rubio, A., & Santin, I. (2021). The Role of lncRNAs in Gene Expression Regulation through mRNA Stabilization. Non-coding RNA, 7(1), 3. https://doi.org/10.3390/ncrna7010003

Sebastian-delaCruz, M., Olazagoitia-Garmendia, A., **Gonzalez-Moro, I**., Santin, I., Garcia-Etxebarria, K., & Castellanos-Rubio, A. (2020). Implication of m6A mRNA Methylation in Susceptibility to Inflammatory Bowel Disease. Epigenomes, 4(3), 16. https://doi.org/10.3390/epigenomes4030016

TABLE OF CONTENTS

Chapter 1

Abbreviations

Laburdurak

ABBREVIATIONS

APC: antigen-presenting cells

ARGI: Antiviral Response Gene Inducer

BIM: Bcl-2 Interacting Mediator of cell death

CDS: Cell Detachment Solution

CeD: celiac disease

ceRNA: competing endogenous RNA

ChIP: chromatin immunoprecipitation

CTCF: CCCTC-binding factor

CTLA4: Cytotoxic T-lymphocyte-associated antigen 4

CVB: coxsackievirus B

CVB5: Coxsackievirus B5

DsRNA: double strand RNA

DM: Diabetes mellitus

DTT: dithiothreitol

ELISA: Enzyme-Linked ImmunoSorbent Assay

ER: endoplasmic reticulum

ERBB2: Erb-B2 receptor tyrosine kinase 3

FC: fold change

GAS5: Growth arrest-specific 5

GO: Gene Ontology

GWAS: Genome Wide Association Study

HBSS: Hanks' Balanced Salt Solution

HLA: Human Leukocyte Antigen

IDF: International Diabetes Federation

Chapter 1: Abbreviations

IDIN: IRF7-driven inflammatory network

IFNAR: IFN-α/β receptor

IGF2: insulin-like growth factor 2

INS: insulin

iPS: induced pluripotent stem

ISG: interferon stimulated genes

JAK1: Janus kinase 1

JNK1: c-Jun N-terminal kinase 1

LD: linkage disequilibrium

lncRNA: long non-coding RNA

LYP: lymphoid tyrosine phosphatase

MALAT1: Metastasis-associated Lung Adenocarcinoma Transcript 1

MDA-5: melanoma differentiation-associated gene 5

MEG3: maternally expressed gene 3

MHC: Mayor Histocompatibility Complex

miRNA: micro RNA

MODY: Maturity onset diabetes of the young

MOI: multiplicity of infection

ND: Neonatal diabetes

NLS: Nuclear Lysis Solution

NT: non-transfected

PI3K: phosphoinositide 3-kinase

PIC: Poly(I:C); polyinosinic:polycytidylic acid

PMSF: phenyl methyl sulfonyl fluoride

PND: permanent neonatal diabetes

pSTAT1: phosphorylated STAT1

PTPN22: protein tyrosine phosphatase non-receptor type 22

Chapter 1: Abbreviations

qPCR: quantitative PCR

RAP: RNA antisense purification

RBPs: RNA binding proteins

RIG-I: retinoic acid-inducible gene-I

RIP: RNA immunoprecipitation

RTH: RNA in Technology and Health

sgRNA: single guide RNA

SNP: single nucleotide polymorphism

STAT1: Signal Transducer and Activator of Transcription 1

T1D: type 1 diabetes

T2D: type 2 diabetes

TCEP: tris (2- carboxyethyl) phosphine

TCR: T cell receptor

TEDDY: The Environmental Determinants of Diabetes in the Young

TLR3: Toll-like receptor 3

TND: transient neonatal diabetes

TPM: transcripts per million

TSS: transcription starting site

tSTAT1: total STAT1

TUG1: Taurin upregulated gene 1

TYK2: tyrosine kinase 2

UPR: unfolded protein response

UTR: untranslated region

VNTR: variable number of tandem repeats

WHO: World Health Organization

Chapter 2

Introduction

Sarrera

1 Diabetes

Diabetes mellitus (DM) is a chronic metabolic disease frequent in the childhood (Atkinson et al., 2014; Grulich-Henn & Klose, 2018). This disease is a heterogeneous condition characterized by a compendium of diverse metabolic disorders derived from persistent high blood glucose (hyperglycemia) as a result of defects in insulin secretion of pancreatic β cells, an abnormal insulin sensitivity of the receptor tissues (mainly liver, muscle and adipocytes) or a combination of both factors (Ize-Ludlow & Sperling, 2005). Nowadays diabetes is an important cause of death worldwide. Indeed, in 2021, 6.7 million died from diabetes; 1 person every 5 seconds (International Diabetes Federation, 2021). In addition, World Health Organization (WHO) have stated that diabetes is the main cause of blindness, kidney failure, heart attacks, stroke and lower limb amputation with a global incidence dramatically increasing with the time in countries of all income levels (International Diabetes Federation, 2021; World Health Organization, 2016).

1.1 Epidemiology of diabetes

In 2021, the International Diabetes Federation (IDF) estimated 537 million adults (20-79 years) were living with diabetes worldwide, meaning 1 diabetic individual out of 10 people (International Diabetes Federation, 2021). The number of cases has been increasing over the past few decades, and the IDF predicts that 783 million people will have diabetes by 2045, an increase of 46% in approximately 20 years. However, depending on the country, diabetes prevalence is different. A worldwide population analysis demonstrated that the type 2 diabetes prevalence is doubled in poor and low-income countries, when compared with high-income areas; potentially linked with the limited access to healthy food and adequate healthcare (Beckles & Chou, 2016; Rabi et al., 2006). Additionally, around half of the people living in poor areas are not diagnosed and are less likely to get a treatment early enough, resulting in adverse outcomes (International Diabetes Federation, 2021). With respect to type 1 diabetes (T1D), both the incidence and the prevalence are slightly increased in high-income populations (X. Lin et al., 2020). Although the association between high socio-economic lifestyle and

T1D prevalence is not fully described, some hypothesis suggest that the improvement of the hygiene and the subsequent low rates of infection in the childhood leads to a weaker immune system and more propensity to autoimmune complications (Gale, 2002; Gomez-Lopera et al., 2019). Finally, regarding monogenic types of diabetes, their prevalence is not correlated with the lifestyle or economy; but an increased prevalence of the maturity onset diabetes of the young (MODY) is found where consanguineous marriages are more usual (generally in low-income countries) (Riddle et al., 2020).

1.2 Classification of diabetes types

Hyperglycemia is a common feature of all types of diabetes but they can be classified according to their origin and etiopathogenesis. The main classification however, is based on the genetics of the disease: Monogenic and polygenic diabetes.

1.2.1 Monogenic diabetes

Monogenic diabetes represent around 5% of all cases of diabetes and is developed as the result of specific mutations in a single gene. In monogenic diabetes, highly penetrant variants cause diabetes regardless of other risk factors (H. Zhang et al., 2021). Most of the monogenic diabetes forms are inherited from a parent, as they are transmitted in an autosomal dominant pattern; however, sometimes the mutation can appear *de novo* (Misra & Owen, 2018; Salzano et al., 2019).

The molecular mechanisms underlying the development of this pathology are related to a diversity of intracellular processes that are essential for pancreatic β cell development and function. For example, alterations in the expression of genes related to insulin synthesis and secretion, defects in glucose level detection, increase in endoplasmic reticulum stress or defects in β cell differentiation (Murphy et al., 2008; Porter & Barrett, 2005). Monogenic forms of diabetes are normally flourished at early ages (people younger than 25 years) (Murphy et al., 2008; Sperling & Garg, 2018; Tattersall, 1974), although it can be classified into two subcategories depending on the age of diagnosis:

- **Neonatal diabetes (ND)** is a rare form of diabetes which is diagnosed in the first 6 months of life (American Diabetes Association, 2019; Beltrand et al., 2020; Busiah et al., 2013). Infants with ND are not capable of producing insulin leading to increased blood glucose levels. About half of the kids suffering from ND continue with the condition lifelong (permanent neonatal diabetes, PND) (American Diabetes Association, 2019; Kanakatti Shankar et al., 2013). In the resting 50% of the ND cases, the disease is temporary and disappears during the first years of life (transient neonatal diabetes, TND) (Beltrand et al., 2020; Busiah et al., 2013).

  Neonatal diabetes has been linked to several mutations in different genes. The 70% of transient neonatal diabetes cases are linked to genetic and epigenetic anomalies in chromosome 6q24 (Docherty et al., 2013; Temple et al., 2000). On the other hand, mutations in the *ABCC8, KCNJ11*, and *INS* genes are linked to both, permanent and transient forms of ND (Babenko et al., 2006; Beltrand et al., 2020; Bonnefond et al., 2011; Støy et al., 2007).

- **Maturity onset diabetes of the young** (MODY) appears later on life, typically before the age of 25 years (Thanabalasingham & Owen, 2011; Urakami, 2019). Currently, mutations in 13 genes have been linked to different forms of MODY. While mutations in *HNF4A* lead to the development of MODY1; the gene *GCK* is responsible of MODY2; alterations in *HNF1A* trigger MODY3; MODY4 is linked to *PDX1* gene modifications; *HNF1B* generates MODY5; and MODY6 is produced by defects in *NEUROD1* (Urakami, 2019). These alterations can be generated *de novo* or inherited from one of the progenitors (Thanabalasingham & Owen, 2011). Mutations in *HNF1A, GCK* and *HNF4A* genes are responsible for most of the MODY cases (Urakami, 2019).

1.2.2 Polygenic diabetes

In contrast to monogenic diabetes, the development of polygenic diabetes is linked to the presence of different risk genetic variants in multiples genes. In addition, the development of this type of diabetes depends on the exposure to environmental

factors that still need to be fully identified. Polygenic diabetes can be classified into two categories depending on the etiopathogenesis of the disease:

- **Type 1 diabetes** (T1D) often appears early in life and is triggered by the autoimmune destruction of pancreatic β cells (Atkinson & Maclaren, 1994). The destruction of pancreatic β cells has a multifactorial origin, and the interplay between genetic susceptibility and predisposing environmental factors result in the development of insulitis (pancreatic β cell inflammation) and the activation of the autoimmune assault against pancreatic β cells. Once the first symptoms are detectable, around 80-90% of the total β cell mass has already been destroyed (Lam et al., 2017; Meier, 2008). The deficiency of β cells and the total loss of insulin production, obligate patients to take exogenous insulin for life (Holt et al., 2021). Insulin demand is not the only complication for patients since a bad control of glucose levels in type 1 diabetes implies the development of long-term difficulties, such as retinopathy, diabetic foot, nephropathy, neuropathy and cardiovascular complications (Ciężki et al., 2022; Kern et al., 2022; Yapanis et al., 2022).

- **Type 2 diabetes** (T2D) is more common than T1D and it is typically caused by both genetic and environmental factors that affect β cell function and insulin sensitivity (Scheen, 2003). Until recently, its onset was typically in the adulthood, but its frequency is now arising among young people and children (World Health Organization, 2016).

  Type 2 diabetes development depends on the interaction between genetic factors and other elements, such as aging, sedentary lifestyle and unhealthy diet (Demir et al., 2021; Udler, 2019; Zheng et al., 2018). Normally, individuals with T2D still produce a small amount of insulin, although it might not be functional enough. Additionally, T2D pathogenesis is also linked to insulin resistance or impaired insulin sensitivity in peripheral tissues, such as muscle, brain and liver (Batista et al., 2021; Rubio Cabezas & Argente, 2012; Scheen, 2003).

In type 2 diabetes, initially β cells are functional and secrete insulin upon high blood glucose (Eizirik et al., 2020). Over time, insulin-sensitive peripheral tissues develop insulin resistance and do not respond to insulin normally (Schofield & Sutherland, 2012; Stumvoll et al., 2005). As a result, the pancreatic β cells increase insulin synthesis and secretion to overcome the increasing blood glucose levels, but are incapable to solve the glucose increase, and glucose is concentrated in the plasma of the patient generating hyperinsulinemia, which exacerbates β cell malfunction (Demir et al., 2021; Eizirik et al., 2020).

2 Pathogenesis of Type 1 Diabetes

There is a global concern about type 1 diabetes as its prevalence is rapidly increasing, especially in resource-limited areas (Beckles & Chou, 2016; Rabi et al., 2006). Indeed, half a million of newly cases were diagnosed worldwide in 2021, and for 2040, it is predicted that the number of T1D cases will be doubled, reaching up to 15 million people worldwide (Gregory et al., 2022).

In order to prevent or even cure this disease, a better knowledge of the factors that lead to the activation of the autoimmune attack against pancreatic β cells is required. In this sense, this thesis is focused on the characterization of the molecular mechanisms that lead to pancreatic β cell inflammation in T1D, and thus, in the following section a description of T1D pathogenesis will be exposed.

2.1 Pathogenesis of T1D

Even though all patients with T1D evidence pancreatic β cell deficit and dysregulated blood glucose levels; the triggering factor in each patient is different, and so it is the development of the disease. Type 1 diabetes is a progressive disease, which means that the clinical symptoms are gradually aggravated as the disease progresses. As a result, a classification system for T1D is designed to categorize the patients depending on their pathophysiological phase: pre-clinical T1D, immune system activation and the clinical stages (from stage 1 to stage 3) (**Figure 1**).

2.1.1 Pre-clinical T1D

Although T1D is a highly heterogeneous condition, the vast majority of cases undergo a period of time (longer or shorter) called pre-clinical T1D in which the disease is "clinically silent". The meaning of "clinically silent" is that clinical manifestations are not yet detectable as the disease is starting to develop and a significant proportion of β cell mass is still functional. Thus, the characteristics of this pre-clinical stage are perfect for the implementation of clinical interventions to delay or arrest pancreatic β cell destruction, and eventually, T1D development. Type 1 diabetes is considered as a complex disease, which means that a single risk factor is not enough to develop the condition. Indeed, T1D is developed upon the interaction of genetic factors and environmental factors, which act as triggers of the autoimmune activation.

2.1.2 Immune system activation

In genetically susceptible individuals, environmental factors activate the immune system to specifically attack the pancreatic β cells. Pancreatic β cell autoantigens are the targets of the immune system (Yoon & Jun, 2005). Autoantigens are processes by the macrophages, dendritic cells or B cells in the pancreatic islets and presented by the MHC class I molecule. Once autoantigens are recognized by the autoreactive T cells, autoantibodies are generated. These autoantibodies include insulin, GAD, IA-2 and ZnT8 (S. Han et al., 2013; Roep et al., 2021). Autoantibodies are not believed to have a pathogenic role against β cells, but the risk of developing T1D is strongly related to the number of autoantibody markers (Yoon & Jun, 2005). The presence of two or more autoantibodies implies higher probability of developing T1D than the presence of a single one. Indeed, in a high-risk birth cohort studied in 2013, only the 12.7% of people with one autoantibody developed T1D-associated clinical symptoms 15 years after autoantibody detection. In the case of children positive for 2 antibodies, 61.6% developed clinical symptoms, and for those positive for three autoantibodies, 79.1% developed symptoms (Ziegler et al., 2013).

2.1.3 Clinical onset of T1D

The initial stage of the clinical onset is characterized by the presence of two or more autoantibodies plus the development of glucose intolerance or dysglycemia that is caused due to the β cell loss (Insel et al., 2015). At this point, blood glucose levels start to fluctuate and there is an accelerated diminish on insulin response and glucose tolerance (Sosenko et al., 2010). At this stage, a gradual and prolonged deterioration of β cell function happens, leading to a total loss of β cell mass. The dysregulation of blood glucose levels leads to the development of common symptoms that include polyuria (extreme urination), polydispsia (extreme thirst), unexpected bodyweight loss, blurred vision and fatigue (Insel et al., 2015; Yapanis et al., 2022). Moreover, during clinical onset, diabetic ketoacidosis and cardiovascular problem also arise in most of diabetic patients (Howard et al., 2021).

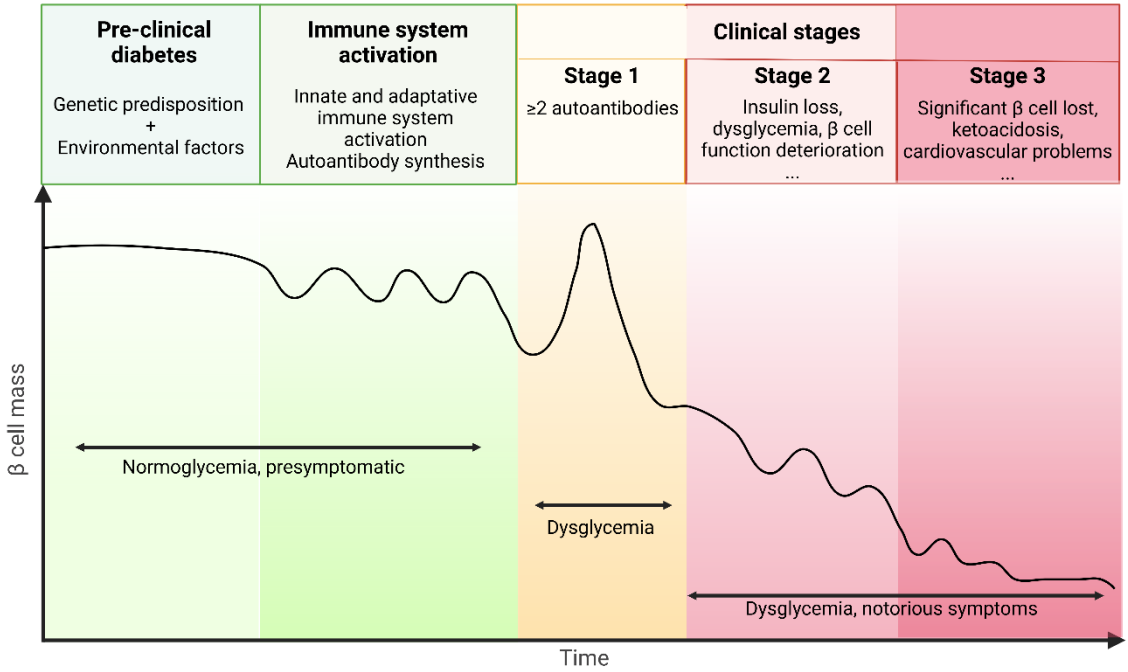


Figure 1. Stages of T1D development. The preclinical diabetes starts when genetically predisposed individuals are exposed to a triggering factor. As a result, inflammation starts and the immune system is activated. Autoantibodies against islets autoantigens are synthetized but the blood glucose levels remain normal and there are no evident symptoms of the disease. When autoreactive T cells start to infiltrate the islets, glucose levels start to dysregulate because of the inability of the β cells to produce enough insulin. Stage 2 starts when T1D is diagnosed in the majority of the patients, is characterized by the decline in β cell mass, insulin insufficiency and the impossibility to control glucose levels. As the disease progresses, complications like ketoacidosis or cardiovascular problems are more frequent.

2.2 Etiology of type 1 diabetes

2.2.1 Genetic factors

As explained earlier, T1D is a polygenic disease, meaning that genetic variants in different genes contribute to the development of the condition (Steck & Rewers, 2011). The genetic risk for people without any diabetic family history is about 1 in 300, whereas for those with diabetic relatives is 15 times higher, 1 out of 20 (Wherrett, 2021).

According to the Genome Wide Association Study (GWAS) catalog, there are more than 60 susceptibility regions across the human genome (Klak et al., 2020; Ram et al., 2016). Even if there are so many genes associated with type 1 diabetes predisposition, approximately 50% of the genetic risk for type 1 diabetes is attributable to the Human Leukocyte Antigen (HLA) region (Grant et al., 2020; Primavera et al., 2020; Steck et al., 2005).

2.2.1.1 *HLA genes*

The first susceptible genes associated with T1D were reported in the '70s (Cudworth & Woodrow, 1975; Singal & Blajchman, 1973). These genes are named *Human Leukocyte Antigen* (*HLA*) and are located within the Mayor Histocompatibility Complex (MHC) in chromosome 6q21, a region characterized by a strong linkage disequilibrium (LD). As a consequence of the LD, genes in this genomic region are normally inherited as haplotype blocks.

The genes in the MHC complex are divided in 3 different genomic regions (**Figure 2**):

- **HLA class I**: a group of genes able to encode very polymorphic proteins that are attached to the surface of every nucleated cell (Choo, 2007; Wieczorek et al., 2017). Their function is to display peptides of protein fragments belonging to the cell itself to introduce them to the cytotoxic T cells ($CD8^+$) (Gibadullin et al., 2021; Marshall et al., 2018). Thus, cells with a MHC class I antigen are targeted for destruction by T cells (Choo, 2007; Gibadullin et al., 2021).

- **HLA class II**: the molecules synthesized from these genes are expressed primarily in the surface of antigen-presenting cells (APCs) (Holling et al., 2004; Roche & Furuta, 2015). The objective of these molecules is to present peptides derived from both extracellular proteins and self-proteins to helper T-cells ($CD4^+$) (Couture et al., 2019; Roche & Furuta, 2015). The exhibition of MHC class II antigens attracts immune cells and induces the production of specific antibodies (Couture et al., 2019; Hickey et al., 2016).
- **HLA class III**: genes encoding inflammation-related molecules as cytokines and heat shock proteins (Deakin et al., 2006; Srivastava et al., 2016).

All these molecules are essential for the organism to distinguish between self and foreign molecules (Shiina et al., 2009; Wieczorek et al., 2017). For the recognition of antigens implicated in T1D, it is necessary to generate a complex between the T cell receptors, the antigenic peptide and the HLA class II molecules (Couture et al., 2019; Morran et al., 2015; Noble & Erlich, 2012; Sundberg et al., 2007). For this reason, any allelic variation on the HLA may alter the antigen-binding groove and have an impact on antigen presentation and recognition (Morran et al., 2015; Noble & Erlich, 2012).

At the same time, HLA has been observed to be associated with other autoimmune diseases including celiac disease, systemic lupus erythematosus, multiple sclerosis and rheumatoid arthritis; highlighting the importance of HLA in immune regulation and in the development of autoimmune conditions (Degos & Dausset, 1974; Isomäki et al., 1974; Mulder, 1974; Suciu-foca et al., 1974).

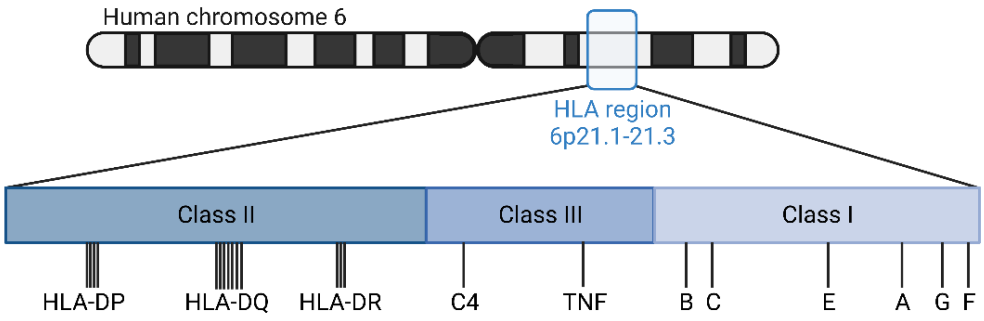


Figure 2. HLA genes distribution. The HLA region is located in the chromosome 6 of the human genome; particularly in the 6p21.1-21.3 domain. The genes conforming the HLA region are divided in 3 groups, the class I, the class II and the class III.

Alleles found on these genes strongly determine the susceptibility to T1D ranging from protection to high risk (Erlich et al., 2008; Nguyen et al., 2013; Santana del Pino et al., 2017). In fact, DRB1*03:01-DQA1*05:01-DQB1*02:01 (also known as DR3-DQ2) and DRB1*04-DQA1*03:01-DQB1*03:02 (also known as DR4-DQ8) alleles are the ones conferring the higher risk for T1D compared to the other possible genotypes (Aly et al., 2006; Schweiger et al., 2016; Vidan-Jeras, 2018). These risk alleles are present in around 40-50% of diagnosed diabetic patients (Redondo et al., 2006; Schweiger et al., 2016). These percentages depend on the prevalence of each haplotype in a given ethnic group. For example, the HLA-DR3/DQ2 and HLA-DR4/DQ8 haplotypes are very frequent in the Nordic countries, leading to increased genetic risk to develop T1D in countries like Sweden or Norway (Sanjeevi et al., 2008; Undlien et al., 1999). On the other hand, some haplotypes have also been described to diminish the risk to suffer from T1D, as it is the case for the haplotype DRB1*15:01-DQA1*01:02-DQB1*06:02 (also known as DR2) (Todd et al., 1987).

2.2.1.2 *Non-HLA genes*

Apart from the HLA genes, GWAS have described a list of multiple other genes potentially implicated in T1D. After the HLA region, the genes with the strongest association with T1D are the insulin gene (*INS*), the protein tyrosine phosphatase non-receptor type 22 (*PTPN22*), the Erb-B2 receptor tyrosine kinase 3 (*ERBB3*) and the Cytotoxic T-lymphocyte-associated antigen 4 (*CTLA4)* gene (Pociot et al., 2010; Redondo et al., 2018).

In 1982, it was described a region near the ***insulin*** gene with a variable number of tandem repeats (VNTR) of 14-15bp (Bell et al., 1982). According to the number of repeats, the risk for developing the disease is different. The highest risk is conferred by homozygosis for shortest repeats (26-63 repeats) (Durinovic-Belló et al., 2010). In fact, these shortest repeats appear in homozygosis in 75-85% of type 1 diabetic people, whereas the percentage is diminished to 50-60% among individuals from the general population (Undlien et al., 1995). Even though the exact mechanism by which the VNTR is implicated in T1D development is still unknown, it seems to affect the transcription

of the *insulin* and the insulin-like growth factor 2 (*IGF2)* genes (Bennett et al., 1995; Paquette et al., 1998). The VNTR is suggested to influence the structure of the chromatin, modulating the accessibility of transcription factors to the nearby genes, *INS* and *IGF2* (Paquette et al., 1998).

PTPN22 is located on chromosome 1p13 and encodes for a protein named lymphoid tyrosine phosphatase (LYP). LYP protein is in charge of inhibiting T cell activation by dephosphorylating the main 3 kinases involved in proximal T cell receptor (TCR) signaling (Castro-Sanchez et al., 2020; Stanford et al., 2012). Additionally, LYP also plays a role in innate immunity and inflammasome activation due to its dephosphorylating capacity (Spalinger et al., 2016; Y. Wang et al., 2013).

The ***ERBB3*** gene encodes a protein with the same name, which is expressed on the surface of $CD11c^+$ antigen presenting cells and participates in antigen presentation (H. Wang et al., 2010). This protein is able to interact with PI3K (Phosphoinositide 3-kinase) regulatory subunits, allowing the activation of the ERBB3-PI3K-Akt cascade (Song et al., 2015). Low ERBB3 levels can lead to lower Th cell activation and autoimmunity (Kaur et al., 2016; H. Wang et al., 2010). In addition, there are several studies supporting the implication of this gene in immune regulation and cytokine-induced pancreatic β cell apoptosis (Kaur et al., 2016; Lemos et al., 2018). An international cohort study called TEDDY (The Environmental Determinants of Diabetes in the Young), which is focused on unraveling the causes of T1D, has described an association between the rs2292239 SNP (single nucleotide polymorphism) located in the intron 7 of the *ERBB3* gene with T1D susceptibility (Krischer et al., 2017; Törn et al., 2015). *ERBB3* levels are significantly lower in the presence of the susceptible genotype conferring increased risk for autoimmunity, β cell apoptosis, and thus, T1D (D. Wang & Pan, 2019; H. Wang et al., 2010).

The protein encoded by the ***CTLA4*** gene is a vital molecule for a correct negative regulation of immune responses (Chikuma, 2017; Waterhouse et al., 1995). *CTLA4* is induced after T cell activation and binds to CD80/CD86 ligands with a stronger affinity than CD28, providing a competitive inhibition (Chikuma, 2017). CD28 binding to

CD80/CD86 leads to increased anti-apoptotic Bcl-xL protein synthesis (Boise et al., 1995) and stimulates glucose uptake by the activation of the glucose transport and glycolysis (Frauwirth et al., 2002). CTLA4 provides negative feedback on the CD28-mediated activation of polyclonal, self-reactive T cells (Chikuma, 2017). However, CTLA-4 deficiency can lead to CD28-mediated autoimmunity due to self-reactive T cell expansion.

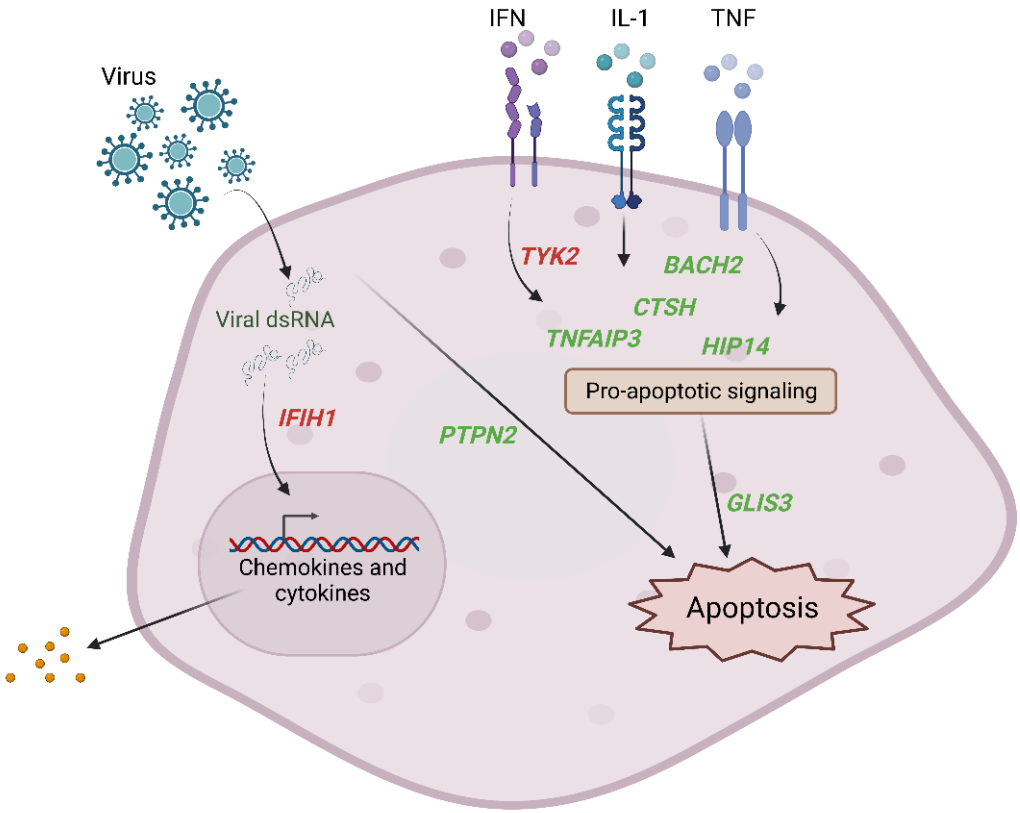


Figure 3. T1D candidate genes modulate β cell inflammation and apoptosis. After a viral infection, there are several genes implicated in the modulation of the pancreatic β cell response. *IFIH1* and *PTPN2* genes are directly activated by the viral infection. *IFIH1* regulates the chemokine and cytokine production and *PTPN2* is responsible of the apoptotic response. In response to the presence of cytokines, genes including *HIP14, TNFAIP3, TYK2, BACH2* and *CTSH* modulate the pro-apoptotic signaling transduction. *GLIS3* is in charge of the modulation the apoptosis mediated by this pro-apoptotic signals. Genes in green illustrate protective roles in the β cell function and red labels, deleterious functions.

Although candidate genes for T1D have been linked to T1D pathogenesis due to their impact in immune regulation, nowadays we know that many of the genes associated with T1D are expressed in pancreatic β cells and modulate the function of these cell (Santin & Eizirik, 2013). T1D candidate genes, including *IFIH1*, *GLIS3* and *PTPN2* genes, modulate the mechanisms leading to β cell dysfunction. While some of them (e.g. *IFIH1* and *PTPN2*) participate in virus-induced β cell inflammation, others (e.g. *GLIS3*) participate in basal β cell function (Colli et al., 2010; Nogueira et al., 2013; Santin et al., 2011) (**Figure 3**).

2.2.2 Environmental factors

Apart from genetic susceptibility, environmental factors are also responsible for the triggering of islet autoimmunity (Rewers & Ludvigsson, 2016). Very diverse effectors have been proposed as environmental triggers in diabetes, including drugs (Sheen et al., 2016), statins (Fève & Scheen, 2022), nitrites (Muntoni et al., 2006), cow's milk (Virtanen et al., 2000), high quantities of polyunsaturated fatty acids (Norris et al., 2007), birthweight (Harder et al., 2009), gluten (Beyerlein et al., 2014), deficient vitamin D (Zipitis & Akobeng, 2008), gut microbiota (Wen et al., 2008) and viral infections (Nekoua et al., 2022; Richardson et al., 2013; Simonsen et al., 2015). Viral infections are the most promising candidate factors, since accumulating clinical and epidemiological evidence suggest their implication in T1D development.

2.2.2.1 Viral infections

Several epidemiological studies have described the potential involvement of enteroviruses, and especially coxsackievirus B (CVB) subtype in the pathogenesis of T1D (Coppieters et al., 2011; Stene & Rewers, 2011). A pooled analysis performed in 2021 consisting of 5921 subjects from all over the world, demonstrated that enterovirus infections are associated with T1D, with almost 8-fold increase in risk when comparing with the controls (e.g. non infected people) (K. Wang et al., 2021).

CVB are small enteroviruses with the ability to directly infect their target cells and lyse them (as cytolytic viruses), but also capable to settle a persistent infection for several months (See & Tilles, 1995). Indeed, signs of persistent CVB infection have been

detected in pancreases of type 1 diabetic donors (Dotta et al., 2007; Krogvold et al., 2015; Morgan & Richardson, 2014; Nigi et al., 2022; Richardson et al., 2009). In particular, there is a strong association between persistent infection of CVB virus and the development of islet autoimmunity in genetically predisposed children (Honkanen et al., 2017; K. W. Kim et al., 2019).

Although the complete virus is not always detectable, markers of enterovirus infections are usually found in pancreata and blood of T1D patients (viral proteins, RNA or specific antibodies). Multiple evidences confirm the presence of enteroviral RNA and anti-enterovirus antibodies in blood serum samples from children genetically predisposed to T1D (Laitinen et al., 2014; S. Oikarinen et al., 2014). Moreover, further analysis have demonstrated there are signs of enteroviral RNA and proteins (in particular the major capsid protein VP1) on pancreas and gut biopsies from newly diagnosed patients (Krogvold et al., 2015; M. Oikarinen et al., 2008; S. Oikarinen et al., 2021).

2.3 Molecular mechanisms of viral infection in pancreatic β cells and their relation with β cell damage

The process mediated by viral infections in pancreatic β cells that might result in T1D development can be divided in three main phases (**Figure 4**):

1. **Induction**: The pancreatic β cells recognize the invading microorganism and activate both, the innate and the adaptive immune system.
2. **Amplification**: In this stage, there is a "dialog" between β cells and the immune system (mainly mediated by chemokines and cytokines). This interaction triggers a cascade of intracellular signaling pathways in β cells.
3. **Resolution**: Finally, in non-susceptible people, the pancreatic β cells manage to eliminate any residue of the viral infection thanks to a balanced antiviral response. However, in people susceptible to T1D, this response is uncontrolled resulting in an excessive inflammation named insulitis. Consequently, there is a progressive death of the pancreatic β cells.

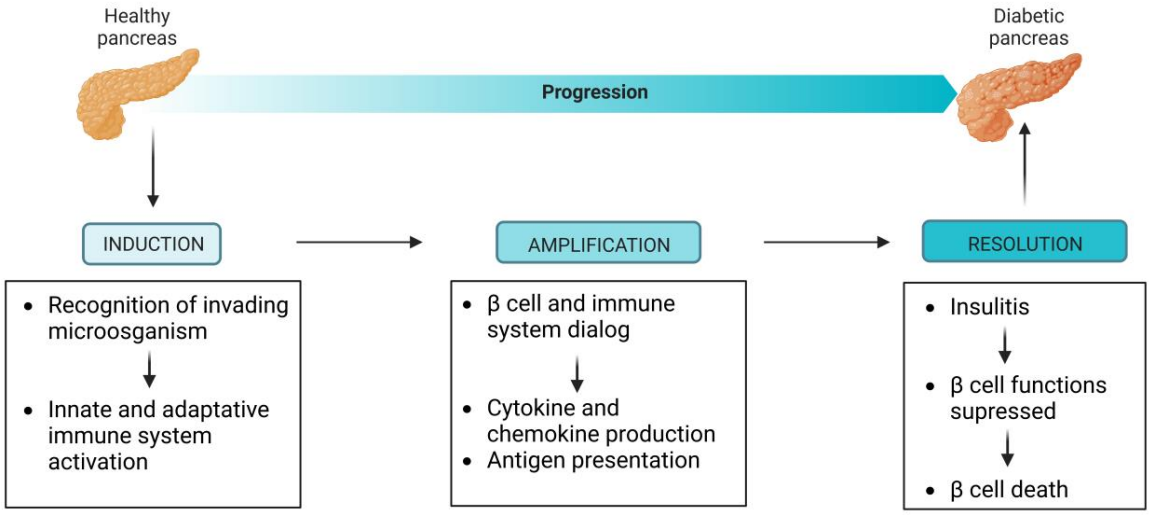


Figure 4. Stages of the progression from a healthy pancreas to a diabetic pancreas after a viral infection.

In detail (and illustrated in **Figure 5**), when a virus infects pancreatic β cells, the viral dsRNA (double stranded RNA, a by-product of viral infections) activate the Toll-like receptor 3 (TLR3) and in some cases, the cytoplasmic viral receptors called retinoic acid-inducible gene-I (RIG-I) and the melanoma differentiation-associated gene 5 (MDA-5) (Assmann et al., 2015; Dixit & Kagan, 2013; Hernández et al., 2007; J. P. Wang et al., 2010). The activation of these receptors leads to the activation of a complex molecular antiviral response, in which a critical step is the activation of several transcription factors (e.g. NFκB, IRF7 and STAT1) that promote the production of type 1 interferons, such as, IFNα and IFNβ (Allred et al., 2021; Goubau et al., 2013). These interferons are key factor in the whole process, as they are the major regulators of the antiviral and inflammatory response, acting both in a paracrine and autocrine way. **Insulitis** is the name adopted for the resulting pancreatic islet inflammation (Op de beeck & Eizirik, 2016; Pugliese, 2017; Yoon et al., 1977). The produced type I interferons are released outside the pancreatic β cell. IFNAR (from IFN-α/β receptor), a receptor located in the cell surface of β cells is able to recognize the IFNα and IFNβ molecules and via TYK2 (tyrosine kinase 2) and JAK1 (Janus kinase 1), promotes the activation of JNK1 (c-Jun N-terminal kinase 1). JNK1 activates both the pro-apoptotic protein BIM (Bcl-2 Interacting Mediator of cell death), and its phosphorylated form (P-BIM), inducing the intrinsic apoptotic pathway.

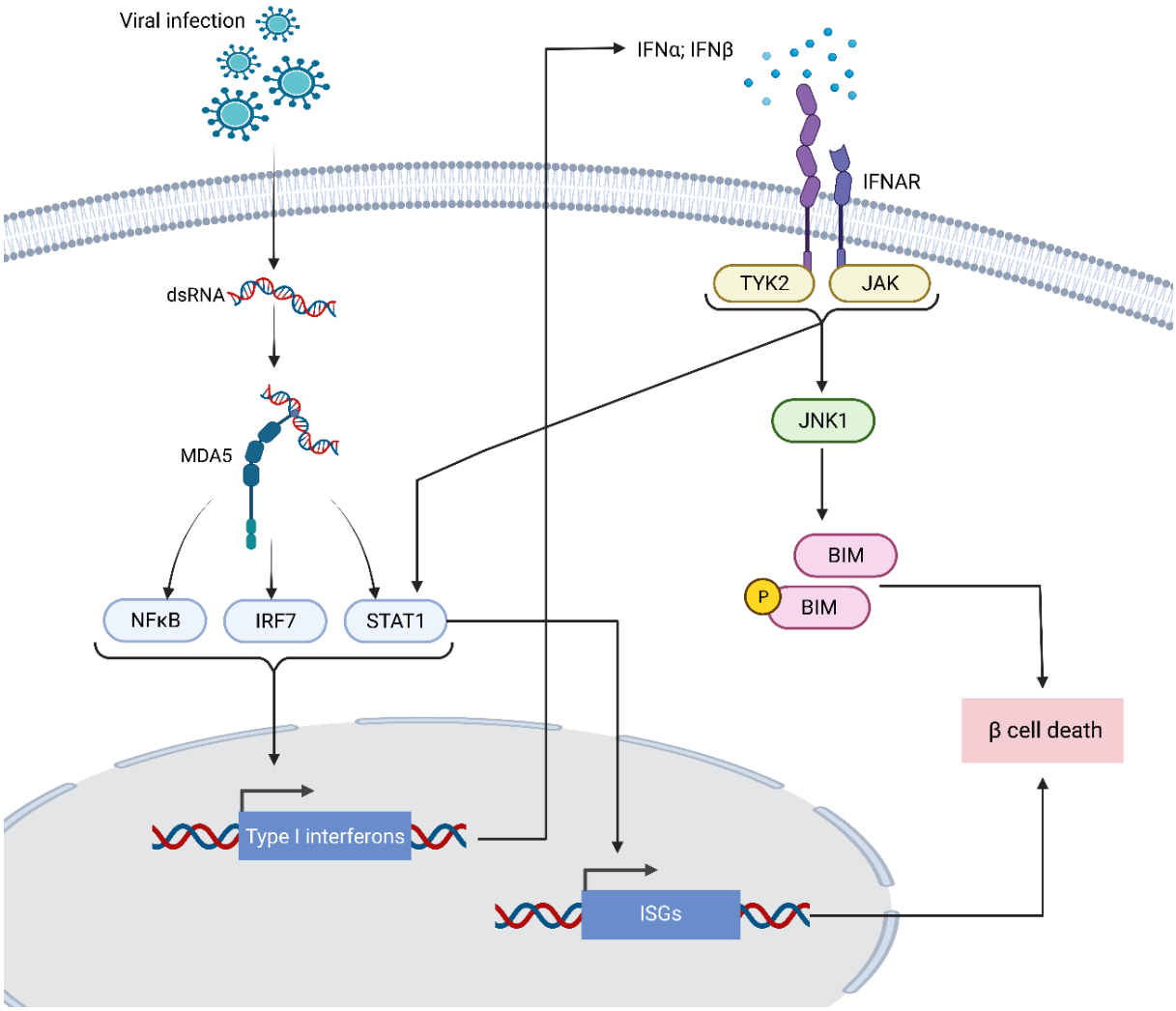


Figure 5. Illustration of the signaling cascade generated in pancreatic β cells after a viral infection.

Additionally, the pro-inflammatory recognition in the cell surface also triggers the activation of STAT1 (Signal Transducer and Activator of Transcription 1) (Colli et al., 2020; Fabbri et al., 2019). STAT1 is in charge of the transcriptional activation of a multitude of apoptotic, antiviral and pro-inflammatory genes. STAT1 activation provokes not only the synthesis of type I IFNs but also helps in the transcriptional activation of interferon stimulated genes (ISGs) (Akhbari et al., 2020; Mostafavi et al., 2016). ISG production promote inflammation, endoplasmic reticulum (ER) stress and unfolded protein response (UPR), which can finally lead to drastic circumstances like cell death (Akhbari et al., 2020; Eizirik et al., 2009).

3 Long non-coding RNAs

The sequencing of the human genome and the development of new RNA sequencing techniques have revealed that while over 85% of the genome is transcribed, only a small portion of RNAs (<2%) encode protein-coding genes (Hangauer et al., 2013; The ENCODE Project Consortium, 2007). Huge efforts are made attempting to characterize every coding and non-coding element in the human genome, which are compiled in freely available big databases and repositories (Derrien et al., 2012; The FANTOM Consortium et al., 2014). Non-coding RNAs display a big family of genes unable to produce any protein, and **long non-coding RNAs** (lncRNAs) are one of the most representative subcategories of this family.

LncRNAs are defined as RNA transcripts longer than 200 nucleotides lacking any protein-coding ability (Derrien et al., 2012; Hangauer et al., 2013; Mercer et al., 2009). In the last version of GENCODE (GENCODE Release 41) there are 19,095 lncRNAs annotated (similar number compared to the annotated 19,370 protein-coding genes). Annotation and characterization of lncRNAs is arising in the 21st century but there is still little information on their function. Up to day, the characterization of lncRNAs have revealed their relation with several crucial cellular and molecular processes, including cell differentiation (Rosa & Ballarino, 2016), cell cycle (Ghafouri-Fard et al., 2020; Kitagawa et al., 2013) and metabolism (W. Lin et al., 2020; Lu et al., 2019).

The ability of lncRNAs to participate in so different biological processes resides in their capacity to interact with proteins, RNAs, DNA or even a combination of these macromolecules (Kazimierczyk et al., 2020; Kuo et al., 2019; Peng et al., 2020; H. Zhang et al., 2019). The versatility of lncRNAs is also linked to their cellular localization, as they can be located anywhere in the cell (Bridges et al., 2021). It is interesting to emphasize that the determination of the subcellular localization of a lncRNA is essential for the understanding of its function.

While lncRNAs are expressed in lower levels compared to mRNAs, they exhibit a strong tissue-, cell- and specie-specificity (Derrien et al., 2012; Fernandes et al., 2019). The lack of conservation among species also hinders their characterization, due to the

impossibility to compare the same lncRNA in different animal models. Moreover, lncRNA sequence conservation is not always translated into conserved functional roles, subcellular distribution or molecular interactions (Bridges et al., 2021; M. Chen et al., 2016).

3.1 Subcellular localization and function of lncRNAs

Similar to proteins, the function of lncRNAs is linked to their subcellular localization.

3.1.1 Nuclear and chromatin bound lncRNAs

The roles of lncRNAs localized in the nuclei are generally linked to transcriptional modulation and chromatin interaction and remodeling (P. Han & Chang, 2015; Statello et al., 2021). Indeed, gene expression can be altered by the interaction of lncRNAs with chromatin-modifying enzymes. LncRNAs can act as kidnappers or helpers; directly binding and sequestering chromatin modifiers or, attaching the chromatin-modifying enzymes serving as a guide to carry these enzymes to their target genomic region (P. Han & Chang, 2015; Saxena & Carninci, 2011). Long non-coding RNAs can also be part of a chromatin complex playing a role on chromatin modification or acting as a scaffold to help in the assembly of the required complex.

In summary, nuclear lncRNAs are able to function in very different ways, but can be categorized in four subtypes (**Figure 6**):

- **Transcriptional modulation:** These lncRNAs can serve as co-factors resulting in the modification of the transcriptional activation of a target gene (Bhat et al., 2016; Faghihi et al., 2010; Hall et al., 2015). The lncRNA itself is able to induce or repress the transcription of the targeted gene. Their adaptability in response to certain stimuli is essential for a finely regulated transcriptional regulation (X. Ji et al., 2022).
- **Guide**: Essential for a proper localization of factors at specific genomic loci. These lncRNAs bind to regulatory active proteins, such as transcription factors and chromatin modifiers, to direct them to a precise genomic location (Balas & Johnson, 2018; Hung & Chang, 2010; Lee, 2009). The mechanism of action of

this kind of lncRNAs is to recruit the protein of interest and guide it to the target genomic position (Bhat et al., 2016). Flexibility and adaptability is what made lncRNAs interesting for molecule guiding. In addition, the ability to bind distinct molecules allow guide long non-coding RNAs to interact and address other molecules to their targeted domains (Bonasio et al., 2010; Hung & Chang, 2010; Lee, 2009).

- **Decoy**: Contrary to the guide lncRNAs, there are some other lncRNAs acting as decoys (Guenther et al., 2007). The main function is to bind their target molecule and clear it away from its location; thus, acting as negative regulatory factors (Balas & Johnson, 2018). Normally, decoy lncRNAs interact with transcription factors and chromatin modifiers blocking their communication with the chromatin (Aguilo et al., 2011; Kotake et al., 2011). Colloquially, these lncRNAs "kidnap" the objective molecule avoiding their function.
- **Scaffold**: Thanks to the capacity of lncRNAs to interact with proteins, DNA and RNA; scaffold lncRNAs are crucial in the assembly of complexes (Spitale et al., 2011; Zhao et al., 2008). These transcripts play a structural role by providing a platform for the assembly of multiple molecules. Scaffold lncRNAs can bring several components together, guiding them to enhance their function (Spitale et al., 2011). To accomplish this purpose, scaffold lncRNAs present different domains to interact with several molecules simultaneously (Spitale et al., 2011; Zhao et al., 2008).

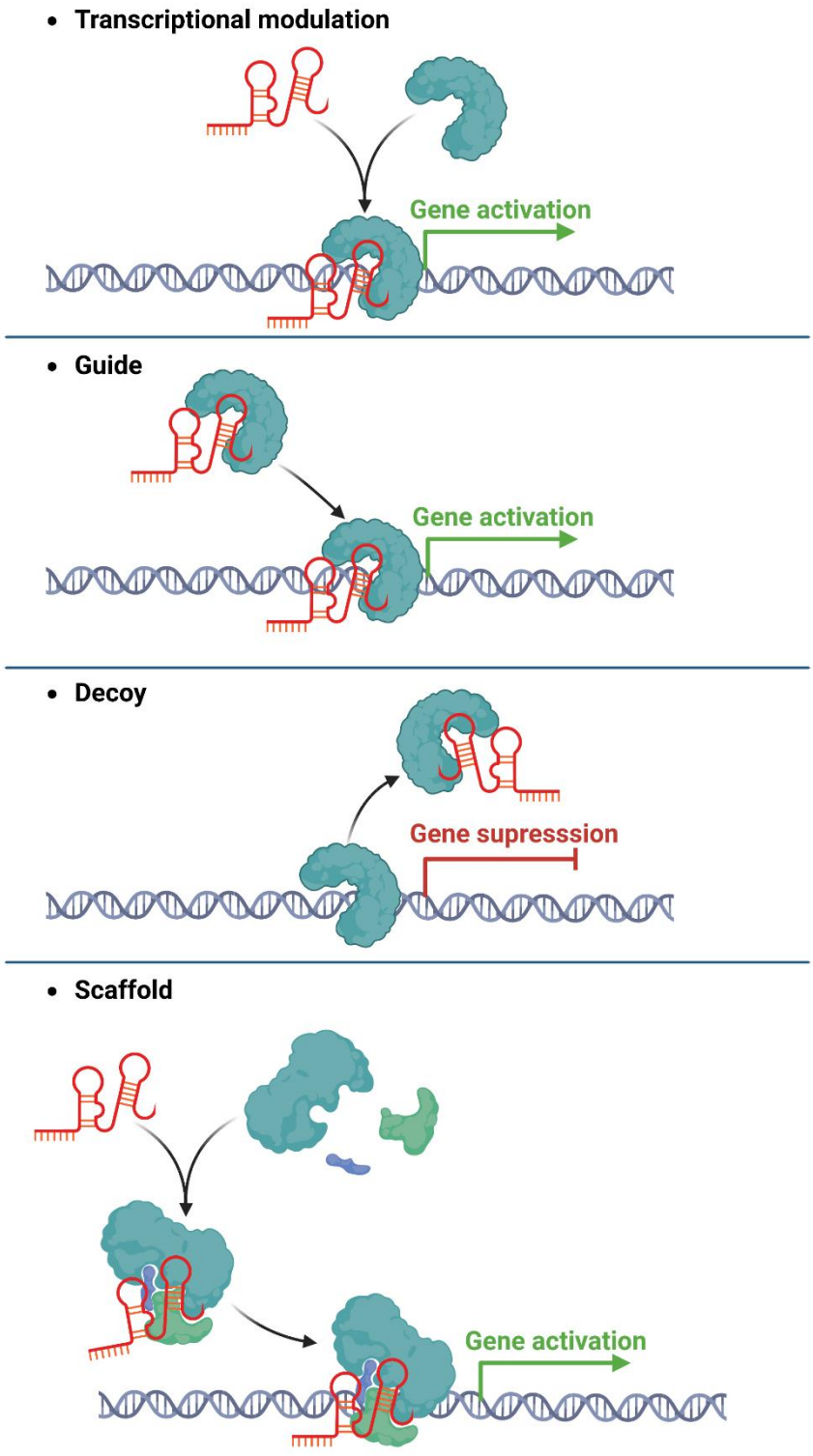


Figure 6. The mechanisms of action of nuclear lncRNAs.

3.1.2 Cytoplasmic lncRNAs

On the other hand, cytoplasmic lncRNAs are less understood compared to nuclear ones. However, cytoplasmic RNAs can also govern essential functions to maintain cellular structure, function and homeostasis (**Figure 7**). The main functions of cytoplasmic lncRNAs are the following:

- **mRNA stability**: One of the most studied function of cytoplasmic lncRNAs is their ability to control the stability of mRNA molecules. In the cytoplasm, lncRNAs can regulate mRNA stability by a number of different mechanisms (Yu & Wang, 2018). One mechanism of action is the direct interaction of the lncRNA with the mRNA (Noh et al., 2018; X. Wu et al., 2017). Normally an antisense lncRNA binds the sense mRNA, generating a duplex, and increasing its stability (Faghihi et al., 2008; Yu & Wang, 2018). Additionally, lncRNAs in combination with RNA binding proteins (RBPs) are also able to modulate the stability of mRNAs (Sebastian-delacruz et al., 2021; Xin et al., 2011).
- **Translation modulation:** Some lncRNAs are involved in the modulation of protein translation, either repressing (Hall et al., 2015) or promoting it (Carrieri et al., 2012). The most studied procedure for translation modulation resides in the interaction between lncRNAs and translation initiating or blocking factors (Rashid et al., 2016). The interaction can allow either the recruitment or degradation of these translational factors, or their guiding into the target mRNA (Noh et al., 2018; Yu & Wang, 2018).
- **Competing endogenous lncRNAs**: Both coding and non-coding RNAs have predisposition to interact with microRNAs (miRNAs) (Salmena et al., 2011). The main function of miRNAs is to interact with mRNAs, blocking its translation or even leading to its degradation (Rashid et al., 2016). Several lncRNAs harbor multiple binding sites for miRNAs and are named as competing endogenous RNAs (ceRNAs) (Noh et al., 2018; Yu & Wang, 2018). CeRNAs can competitively bind the target mRNA and alter the effect of the miRNA on its expression (Gao et al., 2020).

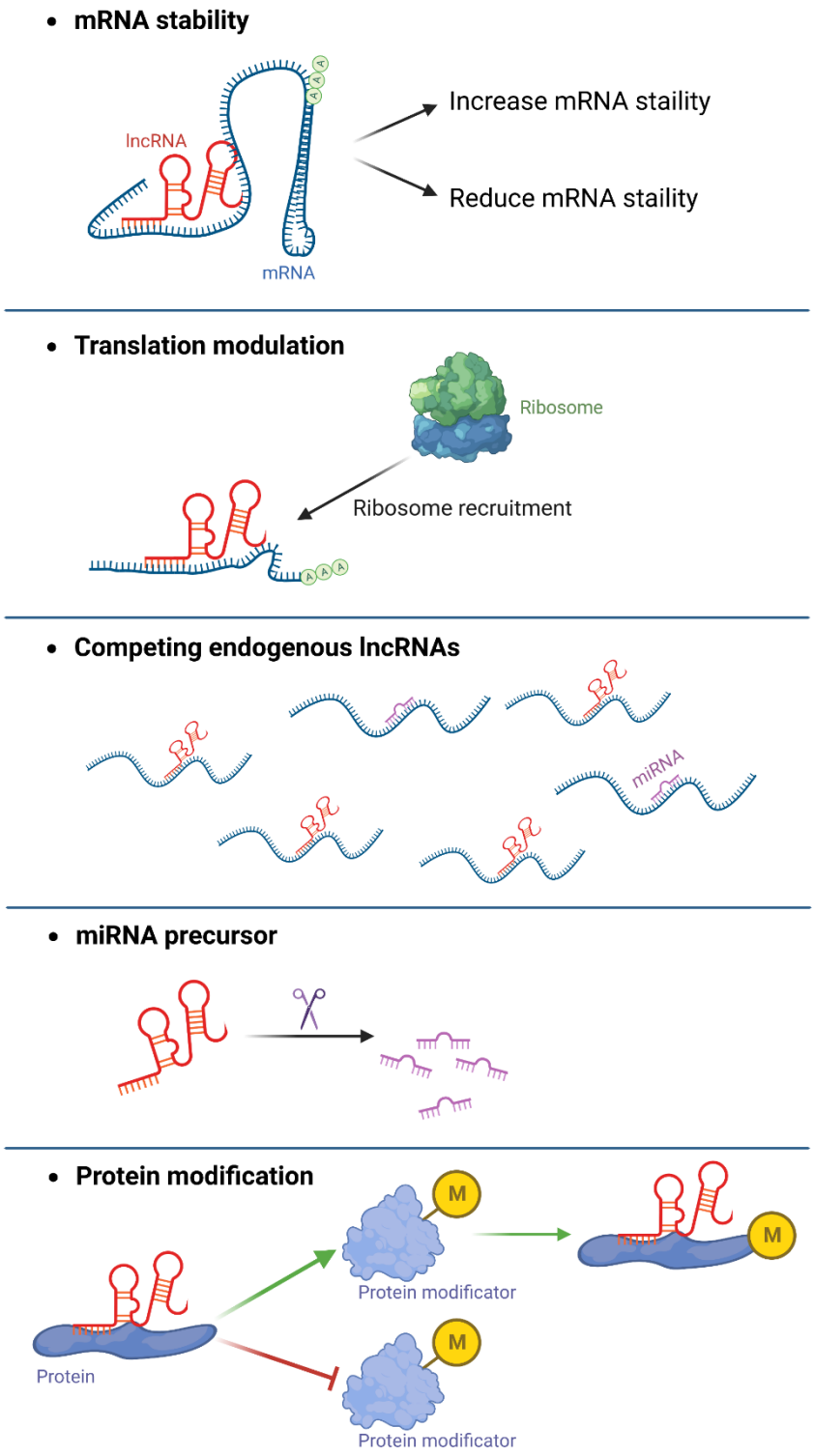


Figure 7. Mechanisms of action of cytoplasmic lncRNAs.

- **miRNA precursor**: lncRNAs can function as precursors of miRNA molecules. In 2008, Shunmin He and colleagues predicted around 100 lncRNAs with ability to encode for a miRNA (S. He et al., 2008). The base of this regulatory process is

simple; lncRNA are cut into smaller pieces in a delicate manner to form miRNAs (Dey et al., 2014). LncRNAs of this kind cannot be defined as purely cytoplasmic as they are processed in both, the nucleus and the cytoplasm to give rise to functional miRNAs (Lanzillotti et al., 2021).

- **Protein modification:** Finally, in the most recent years efforts are made to describe lncRNAs modulating protein modifications (B. Liu et al., 2015; Tan et al., 2021), such as ubiquitination/deubiquitination, acetylation/deacetylation or phosphorylation/dephosphorylation (K. Zhang et al., 2019). Although it is still unknown how lncRNAs control protein modifications; what it is proposed is that in most cases, lncRNAs directly interact with the target protein resulting as a signature to attract protein modifiers (e.g. phosphorylases, ubiquitinases or acetylases) (Su et al., 2022).

3.2 LncRNAs in β cell function

During the last decade, whole transcriptome studies performed in pancreatic islets have revealed the presence of thousands of lncRNAs particularly expressed in β cells (Moran et al., 2012). A comparison in RNA expression of pancreatic islets with other 18 human tissue samples revealed that only 9.4% of the coding genes were specific of pancreatic islets, whereas the percentage of islet-specific lncRNAs was higher than 50% (Moran et al., 2012). In the same study, it was described that islet-specific lncRNAs are dynamically regulated by glucose alterations, suggesting an implication in β cell function (Moran et al., 2012). Other study identified 132 lncRNAs that were upregulated in β cells compared to the whole islet, suggesting a specific role of these lncRNAs in β cells (Nica et al., 2013).

The expression level of most lncRNAs is very low (Derrien et al., 2012; Mercer et al., 2012; Seiler et al., 2017). Nevertheless, their expression changes upon the exposure to different stimuli such as pro-inflammatory cytokines, glucose, specific drugs or under pathophysiological conditions (Fadista et al., 2014; Kutty et al., 2017; Lu et al., 2018; M. M. Perry et al., 2014; Sun et al., 2016). Regarding their regulation in pancreatic islets

and diabetic conditions, a whole RNA sequencing was performed in pancreatic islet of mice fed with a high-fat diet (simulating a model of mild diabetes associated with obesity) and a set of more than 1,500 novel lncRNAs was discovered (Motterle et al., 2017). Additionally, it was observed that 23 lncRNAs were differentially upregulated and 104 downregulated in high-fat diet mice compared with mice fed with a regular diet. This study demonstrated that the diet can alter the expression of lncRNAs in pancreatic islets, suggesting their potential implication in lipotoxicity-induced β cell dysfunction and diabetes development (Motterle et al., 2017).

The main function of pancreatic β cells is to synthetize and secrete insulin to regulate blood glucose levels. Although it is widely described that the expression of the insulin gene is regulated by several transcription factors, during the last years islet-specific lncRNAs have emerge as novel regulator of insulin modulation (Melloul et al., 2002; Singer et al., 2015). The next paragraphs describe examples of lncRNAs closely related to the pancreatic β cell function.

- ***MEG3* (Maternally expressed gene 3):** This imprinted lncRNA is transcribed from the maternal allele located on human chromosome 14 (its homolog in mice resides on chromosome 12). Its expression is quite ubiquous but it is lost in many human tumors (X. Zhang et al., 2010; Y. Zhou et al., 2012). In terms of pancreatic islets, *MEG3* is 20 times more expressed in human pancreatic β cells than in glucagon-producing α cells (Dorrell et al., 2011), suggesting a specific implication in β cell function. *MEG3* expression is modulated by glucose and its inhibition by specific siRNAs in mouse β cells results in insulin synthesis and secretion decrease, and increased β cell apoptosis (You et al., 2015). Although the exact mechanism by which *MEG3* regulates insulin production remains to be clarified, its nuclear localization suggests a role in transcriptional regulation.
- ***HI-LNC78* or *TUNAR*:** This lncRNA is a glucose-regulated islet transcript (Moran et al., 2012). Knockdown of *HI-LNC78* in T antigen-excised EndoC-βH3 cells resulted in insulin reduction and impaired glucose-stimulated insulin secretion (Akerman et al., 2017). The inhibition of this lncRNA is also correlated with

expression changes of crucial genes for pancreatic β cell function. In fact, the transcriptional alterations found after *HI-LNC78* inhibition are similar to the ones observed after the inhibition of key transcription factors (e.g. HNF1A or MAFB) in β cells.

- ***TUG1* (Taurin Upregulated Gene 1):** It is highly expressed in the pancreatic tissue and particularly in rodent β cells (Yin et al., 2015). The knockdown of *Tug1* in mouse pancreatic β cells provoke a reduction in insulin content and secretion and an increase in blood glucose levels. In addition, *Tug1*-inhibited mice presented decreased levels of *PDX1, NEUROD1* and *MAFA* (key genes for the regulation of insulin on β cells), proposing *TUG1* as insulin expression mediator through the regulation of these key transcription factors (Wilson & Pullen, 2021; Zhu et al., 2017).
- ***MALAT1* (Metastasis-associated Lung Adenocarcinoma Transcript 1):** This highly conserved lncRNA is dysregulated in many tumors (P. Ji et al., 2003) and present in many cell types including pancreatic β cells (Dorrell et al., 2011). *MALAT1* is encoded within an active enhancer cluster with several binding domains for islet-transcription factors (Pasquali et al., 2014). *MALAT1* is also regulated by hypoxia (Brock et al., 2017; Lelli et al., 2015), a major determinant of islet post-isolation quality in transplantation; and thus *MALAT1* may influence islet survival under hypoxic conditions. This lncRNA nowadays is suggested as a biomarker to predict islet isolation and to further improve the clinical prediction of islet transplantation outcome (Wong et al., 2019). Recently, *MALAT1* is described to induce β cell dysfunction by reducing the histone acetylation of the *PDX1* promoter (Ding et al., 2020). This study revealed how IL-1β stimulation promoted the expression of *MALAT1* and decreased the expression of *PDX1* in diabetic mouse islets (Ding et al., 2020). In fact, the overexpression of this lncRNA inhibited H3 histone acetylation of the *PDX1* promoter inhibiting its transcription. Taking together, *MALAT1* is suggested as a fundamental lncRNA in pancreas and β cell function.

- ***GAS5* (Growth arrest-specific 5):** It is highly expressed in mouse pancreas and downregulated in db/db mice (T2D mouse model) (Jin et al., 2017). *Gas5* knockdown in mouse β cells inhibited cell proliferation and damaged insulin production and secretion. Additionally, this knockdown also diminished the expression of *PDX1* and *MAFA*, key islet transcription factors; and the expression of *GLUT2,* essential glucose transporter (Jin et al., 2017). Upregulation of this lncRNA alleviated insulin secretion deficit, evidencing a role of *Gas5* on insulin synthesis and secretion (Esguerra et al., 2020).
- ***p3134:*** This lncRNA is mainly expressed in adipose tissue and circulating exosomes secreted by pancreatic β cells. It is described to enhance insulin synthesis and release by the regulation of islet-specific transcription factors (Ruan et al., 2018). Interestingly, *p3134* overexpression resulted in decreased apoptosis and lower glucotoxicity levels on glucose-stimulated mouse pancreatic β cells. *In vivo* upregulation of this lncRNA in db/db mice (T2D model) enhanced insulin production and secretion in response to glucose (Wilson & Pullen, 2021). Additionally, this upregulation in mice increased the mRNA and protein levels of PDX1, MAFA, GLUT2 and TCFL2, of great importance on glucose homeostasis (Ruan et al., 2018).

3.3 LncRNAs in T1D

Genome wide association studies performed in pancreatic β cells have determine plenty of T1D-associated SNPs located in lncRNAs (Mirza et al., 2014). However, the characterization of these lncRNAs is still weak and there is little information about these molecules and their implication in the disease. These disease-linked SNPs usually disturb the function of lncRNAs by altering their secondary structure, an essential feature for their correct function. A recent study of an antisense lncRNA (***NONHSAG011351***) located 5kb from the T1D candidate *loci* named *ERBB3* have revealed that the lncRNA modulates the expression of its neighboring gene (Kaur et al., 2016). This work showed how a SNP located in *ERBB3* has a cis-eQTL effect both on the *ERBB3* gene and in the lncRNA named *NONHSAG011351*. The protective genotype for

T1D is associated with lower levels of both genes but also with reduced β cell apoptosis due to the low abundance of pro-inflammatory cytokines.

Another apoptosis-linked lncRNA is ***GLIS3***, a lncRNA that has been described to modulate cytokine-induced pancreatic β cell destruction by the regulation of a splice variant of the pro-apoptotic protein Bim (Nogueira et al., 2013). Interestingly, *GLIS3* is also capable of altering the expression of other lncRNAs. *G3R1* (GLIS3 regulated 1) is the name of a lncRNA regulated by *GLIS3* in mouse β cells although there is no evidence of its implication in β cell impairment (Scoville et al., 2020). Furthermore, the deletion of the lncRNA named *Lnc25* in the human cell line EndoC-βH1 decreased *GLIS3* expression but did not affect glucose-stimulated insulin secretion (Moran et al., 2012).

Another T1D candidate gene that overlaps with a sense lncRNA (***NONHSAG044354***) is the coding gene named *BACH2*, a transcription factor key in cytokine-triggered pancreatic β cell apoptosis (Marroquí et al., 2014). The T1D-linked SNP in this lncRNA is predicted to disrupt the secondary structure of the lncRNA, suggesting a disruption of its function (Mirza et al., 2014).

Chapter 3

Hypothesis and aims of the study

Hipotesi eta helburuak

Ikerketaren hipotesi eta helburuak

Genoma osoko asoziazio ikerketek T1D pairatzeko arriskuari loturiko polimorfismoak aurkitu dituzte RNA luze ez-kodetzaileetan (lncRNA). Halere, oraindik ezezagunak dira T1D garapena bultzatzen duten lncRNA-en mekanismo molekularrak.

Azken urteetan pankreako β zelula mailako infekzio biralek T1Darekin erlazionatutako zenbait lncRNA-en adierazpena aldatzen dutela egiaztatu dute. Gainera, ikusi egin da nola T1Dari loturiko zenbait lncRNA molekulek berebiziko funtzio zelularrak betetzen dituztela infekzio biral baten ostean.

Proiektu honen **hipotesia** honakoa da; T1Darekin erlazioa duten eta β zelulen berezko immunitatean parte hartzen duten lncRNA-tan ager daitezken aldaketa funtzionalek zelula hauek duten infekzio biralei aurre egiteko gaitasunean eragina dutela. Honek, desorekatutako erantzun antibirala faboratzen du, zeina β zelulen suntsiketa dakarrena eta ondorioz, T1Daren garapena.

Ikerketa lan honen **helburu** nagusia RNA luze ez-kodetzaileetan kokaturik dauden T1D-aren patogenesiarekin loturiko arrisku aldaerek duten eragina aztertzea da; pankreako β zeluletan lan eginda.

Helburu nagusi hau burutzeko, hiru helburu espezifiko definitu dira:

1) IDENTIFIKATU. Pankreako β zeluletan infekzio biral baten ondorioz adierazten eta modulatzen diren lncRNAk identifikatzea; zeintzuek gainera, T1-arekin loturiko SNP bat daukaten.

2) ZEHAZTU. T1D gaixotasunarekin bat datozen lncRNA molekulen funtzioa zehaztea pankreako β zeluletan birusek eragindako inflamazio eta disfuntzio egoeran.

3) KARAKTERIZATU. T1D-loturiko aldaera genetikoek lncRNAtan eragiten dituzten mekanismo molekularrak karakterizatzea, pankreako β zelulen disfuntzioan eta T1Daren garapenean.

Hypothesis and aims of the study

Genome wide association studies focused on T1D have identified multitude of polymorphism located in long non-coding RNAs (lncRNA). However, it is still unknown the molecular mechanisms by which these lncRNAs enhance the progression of T1D.

On the last years, viral infections at a pancreatic β cell level are demonstrated to alter the expression of T1D-linked lncRNAs. Moreover, the lncRNA molecules that are linked to T1D are described to perform essential cellular functions after a viral assault.

The **hypothesis** of this project is the following; T1D-linked lncRNAs participating on the immune system modulation of pancreatic β cells are implicated on the ability of the β cells to face and deal with viral infections. The altered immune system triggers an unbalanced antiviral response, provoking the destruction of pancreatic β cells and finally resulting in T1D development.

The **main objective** of the present work was to assess the contribution of risk genetic variants present in long non-coding genes to the pathogenesis of T1D at the pancreatic β cell level.

To this main objective, three specific aims have been defined:

1) IDENTIFY. To identify lncRNAs that are expressed and modulated by viral infections in pancreatic β cells and harbor a T1D-associated SNP.

2) DETERMINE. To determine the functional effect of T1D-associated lncRNAs in virus-induced pancreatic β cell inflammation and dysfunction.

3) CHARACTERIZE. To characterize the molecular mechanisms by which T1D-associated genetic variants in lncRNAs influence pancreatic β cell dysfunction and T1D pathogenesis.

Chapter 4

Materials and methods

Materialak eta metodoak

1 Eredu esperimentalak

1.1 Giza lagin biologikoak

Giza irla pankreatikoetatik lortutako cDNA (DNA osagarria) laginak bi zentroetatik lortu ziren: Cisanello Unibertsitate Ospitaletik, Pisa, Italia (zuzenean Piero Marchetti-ren bitartez, ala Decio L. Eizirik-eri esker ULB Diabetes Ikerkuntza Zentroaren bitartez, Brusela, Belgika) eta Andaluziako Biologia Molekular eta Birsorkuntza Medikuntza Zentrotik-CABIMER, Sevilla, Espainia (Benoit Gauthier-en bidez).

Eizirik eta Marchetti-k eskainitako cDNA laginak, emaile ez-diabetikoen (n=36 lagin; adina: 72 ± 11 urte; BMI: 25 ± 3 kg/m2) irla pankreatikoetatik erauzitako RNA laginen erretrotranskripzioz lortutakoak dira aurretiaz deskribatu bezala (Marchetti, 2007). Laginak Pisako Unibertsitateko komite etikoaren adostasunarekin eskuratu ziren. Gauthier-ek emandako cDNA laginak Tebu-Bio enpresari erositakoak dira (n=7 lagin; Bartzelona, Espainia) eta emaileen ezaugarriak aldez aurretik azaldu ziren (Vuilleumier, 2018). Giza lagin hauekin egindako esperimentuak Andaluziako Osasun Ministerio Erregionaleko Komite Etikoak onartuak izan dira (Lizentzia zenbakia: 2013-04398).

1.2 In vitro esperimentuetarako zelula-ereduak

Egindako *in vitro* esperimentu funtzional gehienak EndoC-βH1 izeneko giza pankreako β zelula lerro hilezkortuan egin dira. Bestelako esperimentuak induzitutako zelula ama pluripotenteetatik (iPS-etatik) eratorritako giza β zelulak, HEK293FT, SHSY5Y, HCT15, NCI-H23 eta Jurkat zelula lerroetan burutu dira.

Zelula lerro guztiak MycoAlert Mycoplasma Detection kitarekin (Lonza) testatu dira egiaztatzeko Mycoplasma gabekoak direla. Halere, Mycoplasma kutsadura posiblerik ekiditeko, Plasmocin Prophylactic (Invivogen) konposatua gehitu zaie hazkuntza medioei.

1.2.1 EndoC-βH1

EndoC-βH1, giza pankreako β zelula lerroa Human Cell Design enpresari erosiak izan dira (https://www.humancelldesign.com). Laburki, EndoC-βH1 zelulak 70.000-75.000

zelula/zm^2–ko dentsitatearekin hazi dira aurretiaz Matrigel-fibronektina nahasturarekin (100 mg/ml eta 2 mg/mL, hurrenez hurren) (Sigma-Aldrich) estalitako hazkuntza plaketan. Zelulak OPTIβ1 medioan (Human Cell Design) hazi dira 37°C-tan eta %5 CO_2-an. Zelulak 7 egunero pasa dira.

Transfekzio protokoloak gauzatzeko %2 FBS, 5.6 mM glukosa, 50 µM 2-merkaptoetanol (Biorad), 10 mM nikotinamida (Calbiochem), 5.5 µg/ml transferrina eta 6.7 ng/ml selenitedun (Sigma-Aldrich) DMEM medioa erabili da.

1.2.2 iPS-etatik eratorritako giza β zelulak

iPS-etatik eratorritako giza β zelulak, Mariana Igoillo-Esteve eta Miriam Cnop (ULB Diabetes Ikerkuntza Zentroa, Brusela, Belgika) doktoreen partaidetzari esker eskuratu ziren. Laburki, emaile osasuntsuetatik erauzitako bi giza iPS zelula lerro [HEL115.6 iPS-ak Helsinkiko Unibertsitatean sortu ziren eta 1023A iPS-ak DM Egli doktoreak (Columbiako Univertsitatea) utzitakoak dira] β zeluletara bereiztu ziren aurretiaz deskribatu zen metodologia erabiliz (Cosentino et al., 2018). iPS zelulek kariotipo normala zuten, zelula ama morfologia eta markatzaile pluripotenteak adierazten ziturzten (Cosentino et al., 2018). iPS zelulak E8 medioan mantendu ziren Matrigelez estalitako plakatan (Corning). Desberdintzapenaren 7. etapa hasi baino 24h lehenago, 2.5–2.8 × 10^6 zelula jarri ziren 3.5 zm-tako putzutan E8 mediora 5 µM ROCK inhibitzaile (STEMCELL Technologies) gaineraturik. Behin zelulak aitzindari pankreatiko bilakatu zirela, zelulak 24 mikroputzutako plakatan hazi ziren 750 zelula/mikroputzu dentsitatean (AggreWell, STEMCELL Technologies) 10 µM ROCK inhibitzaile eta 1µl/ml heparina (STEMCELL Technologies) gehituz agregatuek 3 dimentsiotako egitura erraztuz.

Infekzioak egin aurretik, 7.etapako agregatuak dispertsatu egiten ziren. Honetarako, agregatuak 0.5 mM EDTAn inkubatzen ziren 4 minutuz giro tenperaturan. Ondoren Accumax (Sigma Aldrich) jartzen zaie beste 8 minutuz eta azkenik, pipeta erabiliz dispertsatu egiten ziren (Nair et al., 2019). Sakabanatze prozesua gelditzeko, zelulei knockout seruma (Gibco) gehi 10 µM ROCK inhibitzailea jartzen zaie. Zelula

dispertsatuak 7.etapako medioan hazten ziren 5 × 10^4 zelula jarrita 6.4 mm –tako putzutan (Cosentino et al., 2018).

1.2.3 HEK293FT

HEK293FT zelulak (CRL-1573) American Type Culture Collection (ATCC; https://www.atcc.org) erakundearen bitartez erosi ziren. Zelulak DMEM + %10 FBS + 100 unitate/ml penizilina + 100 µg/ml estreptomizina (Lonza) medioan hazi ziren. Transfekzio esperimentuetarako, antibiotiko gabeko medio berdina erabili zen.

1.2.4 Jurkat

Jurkat zelulak (ATCC, TIB-152) RPMI + %10 FBS + 100 unitate/ml penizilina + 100 µg/ml estreptomizina medioan hazi ziren. FBS gabeko RPMI medioa erabili zen migrazio entseguetan.

1.2.5 SH-SY5Y

SH-SY5Y zelulak (ATCC; CRL-2266) honako hazkuntza-medioan hazi ziren: medioaren erdia RPMI gehi beste erdia MEM Eagle EBSS medioa; + %10 FBS + 100 unitate/ml penizilina + 100 µg/ml estreptomizina (Lonza).

1.2.6 HCT15

HCT15 zelulak (Sigma-Aldrich, #91030712) RPMI + %10 FBS + 100 unitate/ml penizilina + 100 µg/ml estreptomizina medioan hazi ziren.

1.2.7 NCI-H23

NCI-H23 (ATCC, CRL-5800) RPMI + %10 FBS + 100 unitate/ml penizilina + 100 µg/ml estreptomizina medioan hazi ziren

2 Metodoak

2.1 Zelula tratamenduak

Tratamenduak gauzatzeko, hautatutako konposatua zuzenean gehitzen zaio dagokion zelularen hazkuntza medioari.

- Zitokina tratamendua: EndoC-βH1 zelulak 10 U/ml IL-1β gehi 1000 U/ml IFNγ-zitokinekin tratatu ziren 24 orduz.
- STAT seinaleztapen-bidezidorraren inhibizioa: Tratamendua JAK bidearen inhibitzailea den Ruxolitinib konposatuarekin (Cayman Chemicals) egin zen, zuzenean hazkuntza mediora gaineraturik 5 µg/ml-ko kontzentrazio finalean aurretiaz deskribatua izan den ereduari jarraiki (Coomans de Brachène et al., 2018)
- NFκB seinaleztapen-bidezidorren inhibizioa: Bay 11-7082 (Sigma-Aldrich) konposatua erabiliz NFκB bidezidorra inhibitzea lortu zen. Honetarako, hazkuntza mediora 10 µM gehiturik.
- Transkripzio eta itzulpenaren inhibitzaileak: EndoC-βH1 zelulak 5 µg/ml aktinomicina D (Sigma Aldrich) konposatuarekin tratatu dira 6 orduz transkripzioa inhibitzeko, edo 20 µg/ml zikloheximida (Sigma Aldrich) 24 orduz itzulpena blokeatzeko.

2.2 Coxsackievirus infekzioa

Infekzio biralak bai EndoC-βH1-tan bai iPS-etatik eratorritako β zeluletan egin ziren.

EndoC-βH1 zeluletan egindako infekzioa CVB5/Faulkner birusarekin burutu zen. CVB5/Faulkner birusaren diluzioa (Multiplicity of Infection (MOI)=5) Hanks‘ Balanced Salt Solution (HBSS, Invitrogen) medioan egin ondoren, zelulak infektatu egiten dira. Zelulak ordubetez 37°C-tan mantentzen dira adsortzioa emateko; inokulaturiko birusa kendu eta zelulak bitan garbitu ziren HBSS-z. Azkenik, zelulak hazkuntza medioan jarri ziren eta beste 24 orduz mantendu birusaren erreplikazioa ahalbidetzeko.

iPS-etatik eratorritako β zeluletan egindako infekzioa burutzeko, lehenago azaldu bezala, Accumax erabilita zelulak dispertsatu behar ziren (Cosentino et al., 2018; Lytrivi

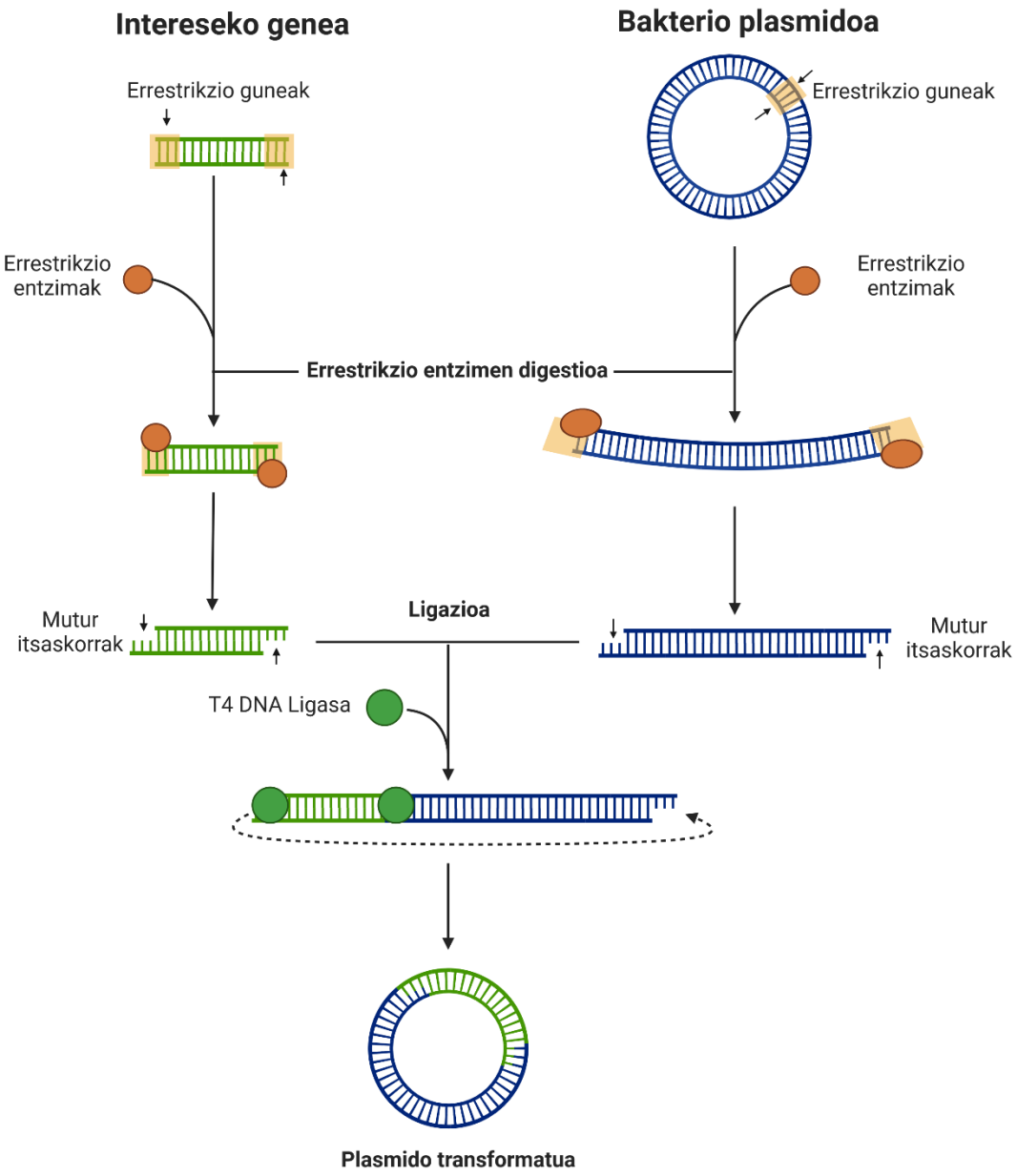
et al., 2021). 24 putzutako plakak Matrigelarekin estali ostean, 300.000 zelula/putzu hazi ziren 10 µM ROCK inhibitzailea zeraman 7.etapako medioan. Gau oso baten ostean, zelulen suspertzea ahalbidetuz; infekzioa gauzatu zen. CVB1/Conn-5 (MOI=0.05) edo CVB4/JVB (MOI=0.5) birusak erabili ziren honako medioa erabiliz: Ham's F-10 Nutrient Mixture (Gibco) + 2 mM GlutaMAX + 50 µM IBMX + %1 FBS. Infekzioa egin eta 2 ordutara infekzio medioa kendu eta Ham's F-10 Nutrient Mixture, %0.75 BSA, 2 mM GlutaMAX, 50 µM IBMX, 50 U/ml penizilina eta 50 µg/ml estreptomizina-dun medioaz ordezkatu zen. Zelulak beste 22 orduz mantendu ziren hazkuntzan.

2.3 Plasmidoen eraikuntza eta transformazioa DH5α E.coli bakterioetan.

- Plasmidoen eraikuntza (**8. Irudia**):

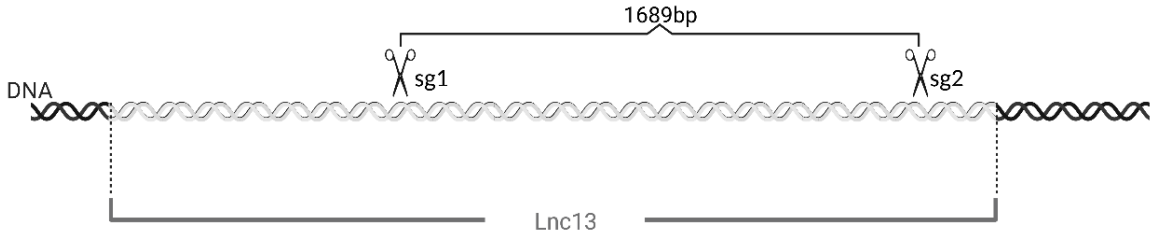
  pLnc13-C eta pLnc13-T plasmidoak sortzeko, rs917997-C edo T alelodun *Lnc13* anplifikatu zen hasle pare espezifiko bat erabiliz (*Lnc13*-overexpression; **1. Taula**) eta giza cDNA eredu bezala izanda. Anplifikaturiko *Lnc13* sekuentziak eta pCMV6 plasmidoa (Cat# PS100001; Origene) FseI (R0588; NEB) eta KpnI (R0142; NEB) errestrizio entzimekin digeritu ziren fabrikatzailearen protokoloa jarraituz. *Lnc13* sekuentzia bakoitza digeritutako pCMV6 bektorean klonatu ziren, horretarako T4 DNA ligasa (M0202; NEB) erabiliz.

  Lnc13-delSNP plasmidoa sortzeko *Lnc13*aren 1. basetik 1771. baserako eta 2278. basetik 2545. baserako sekuentziak anplifikatu ziren. Ondoren gainjarriak diren hasleen (**1.Taula**; *Lnc13*_delSNP) bitartez, bi sekuentzien fusioa egin zen zatiki bakarra lortuz. Lorturiko *Lnc13* mutatuaren sekuentzia eta pCMV6 plasmidoa FseI eta KpnI errstrikzio entzimekin digeritu ziren fabrikatzailearen protokoloa jarraikiz. Amaitzeko, *Lnc13*aren sekuentzia pCMV6 plasmido digerituan klonatu zen.



8. Irudia. Plasmidoa sortzeko erabilitako protokoloaren eskema.

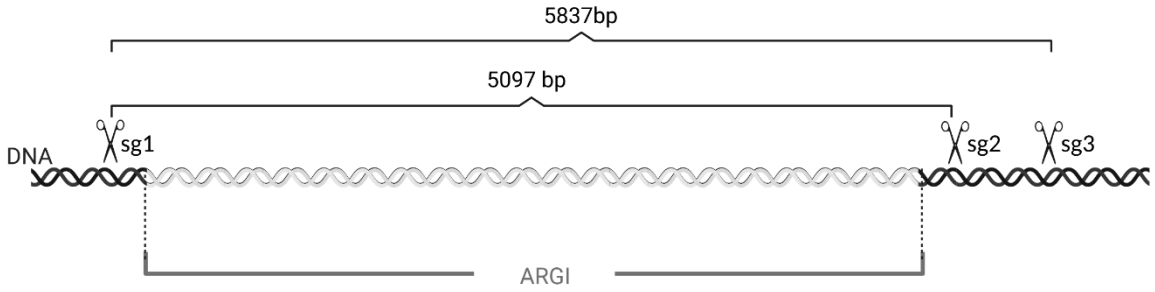
Lnc13 CRISPR-Cas9 esperimentuko RNA gidak (sgRNAs) 1698bp-tako domeinu bat ezabatzeko diseinatu ziren (**1. Taula**). sgRNA bakoitza px459 plasmido (62988; Addgene) batean klonatu ziren BbsI errestrikzio entzimaren erabiliz (R0539; NEB) (**9. Irudia**).



9. Irudia. *Lnc13* egindako delezioaren eskema CRISPR-Cas9 teknologia erabilita. 1698 bp-tako segmentu baten delezioa lortu zen RNA gida pare bat erabilita (Lnc13_KO_1 eta Lnc13_KO_2).

1 motako diabeteserako arrisku aleloa daraman *ARGI* gainadierazpen bektorea (ARGI-R; rs9585056-C) ProteoGenix enpresari erosi zitzaion. Aldiz, babes alelodun *ARGI* gainadierazpen bektorea (ARGI-P; rs9585056-T) zuzenduriko mutagenesiaren bitartez lortu zen, Site-Directed Mutagenesis QuickChange II (Agilent) kitaren bidez. Mutagenesian erabilitako hasleak **1. Taulan** daude zerrendaturik (ARGI _mut_rs9585056).

*ARGI*n egindako CRISPR-Cas9 esperimentuetarako bi RNA gida pare diseinatu ziren. sgRNA_1 eta sgRNA_2 (**1.Taulan**) gida parearekin 5079bp-tako delezio bat sortu zen. Delezioa handiagoa lortu zen, 5387bp-tako hain zuzen, sgRNA_1 eta sgRNA_3 (**1.Taulan**) konbinatu ostean. RNA gida bakoitza px330 bektore (Addgene) batetan klonatu zen BbsI errestrikzio entzimarekin (R0539; NEB) digeritu ondoren (**10. Irudia**).



10.Irudia. *ARGI* genearen delezioaren eskema CRISPR-Cas9 teknologia erabilita. 5097 bp-tako segmento baten delezioa lortu zen RNA gida pare bat erabilita (ARGI_KO_sg1 eta

ARGI_KO_sg2) eta 5837 bp-takoa beste gida pare bat erabili ostean (ARGI_KO_sg1 eta ARGI_KO_sg3).

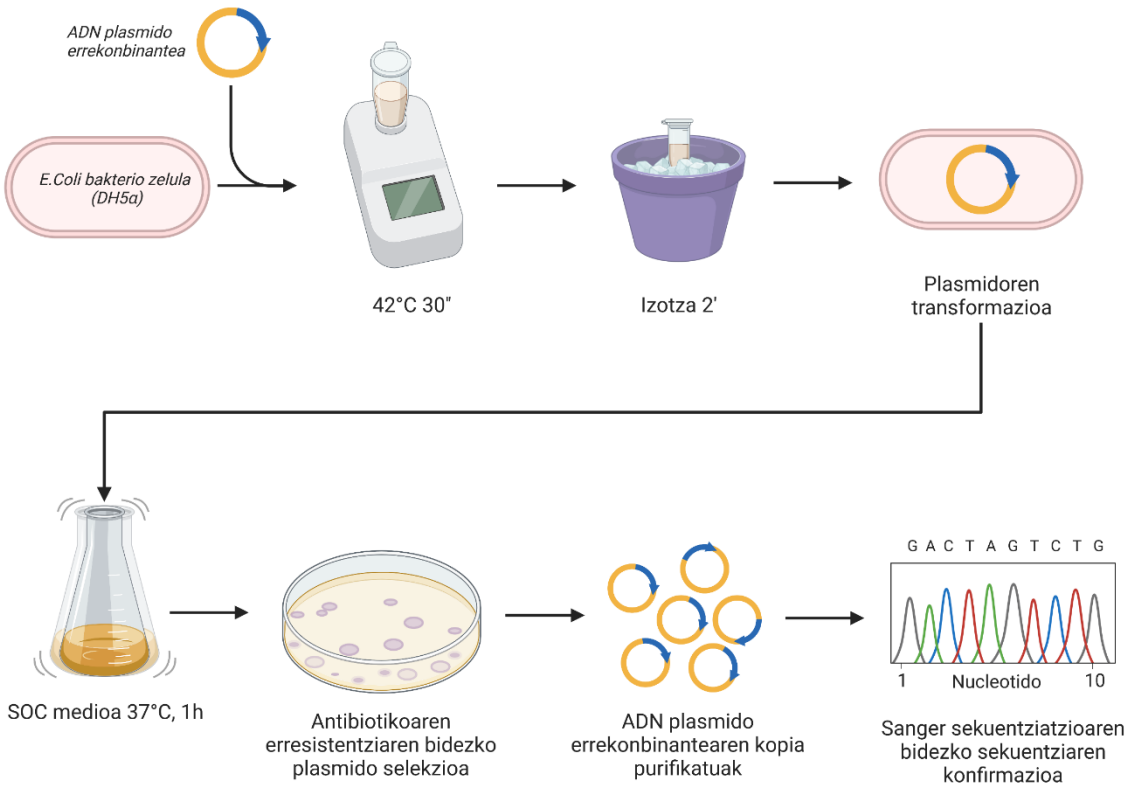
CRISPRi esperimentuak gauzatzeko, RNA gida (ARGI_CRISPRi sgRNA, **1.Taula**) bat klonatu zen pCRISPRi-Cas-Guide-CRISPRi bektore batetan (GE100059; Origene). Klonaketarako BamHI (R0136; NEB) eta BsmBI-v2 (R0739; NEB) errestrikzio entzimak erabili ziren.

- **Plasmidoen transformazioa *E.coli* konpetenteetan (11.Irudia):**

Intereseko plasmido errekonbinanteak aukeratu eta anplifikatzeko, ligaturiko plasmido bakoitza DH5α *E.coli* bakteriatan transformatu zen. Horretarako, DH5α zelulak 10-50ng plasmidorekin nahastu ziren eta izotzetan mantendu 20 minutuz. Txoke termiko baten bitartez (42°C, 30''), bakteriak iragazkortzea lortu zen. Ostean, laginak 2 minutuz izotzetan jarri eta 300 µl SOC mediorekin nahastu ziren. Bakteriak SOC medioan hazi ziren ordubetez 37°C-tan. Hazkuntzaren ondoren, zelulak 6.000 rpm 5 minutuz zentrifugatu ziren zelulak jalkinean gera zitezen. Jalkina 100 µl SOC medioan birsuspenditu eta zuzenean dagokion antibiotikoa daraman agar plaka batean erein ziren. Antibiotikoaren aukeraketa plasmidoak dakarren erresistentzia genearen araberakoa da, egindako esperimentuen kasuan, Kanamizina ala Anpizilina. Ereindako plaka gau osoan zehar hazten utzi zen 37°C-tan.

- **Plasmido errekonbinanteen egiaztapena:**

Hurrengo goizean bakterioen hazkuntza egiaztatu zen. Intereseko plasmidoa barneratu zuten bakterioak soilik hazi ziren (antibiotikoarekiko erresistentzia zeukatelako). Halere, guztiz egiaztatzeko bakterioak barneratutako plasmidoa interesekoa zela, plasmidoak sekuentziatzera bidali ziren. Honetarako, hazitako agar plakatik zenbait kolonia jaso eta indibidualki hazi ziren gau oso batez dagokion antibiotikoa zuen LB broth medioan. Plasmidoa bakterioetatik purifikatzeko NucleoSpin Plasmid EasyPure (Macherey-Nagel) kita erabili zen. Behin plasmidoa purifikaturik, sekuentzia Sanger metodoaren bitartez konprobatu zen.



11.Irudia. Plasmido transformazioa DH5α *E.coli* bakteriotan.

2.4 Transfekzioak

- Polyinosinic:polycytidylic azido (PIC) transfekzioa: EndoC-βH1 zelulak Lipofectamine-2000 (Invitrogen) erreaktiboa erabilita transfektatu ziren fabrikatzailearen gida jarraituz. dsRNA birala simulatzen duen polyinosinic:polycytidylic azidoa (PIC) (InvivoGen) 1 µg/ml-tako kontzentrazio finalean erabili zen.
- CRISPR eta gainadierazpen bektoreen transfekzioa: EndoC-βH1 zelulak 24-putzutako hazkuntza plakatan hazi eta 250ng plasmidorekin transfektatu ziren. Transfekzioak Lipofectamine-2000 (Invitrogen) konposatua erabilita gauzatu ziren fabrikatzailearen protokoloa jarraiki. Gainadierazpen plasmidoen transfekzioa zeluletan 16 orduz mantendu ziren, eta ondoren hazkuntza medio berriaz ordezkatu zen zelulak beste 24 orduz suspertzea baimenduz.

- CRISPRi plasmidoen transfekzioa EndoC-βH1 zeluletan 36 orduz inkubatu zen; ondoren zelulak bestelako esperimentuak egiteko prest egon ziren.
- CRISPR-Cas9 bektoreak Lipofectamine-2000-rekin transfektatu eta 24 ordutara puromizinarekin (2µg/ml) tratatu ziren beste 36 orduz. pSpCas9(BB)-2A-Puro (Px459) V2.0 plasmidoa (#62988; Addgene), puromizinarekiko erresistentzia dauka; eta horri esker puromizina gehitzeak hautaketa espezifikoa eskaintzen dio.

2.5 RNA eta DNA erauzketa

RNA erauzketa NucleoSpin RNA Kita (Macherey Nagel) edo PureLink RNA Mini kita (Invitrogen) erabilita egin zen.

DNA erauzketarako, NucleoSpin Gel eta PCR Cleanup kitak (Machery Nagel) erabili ziren

- **RNA adierazpenaren zehaztapena:** Adierazpen mailak qPCR bitartez kalkulatu ziren; Taqman Gene Expression Assay-ak (Thermo Scientific) edo Sybr Green (Biorad) teknologia erabilita hasle espezifikoen bitartez (**2.Taula**). qPCR neurketa guztiak birritan egin ziren eta gutxienez 3 lagin independenteetan. Kuantifikazioa CFX384 Touch Real-Time PCR Detection System (Bio-Rad Laboratories) makinaren bidez gauzatu eta adierazpen mailak $2^{-\Delta\Delta}$Ct metodoa jarraiki analizatu ziren.
- **SNP genotipaketa:** *Lnc13*n dagoen rs917997 SNParen genotipaketa TaqMan C____345197_1_genotyping assay-a (Thermo Sciencific) erabilita egin zen eta *ARGI*n barne dagoen rs9585056, IDTko rhAmp SNP Assay-a (Hs.GT.rs9585056.T.1) erabiliz.

2.6 Zelulako dauden *Lnc13* molekulen kuantifikazioa

Giza zelula lerro desberdinetan zeuden *Lnc13* kopia kopurua zehazteko, erreferentziazko plasmido bat sortu zen non gene ituaren cDNA sekuentzia barne zeukan. Kuantifikazio absolutua egiteko, erreferentziazko estandarraren diluzio seriatuak egin ziren. Ct balioak diluzio faktorearengatik zatikatu ziren eta base 10eko

eskala semi-logaritmikodun grafikoan irudikatu ziren zuzen lerro bat eratuz. Zuzen lerroa erabilita EndoC-βH1, HEK293, SHSY5Y, HCT15 eta NCI-H23 zelula lerrotan dauden *Lnc13* molekula/zelula kantitatea kalkulatu zen.

2.7 Zatiketa azpizelularra

1. **RNA zatitze azpizelularra:**

- Frakzio nuklearraren isolamendua: Zelulak hauspeatu eta gero, poliki nahastu egin ziren ondorengo tanpoiarekin: 600 µl RNasa gabeko ura + 200 µl PBS + 200 µl C1 lisi disoluzio (osagaiak: 1.28 M sakarosa, 40 mM Tris-HCl pH 7.5, 20 mM MgCl2, %4 Triton X-100). Nahastea izotzetan mantendu zen 15'-z eta zuzenean 2.500 rpm-ko abiaduran zentrifugatu zen 15 minutuz 4°C-tan. Gainjalkina guztiz baztertu eta jalkina RNA erauzketarako bideratu zen. Lagin honetatik lortutako RNAa, RNA nuklearrari dagokio.
- Kuantifikazioa: intereseko genea, *MALAT1* edo *MEG3* (nukleorako kontrola) eta *RPLP0* (zitoplasmarako kontrola) qPCR bidez neurtu ziren. RNA nuklearretik lorturiko qPCR datuak zelula osoko laginetan lorturiko gene berdin horien qPCR datuekin konparatu zen. Emaitzak nukleo/zelula oso ratioaren logaritmoaren bitartez adierazi da. Erabilitako hasleen sekuentziak **2.Taulan** aurki daitezke.

2. **Proteinaren frakzionamendu azpizelularra:**

   EndoC-βH1 zelulen proteina zatitzea honako kit komertzialaren bitartez gauzatu zen: *subcellular protein fractionation kit* (78840; Thermo Scientific). Zatikatzearen efizientzia testatzeko hurrengo proteinak aztertu ziren Western blot bitartez: Hsp90 zitoplasmarako, HDAC1 nukleorako eta H3 kromatinaren frakziorako.

2.8 Western blot analisia

- Zelula lisatuaren elektroforesia: Zelulak PBS hotzarekin garbitu ostean, Laemmli tanpoian (62 mM Tris-HCl, 100 mM dithiothreitol (DTT), %10 glizerol, % 2 SDS, 0.2 mg/ml bromofenol urdina, %5 2-merkaptoetanol) lisatu ziren. Lisatuak %10eko SDS-PAGE geletan kargatu ziren proteinak tamainaren arabera banatzeko.

- Proteinen transferentzia eta antigorputzen inkubazioa: Elektroforesiaren ondoren, proteinak nitrozelulosazko mintzetara transferitu ziren Transblot-Turbo Transfer System (Biorad) makina erabilita. Behin transferentzia eginda, mintza blokeatzeko ordubetez inkubatu zen giro tenperaturan TBSTn (20 mM Tris, 150 mM NaCl eta %0.1 Tween 20) diluitutako %5 esne gaingabetuan. Mintzak blokeaturik, gau osoan zehar inkubatu ziren 4°C-tan antigorputz primarioekin (**3.Taulan**). Mintzak 3 aldiz garbitu ziren PBST diluzioan (PBS gehi %0.1 Tween 20) eta antigorputz sekundarioak ordubetez inkubatu ziren giro tenperaturan (peroxidasarekin konjugaturiko saguaren kontrako antigorputz sekundarioa (1:5.000 diluzioa %5 esne gaingabetuan) edo untxiaren aurkakoa (1:10.000 diluzioa %5 esne gaingabetuan).
- Errebelatu: Banda immunoerreaktiboak *SuperSignal West Femto chemiluminescent* (Thermo Scientific) substratua erabilita errebelatu ziren; Bio-Rad Molecular Imager ChemiDoc XRS makinarekin detektatu eta ImageLab softwarearen (Bio-Rad Laboratories) bidez kuantifikatu.

2.9 ELISA

EndoC-βH1 zelulen gainjalkinak jaso ziren CXCL10 eta CCL5 proteinen zehaztapena egiteko ELISA kitak erabilita (R&D Systems, Abingdon, UK; Cat# DY266-05 eta Cat# DY278-05, hurrenez hurren). Laburki, ELISA protokoloaren prestaketa honakoa da:

- Plakaren prestaketarako 96-putzutako plaka bat estali egiten zen PBSn diluitutako anti-CXCL10 ala anti-CCL5 antigorputzaren 100 µl-rekin (kontzentrazio finala 1 µg/ml izanik). Plaka antigorputzekin gau oso batez utzi zen giro tenperaturan. Hurrengo goizean, putzu bakoitzeko diluzioa xurgatu eta 400 µl garbiketa disoluzioarekin garbitu ziren hirutan. Plaka blokeatzeko putzu bakoitzera 300 µl erreaktibo disolbatzaile gehitu eta giro tenperaturan ordubetez inkubatu zen. Azkenik, putzuak hirutan garbitu ziren garbiketarako soluzioa erabilita.

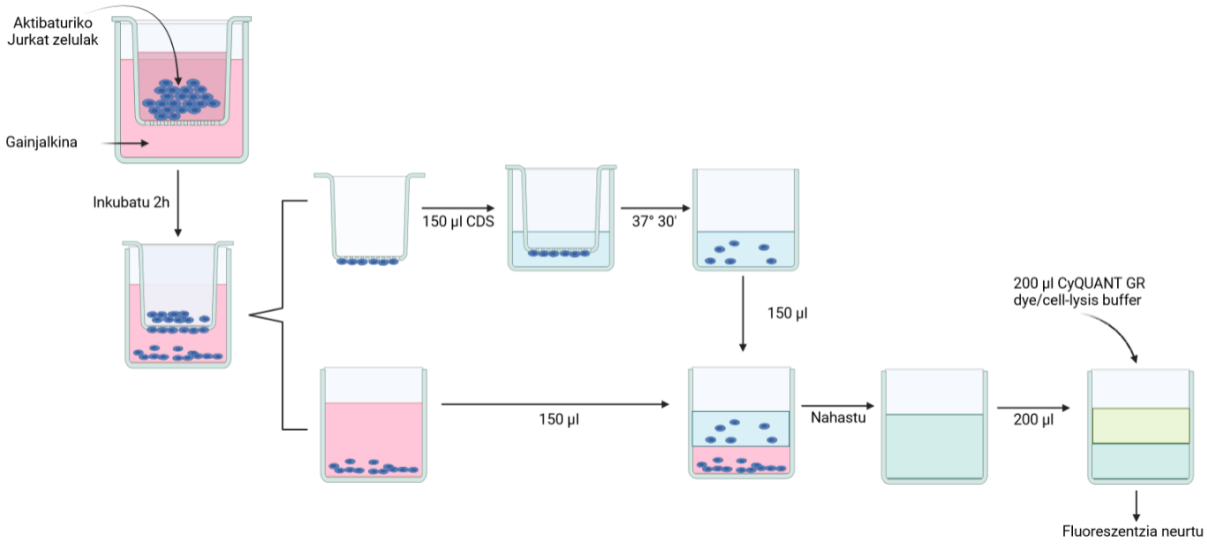
- Detekziorako prozedura: Erreaktibo disolbatzailean nahasturiko lagin edo estandar bakoitzeko 100 µl jarri ziren. Plaka paper eranskailu batez estali eta 2 orduz inkubatu zen giro tenperaturan. Putzuak garbiketa soluzioarekin garbitu ondoren, 100 µl detekzio-antigorputza (20 µg/ml) gehitzen zaio eta beste 2 orduz inkubatu zen giro tenperaturan. Garbitu eta berehala, 100 µl Streptabidina-HRP konposatu gaineratu zitzaion putzu bakoitzari. 20 minutuz inkubatu zen giro tenperaturan argitik babesturik. Segidan, 100µl sustratu-soluzio gehitu eta beste 20 minutuz inkubatu zen giro tenperaturan argitik babesturik. Azkenik, 50 µl Stop soluzioa gaineratu zen putzuetara eta plaka irakurle baten bitartez putzu bakoitzaren dentsitate optikoa neurtu zen; 450nm-tako neurketari 540nm-tako neurketa kenduta.

2.10 Kimiotaxi esperimentua

- *In vitro* kimiotaxi ereduaren prestakuntza: Kimiotaxi esperimentuak Boyden ganbera sistemaren hurbilketa bat (Falasca et al., 2011) eginez. Laburki, EndoC-βH1 zelulak aldez aurretik zegokion tratamenduarekin estimulatu edo transfektatu (*Lnc13*-C edo PIC transfekzioa) ziren. Tratamenduaren ostean, zelula horien gainjalkinak jaso eta 500 µL jarri ziren 24-putzutako plaka bateko putzuetan. Iragazkorrak diren PET-mintzaz egindako euskarriak (#353095; Lonza) jarri ziren gainjalkinak zeuzkaten putzuen gainean. Euskarrien barruan, aldez aurretik aktibaturiko Jurkat zelulak jarri ziren. Jurkat zelulak aktibatzeko, zelulak FBS gabe hazi ziren 16 orduz eta ondoren 50 mg/ml PMA (Sigma-Aldrich) eta 1 µM ionomizina (Sigma-Aldrich) jarri zitzaien 2 orduz. Sorturiko sistema 2 orduz utzi zen 37°C-tan. Kimiotaxi efektuari esker, Jurkat zelulek euskarriaren mintzetik putzuan zeuden gainjalkinetara migratu zuten.
- Lagin bilketa: Iragazkorrak diren PET-mintz euskarrietako medioa baztertu egin zen. Euskarriak, aurretik 37°C-tan berotutako 150 µl zelulak CDS soluzioa (CDS; *cell detachment solution*) daukan zelula jasotzeko erretiluan jarri ziren. Euskarriak 30 minutuz 37°C-tan laga ziren mintzera itsatsitako zelulak askatzeko.

Askaturiko zelulak zituen soluzioa putzuko gainjalkinaren 150 µl-rekin nahastu zen, bertan migratutako Jurkat zelulak zeudelarik.

- Neurketa: Aurreko nahastearen 200 µl 96-putzutako plaka beltz batean jarri ziren. Jarraian 200 µl CyQUANT GR dye/cell-lysis soluzioa gehitu zitzaion putzu bakoitzera. Plaka poliki nahastu eta giro tenperaturan 20 minutuz inkubatu zen argitik babesturik. Baldintza/putzu bakoitzeko fluoreszentzia neurtu zen fluoreszentzia irakurle batekin non kitzikapena 480nm-tan eta emisioa 520nm-tan programatu zen.



12.Irudia. Kimiotaxi esperimentuaren eskema

2.11 RNA immunohauspeatzearen esperimentua (RIP)

RNA immunohauspeatzearen esperimentua bi planteamendu desberdin jarraituz gauzatu zen. A-Agarosa bolatxoak (sc-2001; Santa Cruz Biotechnology) erabili ziren *Lnc13*aren RIP esperimenturako eta magnetikoak diren Dynabeads™ Protein G (10003D; Invitrogen) *ARGI*ren kasuan. Metodologia biak gai dira RNAa eraginkorki immunohauspeatzeko; halere, protokoloa azkarragoa da Dynabeads magnetikoak erabilita (**13. Irudia**).

1. **RIP A-agarosa bolatxoak erabiliz:**

- Zelula bilketa: Zelulak RIP tanpoiarekin (150 mM KCl, 25 mM Tris, 0.5 mM DTT, %0.5 NP-40, proteasa inhibitzaileak) lisatu ziren, 15 minutuz izotzetan

mantendu eta xiringa bat erabilita homogeneizatu ziren. A-agarosa bolatxoak (Santa Cruz) lisatuekin nahastu ziren ordubetez 4°C-tan gurpil irabiatzaile batean lisatuei garbiketa bat egiteko. Laginak zentrifugatu egin ziren bolatxoak kentzeko. Aurretik garbiturik zegoen lisatuaren %10a gorde egin zen input bezala eta gainerako %90a bitan banandu zen erdia intereseko proteina aztertzeko (PCBP2) eta beste erdia kontrol negatiboa egiteko (IgG).

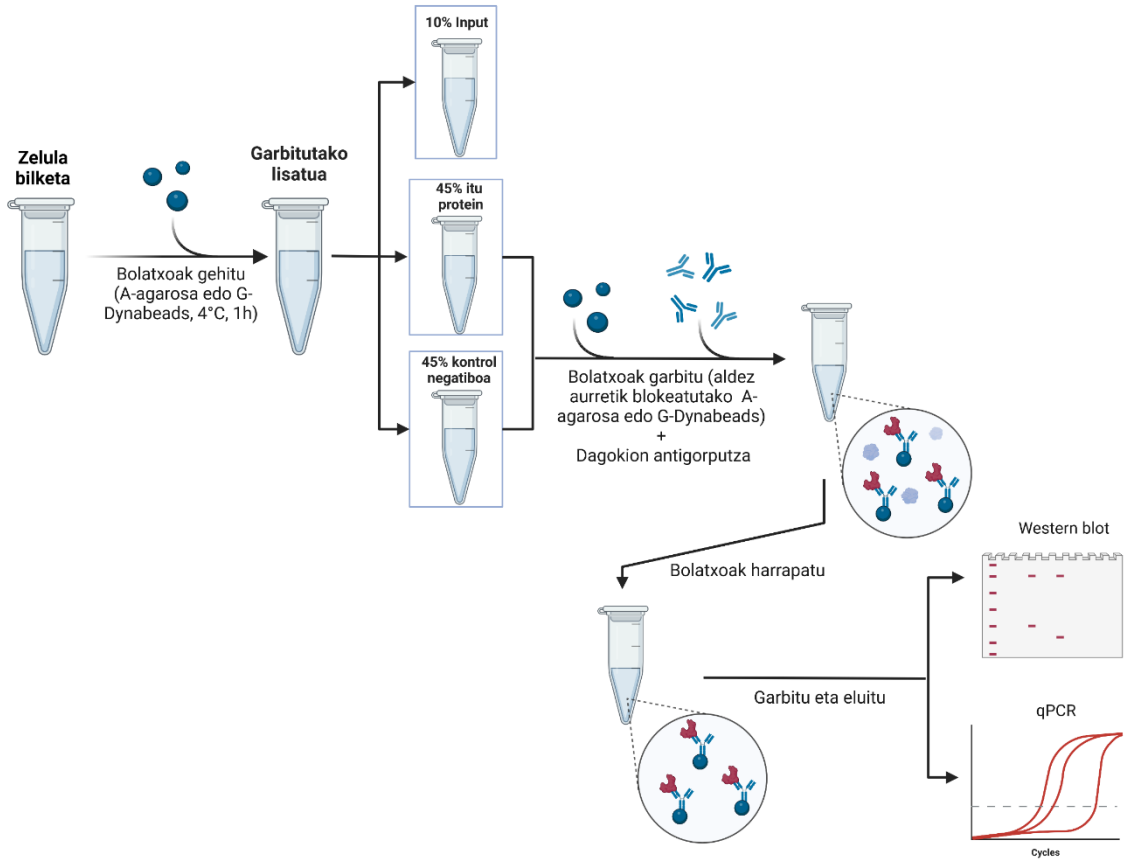
- Bolatxoen blokeoa eta immunohauspeatzea: lagina bakoitzeko 30 µl agarosa bolatxo blokeatu ziren RIPan disolbaturiko %20 BSAn ordubetez 4°C-tan biraka mantenduz. Blokeaturiko bolatxoak RIP tanpoiarekin garbitu ziren eta aurretik garbiturik zeuden lisatuetara gehitu ziren. Antigorputz bakoitzeko 1 µg (PCBP2 edo IgG; **3.Taula**) gaineratu zitzaien dagokien laginei. Nahastea gau oso batez egon zen biraka 4°C-tan antigorputz-proteina-RNA konplexua eratzen laguntzeko.
- Garbiketa eta eluzioa: bolatxoak zentrifugazio bitartez hauspeatu ziren eta gainjalkina baztertu zen. Bolak 3 aldiz garbitu ziren RIP tanpoiarekin, 3 aldiz gatz-baxuko tanpoiarekin (50 mM NaCl, 10 mM Tris-HCl, %0.1 NP-40) eta 3 aldiz gatz-altuko tanpoia (500 mM NaCl, 10 mM Tris-HCl, %0.1 NP-40) erabilita. Behin bolatxoak garbiturik, bitan banatu egiten dira RNAa eta proteina aztertu ahal izateko.
- Analisia: RNA erauzketaren ondoren, *Lnc13* eta *STAT1* geneak qPCR bidez aztertu ziren. Proteina laginak PCBP2 proteinaren immunohauspeatzearen efizientzia testatzeko erabili ziren, Western blot metodologia erabiliz.

2. RIP Dynabeads™ Proteina G erabiliz:

- Zelula bilketa: Zelulak RIP tanpoiarekin (150 mM KCl, 25 mM Tris, 0.5 mM DTT, % 0.5 NP-40, proteasa inhibitzaileak) lisatu ziren izotzetan 15 minutuz mantenduz eta xiringa batez homogeneizatuz. Lisatuak garbitzeko, Dynabeads G bolatxo magnetikoak gehitu zitzaizkien ordubetez 4°C-tan. Garbituriko lisatuen %10 input bezala gorde zen eta laginaren gainerako %90a bitan banandu zen, intereseko proteina (CTCF) edo kontrol negatiboa (IgG) aztertzeko.

- Immunohauspeatzea: Dynabeads G bolak RIP tanpoiarekin garbitu ostean, aurretik garbituriko lisatuetara gaineratu ziren. Antigorputz bakoitzeko 1 µg (**3.Taula**) dagokion lisatura gehitu zen eta antigorputz-proteina-RNA konplexua sortzen laguntzeko gau oso batez inkubatu zen biraka 4°C-tan.
- Garbiketa eta eluzioa: bolatxoak iman batekin harrapatu ziren eta gainjalkina baztertu zen. Bolatxoak 3 aldiz garbitu ziren RIP tanpoiarekin, 3 aldiz gatz-baxuko tanpoiarekin (50 mM NaCl, 10 mM Tris-HCl, %0.1 NP-40) eta 3 aldiz gatz-altuko tanpoia (500 mM NaCl, 10 mM Tris-HCl, %0.1 NP-40) erabilita. Behin bolatxoak garbiturik, bitan banatu egiten ziren RNA eta proteina aztertu ahal izateko.
- Analisia: RNA erauzketa eginda *ARGI* genearen adierazpena aztertu zen qPCR bitartez. Proteina laginak CTCF immunohauspeatzearen efizientzia testatzeko erabili zen Western blot bat eginez.

RNA inmunohauspeatzea (RIP)



13.Irudia. RNA immunohauspeatzearen protokoloa

2.12 Alderantzizko osagarriak diren RNA bitarteko purifikazioa (RAP)

RAP esperimentuak egiteko prozedura Guttman laborategiaren protokoloa (Engreitz et al., 2013) moldatuz egin da (**14.Irudia**):

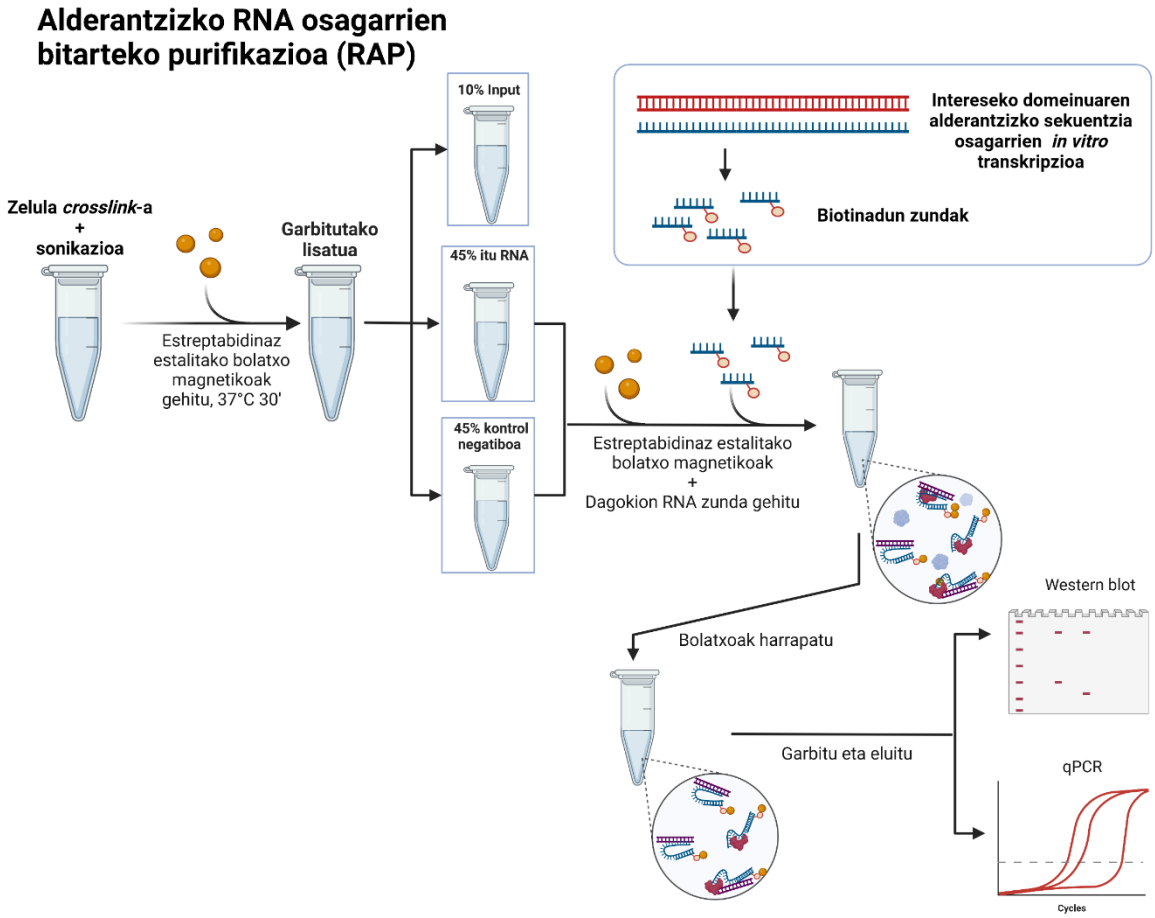
- Alderantzizko osagarriak diren RNA zunden sorkuntza: lehendabizi intereseko RNA molekulen (esaterako *Lnc13* eta *ARGI)* kontrako zunda antisentsu osagarriak sortu behar dira. Honetarako, itu geneen alderantzizko RNA molekulak *in vitro* transkribatu eta biotinilatu ziren Takara enpresatako T7 RNA polimerasa eta rNTPak (Clonthech) erabiliz. Honakoa aurrera eramateko intereseko geneen alderantzizko cDNAren sekuentzia T7 promotorea daukan pCMV6 bektorean klonatu zen. Laburki, DNA ituaren sekuentziaren 200 ng ondorengo konposatuekin nahastu zen: 2 µl x10 T7 RNA polimerasa tanpoia (2540A; Takara Bio), 50U T7 polimerasa (2540A; Takara Bio), 2 µl x10 Biotin RNA Labeling Mix (11685597910; Roche), 2 µl DTT (50 mM, 2540A; Takara Bio) + RNasa gabeko ura bolumen finala 20 µl izan arte. Nahastura 2 orduz inkubatu zen 37°C-tan *in vitro* transkripzioa emateko. Lorturiko RNA PureLink™ RNA Mini Kit (12183025; ThermoFisher Scientific) erabiliz purifikatu zen. Ostera, purifikaturiko RNA zatikatu zen RNA Fragmentation Reagent (AM8740; Invitrogen) erabilita eta fabrikatzailearen argibideak jarraituz. Kontrol negatibo modura, antzeko tamaina daukan lncRNA baten aurkako zundak ere sortu ziren.
- Zelulen gurutzamendua eta bilketa: EndoC-βH1 zeluletara 2 mM DSG (20593; Thermo Scientific) gehitu zitzaien 45 minutuz leunki higituz giro tenperaturan. DSG konposatua kendu eta berehala, zelulak PBSrekin garbitu ziren. Ondoren, zelulei %3 formaldehido gehitu zitzaien 10 minutuz 37°C-tan. Formaldehidoaren eragina gelditzeko 100 mM glizina gaineratu zen 5 minutuz 37°C-tan. Plaka PBSrekin garbitu zen 3 aldiz eta zelulak *scrapping buffer* (PBS + %0.5 BSA fraction V) hotzarekin jaso ziren. Zelulak zentrifugatu (5', 4°C, 1.000 g) eta gero izotzetan mantendu ziren. Gainjalkina kendu eta jalkina 1 ml *scrapping buffer* hotzetan birsuspenditu ziren. 2.000 g-tan 5'-tuz zentrifugatu ondoren, jalkina izotz-hotzetan dagoen 500 µl zelula lisi tanpoian (10 mM HEPES pH 7.5, 20 mM KCl,

1.5 mM MgCl2 eta 0.5 mM EDTA. Erabili aurretik gehitu 1 mM tris (2-carboxyethyl) phosphine (TCEP) eta 0.5 mM phenyl methyl sulfonyl fluoride (PMSF)) birsuspenditu zen. Zelulak 10 minutuz inkubatu ziren eta xiringa batez homogeneizatu. Laginak sonikatu egin ziren 5W minutu batez (ziklo bakoitzeko 0.7 segundo ON, 1.3 segundo OFF). *DNase cofactor solution* eta 50U turbo DNasa (AM2238; Invitrogen) gehiturik 37°C-tan inkubatu zen 10 minutuz. Segituan, laginak izotzetan ipini eta DNase stop soluzioa gehiturik, digestioa eten egin zen. Hibridazio tanpoia (20 mM Tris-HCl pH 7.5, 7 mM EDTA, 3 mM EGTA, 150 mM LiCl, %1 NP-40, %0.2 N -lauroylsarcosine, %0.1 sodium deoxycholate, 3 M guanidine thiocyanate eta 2.5 mM TCEP) jarri eta nahastea abiadura maximoan zentrifugatu zen 10 minutuz. Gainjalkina jaso egin zen lisatuak lortuz.

- Lisatuak aurre-garbitu: Estreptabidinaz estalitako bolatxo magnetikoak hibridazio tanpoiarekin garbitu eta lisatuetara gehitu ziren. Bolatxoak lisatuekin 30 minutuz inkubatu ziren 37°C-tara. Inkubazioa amaiturik, bolatxoak kendu ziren eta laginaren %10 gorde zen input bezala.
- Hibridazioa, harrapaketa eta garbiketa: zunda bakoitzeko 50 ng desnaturalizatu ziren 85°C-tan 3 minutuz eta zuzenean izotzetan jarriz. Zundak dagokien lisatuekin inkubatu ziren 2 orduz 37°C-tan jarriz mugimendu konstantean. Inkubazioaren ostean, aldez aurretik garbitutako estreptabidina bolatxo magnetikoak gehitu eta 30 minutuz inkubatu ziren 37°C-tan. Lisatuak iman batean kokaturik, gainjalkinak baztertu ziren. Bolatxoak 6 aldiz garbitu ziren honako konposatuak dituen soluzioarekin: 20 mM Tris-HCl pH 7.5, 10 mM EDTA, %1 NP-40, %0.2 N-lauroilsarkosina, %0.1 sodio deoxikolato, 3 M guanidinio tiozianato eta 2.5 mM TCEP.
- RNA eta DNA berreskuratzea: bolatxoak nukleorako lisi tanpoian (20 mM HEPES pH 7.5, 50 mM KCl, 1.5 mM $MnCl_2$, %1 IGEPAL CA630 (NP-40), %0.4 sodio, 1 mM TCEP eta 0.5 mM PMSF) eluitu eta 5 minutuz inkubatu ziren 94°C-tan biotina-estreptabidina loturak askatzeko. Bolatxoak iman baten bidez harrapatu ziren

eta eluzioa gorde egin zen. Eluzioak ur-beherako esperimentuen arabera alikuota desberdinetan banandu egin ziren.

- RNA eta DNA kuantifikazioa qPCR bidez: RNA *Lnc13* eta *ARGI* geneen purifikazioaren efizientzia testatzeko erabili ziren. RNA *Lnc13*aren purifikazioaren kasuan *STAT1*en 3'UTRaren agerpena aztertzeko ere erabili zen. Eta DNA *ARGI*ren purifikazioaren kasuan, *IFNβ* eta *ISG15* geneen promotore eta sustatzaileen agerpena analizatu zen. Erabilitako hasle guztiak **2.Taulan** zerrendaturik daude.



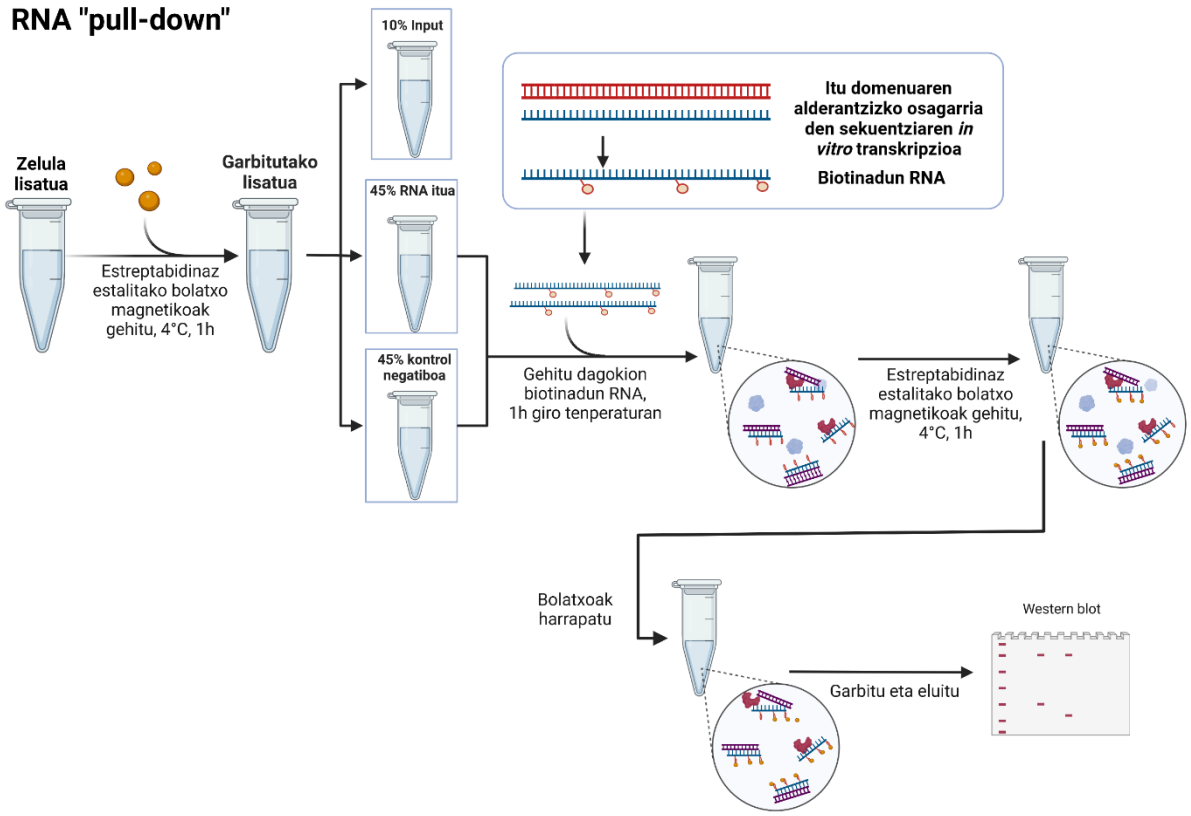
14.Irudia. Alderantzizko RNA osagarrien bitarteko purifikaziorako jarraituriko protokoloa.

2.13 RNA "pull-down" purifikazioa (15.Irudia)

- Biotinadun RNA molekulen sintesia: *STAT1* genearen 3'UTRa giza cDNA batetik anplifikatu zen **1. taulan** zerrendaturiko hasleak erabilita. Lorturiko PCR produktua purifikaturik *in vitro* transkribatu egin zen Takara markako T7 RNA

polymerasa eta rNTPs, eta Sigmako RNA biotin labeling kita erabiliz. Eskuratutako biotinadun RNA “pull-down” egiteko erabili zen.

- RNA “pull-down”: zelula ituak RIP tanpoiarekin (150 mM KCl, 25 mM Tris, 0.5 mM DTT, %0.5 NP-40, proteasa inhibitzaileak) lisatu ziren. Lisatuak Streptavidin Mag Sepharose bolatxo magnetikoak (GE Healthcare) erabiliz aurre-garbitu ziren, ordu batez 4°C-tan gurpil batean mugituz. Lisatuak 3 tutu desberdinetan banandu ziren: %10 inputerako, %45 3’UTR-*STAT1*-erako eta gainerako %45 kontrol negatiboa egiteko. Lisatuen proteina kontzentrazioa kuantifikatu ostean, 1 mg lisatu proteina aldez aurretik prestaturiko biotinadun 3’UTR-*STAT1* edo kontrol negatiboaren 1µg-rekin nahastu zen. Nahastea ordubetez inkubatu zen giro tenperaturan. Ostera, Streptabidina bolatxo magnetikoak gehitu eta beste ordubetez inkubatu zen bolatxoak biotinadun RNArekin (eta honetara elkarturiko edozer) lotu ahal izateko.
- Bilketa: Bolatxo magnetikoak iman baten bitartez harrapatu eta garbitu egin ziren; 3 aldiz RIP tanpoiarekin, beste 3 aldiz gatz-altuko tanpoiarekin (500 mM NaCl, 10 mM Tris-HCl, %0.1 NP-40) eta amaitzeko beste 3 aldiz gatz-baxuko tanpoiarekin (50 mM NaCl, 10 mM Tris-HCl, %0.1 NP-40). Azkenik, bola magnetikoak Laemmli tanpoiarekin nahastu ziren PCBP2 proteinaren mailak Western blot bidez aztertzeko.



15.Irudia. RNA "pull-down" esperimentuaren eskema.

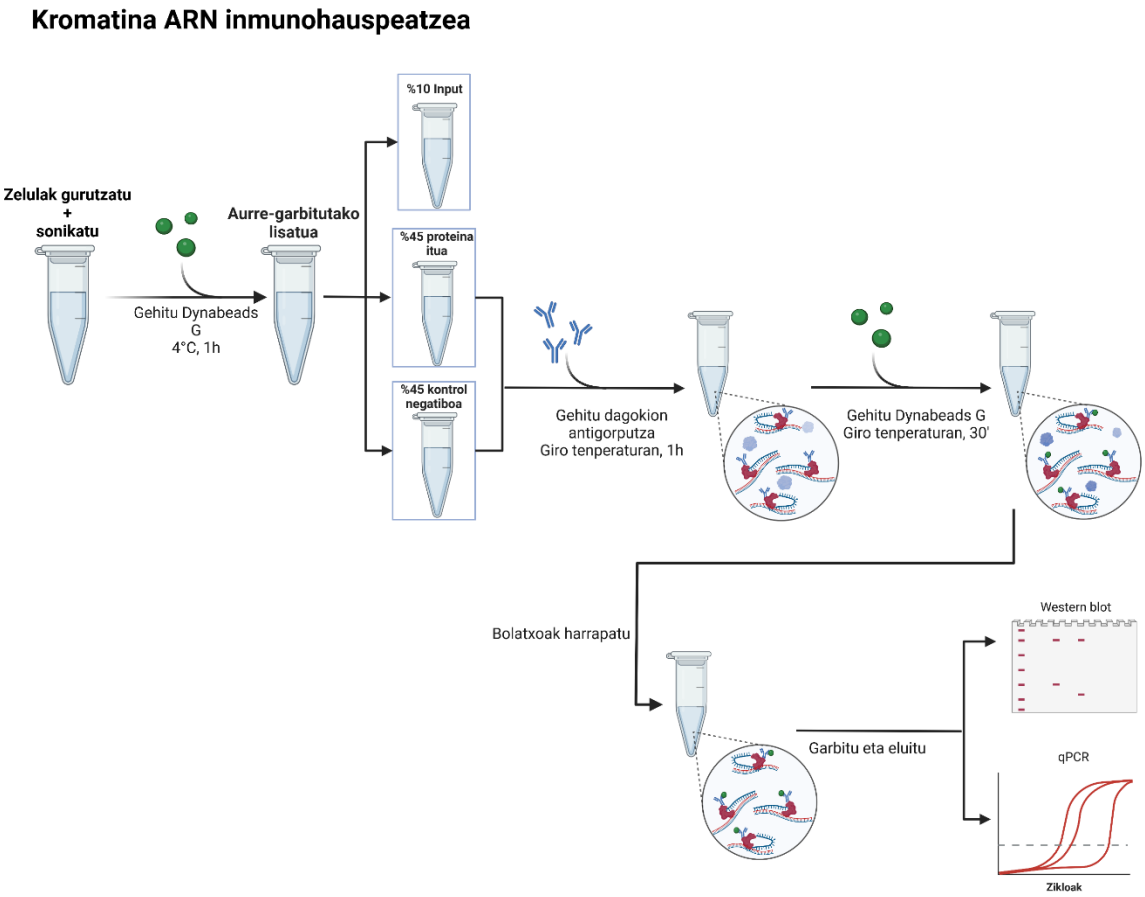
2.14 Kromatina immunohaupeatzea (16.Irudia)

- Zelulen *crosslink*-a eta bilketa: zelulak %1era dagoen formaldehidoaren 500 µl-tan jaso eta 5 minutuz giro tenperaturan inkubatu ziren. Ondoren, zelulak hauspeatu eta PBS hotzarekin garbitu ziren birritan. Jalkinak 500 µl NLS tanpoiarekin (NLS; *Nuclear Lysis Solution*: 20 mM HEPES pH 7.5, 50 mM KCl, 1.5 mM $MnCl_2$, % 1 IGEPAL CA630 (NP-40), %0.4 sodio; gehi 1 mM TCEP eta 0.5 mM PMSF momentuan gehituta) lisatu ziren 10 minutuz izotzetan jarrita. Lisatuak sonikatu egin ziren 5Wtan 10 ziklo eginda (30'' ON, 30'' OFF). Lisatuak aurre-garbitzeko, G Dynabeads magnetikoak RIP tanpoiarekin (150 mM KCl, 25 mM Tris, 0.5 mM DTT, % 0.5 NP-40, proteasa inhibitzaileak) garbitu eta bolatxo magnetiko hauen 10 µl lisatu bakoitzera gaineratu zen ordu batez 4°C-tan.
- Hibridizazioa, harrapaketa eta garbiketa: Aurre-garbituriko lisatuen %10a inputerako gorde zen. Gainerako %90 bitan banandu zen; erdi bat intereseko proteina immunohauspeatzeko (STAT1 *Lnc13* karakterizazioaren kasuan eta

CTCF *ARGI*ren kasuan) eta beste erdia kontrol negatiboa (IgG) egiteko. Dagokien antigorputza (**3.Taula**) kasuan kasuko lisatura gehitu eta ordubetez inkubatu zen giro tenperaturan higidura konstantean. Gero, Dynabeads bolatxoen 20 µl garbitu eta lagin bakoitzean gaineratu zen 30 minutuz mugimenduan egon zitezen bolatxo eta antigorputzen arteko loturak baimenduz. Denbora igarota, bolatxoak 3 aldiz garbitu ziren RIP tanpoiarekin, 3 aldiz gatz-baxuko tanpoiarekin (50 mM NaCl, 10 mM Tris-HCl, %0.1 NP-40) eta 3 aldiz gatz-altuko tanpoiarekin (500 mM NaCl, 10 mM Tris-HCl, %0.1 NP-40).

- RNA, DNA eta proteina biltzea: Bolatxoak NLS tanpoian eluitu eta 10 minutuz inkubatu ziren 94°C-tan molekulak G Dynabeads bolatxoetatik askatzeko. Bolatxoak iman baten bidez erakarri ziren eluitua gorde ahal izateko. Eluitua alikuota desberdinetan banandu zen gerora egin beharreko esperimentuen arabera.
- RNA eta DNA kuantifikazioa qPCR-en bidez: Aztertzeko *ARGI* CTCF proteinara loturik zegoen, *ARGI* lncRNAren kuantifikazioa egin zen CTCF proteinara elkarturiko RNA frakzioa erabiliz. Froga bera egin zen konprobatzeko *Lnc13* STAT1 proteinara lotzen zela.

  CTCF proteinara elkarturiko DNA qPCR bidez analizatu zen CTCFra loturiko kromatinan *IFNβ* eta *ISG15* geneen sustatzaile eta sekuentzia areagotzaileak agertzen ziren zehazteko. STAT1 proteinaren kasuan *CXCL10* sustatzailearen presentzia aztertu zen qPCR-en bitartez (**hasleak 2.Taulan**).



16.Irudia. Kromatina immunohauspeatze esperimentuaren protokoloren eskema.

2.15 RNA-proteina elkarrekintzaren esperimentua

- RNA sintesia eta tolestura: *STAT1*en 3'UTRa anplifikatzeko, T7-3'UTR-STAT1 izeneko hasle parea erabili zen (**1.Taula**). Anplifikaturiko domeinua T7 polimerasaren sistemaren bidez *in vitro* transkribatu zen aurretik azaldu bezala. RNA bigarren mailako egitura hartzea baimentzeko, RNA egitura tanpoia (10 mM Tris pH 7, 0.1 M KCl, 10 mM $MgCl_2$) gehitu eta lagina jarraian ageri diren tenperatura eta denboratan ipini zen: 2 minutuz 90°C-tan berotu, beste 2 minutu izotzetan, eta azkenik 20 minutuz giro tenperaturan.
- RNA-proteina elkarrekintza: tolesturiko RNAren 140 ng proteina lisatuekin nahastu zen eta ordubetez giro tenperaturan inkubatu zen elkarrekintzak sortzea baimenduz.
- Gel elektroforesia: laginak TB tanpoian egindako %1 agarosa gelean kargatu ziren. RNA bandak GelRed® Nucleic Acid Gel Stain (BT-41003-T; Biotium)

tindatzailea erabiliz tindatu ziren. Proteinak nitrozelulosazko mintz batera transferitu ziren eta proteina ituaren (PCBP2) kontrako antigorputzarekin inkubatu zen proteina detektatu ahal izateko.

2.16 Estatistika

Datu esperimentalak grafikatu egin dira batezbesteko±SEM eran, figuren legendetan adierazita dagoen bezala. Baldintza desberdinen artean desberdintasun esanguratsuak dauden aztertzeko Student's t testa egin da edo ANOVA, Student's t testa batez jarraituta eta Bonferroni metodoaren bidez zuzendurik. Grafikoak GraphPad Prism v8.0.1 softwarearen bidez egin dira.

1.taula. DNA anplifikatzeko erabili ziren hasle-sekuentzien zerrenda.

Gene itua	Sekuentzia
Lnc13-overexpression	Fw: AAGGATCATTGCAGGGTCTC Rv: GTGGCCAAAAGAAGTCTGAGTC
*Lnc13*_delSNP	Fw: GCCTTTGATTTCCTGGACTG Rv:TTAAAACCCGAAAAGGACCA
*Lnc13*_KO_sg1	Fw: CACCGAACTCCTGACCTCAGGAGAT Rv: AAACATCTCCTGAGGTCAGGAGTTC
*Lnc13*_KO_sg2	Fw: CACCGTCTGAAAAAGTGTCCTACCT Rv: AAACAGGTAGGACACTTTTTCAGAC
T7-3'UTR-*STAT1*	Fw: TAATACGACTCACTATAGGTGAAACTTAGGTTCTCGCCATC Rv: TTCACATTTGCGAATGGTTC
ARGI _mut_rs9585056:	Fw: AAAAACATTGAGACCAATGCGAGTTT Rv: GCCCTTGCAAATGTTATTTCTGAGCC
ARGI_KO_sg1	Fw: CACCGCTGTAGGGACGTCTTTCCG Rv: AAACCGGAAAGACGTCCCTACAGC
ARGI_KO_sg2	Fw: CACCGGGATCCTTCCAAAATTGACA Rv: AAACTGTCAATTTTGGAAGGATCCC

ARGI_KO_sg3	Fw: CACCGGCCAGTCCCCGATCAGTGTA Rv: AAACTACACTGATCGGGGACTGGCC
ARGI_CRISPRi sgRNA	Fw: GATCGCCGGGGATTCCCAGTTCCCC Rv: AAAACGGGAACTGGGAATCCCCGG

2.Taula. qPCR esperimentuak egiteko erabilitako hasle eta zunden zerrenda.

Sybr Green sistema	
Gene itua	**Sekuentzia**
Lnc13	Fw: AAGGATCATTGCAGGGTCTC Rv: GTGGCCAAAAGAAGTCTGAGTC
MALAT1	Fw: GCTGTGGAGTTCTTAAATAT Rv: TTCTCAATCCTGAAATCCCC
RPLP0	Fw: GCAGCATCTACAACCCTGAAG Rv: CACTGGCAACATTGCGGAC
3'UTR of *STAT1*	Fw: TGAAACTTAGGTTCTCGCCATC Rv: TTCACATTTGCGAATGGTTC
CXCL10 promoter	Fw: TGGATTGCAACCTTTGTTTTT Rv: GTCCCATGTTGCAGACTCG
ISG15 promoter	F1: TCCCTGTCTTTCGGTCATTC R1: ACGGCACAAGCTCCTGTACT
ISG15 proximal enhancer	F2: CACCTGAAGCAGCAAGTGAG R2: CTTTATTTCCGGCCCTTGAT
IFNβ promoter	F1: TCCCACTTTCACTTCTCCCT R1: GCTTTCCTTTGCTTTCTCCCA
IFNβ proximal enhancer	F2: GGGTGGGATGGAGAACTCAG

	R2: ACTTTTCTGTTGTTTGGTCTTGT
IFNβ distal enhancer	F3: GAGAACTCCTGCCCAGAGG R3:AGCACCTCAAGAACACAATAGC

TaqMan sistema		
Gene itua	**Katalogo zenbakia**	**Markaren izena**
ARGI	Custom assay for *ARGI*	(IDT)
MX1_PrimeTime qPCR Assay	Hs.PT.58.26787898	(IDT)
IFIT1_PrimeTime qPCR Assay	Hs.PT.56a.2076909	(IDT)
IFIT3_PrimeTime qPCR Assay	Hs.PT.58.20456374	(IDT)
IFI6_PrimeTime qPCR Assay	Hs.PT.58.4390209	(IDT)
*Lnc13*_TaqMan® Gene Expression Assay	4332078	(ThermoFisher)
MALAT1_TaqMan® Gene Expression Assay	Hs00273907_s1	(ThermoFisher)
RPLP0_TaqMan® Gene Expression Assay	Hs99999902_m1	(ThermoFisher)
CXCL10_TaqMan® Gene Expression Assay	Hs00171042_m1	(ThermoFisher)
CXCL9_TaqMan® Gene Expression Assay	Hs00171065_m1	(ThermoFisher)
CCL5_TaqMan® Gene Expression Assay	Hs00982282_m1	(ThermoFisher)
CXCL1_TaqMan® Gene Expression Assay	Hs00605382_gH	(ThermoFisher)
STAT1_TaqMan® Gene Expression Assay	Hs01013996_m1	(ThermoFisher)
ISG15_TaqMan® Gene Expression Assay	Hs00192713_m1	(ThermoFisher)
IFNβ_TaqMan® Gene Expression Assay	Hs01077958_s1	(ThermoFisher)
Actinaβ_TaqMan® Gene Expression Assay	Hs01060665_g1	(ThermoFisher)
MEG3_TaqMan® Gene Expression Assay	Hs00292028_m1	(ThermoFisher)
RPLP0_TaqMan® Gene Expression Assay	Hs99999902_m1	(ThermoFisher)

3.Taula. Western blot esperimentuak egiteko erabilitako antigorputzen zerrenda.

Proteina itua	Katalogo zenbakia	Markaren izena	Erabiltzeko diluzioa
STAT1	sc-346	Santa Cruz Biotechnologies	1:1.000 %5 BSAn
pSTAT1	#7649	Cell Signaling Technology	1:1.000 %5 BSAn
α-tubulina	#T9026	Sigma-Aldrich	1:5.000 %5 BSAn
Hsp90	# 4877	Cell Signaling Technology	1:1.000 %5 BSAn
HDAC1	#4874	Abcam	1:500 TBSTn
H3	#4499	Cell Signaling Technology	1:500 TBSTn
PCBP2	#83017	Cell Signaling Technology	1:500 %5 esne gaingabetuan
Normal mouse IgG	sc-2025	Santa Cruz Biotechnologies	Esperimentuaren arabera
CTCF	#PA5-17143	Invitrogen	1:1.000 %5% BSAn
GAPDH	sc-365062	Santa Cruz Biotechnologies	1:5.000 %5 BSAn
HRP-conjugated anti-mouse	sc-516102	Santa Cruz Biotechnologies	1:5.000 %5 esne gaingabetuan
HRP-conjugate mouse anti-rabbit	sc-2357	Santa Cruz Biotechnologies	1:10.000 %5 esne gaingabetuan

1 Experimental models

1.1 Human biological samples

cDNA samples from human pancreatic islets were obtained from two centers, namely Cisanello University Hospital, Pisa, Italy (directly from Piero Marchetti or via the ULB Center for Diabetes Research, Brussels, Belgium thanks to Decio L. Eizirik) and Andalusian Center for Molecular Biology and Regenerative Medicine-CABIMER, Seville, Spain (from Benoit Gauthier).

For the cDNA samples provided by Eizirik and Marchetti, human islet isolation from non-diabetic organ donors (n=36 samples; age: 72 ± 11 years; BMI: 25 ± 3 kg/m2) was performed as described (Marchetti et al., 2007) and in accordance with the local Ethical Committee in the University of Pisa, Italy. For cDNA samples provided by Gauthier, human islets were purchased from Tebu-Bio (n=7 samples; Barcelona, Spain) and donor characteristics described elsewhere (Cobo-Vuilleumier et al., 2018). Experiments using human islets were approved by the Ethical Committee of the Andalusian Regional Ministry of Health, Spain (authorization number 2013-04398).

1.2 In vitro experimental cell models

Most of the *in vitro* functional experiments were performed in the immortalized human pancreatic β cell line named EndoC-βH1. Some experiments were also performed in induced pluripotent stem cell (iPS)-derived human β cells, HEK293FT, SHSY5Y, HCT15, NCI-H23 and Jurkat.

All cell lines were free of Mycoplasma, as evaluated by the MycoAlert Mycoplasma Detection kit (Lonza). For the prevention of Mycoplasma contamination, Plasmocin Prophylactic (Invivogen) was added to the culture medium on a regular basis.

1.2.1 EndoC-βH1

EndoC-βH1 pancreatic β cell line was purchased from Human Cell Design (https://www.humancelldesign.com). Briefly, EndoC-βH1 cells were seeded at an approximate density of 70.000-75.000 cells/cm^2 on culture plates pre-coated with Matrigel-fibronectin (100 mg/ml and 2 mg/mL, respectively) (Sigma-Aldrich) for no

more than 4 hours at 37°C, 5% CO_2 in complete OPTIβ1 medium (Human Cell Design). Cells were passed every 7 days.

For transfection protocols a DMEM media containing 2% FBS, 5.6 mM glucose, 50 µM 2-mercaptoethanol (Bio-Rad Laboratories), 10 mM nicotinamide (Calbiochem), 5.5 µg/ml transferrin, 6.7 ng/ml selenite (Sigma-Aldrich) was used.

1.2.2 iPS-derived human β cells

iPS-derived human β cells were obtained thanks to a collaboration with Dr. Mariana Igoillo-Esteve and Dr. Miriam Cnop (ULB Center for Diabetes Research, Belgium). Briefly, two human iPS lines (HEL115.6 and 1023A) derived from healthy donors were cultured and differentiated into β cells as described previously [HEL115.6 iPSs were generated at the University of Helsinki; 1023A iPSs were kindly provided by Dr. DM Egli (Columbia University)](Cosentino et al., 2018). The iPSs had normal karyotype, stem cell colony morphology and expressed pluripotency markers (Cosentino et al., 2018). iPSCs were maintained in E8 medium on Matrigel-coated plates (Corning) and seeded at 2.5–2.8 × 10^6 cells per 3.5 cm well in E8 medium containing 5 µM ROCK inhibitor (STEMCELL Technologies) 24h prior to the 7-stage differentiation. When reaching pancreatic progenitor stage, cells were plated into 24 well microwell plates at a density of 750 cells/microwell (AggreWell, STEMCELL Technologies) with 10 µM ROCK inhibitor and 1µl/ml heparin (STEMCELL Technologies) to allow 3-dimensional formation of aggregates.

Prior to infections, stage 7 aggregates were dispersed. Aggregates were incubated in 0.5 mM EDTA at room temperature for 4 min, exposed to Accumax (Sigma–Aldrich) for 8 min and then dispersed by gentle pipetting (Nair et al., 2019). Knockout serum (Gibco) with 10 µM ROCK inhibitor was added to quench the dissociation process, and cells were seeded at 5 × 10^4 cells per 6.4 mm well in stage 7 medium (Cosentino et al., 2018).

1.2.3 HEK293FT

HEK293FT cells (CRL-1573) were purchased from the American Type Culture Collection (ATCC; https://www.atcc.org). Cells were cultured in DMEM supplemented with 10%

FBS and 100 units/ml penicillin and 100 µg/ml streptomycin (Lonza). The same medium without antibiotics was used for transfection experiments.

1.2.4 Jurkat

Jurkat cells (ATCC, TIB-152) were cultured in RPMI supplemented with 10% FBS and 100 units/ml penicillin and 100 µg/ml streptomycin (Lonza). Non-supplemented RPMI medium was used for migration assays.

1.2.5 SH-SY5Y

SH-SY5Y cells (ATCC; CRL-2266) were cultured in a mix of a half of RPMI medium plus a half of MEM Eagle EBSS medium supplemented with 10% FBS and 100 units/ml penicillin and 100 µg/ml streptomycin (Lonza).

1.2.6 HCT15

HCT15 cells (Sigma-Aldrich, #91030712) were cultured in RPMI supplemented with 10% FBS and 100 units/ml penicillin and 100 µg/ml streptomycin (Lonza).

1.2.7 NCI-H23

NCI-H23 (ATCC, CRL-5800) were cultured in RPMI supplemented with 10% FBS and 100 units/ml penicillin and 100 µg/ml streptomycin (Lonza).

2 Methods

2.1 Cell treatments

For cell treatments, the selected compounds were directly added to the corresponding culture cell medium.

- Cytokine treatment: EndoC-βH1 cells were treated for 24h with IL-1β plus IFNγ at a final concentration of 10 U/ml and 1000 U/ml, respectively.
- STAT signaling pathway inhibitor: The treatment with the JAK inhibitor Ruxolitinib (Cayman Chemicals) was performed adding the compound to the culture medium at a final concentration of 5 µg/ml as previously described (Coomans de Brachène et al., 2018).

- NFκB signaling pathway inhibitor: The treatment with the NFκB inhibitor Bay 11-7082 (Sigma-Aldrich) was performed by the addition of the compound in the cell culture medium at a final concentration of 10 μM.
- Translation and transcription inhibitors: EndoC-βH1 cells were treated with actinomycin D (Sigma-Aldrich) at a final concentration of 5 μg/ml for 6h or with cycloheximide (Sigma-Aldrich) at a final concentration of 20 μg/ml for 24h.

2.2 Coxsackie virus infection

Viral infections were performed in both EndoC-βH1 and iPS-derived β cells.

The viral infection in EndoC-βH1 was performed using the CVB5/Faulkner virus. Cells were infected with a dilution of CVB5/Faulkner (Multiplicity of Infection (MOI)=5) on Hanks' Balanced Salt Solution (HBSS, Invitrogen). After 1h adsorption at 37°C, the inoculum virus was removed and cells were washed twice with HBSS. Then, culture medium was added and the virus allowed to replicated for another 24h.

For viral infection in iPS-derived β cells, firstly cells were dispersed using Accumax (Sigma) as previously described (Cosentino et al., 2018; Lytrivi et al., 2021). Cells were plated at a cell density of 300,000/well in Matrigel-coated 24-well plates in stage 7 medium (Cosentino et al., 2018) supplemented with 10 μM ROCK inhibitor. After overnight recovery, cells were infected with CVB1/Conn-5 (MOI=0.05) and CVB4/JVB (MOI=0.5) in Ham's F-10 Nutrient Mixture (Gibco), supplemented with 2 mM GlutaMAX, 50 μM IBMX and 1% FBS. Two hours after infection medium was replaced by Ham's F-10 Nutrient Mixture, 0.75% BSA, 2 mM GlutaMAX, 50 μM IBMX, 50 U/ml penicillin, 50 μg/ml streptomycin and cells were cultured for an additional 22h as previously described.

2.3 Plasmid construction and transformation into DH5α E.coli.

- Plasmid construction (**Figure 8**):

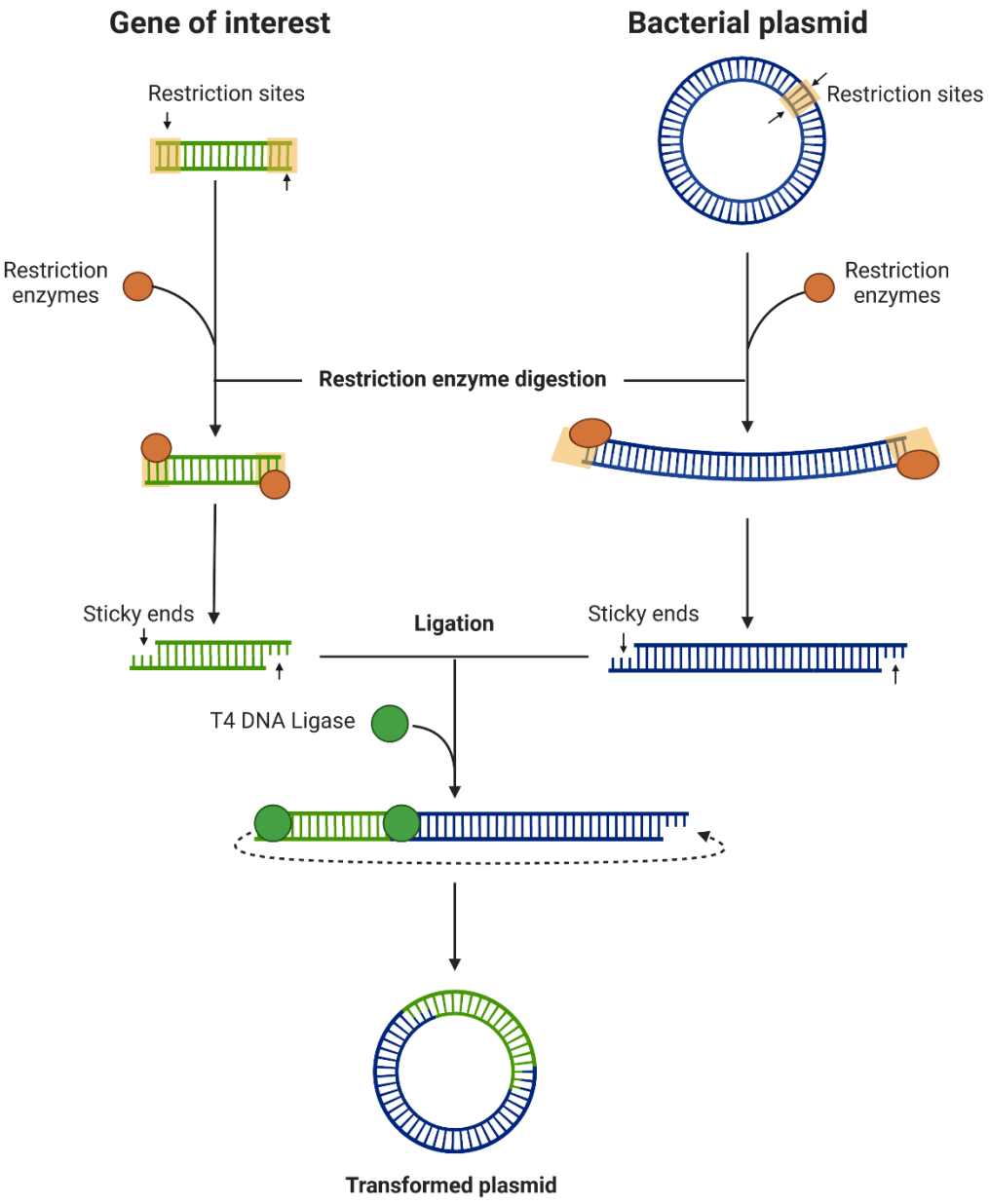


Figure 8. Schematic protocol followed for plasmid construction.

For the synthesis of pLnc13-C and pLnc13-T plasmids, *Lnc13* harboring rs917997-C or T alleles were amplified from human cDNA using a specific pair of primers (*Lnc13*-overexpression; **Table 1**). Amplified *Lnc13* sequences and pCMV6 plasmid (Cat# PS100001; Origene) were digested using FseI (R0588; NEB) and KpnI (R0142; NEB) restriction enzymes following manufacturer's restriction digestion protocols. Each

Lnc13 sequence was cloned into the already digested pCMV6 vector using T4 DNA ligase (M0202; NEB).

Lnc13-delSNP plasmid was constructed by amplifiying the fragments of *Lnc13* ranging from base 1 to 1771 and from 2278 to 2545 using overlapping primers to allow the fusion of both fragments into a unique one (Sequences on **Table 1.** ***List of the primers employed for DNA amplification***; *Lnc13*_delSNP). This *Lnc13* mutated sequence and pCMV6 plasmid were digested by FseI and KpnI restriction enzymes following manufacturer's restriction digestion protocols, and the *Lnc13* mutated sequence was cloned into the digested pCMV6 vector.

For *Lnc13* CRISPR-Cas9 experiments sgRNAs were designed to delete a fragment of 1698 bp using a pair of sgRNAs (**Table 1**). Each sgRNA was cloned into a px459 vector (62988; Addgene) using BbsI restriction enzyme (R0539; NEB) (**Figure 9**).

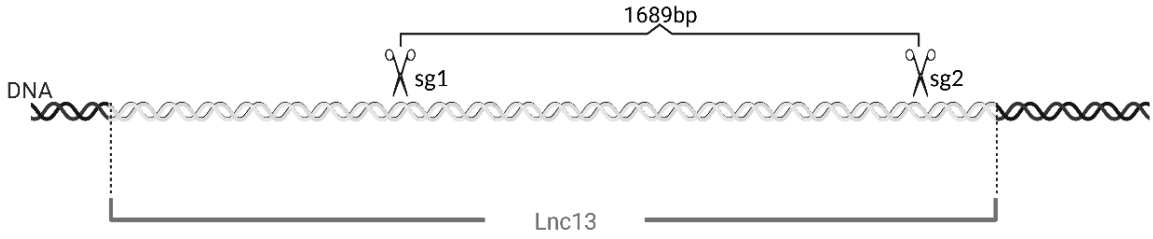


Figure 9. Schema of *Lnc13* disruption using CRISPR-Cas9 technology. A deletion of 1698 bp was performed using pair of guide RNAs (Lnc13_KO_1 and Lnc13_KO_2).

The overexpression vector for *ARGI* harboring the T1D risk allele (ARGI-R; rs9585056-C) was purchased from ProteoGenix. The overexpression vector for *ARGI* harboring the T1D protective allele (ARGI-P; rs9585056-T) was produced by site-directed mutagenesis using the Site-Directed Mutagenesis QuickChange II (Agilent). The primers used for the mutagenesis are listed on **Table 1** (ARGI _mut_rs9585056).

For CRISPR-Cas9 experiments on *ARGI,* sgRNAs were designed to generate a deletion of 5079 bp using a combination of sgRNA_1 and sgRNA_2 (**Table 1**). A bigger deletion

of 5387 bp was also generated combining the sgRNA_1 with the sgRNA_3 (**Table 1**). Each sgRNA was cloned into a px330 vector (Addgene) using BbsI restriction enzyme (R0539; NEB) (**Figure 10**).

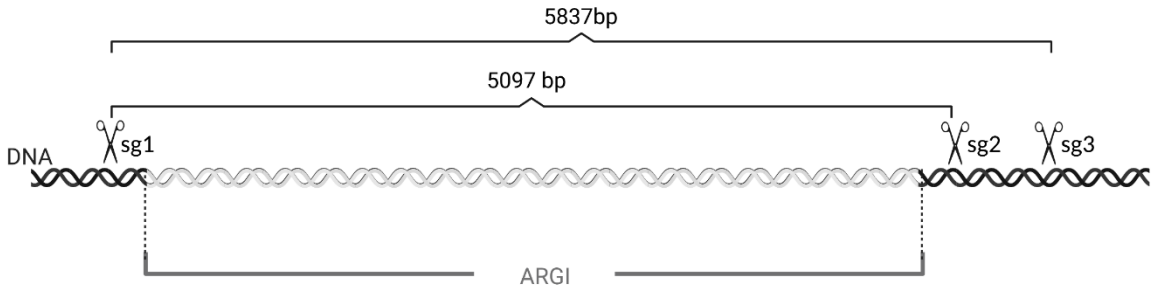


Figure 10. Schema of ARGI disruption using CRISPR-Cas9 technology. A deletion of 5097 bp using single guide RNAs (sgRNAs) 1 and 2, or a deletion of 5837 bp using sgRNAs 1 and 3.

For CRISPRi experiments, a sgRNA (ARGI_CRISPRi sgRNA, **Table 1**) was cloned into pCRISPRi-Cas-Guide-CRISPRi vector (GE100059; Origene) using BamHI (R0136; NEB) and BsmBI-v2 (R0739; NEB) restriction enzymes.

- **Plasmid transformation into competent *E.*coli (Figure 11):**

In order to amplify and select the recombinant plasmids of interest, ligation reactions of each plasmid were transformed into DH5α *E.coli* bacteria. DH5α cells were mixed with 50 ng plasmid and kept on ice for 20'. After a heat-shock (42°C for 30'') to permeabilize the bacteria, they were put on ice for 2' and mixed with 300 µl SOC medium. Bacteria was grown on SOC medium for 1 hour at 37°C. After, cells were pelleted at 6.000 rpm for 5', resuspended in 100 µl SOC medium and directly plated on agar plates containing the corresponding antibiotic (depending on the plasmid containing resistance gene; normally Kanamycin or Ampicillin). Cells were allowed to grow overnight.

- **Plasmid testing:**

Next morning cell growth was checked. Only bacteria with the plasmid inserted (with the antibiotic resistance) must have grown. Some colonies were taken and cultured individually in LB broth medium plus the corresponding antibiotic

overnight. The plasmid was purified from the bacteria by using the kit NucleoSpin Plasmid EasyPure (Macherey-Nagel). To check whether the purified plasmids contained with the sequences of interest, they were sequenced by Sanger.

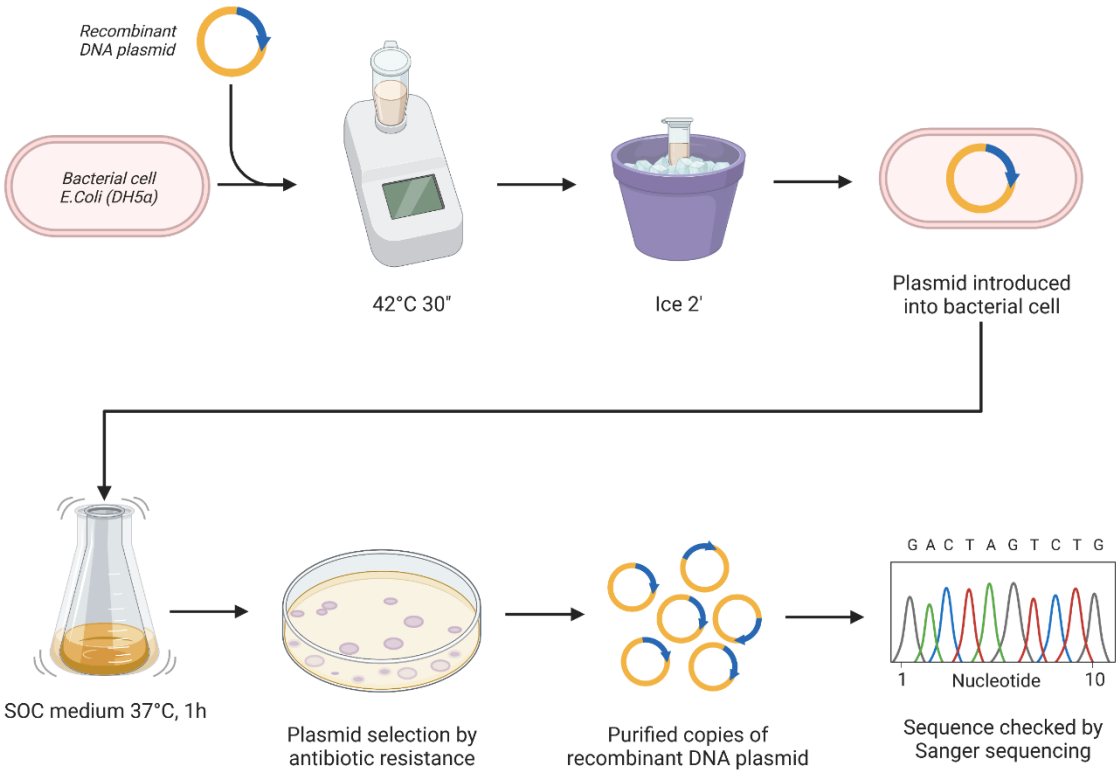


Figure 11. Plasmid transformation into DH5α E.coli.

2.4 Transfections

- Transfection of polyinosinic:polycytidylic acid (PIC): EndoC-βH1 cells were seeded and transfected using Lipofectamine-2000 reagent (Invitrogen) following manufacturer's instructions. The synthetic viral dsRNA mimic polyinosinic:polycytidylic acid (PIC) (InvivoGen) was used at a final concentration of 1 µg/ml.
- Transfection of overexpression and CRISPR vectors: EndoC-βH1 cells were seeded and transfected with 250ng of each overexpression plasmid using

Lipofectamine-2000 (Invitrogen) following the manufacturer's instructions. Transfection mix was left for 16h, and afterwards, it was replaced with fresh medium to allow cell recovery for another additional 24 hours before downstream experiments were performed.

Transfection of CRISPRi plasmids was performed using Lipofectamine-2000 (Invitrogen). After transfection, cells were incubated for 36h before downstream experiments were performed.

CRISPR-Cas9 vectors were transfected using Lipofectamine-2000 (Invitrogen) and left for 24h before puromycin (2µg/ml) was added for 36h. The plasmid named pSpCas9(BB)-2A-Puro (Px459) V2.0 (#62988; Addgene) has puromycin resistance, and thus, puromycin addition allowed to select recombinant EndoC-βH1 cells.

2.5 RNA and DNA extraction

RNA extraction was performed using the NucleoSpin RNA Kit (Macherey Nagel) or PureLink RNA Mini kit (Invitrogen).

For DNA extraction, the NucleoSpin Gel and PCR Cleanup is employed (Machery Nagel).

- **RNA expression determination**: The expression values were determined by quantitative PCR (qPCR) using specific Taqman Gene Expression Assays (Thermo Scientific) or Sybr Green (Biorad) using specific primers (**Table 2**). All qPCR measurements were performed in duplicates in at least 3 independent samples in the CFX384 Touch Real-Time PCR Detection System (Bio-Rad Laboratories) and expression levels were analyzed using the $2^{-\Delta\Delta}$Ct method.
- **SNP genotyping**: The genotyping of the SNP rs917997 in *Lnc13* was performed using the TaqMan C____345197_1_genotyping assay (Thermo Scientific) and for the SNP rs958056 located in *ARGI,* a rhAmp SNP Assay from IDT (Hs.GT.rs9585056.T.1).

2.6 Quantification of Lnc13 molecules per cell

In order to determine the copy number of *Lnc13* in human cell lines, a reference plasmid was generated incorporating the cDNA sequence of the target gene in a PCR blunt vector. Absolute quantification was performed using ten 2-fold serial dilutions of the reference standard. Ct versus the dilution factor was plotted in a base-10 semi-logarithmic graph, fitting the data to a straight line. Plot was then used as a standard curve for extrapolating the number of molecules of the *Lnc13* in the following cell lines: EndoC-βH1, HEK293, SHSY5Y, HCT15 and NCI-H23.

2.7 Subcellular fractionation

1. **Subcellular RNA fractionation:**

- Nuclear fraction isolation: Cell pellet was resuspended carefully in a mixture of 600 µl RNAse free water + 200 µl PBS + 200 µl C1 lysis buffer (composed by: 1.28 M sucrose, 40 mM Tris-HCl pH 7.5, 20 mM MgCl2, 4% Triton X-100). The mix was put on ice for 15' and subsequently centrifuged at 2,500 rpm for 15' at 4°C. The supernatant was fully discarded and the pellet was processed for RNA extraction. The RNA extracted from this pellet corresponded to nuclear RNA.
- Quantification: The amounts of the gene of interest, *MALAT* or *MEG3* (as nuclear control) and *RPLP0* (as cytoplasmic control) were measured by qPCR and compared to the total amount of those RNAs in the whole cell lysate. Results were expressed the logarithm of nuclear/whole ratio. The primers employed are listed in **Table 2**.

2. **Subcellular protein fractionation:**

   The EndoC-βH1 protein fractionation was performed using the commercial subcellular protein fractionation kit (78840; Thermo Scientific). The success of the fractionation efficiency was checked evaluating the expression of the following proteins by Western blot: Hsp90 for cytoplasm, HDAC1 for nuclei and H3 for chromatin fraction.

2.8 Western blot analysis

- Cell lysate electrophoresis: Cells were washed with cold PBS and lysed in Laemmli buffer (62 mM Tris-HCl, 100 mM dithiothreitol (DTT), 10% glycerol, 2% SDS, 0.2 mg/ml bromophenol blue, 5% 2- mercaptoethanol). Proteins were separated by SDS-PAGE on 10% gels.
- Protein transference and antibody incubation: Following electrophoresis, proteins were transferred into nitrocellulose membranes using a Transblot-Turbo Transfer System (Biorad) and blocked in 5% non-fatty milk diluted in TBST (20 mM Tris, 150 mM NaCl and 0.1% Tween 20) at room temperature for 1h. The membranes were incubated overnight at 4°C with primary antibodies (**Table 3**). Membrane was washed 3 times with PBS-T and incubated during 1 hour at room temperature with a horseradish peroxidase-conjugated anti-mouse (1:5,000 dilution in 5% non-fatty milk) or anti-rabbit (1:10,000 dilution in 5% non-fatty milk) secondary antibody.
- Reveal: Immunoreactive bands were revealed using the SuperSignal West Femto chemiluminescent substrate (Thermo Scientific), detected using a Bio-Rad Molecular Imager ChemiDoc XRS and quantified using the ImageLab software (Bio-Rad Laboratories).

2.9 ELISA

Supernatant fractions from EndoC-βH1 cells were collected for determination of human CXCL10 and CCL5 proteins using commercially available ELISA kits (R&D Systems, Abingdon, UK; Cat# DY266-05 and Cat# DY278-05, respectively). Briefly, the ELISA was performed as follows:

- Plate preparation: A 96-well microplate was coated with 100 µl of the capture antibody diluted in PBS at a final concentration of 1µg/ml and let overnight at room temperature. Next morning, each well was aspirated and washed with 400 µl washing buffer three times. The plate was then blocked by adding 300 µl Reagent diluent to each well and incubating for 1 hour at room temperature. Finally, the wells were washed three times with the wash buffer.

- Detection procedure: 100 µl of each sample or standards mixed with the Reagent diluent were added to each well, after the plate was covered it was incubated for 2h at room temperature. The wells were washed with washing buffer, and after, 100 µl of Detection antibody (20 µg/ml) was added and incubated for 2h at room temperature. After washing, 100 µl working dilution of Streptavidin-HRP was added to each well and incubated for 20 minutes at room temperature protected from light. After, 100 µl of Substrate solution was incubated for 20' protected from light. Finally, 50 µl Stop solution was added to each well. The optical density of each well was determined using a plate reader at a wavelength of 450nm after subtracting the reading at 540nm.

2.10 Chemotaxis assay (Figure 12)

- *In vitro* chemotaxis model set-up: Chemotaxis experiments were performed using an adaptation of Boyden chamber system (Falasca et al., 2011). Briefly, supernatants were collected from EndoC-βH1 cells previously stimulated or transfected with the corresponding treatments (overexpression of *Lnc13*-C or PIC transfection). 500 µL supernatant were placed individually in the bottom of the wells of a 24-well plate. Permeable PET-membrane supports for 24-well plates (#353095; Lonza) were placed just over the supernatant-containing wells. Inside these membrane supports, previously activated Jurkat cells were placed. Jurkat cell activation was performed growing the cells without FBS for 16h and adding 50 mg/ml PMA (Sigma-Aldrich) and 1 µM ionomycin (Sigma-Aldrich) for 2h. The system was kept at 37°C for 2 hours to allow Jurkat cell migration from the upper to the bottom chamber through the membrane.
- Sample collection: The media inside the permeable PET-membrane supports was discarded. The supports were moved to the Cell Harvesting Tray containing 150 µl of pre-warmed (37°C) Cell Detachment Solution (CDS) and incubated for 30' at 37°C to allow cells to detach. The solution with the detached cells was

mixed with 150 µl supernatant from the lower part of the chamber system (where the migrated cells were located).

- Measurement: 200 µl of the previous mixture was added to an opaque 96 well plate. In addition, 200 µl of the CyQUANT GR dye/cell-lysis buffer was incorporated to each well. The plate was mixed gently and incubated for 20' at room temperature protected from the light. The fluorescence of each condition was measured in a fluorescence reader at a wavelength of 480nm for excitation and 520nm for emission.

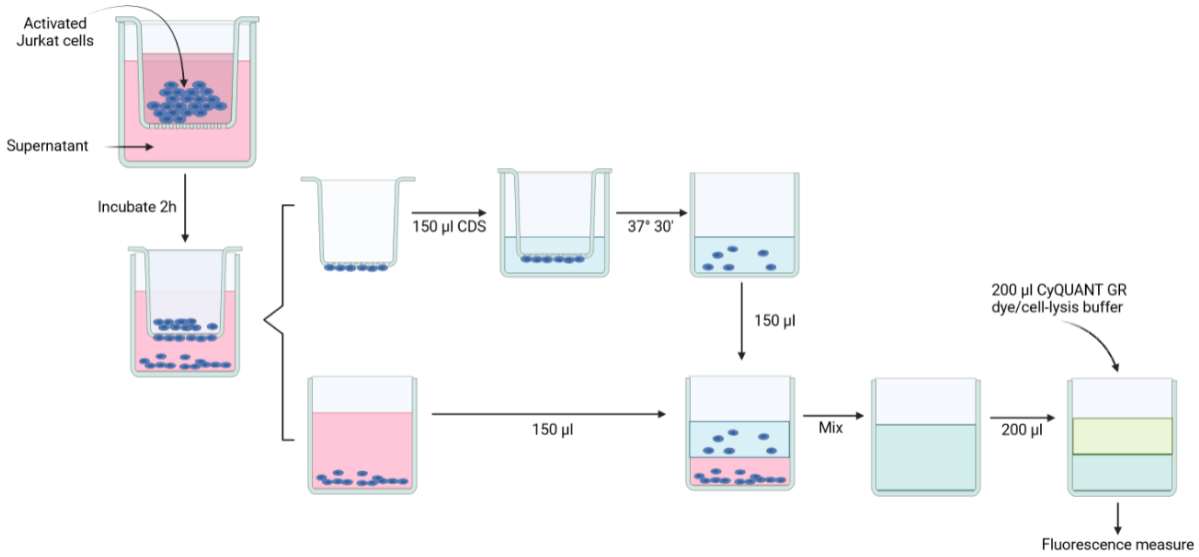


Figure 12. Chemotaxis experiment schema

2.11 RNA immunoprecipitation assay (RIP)

The RNA immunoprecipitation assay was performed following two different approaches. A-Agarose beads (sc-2001; Santa Cruz Biotechnology) were employed for the RIP experiment of *Lnc13* and magnetic Dynabeads™ Protein G (10003D; Invitrogen) in the case of *ARGI.* Both methodologies were able to immunoprecipitate RNA effectively; however, the protocol using magnetic Dynabeads was faster (**Figure 13**).

1. **RIP using A-agarose beads**:

- Cell harvesting: Cells were lysed in RIP buffer (150 mM KCl, 25 mM Tris, 0.5 mM DTT, 0.5% NP-40, protease inhibitors), kept on ice for 15 minutes and homogenized using a syringe. A-agarose beads (Santa Cruz) were mixed with

the lysates for 1h at 4°C in a wheel shaker to allow the lysates to pre-clear. Beads were removed by sample centrifugation. 10% of the pre-cleared lysate was saved as an input and the resting 90% was divided in two halves, one for the analysis of the protein of interest (PCBP2) and the other half for the negative control (IgG).

- Bead blockage and immunoprecipitation: 30 µl of the agarose beads per sample were blocked with 20% BSA diluted in RIP buffer and let rotate for 1 hour at 4°C. Blocked beads were washed once with RIP buffer and added to the pre-cleared lysate. 1 µg of each antibody (for PCBP2 or IgG; **Table 3**) was added to the corresponding sample and the mixture was incubated overnight at 4°C on rotation to allow the formation of the antibody-protein-RNA complex.
- Washing and elution: Beads were precipitated by centrifugation and the supernatant discarded. Beads were washed 3 times with RIP buffer, 3 times with a low salt buffer (50mM NaCl, 10mM Tris-HCl, 0.1% NP-40) and 3 times with the high salt buffer (500mM NaCl, 10mM Tris-HCl, 0.1% NP-40) Washed beads were then divided for the analysis of RNA and protein.
- Analysis: After RNA extraction, the expression of *Lnc13* and *STAT1* were analyzed by qPCR. The protein samples were employed to test the efficiency of the PCBP2 immunoprecipitation by Western blot.

2. **RIP using Dynabeads™ Protein G**:

- Cell harvesting: Cells were lysed in RIP buffer (150 mM KCl, 25 mM Tris, 0.5 mM DTT, 0.5% NP-40, protease inhibitors), kept on ice for 15 minutes and homogenized using a syringe. Lysates were pre-cleared by incorporating Dynabeads G to the lysates for 1 hour at 4°C. A 10% of the pre-cleared lysate was saved as input. The resulting 90% of the sample was divided in two halves, one for the analysis of the protein of interest (CTCF) and the other half for the negative control (IgG).
- Immunoprecipitation: Dynabeads G were washed once with RIP buffer and added to the pre-cleared lysates. 1 µg of each antibody (**Table 3**) was added to

the corresponding lysate and incubated overnight at 4°C on rotation to allow the antibody-protein-RNA binding.

- Washing and elution: Beads were captured by a magnet and the supernatant was wasted. Beads were washed 3 times with RIP buffer, 3 times with a low salt buffer (50 mM NaCl, 10 mM Tris-HCl, 0.1% NP-40) and 3 times with a high salt buffer (500 mM NaCl, 10 mM Tris-HCl, 0.1% NP-40). After washing, the beads were divided in two halves, one for protein analysis and the other for RNA analysis.
- Analysis: After RNA extraction, the expression of *ARGI* was analyzed by qPCR. The protein samples were employed to test the efficiency of the CTFC immunoprecipitation by Western blot.

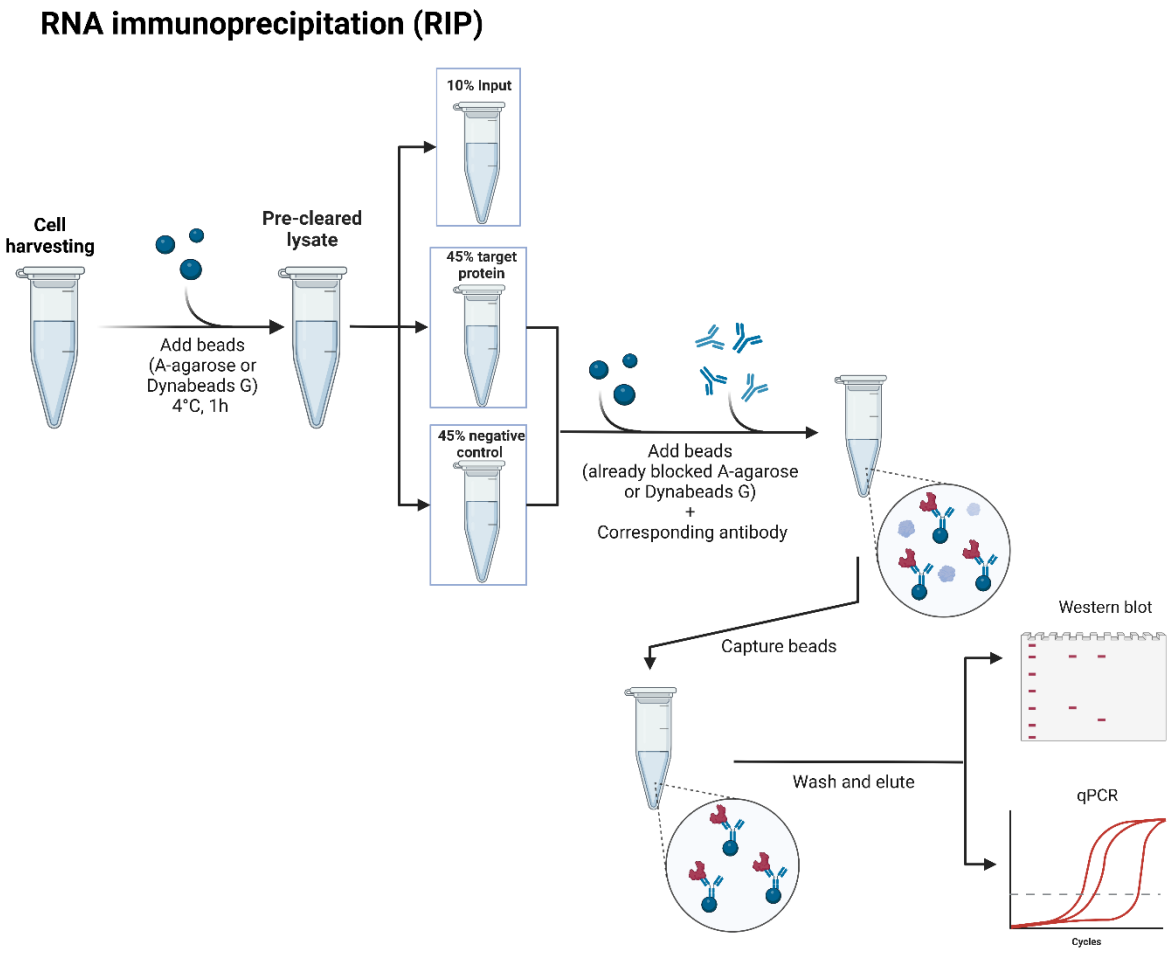


Figure 13. Protocol for RNA immunoprecipitation

2.12 RNA antisense purification (RAP)

RAP was performed following an adapted protocol from the Guttman lab (Engreitz et al., 2013) (**Figure 14**).

- Antisense probe generation: Firstly, antisense complementary probes against the RNA molecule of interest (e.g. *Lnc13* and *ARGI)* needed to be generated. To this aim, biotinylated antisense RNA molecules of the target genes were *in vitro* transcribed using a T7 RNA polymerase from Takara (Clonthech) following manufacturer's instructions. Thus, the antisense cDNA sequence of the genes of interest were cloned into a T7 promoter containing pCMV6 vector. Briefly, 200 ng of the interest DNA sequence was mixed with 2 µl x10 T7 RNA polymerase buffer (2540A; Takara Bio), 50U T7 polymerase (2540A; Takara Bio), 2 µl x10 Biotin RNA Labeling Mix (11685597910; Roche), 2 µl DTT (50 mM, 2540A; Takara Bio) + up to 20 µl RNAse free water; and incubated for 2 hour at 37°C. The RNA obtained from the reaction was purified using the PureLink™ RNA Mini Kit (12183025; ThermoFisher Scientific). Afterwards, the purified RNA was fragmented using the RNA Fragmentation Reagent (AM8740; Invitrogen) following the manufacturer's instructions. Probes against a non-related long noncoding RNA of a similar size were used as negative control.
- Cell crosslinking and harvesting: 2 mM DSG (20593; Thermo Scientific) was added to EndoC-βH1 cells and left gently shaking for 45' at room temperature. DSG was removed from the plate and cells were washed with PBS. After, 3% formaldehyde was added to the cells for 10' at 37°C. 100 mM glycine was added to halt formaldehyde and incubated for 5' at 37°C. Plate was washed with PBS three times, harvested in cold scraping buffer (PBS+0.5% BSA fraction V) and centrifuged at 1,000 g for 5' at 4°C. Cells were onward kept on ice. Supernatant was removed and the pellet resuspended in 1ml of cold scraping buffer. It was then centrifuged at 2,000 g for 5' and the pellet resuspended in 500 µl ice-cold total cell lysis buffer (10 mM HEPES pH 7.5, 20 mM KCl, 1.5 mM MgCl2 and 0.5 mM EDTA. Before use, freshly prepared 1 mM tris (2- carboxyethyl) phosphine

(TCEP) and 0.5 mM phenyl methyl sulfonyl fluoride (PMSF) was added. Cells were incubated for 10' and homogenized using a syringe. The samples were sonicated at 5W during 1' (Cycle 0.7s ON, 1.3s OFF). DNase cofactor solution and 50 units of turbo DNasa (AM2238; Invitrogen) were added and incubated at 37°C for 10'. Immediately, they were put on ice and DNase stop solution was incorporated to stop the digestion. Hybridization buffer was added and the mixture was centrifuged at maximum speed for 10'. The supernatant was taken to obtain a clear lysate.

- Lysate pre-clearing: Streptavidin-coated magnetic beads were washed in hybridization buffer (20 mM Tris-HCl pH 7.5, 7 mM EDTA, 3 mM EGTA, 150 mM LiCl, 1 % NP-40, 0.2 % N -lauroylsarcosine, 0.1 % sodium deoxycholate, 3 M guanidine thiocyanate, and 2.5 mM TCEP), and immediately added to the lysates. Beads were incubated with the lysate shaking for 30' at 37°C. After beads removal, input samples were collected (10% of the whole sample volume).
- Hybridization, capture and wash: 50 ng of each probe was denatured at 85°C for 3' and immediately transferred to ice. Probes were incubated with the corresponding lysate for 2 hours at 37°C shaking. After incubation, pre-washed streptavidin magnetic beads were added and incubated for 30' at 37°C. Samples were then placed in a magnet and the supernatant was discarded. Beads were washed 6 times in a solution containing 20 mMTris-HCl pH 7.5, 10 mM EDTA, 1 % NP-40, 0.2 % N -lauroylsarcosine, 0.1 % sodium deoxycholate, 3 M guanidine thiocyanate, and 2.5 mM TCEP.
- RNA and DNA recovery: Beads were eluted in the nuclear lysis buffer (20 mM HEPES pH 7.5, 50 mM KCl, 1.5 mM $MnCl_2$, 1 % IGEPAL CA630 (NP-40), 0.4 % sodium, 1 mM TCEP and 0.5 mM PMSF) and incubated at 94°C for 5' to release the biotin bound to the streptavidin. Beads were then retrained using a magnet and the eluate saved. The eluate was divided in different aliquots according to the needs of downstream experiments.
- RNA and DNA quantification by qPCR: RNA was used to check for the correct purification of *Lnc13* and *ARGI*, and to determine the presence of the 3'UTR of

STAT1 in the case of *Lnc13* purification. The DNA was used in for the determination of the promoter and enhancer regions of *IFNβ* and *ISG15* in ARGI-bound chromatin fragments. All the primer used are listed on **Table 2**.

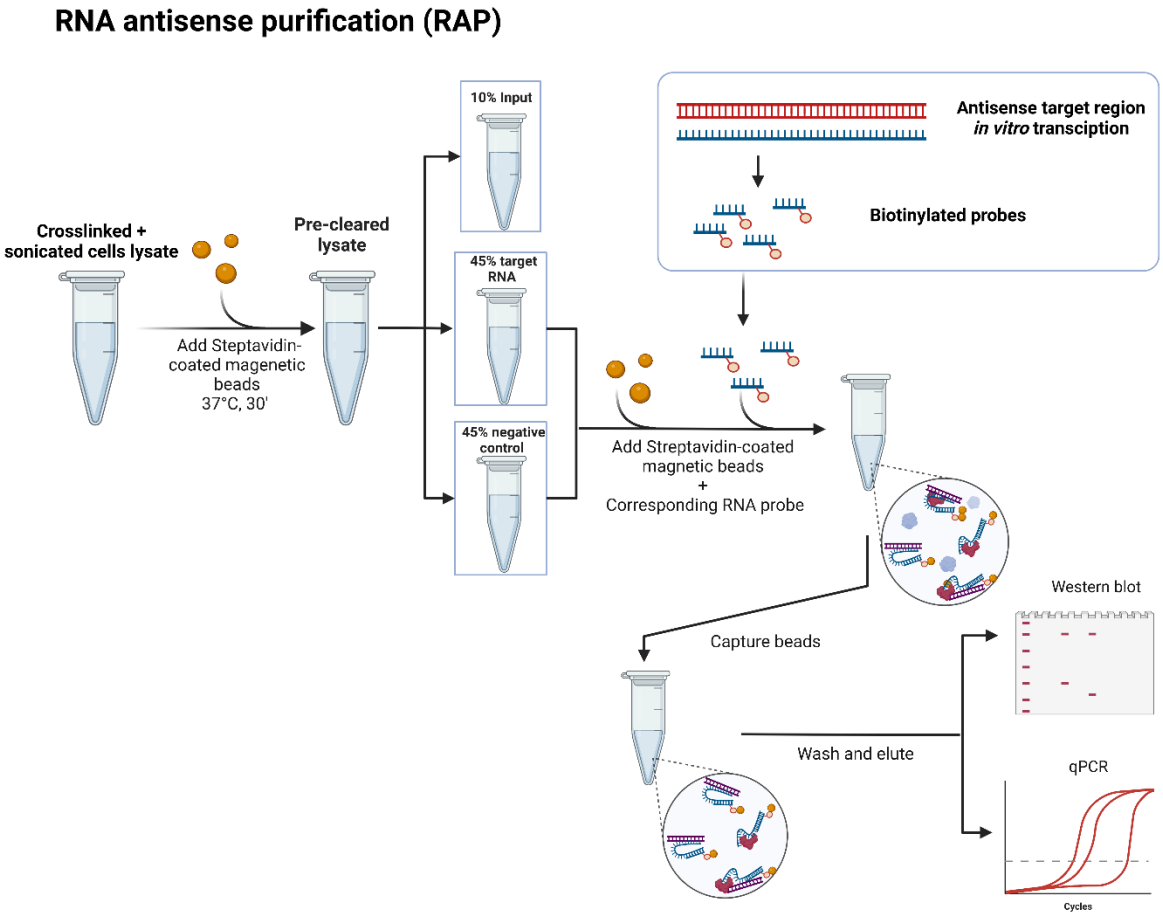


Figure 14. Protocol followed for RNA antisense purification.

2.13 RNA pull-down (Figure 15)

- Biotinylated RNA synthesis: the 3'UTR of *STAT1* gene was amplified from human cDNA using the primer listed on (**Table 1**). The PCR product was purified and in vitro transcribed using the T7 RNA polymerase from Takara and the RNA biotin labeling kit (Roche) as described previously. The resulting biotinylated RNA was employed for the RNA pull-down.
- RNA pull-down: target cells were lysed in RIP buffer (150 mM KCl, 25 mM Tris, 0.5 mM DTT, 0.5% NP-40, protease inhibitors). Lysates were pre-cleared using Streptavidin Mag Sepharose magnetic beads (GE Healthcare) for 1h at 4°C in a

wheel shaker. Lysates were divided in 3 different tubes: 10% of the lysate for the input, 45% for the 3'UTR-*STAT1* pull-down and 45% for the negative control RNA. The protein concentration of the already pre-cleared lysates were quantified and 1 mg lysate protein was mixed with 1 µg of the already biotinylated 3'UTR-*STAT1* or negative control RNA. Mixture was incubated for 1h at room temperature. After incubation, Streptavidin magnetic beads were added to the mixture for another 1h for capturing the biotinylated RNA, and everything bound to it.

- Collection: Magnetic beads were captured with the help of the magnet and washed 3 times with RIP buffer, another 3 times with high salt buffer (500 mM NaCl, 10 mM Tris-HCl, 0.1% NP-40) and lastly, 3 times with low salt buffer (50 mM NaCl, 10 mM Tris-HCl, 0.1% NP-40). Finally, beads were mixed with laemmli buffer and PCBP2 protein levels were analyzed by Western blot.

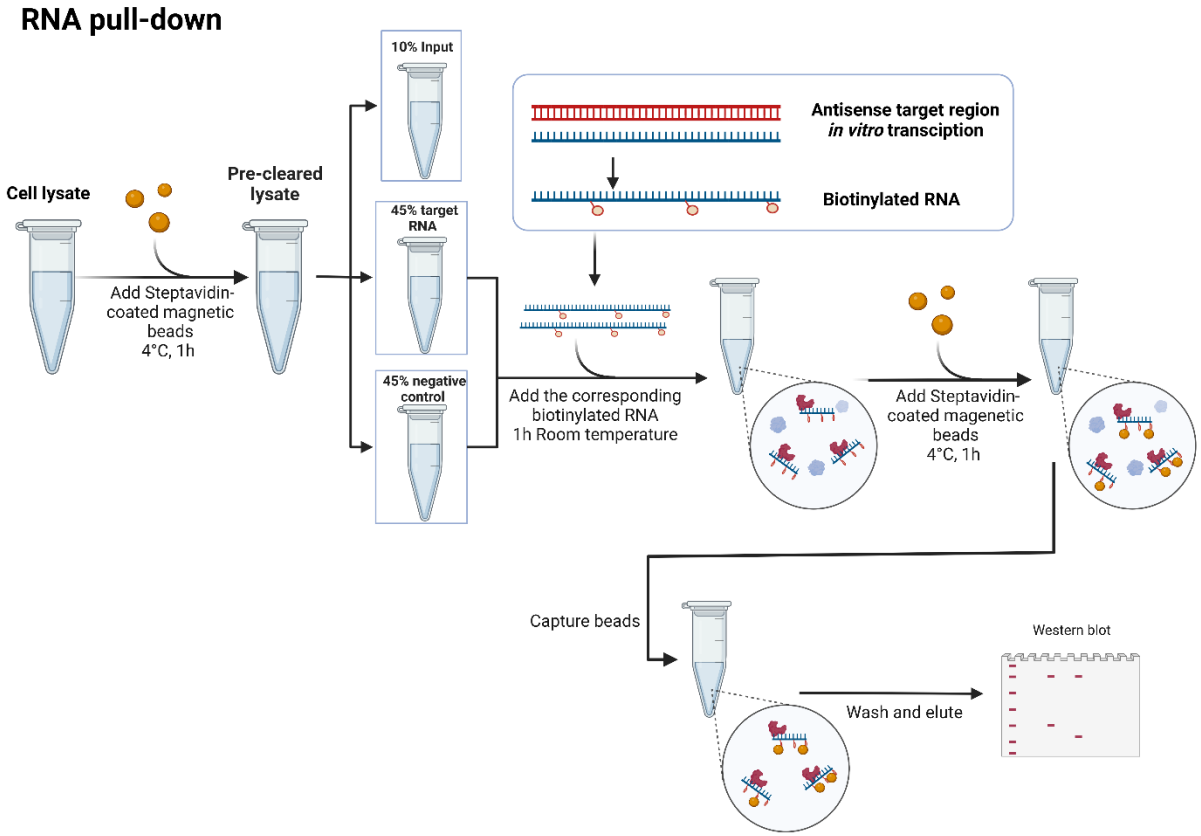


Figure 15. Schema of the RNA pull-down experiment

2.14 Chromatin immunoprecipitation (Figure 16)

- Cell crosslinking and harvesting: Cells were harvested in 500 µl of 1% formaldehyde and incubated 5' at room temperature. After incubation, cells were pellet and washed with cold PBS twice. Pellet was lysed in 500 µl of Nuclear Lysis Buffer for 10' on ice (NLS: 20 mM HEPES pH 7.5, 50 mM KCl, 1.5 mM $MnCl_2$, 1 % IGEPAL CA630 (NP-40), 0.4 % sodium fresh 1 mM TCEP and 0.5 mM PMSF). Lysates were sonicated at 5W for 10 cycles (30'' ON, 30'' OFF). For pre-clearing of the lysates, magnetic Dynabeads G were washed with RIP buffer (150 mM KCl, 25 mM Tris, 0.5 mM DTT, 0.5% NP-40, protease inhibitors), and 10- µl of these washed beads were incorporated to each lysate for 1 hour at 4°C.
- Hybridization, capture and wash: Input was saved taking a 10% of the pre-cleared lysate. The resulting 90% of the sample was divided in two halves, the first for immunoprecipitation of the protein of interest (STAT1 for *Lnc13* characterization and CTCF for *ARGI*) and the other half, for the negative control (IgG). The corresponding antibody (**Table 3**) was added to the corresponding lysates and incubated for 1 hour at room temperature on continuous movement. After, 20 µl of washed Dynabeads G were incorporated into each sample and incubated for 30' on movement at room temperature to allow the binding between the magnetics beads and the specific antibodies. Then, beads are washed with 3 times with RIP buffer, 3 times with low salt buffer (50 mM NaCl, 10 mM Tris-HCl, 0.1% NP-40), and 3 times with high salt buffer (500 mM NaCl, 10 mM Tris-HCl, 0.1% NP-40).
- RNA, DNA and protein recovery: Beads were eluted in the Nuclear Lysis Buffer and incubated at 94°C for 10' to release the molecules from the Dynabeads. G Beads were retained with a magnet to save the eluate in different aliquots according to the needs of downstream experiments.
- RNA and DNA quantification by qPCR: To check whether *ARGI* was bound to the protein CTCF, the quantification of *ARGI* was performed in the RNA fraction bound to CTCF. The same approach was followed to confirm whether *Lnc13* was interacting with STAT1 protein.

CTCF-bound DNA was analyzed by qPCR using specific primers (**Table 2**) to determine the presence of the promoter and enhancer regions of *IFNβ* and *ISG15* in CTCF-bound chromatin fragments. In addition, STAT1-bound chromatin was used to determine the presence of the *CXCL10* promoter region by qPCR (primers on **Table 2**).

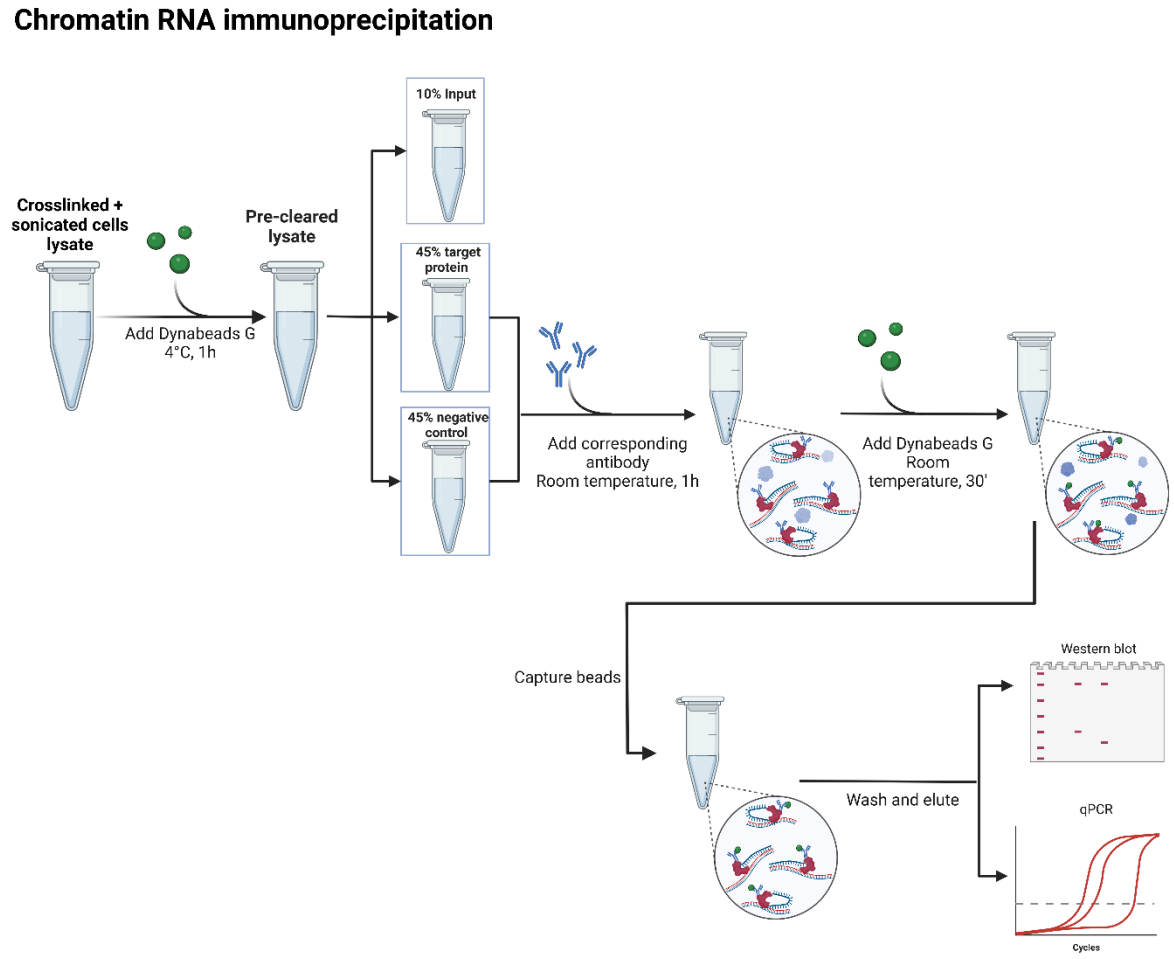


Figure 16. Schematic chromatin immunoprecipitation protocol

2.15 RNA-protein interaction assay

- RNA production and folding: 3'UTR of *STAT1* was amplified using the T7-3'UTR-STAT1 primer pair (Listed on **Table 1**). The amplified region was *in vitro* transcribed using a T7 polymerase system as explained previously. To allow the RNA to fold into its native secondary structure, the generated RNA was supplied with a RNA structure buffer (10 mM Tris pH 7, 0.1 M KCl, 10 mM $MgCl_2$) and

heated to 90°C for 2', put on ice for another 2', and finally shifted to room temperature for 20'.

- RNA-protein interaction: 140 ng of folded RNA was mixed with protein lysates and incubated for 1 hour at room temperature to allow the formation of RNA-protein interaction.
- Gel electrophoresis: Samples were run in a 1% agarose gel with Tris-Borate buffer. RNA bands were stained using GelRed® Nucleic Acid Gel Stain (BT-41003-T; Biotium). Proteins were transferred to a nitrocellulose membrane and the antibody against the target protein (PCBP2) was incubated for protein visualization.

2.16 Statistics

Experimental data are displayed as means ±SEM as indicated in the figure legends. A significant difference between experimental conditions was assessed by Student's t test or ANOVA followed by Student's t test with Bonferroni correction using GraphPad Prism v8.0.1.

Table 1. List of the primers employed for DNA amplification

Target gene	Sequence
Lnc13-overexpression	Fw: AAGGATCATTGCAGGGTCTC Rv: GTGGCCAAAAGAAGTCTGAGTC
*Lnc13*_delSNP	Fw: GCCTTTGATTTCCTGGACTG Rv:TTAAAACCCGAAAAGGACCA
*Lnc13*_KO_sg1	Fw: CACCGAACTCCTGACCTCAGGAGAT Rv: AAACATCTCCTGAGGTCAGGAGTTC
*Lnc13*_KO_sg2	Fw: CACCGTCTGAAAAAGTGTCCTACCT Rv: AAACAGGTAGGACACTTTTTCAGAC
T7-3'UTR-*STAT1*	Fw: TAATACGACTCACTATAGGTGAAACTTAGGTTCTCGCCATC Rv: TTCACATTTGCGAATGGTTC

ARGI _mut_rs9585056	Fw: AAAAACATTGAGACCAATGCGAGTTT Rv: GCCCTTGCAAATGTTATTTCTGAGCC
ARGI_KO_sg1_	Fw: CACCGCTGTAGGGACGTCTTTCCG Rv: AAACCGGAAAGACGTCCCTACAGC
ARGI_KO_sg2	Fw: CACCGGGATCCTTCCAAAATTGACA Rv: AAACTGTCAATTTTGGAAGGATCCC
ARGI_KO_sg3	Fw: CACCGGCCAGTCCCCGATCAGTGTA Rv: AAACTACACTGATCGGGGACTGGCC
ARGI_CRISPRi sgRNA	Fw: GATCGCCGGGGATTCCCAGTTCCCC Rv: AAAACGGGAACTGGGAATCCCCGG

Table 2. List of the primers and assays employed all along this thesis for qPCR experiment

Sybr Green system	
Target gene	**Sequence**
Lnc13	Fw: AAGGATCATTGCAGGGTCTC Rv: GTGGCCAAAAGAAGTCTGAGTC
MALAT1	Fw: GCTGTGGAGTTCTTAAATAT Rv: TTCTCAATCCTGAAATCCCC
RPLP0	Fw: GCAGCATCTACAACCCTGAAG Rv: CACTGGCAACATTGCGGAC
3'UTR of *STAT1*	Fw: TGAAACTTAGGTTCTCGCCATC Rv: TTCACATTTGCGAATGGTTC
CXCL10 promoter	Fw: TGGATTGCAACCTTTGTTTTT Rv: GTCCCATGTTGCAGACTCG
ISG15 promoter	Fw1: TCCCTGTCTTTCGGTCATTC Rv1: ACGGCACAAGCTCCTGTACT

ISG15 proximal enhancer	Fw2: CACCTGAAGCAGCAAGTGAG Rv2: CTTTATTTCCGGCCCTTGAT
IFNβ promoter	Fw1: TCCCACTTTCACTTCTCCCT Rv1: GCTTTCCTTTGCTTTCTCCCA
IFNβ proximal enhancer	Fw2: GGGTGGGATGGAGAACTCAG Rv2: ACTTTTCTGTTGTTTGGTCTTGT
IFNβ distal enhancer	Fw3: GAGAACTCCTGCCCAGAGG Rv3:AGCACCTCAAGAACACAATAGC

TaqMan system		
Target gene	**Catalog number**	**Brand name**
ARGI	Custom assay for *ARGI*	(IDT)
MX1_PrimeTime qPCR Assay	Hs.PT.58.26787898	(IDT)
IFIT1_PrimeTime qPCR Assay	Hs.PT.56a.2076909	(IDT)
IFIT3_PrimeTime qPCR Assay	Hs.PT.58.20456374	(IDT)
IFI6_PrimeTime qPCR Assay	Hs.PT.58.4390209	(IDT)
*Lnc13*_TaqMan® Gene Expression Assay	4332078	(ThermoFisher)
MALAT1_TaqMan® Gene Expression Assay	Hs00273907_s1	(ThermoFisher)
RPLP0_TaqMan® Gene Expression Assay	Hs99999902_m1	(ThermoFisher)
CXCL10_TaqMan® Gene Expression Assay	Hs00171042_m1	(ThermoFisher)
CXCL9_TaqMan® Gene Expression Assay	Hs00171065_m1	(ThermoFisher)
CCL5_TaqMan® Gene Expression Assay	Hs00982282_m1	(ThermoFisher)
CXCL1_TaqMan® Gene Expression Assay	Hs00605382_gH	(ThermoFisher)
STAT1_TaqMan® Gene Expression Assay	Hs01013996_m1	(ThermoFisher)
ISG15_TaqMan® Gene Expression Assay	Hs00192713_m1	(ThermoFisher)
IFNβ_TaqMan® Gene Expression Assay	Hs01077958_s1	(ThermoFisher)
Actinaβ_TaqMan® Gene Expression Assay	Hs01060665_g1	(ThermoFisher)
MEG3_TaqMan® Gene Expression Assay	Hs00292028_m1	(ThermoFisher)

RPLP0_TaqMan® Gene Expression Assay	Hs99999902_m1	(ThermoFisher)

Table 3. List of the antibodies used for Western blot experiments.

Target protein	Catalog number	Brand name	Working dilution
STAT1	sc-346	Santa Cruz Biotechnologies	1:1,000 in 5% BSA
pSTAT1	#7649	Cell Signaling Technology	1:1,000 in 5% BSA
α-tubulin	#T9026	Sigma-Aldrich	1:5,000 in 5% BSA
Hsp90	# 4877	Cell Signaling Technology	1:1,000 in 5% BSA
HDAC1	#4874	Abcam	1:500 in 1X TBS, 0.1% Tween
H3	#4499	Cell Signaling Technology	1:500 in 1X TBS, 0.1% Tween
PCBP2	#83017	Cell Signaling Technology	1:500 in 5% non-fat milk
Normal mouse IgG	sc-2025	Santa Cruz Biotechnologies	Depends on the experiment
CTCF	#PA5-17143	Invitrogen	1:1,000 in 5% BSA
GAPDH	sc-365062	Santa Cruz Biotechnologies	1:5,000 in 5% BSA
HRP-conjugated anti-mouse	sc-516102	Santa Cruz Biotechnologies	1:5,000 in 5% non-fat milk
HRP-conjugate mouse anti-rabbit	sc-2357	Santa Cruz Biotechnologies	1:10,000 in 5% non-fat milk

Chapter 5

Results

Emaitzak

Results 1

Identification of lncRNAs associated to type 1 diabetes

Accumulating evidence suggest that lncRNAs are important for the regulation of several biological and cellular processes, however, most of the annotated lncRNAs have not been characterized yet and their function remain unknown. Since lncRNAs seem to have a key role in gene expression regulation, disruption of their function may affect the regulation of potentially pathogenic gene pathways. Interestingly, around 10% of the SNPs associated with immune disorders lie in lncRNAs (Ricaño-Ponce & Wijmenga, 2013). The presence of these SNPs in exonic regions of lncRNAs usually disrupt their secondary structure, altering their function and eventually, dysregulating gene expression networks important for pancreatic β cell function and T1D development. For this reason, the first step of this work was to identify long non-coding RNAs harboring a polymorphism associated with type 1 diabetes that additionally are regulated by viral infections in pancreatic β cells.

To this aim, the first step was stablish a procedure to identify candidate lncRNAs potentially implicated on T1D development. The workflow was divided in four main steps (**Figure 17**):

1. Identification of lncRNAs harboring a T1D-associated SNP:

   The genomic location of all human lncRNAs annotated in NONCODE version 6 was intersected with the genomic position of all annotated T1D-associated SNPs listed in the NHGRI-EBI Catalog of Human Genome Association Studies. The intersection of both data sets resulted in the identification of 69 lncRNAs that harbored at least one T1D-associated SNP (**Appendix 1**).

2. Identification of lncRNAs harboring an exonic T1D-associated SNP:

   The specific localization of each SNP inside the lncRNA was analyzed one by one using the UCSC Genome Browser Home. This search led to the identification of 19 lncRNAs harboring a T1D-associated SNP located in an exon (**Appendix 2**).

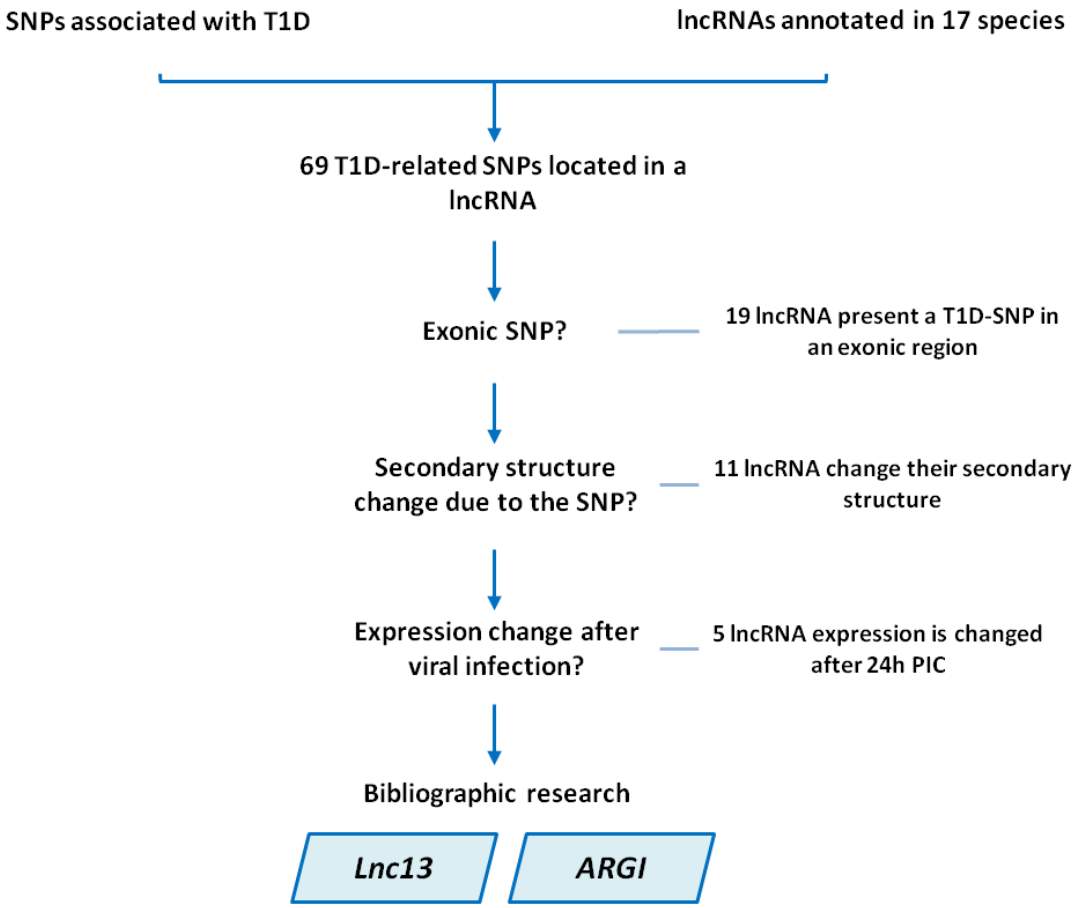


Figure 17. Schematic outline of the workflow used for the identification and selection of T1D-associated lncRNAs. Genomic positions of T1D-associated SNPs annotated in the NHGRI-EBI Catalog of Human Genome Association Studies were intersected with the genomic localization of all lncRNAs annotated in NONCODE version 6. 19 T1D-associated lncRNAs harboring an exonic SNP were analyzed in the RNAsnp software from the Center for non-coding RNA in Technology and Health (RTH) to determine potential changes in their secondary structure. Eleven lncRNAs were predicted to suffer secondary structure changes due to the T1D-associated SNP. The expression of the 11 lncRNAs was analyzed in EndoC-βH1 cells in basal condition and after PIC transfection. The expression of 5 lncRNAs was altered in EndoC-βH1 cells exposed to PIC for 24h. The final prioritization of lncRNAs for functional characterization was based on the information found upon a bibliographic search. Finally, *Lnc13* and *ARGI* were selected for further functional studies.

3. *In silico* prediction of the secondary structure of lncRNAs associated with T1D:

   In order to evaluate whether the T1D-asssociated SNPs disrupted the secondary structure of the previously identified 19 lncRNAs, the RNAsnp

software from the Center for non-coding RNA in Technology and Health (RTH) was used. This software provides the prediction of the most probable RNA folding and compares the wild-type lncRNA secondary structure (protective allele for T1D) to the lncRNA structure when the T1D-risk allele is present. The software provides a p-value according to the significance of the structural change and a p-value lower than 0.2 is considered significant. From the 19 lncRNAs listed on the previous step, there were 11 lncRNAs in which the presence of the T1D-associated SNP was predicted to disrupt the secondary structure of the lncRNA (**Appendix 2**).

4. Identification of T1D-associated lncRNAs expressed in EndoC-βH1 cells and modulated by viral infections: To clarify whether the previously selected 11 lncRNAs were expressed in pancreatic β cells and modulated by viral infections, their expression was measured in the EndoC-βH1 cells in basal condition and after PIC transfection for 8h and 24h (**Figure 18**). PIC is a synthetic analog of double-stranded viral RNA, which simulates a viral infection. Five out of the 11 lncRNAs exhibited an increased expression after 24h of intracellular PIC exposure when compared to basal condition (e.g. 0h). One out of the 11 lncRNAs measured was not expressed in EndoC-βH1 cells (this lncRNA is not listed in **Figure 18**).

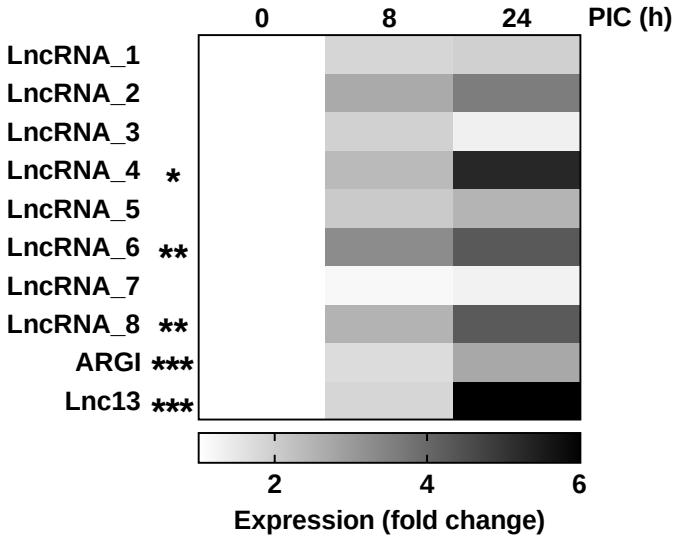


Figure 18. T1D-associated LncRNA expression in basal and PIC-transfected EndoC-βH1 cells. EndoC-βH1 cells were exposed to intracellular PIC (1 µg/mL) for 8h or 24h, and the expression of eleven T1D-associated lncRNAs was determined by qPCR. One out of eleven lncRNAs was not expressed in EndoC-βH1 cells, and the expression of the other 10 lncRNAs is represented as a heatmap (fold change vs. PIC 0h). Results are means ± SEM of 3 independent experiments; ***p < 0.001, **p < 0.01 and *p < 0.05 as indicated; Student's t test comparing 0h PIC versus 24h PIC.

Once a list of 5 lncRNAs was obtained, I decided to prioritize a couple of them for posterior functional characterization. In order to select the most interesting candidates, a deep bibliographic research was performed to find previous information on the function of the lncRNA or on the potential functional effect of the T1D-associated SNP. This process resulted in the selection of two lncRNAs named *Lnc13* and *ARGI*.

On one hand, ***Lnc13*** is a lncRNA located on chromosome 2q12. It harbors a SNP (rs917997) that has been previously associated with several autoimmune diseases, including T1D (Smyth et al., 2008), celiac disease (Hunt et al., 2009) and inflammatory bowel disease (Zhernakova et al., 2008). A previous study characterized the role of *Lnc13* in the pathogenesis of celiac disease and described its participation in the regulation of pro-inflammatory gene expression in immune cells through direct interaction with the chromatin and the multifunctional protein hnRNPD (Castellanos-Rubio et al., 2016). Interestingly, the risk allele for celiac disease in *Lnc13* (rs917997-T)

is the opposite of the risk allele for T1D (rs917997-C), suggesting that this SNP may alter the function of this lncRNA in a disease- and tissue-specific manner.

On the other hand, lncRNA ***ARGI*** (Antiviral Response Gene Inducer) is located on chromosome 13q32 and harbors a SNP (rs9585056) that has been previously described as an eQTL that regulates the expression of an antiviral gene network named IRF7-driven inflammatory network (IDIN) in monocytes and macrophages (Heinig et al., 2010). Interestingly, when this eQTL was described, the T1D-associated SNP was annotated as intergenic; however, and thanks to this work, now it is known that is located inside an exon of *ARGI*.

Results 2

Functional characterization of *Lnc13* in pancreatic β cells

Atal honetan aurkezten diren emaitzak hurrengo artikuluan izan dira publikatuak:

The results presented in this section have been published in:

> **Gonzalez-Moro, I.**, Olazagoitia-Garmendia, A., Colli, M. L., Cobo-Vuilleumier, N., Postler, T. S., Marselli, L., Marchetti, P., Ghosh, S., Gauthier, B. R., Eizirik, D. L., Castellanos-Rubio, A., & Santin, I. (2020). The T1D-associated lncRNA Lnc13 modulates human pancreatic β cell inflammation by allele-specific stabilization of STAT1 mRNA. Proceedings of the National Academy of Sciences of the United States of America, 117(16), 9022–9031. https://doi.org/10.1073/pnas.1914353117

Laburpena:

1 motako diabetesarekin (T1D) loturiko marka genetikoen gehiengoa giza genomaren domeinu ez-kodetzaileetan kokatzen dira. Hauetako askotan aurreikusi da RNA luze ez-kodetzaileen (lncRNA) bigarren mailako egituran eta adierazpenean eragiten dute, baina oraindik argitzeko dago lncRNA hauek duten efektua T1D gaixotasunaren garapenean. Hemen, T1D gaixotasunarekin loturiko nukleotido bakarreko aldaera (SNP) bat daukan lncRNA baten funtzio osoaren karakterizazioa egin da, *Lnc13* deritzona. T1Derako arrisku genotipoa daukan *Lnc13* molekula (rs917997*CC) daukaten giza pankreako irletan ikusi da *STAT1* gene maila altuagoak dituela heterozigotoak diren irlekin alderaturik (rs917997*CT). *Lnc13* gainadierazteak *STAT1*aren bidezidorraren aktibazioa ekartzen du, zeina kemokinen ekoizpenaren emendioa dakarrena era alelo-espezifiko batean. Ispilu-irudi modura, *Lnc13* adierazpenaren geldialdiak partzialki indargabetzen du PIC molekulak (mimetiko birala) eragindako *STAT1* eta kemokina pro-inflamatorioen adierazpena. Gainera, ikusi dugu PICak *Lnc13* nukleotik zitoplasmara translokatzea sorrarazten duela, *STAT1* mRNA molekula PCBP2 proteinarekin interakzionatzea baimenduz. Interesgarria da nola *Lnc13*-PCBP2 loturak *STAT1* mRNA molekularen estabilitatea modulatzen duela, β zeluletan ematen den inflamazioa bermatuz era alelo-espezifiko batean. Gure emaitzek T1Dari loturiko *Lnc13* molekulak gaixotasunean efektua izan dezakela pankreako β zelulen inflamazioa emendatzen duelarik. Aurkikuntza hauek informazio berria ematen dute lncRNAtan dauden gaixotasunei-loturiko SNPek gaixotasunaren garapenean duten eraginean eta atea irekitzen du lncRNAk xede izanda, aurrerapen diagnostiko eta terapeutikoak garatzeko.

1. *Lnc13* is expressed ubiquitously but more abundantly in human pancreatic β cells.

Firstly, *Lnc13* expression was measured in a set of different human tissue samples and compared to its expression level in the human β cell line EndoC-βH1 and in human pancreatic islets. As previously described, *Lnc13* was ubiquitously expressed in human tissues (Castellanos-Rubio et al., 2016). Interestingly, the expression of this lncRNA in EndoC-βH1 was around 20 times higher than in the most *Lnc13*-expressing human tissues (heart, liver and muscle), most probably due to the embryonic nature of the EndoC-βH1 cell line (**Figure 19**). In contrast, the expression of *Lnc13* in human pancreatic islets, which were on average composed by 50-60% β cells (Cabrera et al., 2006), was similar to that of the thymus, colon and lung.

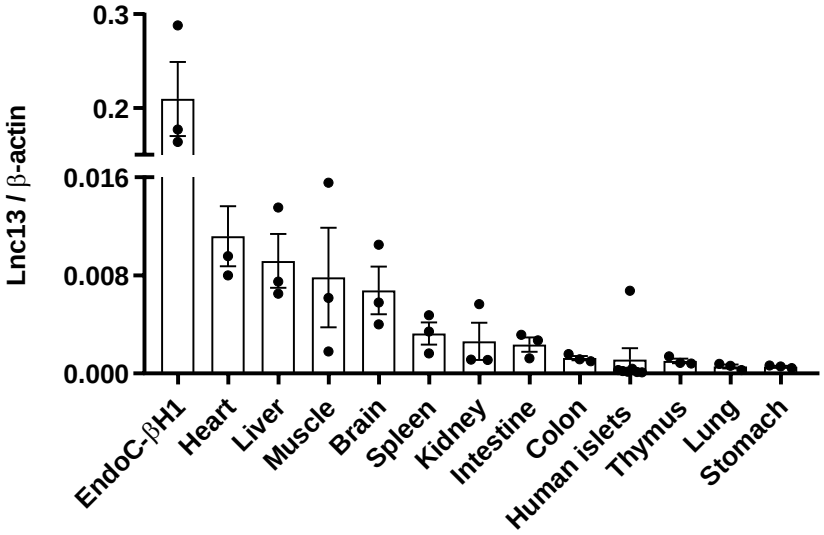


Figure 19. *Lnc13* is expressed in pancreatic β cells and in a set of human tissues. *Lnc13* expression was analyzed in the human β cell line EndoC-βH1, in human pancreatic islets, and in a set of human tissues (heart, liver, muscle, brain, spleen, kidney, intestine, colon, thymus, lung, and stomach). *Lnc13* expression was determined by qPCR and normalized by the housekeeping gene *β-actin*. Results are means ± SEM of three experimental replicates.

Publicly available RNA sequencing data of a set of human primary cells confirmed these results. As shown in **Figure 20**, *Lnc13* was ubiquitously expressed in most of the cell types, although its expression was very low (less than two transcripts per million).

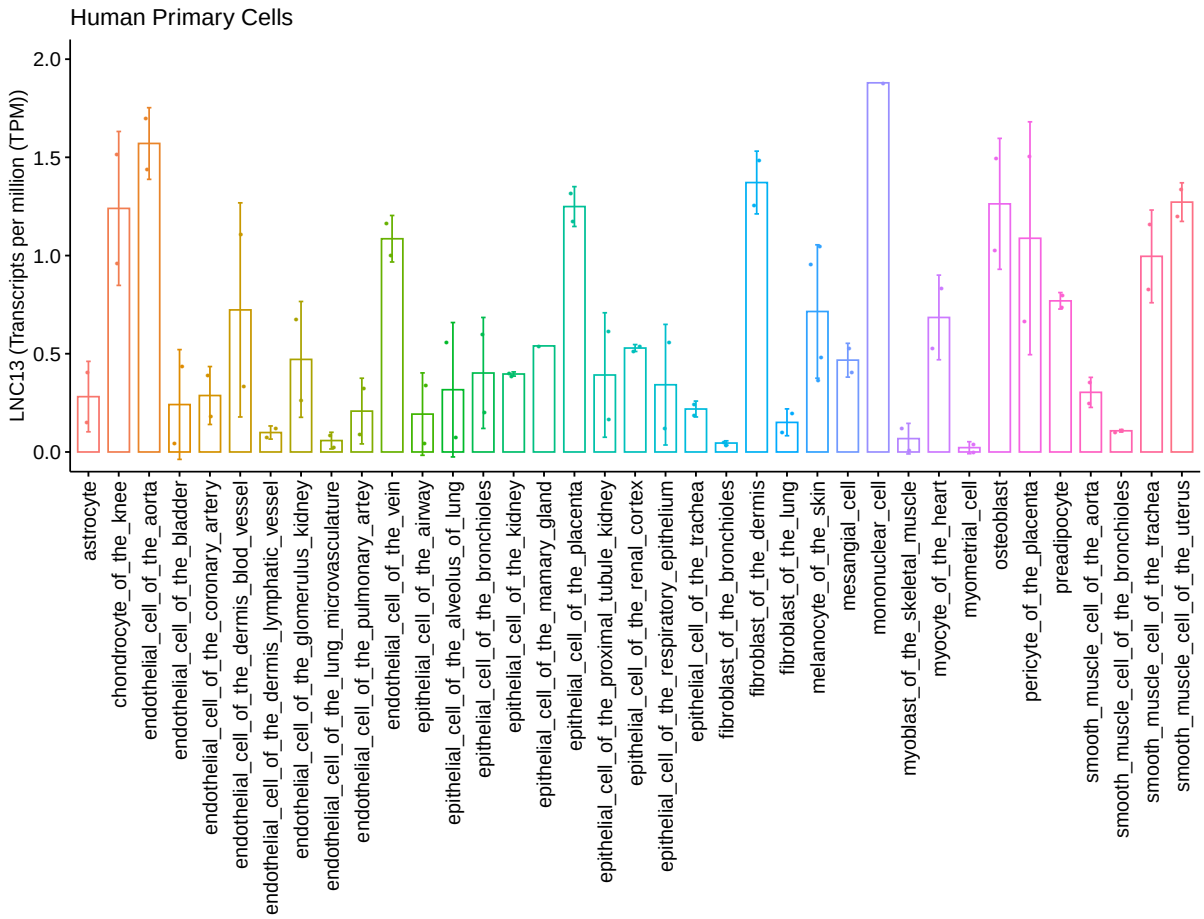


Figure 20. *Lnc13* is ubiquitously expressed in several human cell types. Expression of *Lnc13* in a set of human primary cells obtained from RNA-sequencing data generated by the ENCODE Project Consortium. The relative expression level is expressed as transcripts per million (TPM) units.

In order to translate the expression level to the specific number of *Lnc13* copies per cell, I determined the number of *Lnc13* RNA molecules in different cell types, including EndoC-βH1, HEK293FT (kidney cells), SHSY5Y (brain cells), HCT15 (cells from the colon) and NCI-H23 (lung cells). EndoC-βH1 cells presented the higher abundance of *Lnc13* per cell, being around 2.8 copies per cell (**Figure 21**). In the other cell types, there was less than 1 copy of *Lnc13* per cell. However, these numbers were under the average for lncRNAs, which normally range between dozens to hundreds of copies per cell, or between 0.3-100 copies per cell for chromatin-associated lncRNAs (Seiler et al., 2017; M. Wu et al., 2021).

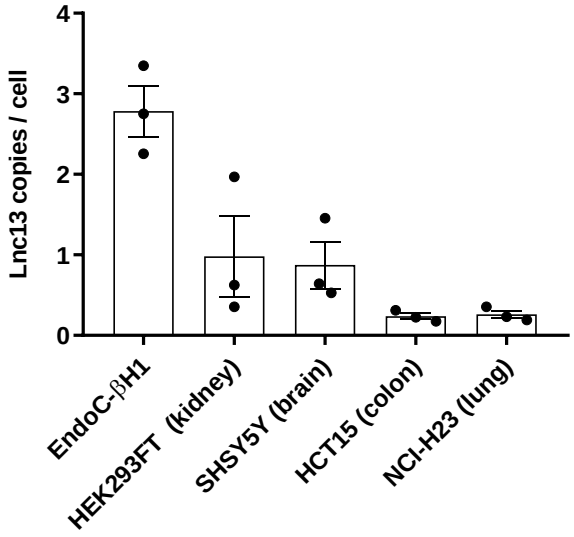


Figure 21. *Lnc13* RNA copy number per cell in different cell types. The number of *Lnc13* molecules per cell was determined in EndoC-βH1, HEK293FT, SHSY5Y, HCT15 and NCI-H23 cell lines by qPCR and using a standard curve generated from dilutions of a reference amplicon. Results are means ± SEM of 3 independent experiments.

2. *Lnc13* is upregulated by viral Infections, and correlates with *STAT1* expression in human pancreatic islets.

As previously explained, several environmental factors, including viral infections, have been linked to T1D development (Coppieters et al., 2012; Dotta et al., 2007; Filippi & Von Herrath, 2008; Op de beeck & Eizirik, 2016). Recent studies indicate that lncRNAs may participate in the regulation of virus-induced immune response in parallel or in combination to the activation driven by specific transcription factors, such as STAT1, IRF7 or STAT2, among others (Huang et al., 2020; Kambara et al., 2014; Nishitsuji et al., 2016). In order to check whether the infection with a diabetogenic strain of Coxsackievirus affected *Lnc13* expression in pancreatic β cells, EndoC-βH1 were infected with the Coxsackievirus B5 (CVB5) and *Lnc13* expression was determined. As shown in **Figure 22**, CVB5 infection for 24h resulted in a 2.5-fold increase in the expression of the lncRNA. These results were in line with the *Lnc13* expression pattern observed on the previously explained heatmap after intracellular PIC transfection, in which intracellular PIC exposure induced a 5-fold increase in *Lnc13* expression (**Figure 18**).

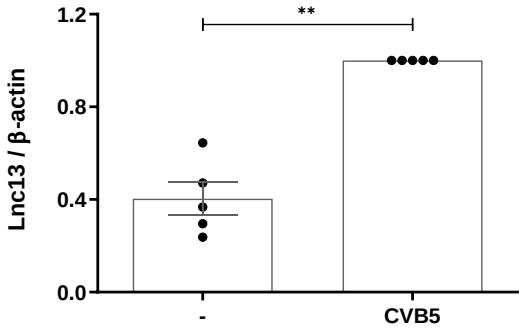


Figure 22. *Lnc13* expression in response to CVB5 infection. EndoC-βH1 cells were left uninfected (-) or infected with the diabetogenic Coxsackievirus B5 (CVB5) for 24h using a multiplicity of infection (MOI) of 5. Expression of *Lnc13* was assessed by qPCR and normalized by the housekeeping gene *β-actin*. Results are means ± SEM of 5 independent experiments.**p<0.01; Student's t test.

Virus induced inflammation in pancreatic β cells is partially mediated by the activation of type I IFN and the STAT1 signaling pathway (Eizirik et al., 2020; Marroqui et al., 2015; Santin et al., 2011). In concordance with this, Coxsackie virus B5 (CVB5) infection and intracellular PIC exposure increased *STAT1* RNA expression by around 3- and 10-fold, respectively (**Figure 23A-B**). Interestingly, the expression of *Lnc13* and *STAT1* was tightly correlated in pancreatic β cells upon exposure to intracellular PIC exposure (Pearson' s r^2 = 0.99; $p < 0.05$) (**Figure 23C**). In line with these results, the expression levels of *STAT1* and *Lnc13* were also correlated in a set of human pancreatic islets from 43 non-diabetic individuals. In general, the islets that expressed higher amounts of *Lnc13* also had higher levels of *STAT1* (Spearman's R= 0.51 [0.24–0.71]; p<0.001) (**Figure 23D**).

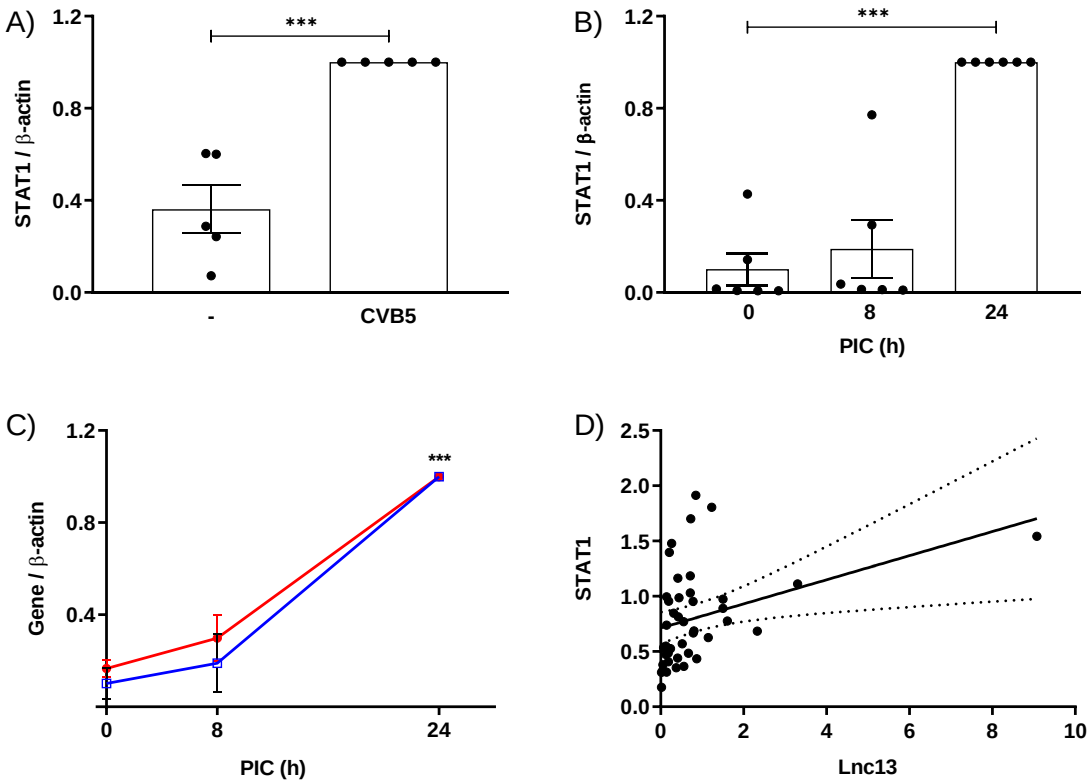


Figure 23. *Lnc13* expression correlates with *STAT1* expression in pancreatic β cells. (A) Human EndoC-βH1 cells were left uninfected (-) or infected with the CVB5 (CVB5; MOI=5) for 24 h. *STAT1* expression was determined by qPCR and normalized by the housekeeping gene *β-actin*. Results are means ± SEM of five independent experiments. **p<0.01; Student's t test. (B) EndoC-βH1 cells were left untransfected (0h) or transfected with PIC (1µg/ml) for 8 or 24h. Expression of *STAT1* was assayed by qPCR and normalized by the housekeeping gene *β-actin*. Results are means ± SEM of 6 independent experiments; ***p<0.001 vs. time 0h; Student's t test. (C) EndoC-βH1 cells were kept untreated (0h) or treated with PIC for 8 or 24 h. Relative *Lnc13* (red) and *STAT1* (blue) expressions were determined by qPCR and normalized by the housekeeping gene *β-actin*. Results are means ± SEM of six independent experiments; ***p<0.001 vs. time 0 h; Student's t test.(D) *Lnc13* and *STAT1* expressions were determined in 43 human pancreatic islets of non-diabetic individuals. Expression values were normalized by the housekeeping gene β-actin. Spearman's correlation analysis was performed to check for correlation between *Lnc13* and *STAT1* expression values; Spearman's R = 0.51 (0.24–0.0.71); p< 0.001.

3. The T1D-associated SNP risk genotype in the *Lnc13* gene correlates with increased *STAT1* expression in human pancreatic islets.

In order to determine whether the T1D-associated SNP genotype in *Lnc13* affected *STAT1* expression levels, rs917997 was genotyped in 43 human pancreatic islet cDNA samples from non-diabetic individuals and the expression of *STAT1* was determined in

the same samples. Out of the 43 cDNA samples, 15 were homozygous for the risk allele (rs917997-CC), 2 samples were homozygous for the protective allele (rs917997-TT) and the other 26 samples were heterozygous. Regarding *STAT1* levels, as shown in **Figure 24**, islets homozygous for the rs917997 risk allele presented higher expression levels of *STAT1* than the islets heterozygous or homozygous for the T1D protective allele.

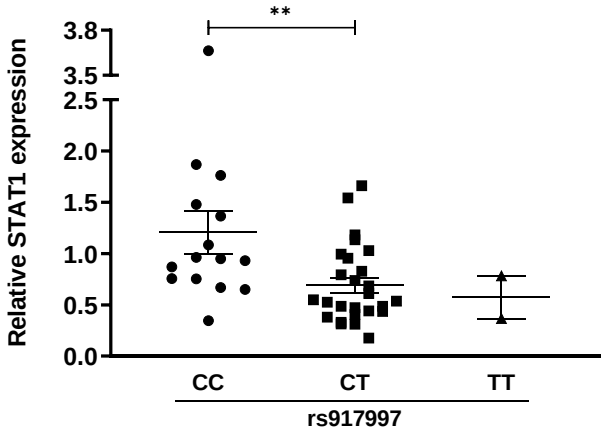


Figure 24. The T1D-associated SNP genotype in the *Lnc13* gene correlates with *STAT1* expression in human pancreatic islets. Expression of *STAT1* in human pancreatic islets stratified by the genotype of the T1D-associated SNP rs917997 in the *Lnc13* gene. Results are means ± SEM of 15 samples with the homozygous risk genotype (CC), 26 samples with the heterozygous genotype (CT), and two samples with the homozygous protective genotype (TT). **P < 0.01; Student's t test.

4. Lnc13 overexpression activates the STAT1 signaling pathway and increases production of pro-inflammatory chemokines in an allele-specific manner.

To follow up with the functional characterization of *Lnc13*, the next step was to clarify the molecular mechanisms by which the T1D-associated SNP rs917997 in *Lnc13* affect the expression level of *STAT1* in β cells .To this aim, allele specific *Lnc13*-overexpressing vectors were generated (one plasmid harboring the T1D-risk allele (pLnc13-C) and one harboring the protective allele (pLnc13-T)). As shown in **Figure 25A**, transfection of both plasmids efficiently upregulated the expression of *Lnc13*, however while pLnc13-C resulted in a 1000-fold overexpression, pLnc13-T increased *Lnc13* expression by around 2.500-fold. Taking into account that pLnc13-T and pLnc13-C expression was under the control of the same promoter (CMV6 promoter), the differences in the

overexpression levels obtained by the transfection of each plasmid could be due to the effect of the SNP allele in the stability of the RNA molecule.

After establishing the conditions for the allele-specific overexpression of *Lnc13*, the expression of *STAT1* was determined in *Lnc13*-overexpressing β cells. As observed in **Figure 25B**, the expression levels of *STAT1* were very similar in both, pLnc13-C- and pLnc13-T- transfected cells. However, taking into account the differences in *Lnc13* expression levels obtained by the transfection of pLnc13-C and pLnc13-T overexpression plasmids, the expression of *STAT1* was corrected by the expression of *Lnc13*. This correction revealed that *STAT1* expression was higher in β cells overexpressing the lncRNA harboring the risk allele for T1D (Lnc13-C) (**Figure 25C**).

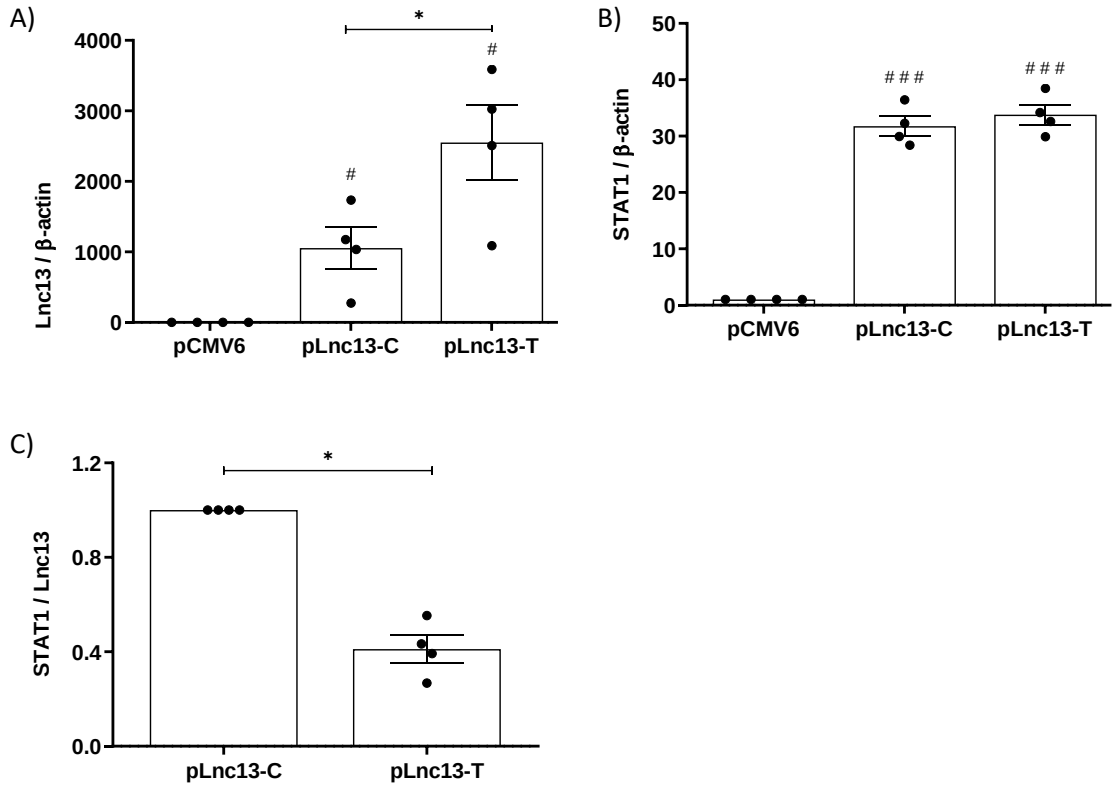


Figure 25. *Lnc13* upregulation increases *STAT1* expression in β cells in an allele-specific manner. EndoC-βH1 cells were transfected with pCMV6, pLnc13-C, or pLnc13-T. After 48h, expression levels of *Lnc13* (A) and *STAT1* (B) were determined by qPCR and normalized by the housekeeping gene *β-actin*. Results are means ± SEM of four independent experiments; # # # $p < 0.001$ and # $p < 0.05$ when compared the indicated sample with the pCMV6 *$p<0.05$; Student's t test. (C) *STAT1* mRNA expression in pLnc13-C- and pLnc13-T-transfected β cells corrected by Lnc13 expression values to control for differences in *Lnc13* allele stability. Results are means ± SEM of four independent experiments;*$p<0.05$; Student's t test.

Taking into account that *Lnc13* upregulation increased the expression of the pro-inflammatory transcription factor STAT1, I decided to check the expression of several pro-inflammatory chemokines known to be regulated by STAT1 (e.g. *CXCL10, CXCL9* and *CCL5)*. Overexpression of *Lnc13* in EndoC-βH1 cells led to an increase in the expression of the three chemokines under study (**Figure 26A-C**). After correction by the expression level of *Lnc13,* a significant increment of *CXCL10, CXCL9* and *CCL5* expression was observed in pLnc13-C-transfected β cells in comparison with β cells transfected with pLnc13-T (**Figure 26D-F**). These data suggested that *Lnc13* upregulation promoted the expression of pro-inflammatory chemokines in β cells in an allele-specific manner, meaning that the upregulation of the *Lnc13* harboring the risk allele for T1D induced a higher increase in chemokine expression than the *Lnc13* harboring the protective T1D allele.

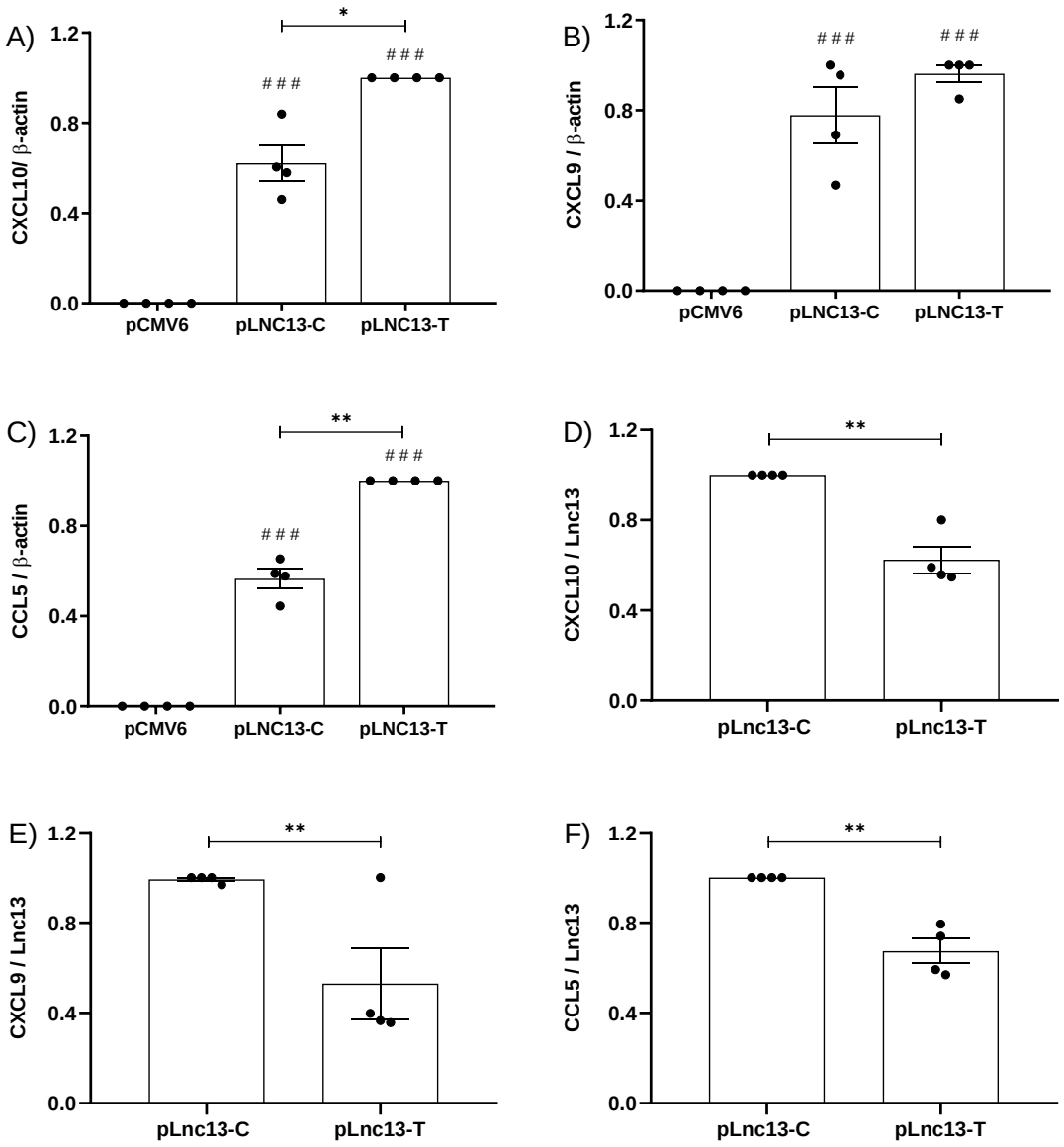


Figure 26. Overexpression of *Lnc13* in human pancreatic β cells led to increased pro-inflammatory chemokine expression in an allele-specific manner.(A-C) EndoC-βH1 cells were transfected with pCMV6 (empty vector used as control), pLnc13-C or pLnc13-T. After 48h, mRNA of *CXCL10* (A), *CXCL9* (B) and *CCL5* (C) was determined by qPCR and normalized by the housekeeping gene *β-actin*. (D–F) *CXCL10*, *CXCL9* and *CCL5* expression levels in pLnc13-C- or pLnc13-T-transfectred cells were corrected by *Lnc13* expression values to control for differences in *Lnc13* allele stability. Results are means ± SEM of 4 independent experiments; # # # $p < 0.001$ when compared the indicated sample with the pCMV6; *$p<0.05$ and **$p<0.01$; Student's t test.

I next decided to check CXCL10 and CCL5 protein levels secreted to the medium under the same conditions using an Enzyme-Linked ImmunoSorbent Assay (ELISA). Similar to

the observations made at the mRNA level, *Lnc13* upregulation led to increased secretion of CXCL10 and CCL5 when compared to control cells (pCMV6-transfected cells) (**Figure 27A-B**). Moreover, as shown in **Figure 27C-D**, EndoC-βH1 cells secreted higher amounts of CXCL10 and CCL5 when the T1D risk allele was present in *Lnc13*, suggesting that β cells harboring the risk allele in *Lnc13* secreted more pro-inflammatory mediators to the medium that cells harboring the T1D protective allele.

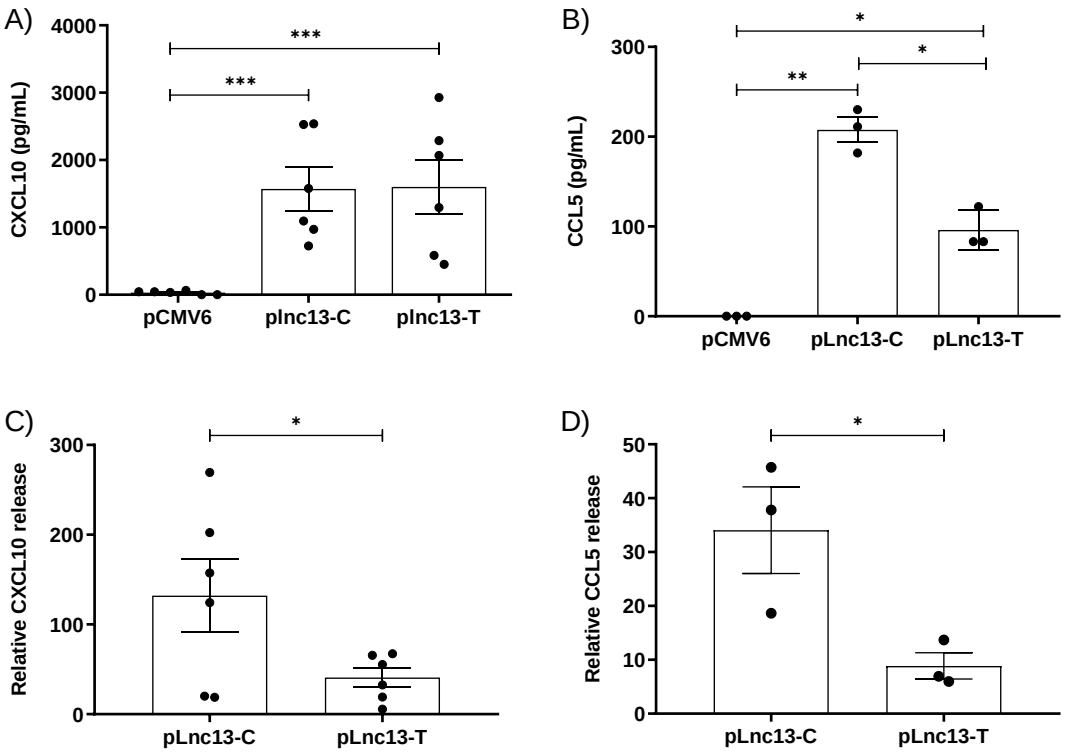


Figure 27. Overexpression of *Lnc13* in human pancreatic β cells led to an allele-specific increase of pro-inflammatory chemokine secretion in EndoC-βH1.(A-B) EndoC-βH1 cells were transfected with pCMV6 (empty vector used as control) with pLnc13-C or pLnc13-T. After 48h, protein levels were determined by ELISA in cell supernatants. (C-D) CXCL10 and CCL5 protein levels were corrected by *Lnc13* expression values to control differences in *Lnc13* allele stability. Results are means ± SEM of 3 or 6 independent experiments; *p<0.05, **p<0.01 and ***p<0.001; Student's t test.

Next, to determine whether *Lnc13* affects PIC-induced pro-inflammatory chemokine expression through modulation of the STAT signalling pathway, I exposed *Lnc13*-C-overexpressing EndoC-βH1 cells to Ruxolitinib, a JAK inhibitor. As shown in **Figure 28A**, *Lnc13*-C overexpression enhanced PIC-induced *CXCL10* mRNA expression, and this

effect was counteracted by the presence of Ruxolitinib. These data suggest that *Lnc13* upregulation induces *CXCL10* expression via activation of the STAT1 signalling pathway. In order to clarify whether *Lnc13* participated in the regulation of *CXCL10* expression by facilitating the binding of the STAT1 transcription factor to the *CXCL10* promoter, I next performed a chromatin immunoprecipitation (ChIP) assay. The main goal of this experiment was to analyze the capacity of STAT1 protein to bind to *CXCL10* promoter in the presence or absence of *Lnc13* upregulation. According to the data shown in **Figure 28B**, binding of STAT1 to *CXCL10* promoter was stronger in *Lnc13*-overexpressing β cells than in pCMV6-transfected control cells. These results confirmed that the upregulation of *CXCL10* expression in EndoC-βH1 cells overexpressing *Lnc13* was partially driven by an increased activation of the STAT1 signaling pathway.

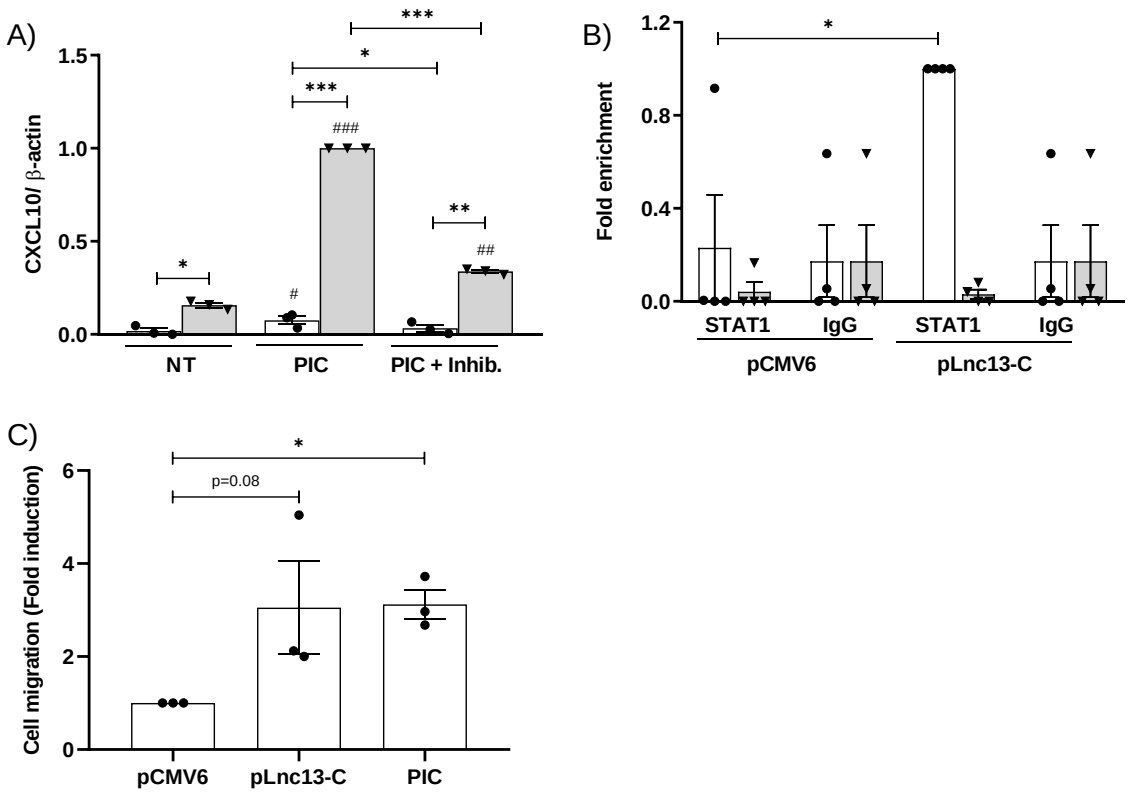


Figure 28. Upregulation of pro-inflammatory chemokine expression in *Lnc13*-overexpressing EndoC-βH1 cells is driven by an increased activation of the STAT1 signaling pathway. (A)Human EndoC-βH1 cells were transfected with pCMV6 (white bars) or with pLnc13-C (grey bars). Subsequently, cells were left untreated (NT), treated with intracellular PIC (1 µg/mL) for 24 h (PIC), or treated with PIC and Ruxolitinib for24 h (PIC + Inhib). *CXCL10* mRNA expression

was measured by qPCR and normalized by the housekeeping gene *β-actin*. The results are means ± SEM of three independent experiments; # p< 0.05, ## p< 0.01 and ### p<0.001 for. NT vs PIC or PIC + Inhib transfected with the same plasmid; *p< 0.05, **p< 0.01 and ***p< 0.001 as indicated; ANOVA followed by Student's t test followed with Bonferroni correction. (B) EndoC-βH1s were transfected with pCMV6 or pLnc13-C, chromatin was fragmented and precipitated with anti-STAT1 or anti-IgG (as the negative control), and *CXCL10* promoter or a control region (*OCT4* gene body) was amplified by qPCR. Results are means ± SEM of four independent experiments; *p< 0.05 as indicated; ANOVA followed by Student's t test followed with Bonferroni correction. (C) Supernatants of pCMV6- or pLnc13-C-transfected EndoC-βH1 cells were used to determine chemotactic migration of Jurkat cells using a transwell system and a fluorescence-based assay. Supernatant of PIC-transfected cells was used as the positive control. The results are means ± SEM of three independent experiments.*p< 0.05 as indicated; Student's t test.

To determine the biological relevance of *Lnc13*-driven chemokine production, I performed a chemotactic migration assay using supernatants of pCMV6- or pLnc13-C-transfected β cells. As shown in **Figure 28C**, supernatants of *Lnc13*-C-overexpressing β cells induced a higher migration of Jurkat T cells than the control supernatants (e.g. pCMV6-transfected control cells).

5. *Lnc13* disruption by CRISPR-Cas9 counteracts PIC-induced chemokine expression.

In order to obtain a mirror image of the overexpression experiments, using the CRISPR-Cas9 technique I partially disrupted the *Lnc13* gene in the EndoC-βH1 cell line by generating a deletion of 1698 bp (**Figure 29A**). However, due to the low proliferation rate and fragility of the EndoC-βH1 cell, I did not obtain a homogenous *Lnc13*-disrupted cell line. However, the mix population obtained (a combination of both disrupted *Lnc13* and wild type *Lnc13*) showed about a 60% reduction of *Lnc13* expression when compared to wild type cells (**Figure 29B**). Under this condition, the expression of *STAT1, CXCL10* and *CCL5* was measured in basal and PIC-transfected cells. As shown in **Figure 29C-E**, as a consequence of *Lnc13* reduction, PIC-induced *STAT1, CXCL10* and *CCL5* expression was reduced by 3.5-, 3.5- and 3.7-fold, respectively.

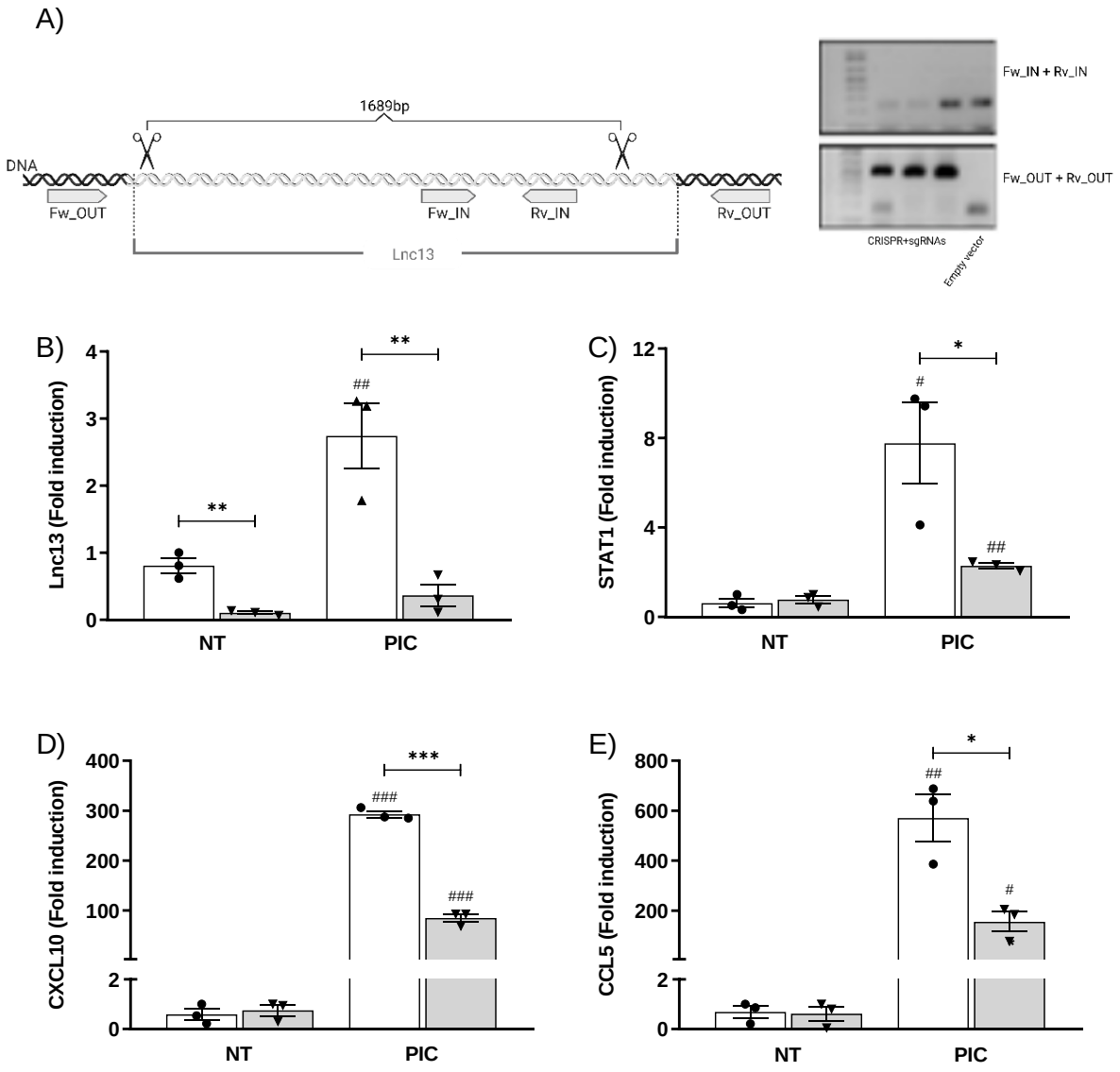


Figure 29. *Lnc13* gene disruption in pancreatic β cells partially counteracts the effect of PIC in *STAT1* and pro-inflammatory chemokine expression upregulation. (A) *Lnc13* disruption was performed by generating a deletion of 1,698 bp using the CRISPR-Cas9 technique and single guide RNAs (sgRNAs) targeting the *Lnc13* gene. The presence of the deletion was confirmed by PCR using a pair of primers located inside the deleted region (for detection of unedited cells; wild type forward (WTF) and wild type reverse (WTR)) and a pair of primers located outside the deleted region (for detection of edited cells; mutF and mutR). (B) EndoC-βH1 cells were transfected with an empty px459 vector (white bars) or with a vector harboring the Lnc13 targeting sgRNAs (grey bars). Lnc13 expression was determined by qPCR and normalized by the housekeeping gene *β-actin*. The results are means ± SEM of three independent cell populations.##p< 0.01 vs. untreated cells; **p< 0.01 as indicated; Student's t test (C–E) Control and *Lnc13*-disrupted mixed cell populations were exposed to intracellular PIC for 24 h, *STAT1* (C), *CXCL10* (D), and *CCL5* (E) expressions were determined by qPCR and normalized by the housekeeping gene *β-actin*. The results are represented as fold induction and are means ± SEM of three independent cell populations. ###p< 0.001, ##p< 0.01 and #p< 0.05 vs. untreated cells; ***p< 0.001 and *p< 0.05 as indicated; Student's t test.

6. *Lnc13* does not directly regulate the transcription of the *STAT1* gene, but participates in its mRNA stabilization in an allele-specific manner.

The next step was to determine the molecular mechanisms by which *Lnc13* activates the STAT1 signaling pathway in pancreatic β cells. The first attempt was to check the potential effect of *Lnc13* on the transcriptional regulation of *STAT1*. The regulation of *STAT1* transcription has been well characterized and is mainly mediated by the binding of several transcription factors to its promoter, including homodimers of phosphorylated STAT1 and heterodimers of phosphorylated STAT1/STAT2 or STAT1/STAT3 (Butturini et al., 2020; F. He et al., 2005; Yuasa & Hijikata, 2016).

Having this autoregulatory mechanism in mind, firstly I analyzed the potential implication of *Lnc13* in STAT1 phosphorylation (pSTAT1). As shown in **Figure 30A-B**, *Lnc13* upregulation led to increased pSTAT1 levels. However, the different T1D alleles in *Lnc13* did not affect the phosphorylation level, suggesting that pSTAT1 levels were not responsible for the differences observed in *Lnc13* allele-specific upregulation of chemokine expression.

The next step was to analyze the potential implication of *Lnc13* in the transcriptional activation of *STAT1*. To this aim, cycloheximide (an inhibitor of protein translation) was added to *Lnc13*-overexpressing EndoC-βH1 cells for 24h to inhibit the synthesis of the STAT1 protein. The protein fraction of the cells were taken to check whether the treatment with cycloheximide efficiently decreased STAT1 protein levels (**Figure 30C-D**). Once checked, the RNA from the same samples was extracted to determine both *STAT1* mRNA expression levels, and *STAT1* primary transcript using specific primers. Cycloheximide treatment decreased the expression of *STAT1* primary transcript levels by approximately 5 times, whereas it did not affect *STAT1* total mRNA levels as they maintained similar in non-treated and cycloheximide-treated cells (**Figure 30D**). This information suggested that in the absence of pSTAT1 protein synthesis *Lnc13* was not modulating *STAT1* gene transcription.

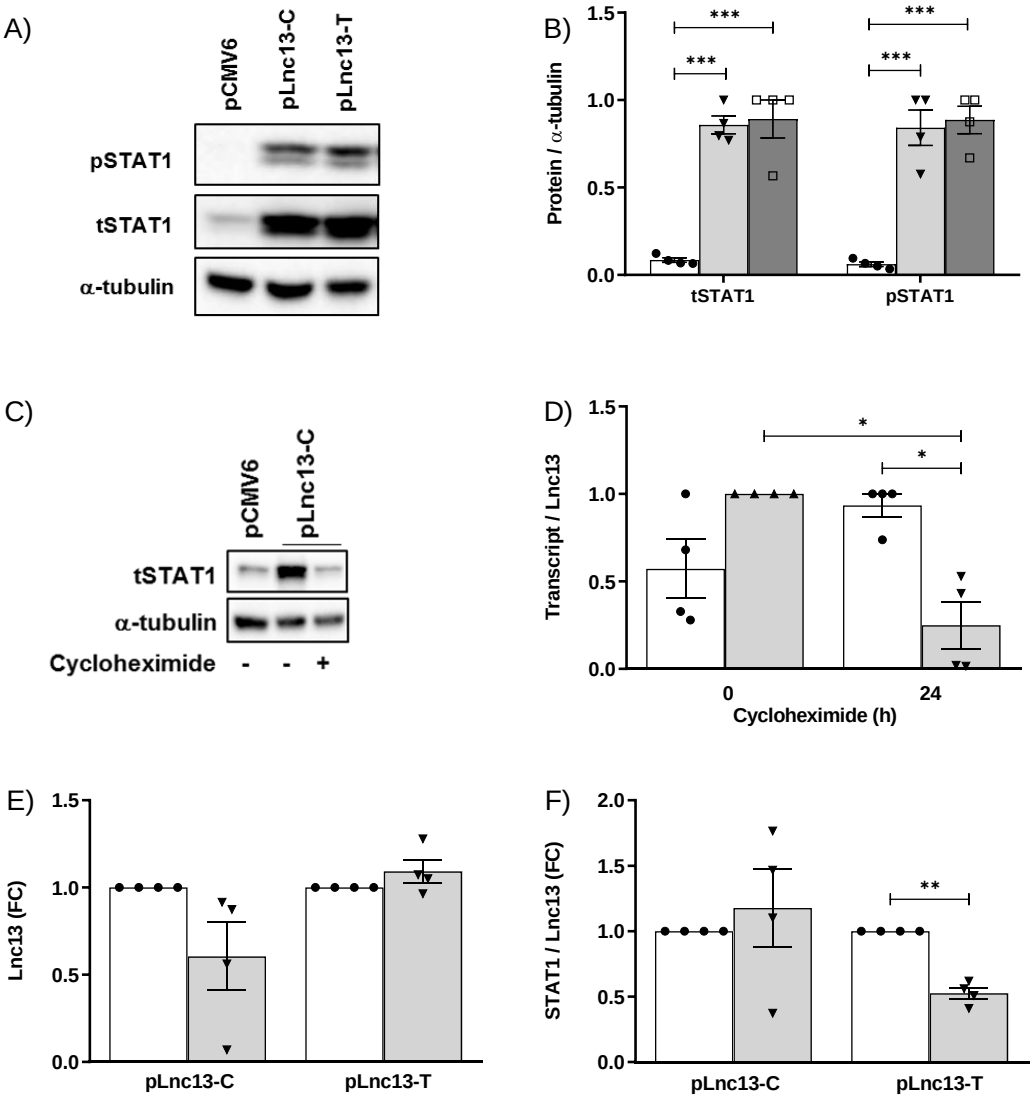


Figure 30. Overexpression of *Lnc13* in β cells increases *STAT1* signaling pathway activation. (A) EndoC-βH1 cells were transfected with pCMV6 (empty vector used as control, white bars), pLnc13-C (light grey) or pLnc13-T (dark grey) and protein levels of phospho-STAT1 (pSTAT1), total STAT1 (tSTAT1) and α-tubulin (as loading control) were determined by Western blot. The results are representative of 4 independent experiments. (B) Densitometry results for pSTAT1 and tSTAT1 are represented as means ± SEM of 4 independent experiments. ***p<0.001 as indicated; Student's t test. (C) EndoC-βH1 cells were transfected with pCMV6 or pLnc13-C and kept untreated or treated with cycloheximide (50µg/ml) for 24h as indicated. Total *STAT1* (tSTAT1) and *α-tubulin* (as loading control) were determined by Western blot. Results are representative of two independent experiments. (D) EndoC-βH1 cells were transfected with pLnc13-C and treated with cycloheximide (50µg/ml) for 24h. Expression values of *STAT1* total mRNA (white bars) and *STAT1* primary transcript (grey bars) were determined by qPCR and normalized by the housekeeping gene *β-actin* and corrected by *Lnc13* expression values. *STAT1* total mRNA was determined using a Taqman gene expression assay in which the probe spans an exon-exon junction and STAT1 primary transcript was determined using primers located in intron 11 of *STAT1* gene in RNA samples treated with DNAse to avoid detection of genomic DNA. Results are means ± SEM of 4 independent experiments; *p<0.05 as indicated; Student's t test. (E) EndoC-βH1 cells were transfected with pLnc13-C and pLnc13-T and kept

untreated (white bars) or treated with actinomycin D for 6h (grey bars). Expression of *Lnc13* was determined by qPCR. Results are represented as fold change (FC) and are means ± SEM of 4 independent experiments and are expressed as fold change against non-treated samples. (F) EndoC-βH1 cells were transfected with pLnc13-C and pLnc13-T and kept untreated (white bars) or treated with actinomycin D for 6h (grey bars). Expression of *STAT1* was determined by qPCR and corrected by *Lnc13* expression to control for differences in *Lnc13* allele stability. Results are represented as fold change (FC) against non-treated samples and are means ± SEM of 4 independent experiments **p<0.01 as indicated; Student's t test.

Taking into account that *Lnc13* upregulation was not promoting increased transcription of *STAT1*, I next checked whether *Lnc13* was implicated in the stabilization of *STAT1* mRNA. To this aim, transcription was inhibited using actinomycin D in pLnc13-C- or pLnc13-T-overexpressing β cells. Under these conditions, total RNA was extracted and the expression of *Lnc13* and *STAT1* measured by qPCR. As shown in **Figure 30E**, *Lnc13* harboring the risk allele for T1D (*Lnc13*-C) was less stable than *Lnc13*-T. In fact, after 6h actinomycin D exposure, *Lnc13* expression was reduced by around a 40% in pLnc13-C-transfected EndoC-βH1 cells, but did not change in pLnc13-T-transfected cells. Regarding *STAT1* mRNA levels, in pLnc13-T-transfected EndoC-βH1 cells, *STAT1* expression showed a 1.9-fold decrease (**Figure 30F**), while in Lnc13-C-overexpressing β cells, *STAT1* mRNA expression remained similar in the presence or absence of actinomycin D. Regardless of Lnc13-C being less stable, it was capable to stabilize *STAT1* mRNA molecules more efficiently that *Lnc13*-T.

7. *Lnc13* interacts with PCBP2 in the cytoplasm of EndoC-βH1 cells leading to enhanced *STAT1* gene expression in response to viral dsRNA.

The cellular localization of lncRNAs often determines their functional role (Ransohoff et al., 2018), so that to further characterize the function of *Lnc13*, its subcellular localization was assessed in basal and PIC-transfected EndoC-βH1 cells. As shown in **Figure 31A.** *Lnc13* in basal condition (e.g. non-treated cells) was preferentially expressed in the nuclei of pancreatic β cells, while PIC transfection induced its translocation into the cell cytoplasm.

Previous studies have shown that in mouse macrophages *Lnc13* is able to bind a protein called PCBP2 (Castellanos-Rubio et al., 2016). This protein is a RNA-binding protein that participates in antiviral cellular response by stabilizing the 3'UTR of the *STAT1* mRNA molecules (Xin et al., 2011). Having this in mind, firstly the effect of intracellular PIC on PCBP2 protein expression in β cells was assessed. As shown in **Figure 31B**, as observed with *Lnc13*, PIC induced an increase in PCBP2 protein expression, both in the nucleus and in the cytoplasm, of EndoC-βH1 cells.

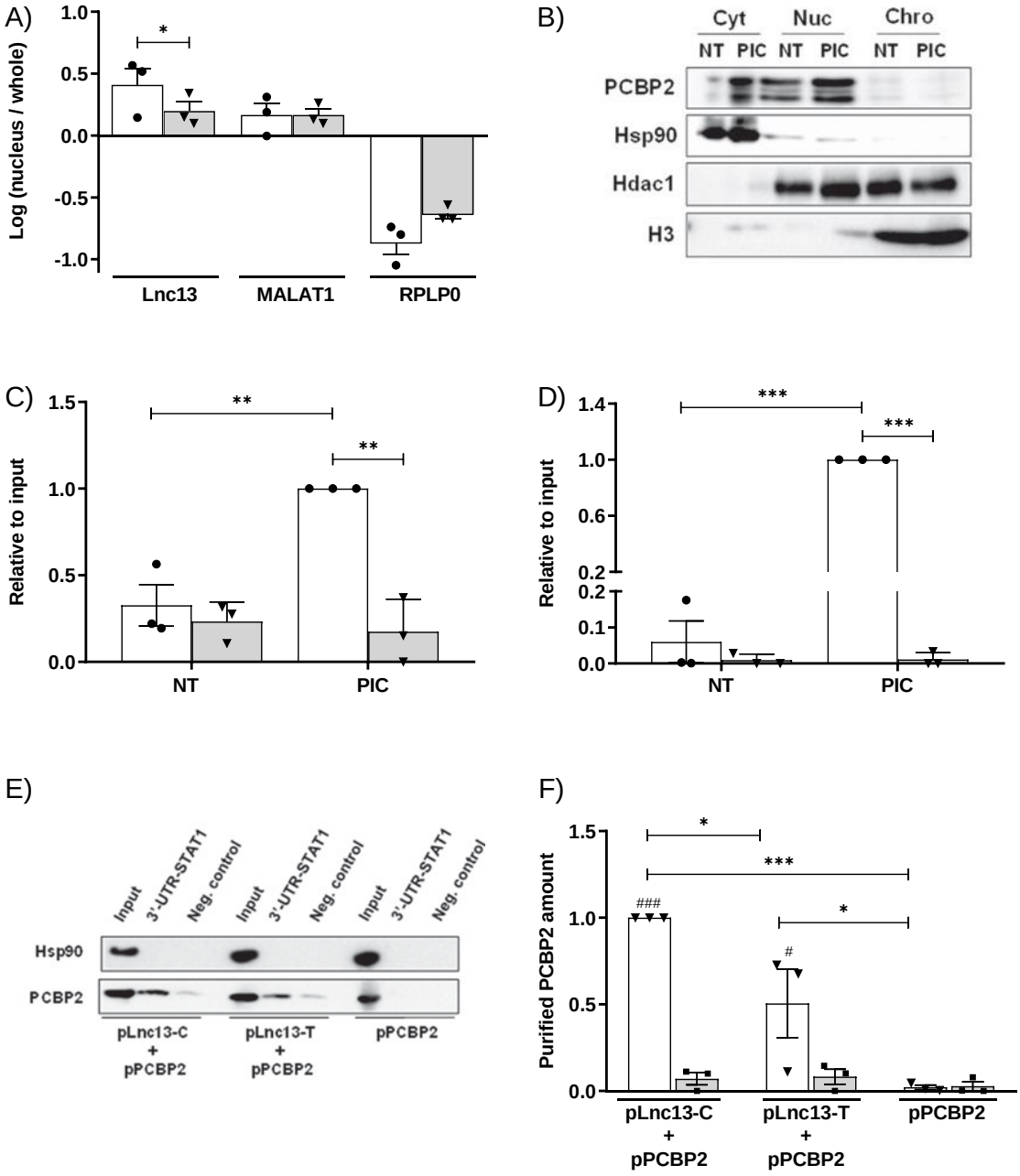


Figure 31. After PIC exposure, *Lnc13* translocates from the nucleus to the cytoplasm where it facilitates the interaction between PCBP2 and the 3'UTR of the gene *STAT1*. (A) EndoC-βH1 cells were left non-transfected (white bars) or exposed to PIC (1 μg/mL) for 24 h (grey bars), and relative *Lnc13* expression was determined in nucleus and whole extracts. Expression of *MALAT1* and *RPLP0* was used as a control for nucleus and whole cell fractions, respectively. Amounts of specific nuclear RNA were measured by qPCR and compared to the total amount of RNA in the whole cell. Results are represented as a logarithm (nucleus/whole) and are means ± SEM of three independent experiments;*p< 0.05; Student's t test. (B) EndoC-βH1 cells were exposed to PIC (1 μg/mL) for 24 h and cytoplasmic (Cyt), nuclear (Nuc), and chromatin (Chro) fractions were purified. PCBP2 protein expression was determined in all cellular compartments by Western blot and protein expression of Hsp90, HDAC1, and H3 was used as controls for cytoplasmic, nuclear, and chromatin fractions, respectively. The results are representative of three independent experiments. (C) EndoC-βH1 cells were left non-transfected (NT) or transfected with PIC for 24 h. RNA immunoprecipitation was performed using a specific antibody for PCBP2. *Lnc13* expression was determined in PCBP2-bound RNA by qPCR. Results are means ± SEM of three independent experiments, and the amounts of *Lnc13* are expressed as relative to the input (white bars). IgG was used as the negative control (grey bars); **p< 0.01 as indicated; ANOVA followed by Student's t test. (D) EndoC-βH1 cells were left non-transfected (NT) or transfected with PIC for 24 h. RNA immunoprecipitation was performed using a specific antibody for PCBP2. *STAT1* expression was determined in PCBP2-bound RNA by qPCR. Results are means ± SEM of three independent experiments, and the amounts of STAT1 are expressed as relative to the input (white bars). IgG was used as the negative control (grey bars); ***p< 0.001, as indicated; ANOVA followed by Student's t test. (E) In vitro transcribed biotinylated 3'-UTR region of *STAT1* was incubated with cellular extracts overexpressing Lnc13-C+PCBP2, Lnc13-T+PCBP2, or PCBP2. Afterward, 3'-UTR-STAT1-bound proteins were purified using streptavidin beads, and PCBP2 and Hsp90 (as the negative control) were detected by Western blot. Incubation with streptavidin beads alone was used as the negative control. The results are representative of three independent experiments. (F) Densitometry results for purified PCBP2 amounts in RNA pull-down experiments are represented as means ± SEM of three independent experiments. PCBP2 protein bound to the 3'UTR-STAT1 is represented in white bars, whereas PCBP2 bound to the negative control is represented in grey bars. ANOVA followed by Student's t test with Bonferroni correction; ***p< 0.001 and *p< 0.01 as indicated, ### p< 0.001 and # p< 0.05 compared versus the negative control.

Next, to test whether there was a physical interaction between PCBP2 and *Lnc13,* I decided to perform a RNA immunoprecipitation assay using cellular extracts from non-treated and PIC-transfected EndoC-βH1 cells. To this aim, PCBP2-bound RNA molecules were purified and *Lnc13* expression was determined by qPCR. As shown in **Figure 31C**, in non-treated cells, *Lnc13* slightly interacted with PCBP2, but after PIC transfection, the binding between *Lnc13* and PCBP2 was significantly increased. In addition, *STAT1* expression was also analyzed in these samples, and as observed for *Lnc13*, *STAT1*

mRNA was poorly bound to PCBP2 under basal condition, but after PIC exposure, the expression level of PCBP2-bound *STAT1* increased by 17-fold (**Figure 31D**).

Once the interaction between *Lnc13* and PCBP2, as well as the interaction between *STAT1* and PCBP2, were confirmed, I decided to determine whether *Lnc13* was an auxiliary molecule for the binding between PCBP2 and the 3'UTR of STAT1 RNA molecule and whether the T1D allele in *Lnc13* affected this process. To this aim, I performed a RNA pull-down experiment to determine whether PCBP2 was differentially bound to the 3'UTR of *STAT1* in the presence of *Lnc13*-C or *Lnc13*-T. As shown in **Figure 31E-F**, the presence of *Lnc13* was necessary to enable the binding between PCBP2 and the 3'UTR of *STAT1.* Furthermore, *Lnc13* harboring the risk allele for T1D (pLnc13-C) provoked an increased interaction between PCBP2 and the 3'UTR of *STAT1,* comparing to *Lnc13* harboring the protective allele (pLnc13-T) (**Figure 31E-F**). These data confirm that *Lnc13* acted as a linker between PCBP2 and *STAT1.* Moreover, the data suggested that the risk allele for T1D present in *Lnc13* promotes a more robust binding between PCPB2 and the 3'UTR of *STAT1*, which in turn, leads to an increased stabilization of *STAT1* RNA molecule.

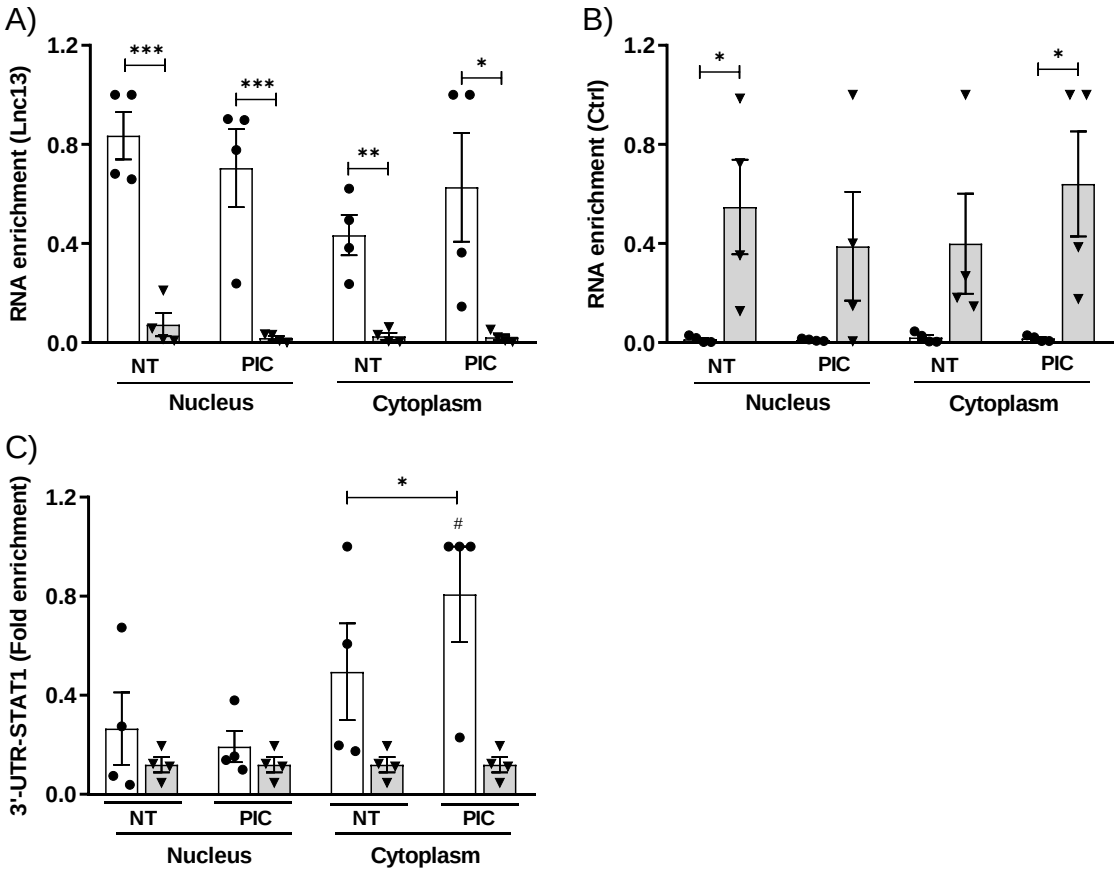


Figure 32. After PIC exposure, *Lnc13* facilitates the interaction between the protein PCBP2 and the 3'UTR of the gene *STAT1* in the cytoplasm of EndoC-βH1. *Lnc13* (A) or a control lncRNA (B) were purified using biotinylated specific antisense oligonucleotides in non-treated (NT) and PIC-transfected nuclear and cytoplasmic fractions of EndoC-βH1 cells, and RNA enrichment of *Lnc13* (white bars) or control lncRNA (grey bars) was determined by qPCR. The results are means ± SEM of 4 independent experiments. ***p<0.001, **p<0.01 and *p<0.05 as indicated; ANOVA followed by Student's t test with Bonferroni correction.(C) *Lnc13* antisense purification was performed in non-treated (NT) and PIC-transfected nuclear and cytoplasmic fractions of EndoC-βH1 cells. *Lnc13*-bound 3'-UTR-STAT1 amounts were determined by qPCR using specific primers (white bars). A nonrelated similar lncRNA was used as the negative control (grey bars).Results are represented as fold enrichment and are means ± SEM of four independent experiments. *p< 0.05 as indicated; ANOVA followed by Student's t test with Bonferroni correction.

To clarify whether viral dsRNA (e.g. PIC) promoted the interaction between *Lnc13* and *STAT1*, a RNA antisense purification (RAP) of *Lnc13* was performed in nuclear and cytoplasmic fractions of untreated and PIC-transfected EndoC-βH1 cells. The RNA antisense purification was performed using biotinylated antisense oligos complementary to *Lnc13* (**Figure 32A**), and also complementary oligos to a lncRNA with

similar length used as negative control (**Figure 32B**). As shown in **Figure 32C**, in the nuclei of β cells, there was not interaction between the purified *Lnc13* molecules and the 3'-UTR of *STAT1,* either in non-treated or PIC-transfected cells. In contrast, in the cytoplasmic fraction of EndoC-βH1 cells, *Lnc13* interacted with the 3'UTR of *STAT1,* especially after intracellular PIC exposure.

8. The region containing the T1D-associated SNP in *Lnc13* is crucial for the binding between PCBP2 and *STAT1*.

Finally, to clarify whether the region containing the T1D associated SNP in the *Lnc13* molecule was implicated in the binding between the PCBP2 and the 3'-UTR of *STAT1* RNA molecule, a mutant *Lnc13* (Lnc13-delSNP) that lacked a 507 bp region (1,771–2,278 bp; including the SNP for T1D) was generated and a RNA-protein interaction assay was performed. To this aim, β cells were transfected with pPCBP2 alone, pPCBP2+pLnc13C, pPCBP2+pLnc13-T or pPCBP2+pLnc13-delSNP and the cell lysates were incubated with *in vitro* transcribed 3'UTR-*STAT1* RNA molecules. The incubated lysates were run onto a native agarose gel and after electrophoresis, protein were transferred into a nitrocellulose membrane for PCBP2 visualization. As shown in **Figure 33A**, in the absence of *Lnc13*, there was not binding between PCBP2 and the 3'UTR of *STAT1*, as previously observed in *Lnc13* pull-down assays. While the migration pattern was similar between the PCBP2+Lnc13-C- and the PCBP2+Lnc13-T-overexpressing cells, cells expressing the mutant *Lnc13* (Lnc13-delSNP) presented a different migration profile, suggesting that the T1D associated SNP region was implicated in the configuration of the *Lnc13*-*STAT1*-PCBP2 complex.

The implication of the T1D-associated SNP region in *Lnc13* was further confirmed by analyzing the expression of *STAT1, CXCL10*, and *CCL5* in Lnc13-C- and Lnc13-delSNP-overexpressing EndoC-βH1 cells. As shown in **Figure 33B**, cells overexpressing the mutant *Lnc13* expressed less *STAT1, CXCL10* and *CCL5* than the cells overexpressing the wild-type *Lnc13* harboring the risk allele for T1D.

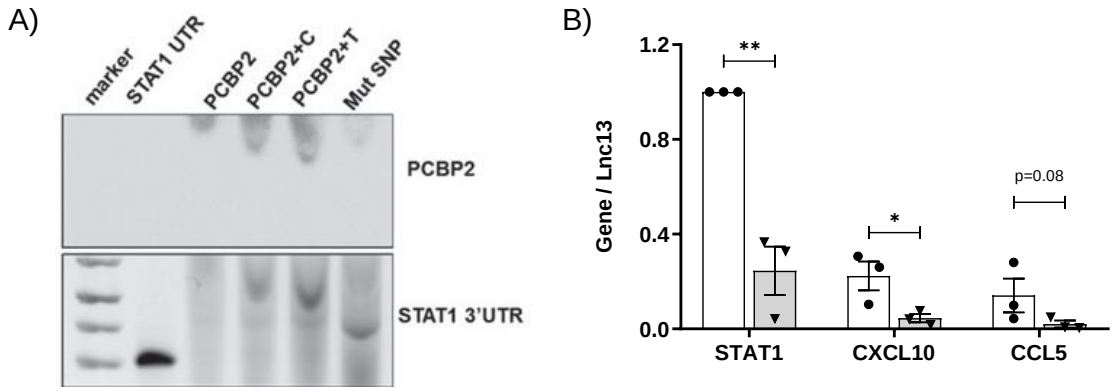


Figure 33. The rs917997 SNP region in *Lnc13* affects to the interaction between PCBP2 and *STAT1*. (A) RNA-protein interaction assay. Cells were transfected with pPCBP2 alone, pPCBP2+pLnc13-C, pPCBP2+pLnc13-T, or pPCBP2+pLnc13-delSNP. Cell lysates were incubated with in vitro transcribed 3'-UTR-*STAT1* molecules, and a native agarose gel electrophoretic mobility shift assay was performed (Lower image). After electrophoresis, proteins were transferred to a nitrocellulose membrane for PCBP2 visualization (Upper image). The 3'-UTR-*STAT1* RNA molecule alone was loaded as a control. The results are representative of three independent experiments. (B) EndoC-βH1 cells were transfected with a plasmid encoding *Lnc13*-C (white bars) or a mutant *Lnc13* in which the region containing the T1D-associated SNP was deleted (grey bars). *STAT1*, *CXCL10*, and *CCL5* expressions were determined by qPCR and normalized by the housekeeping gene *β-actin* and corrected by *Lnc13* expression values. Results are means ± SEM of four independent experiments. **p<0.01 and *p< 0.05 as indicated; Student's t test.

Gathering all these results together, these data demonstrated that *Lnc13* is able to enhance *STAT1* mRNA stability in the cytoplasm of β cells, by promoting the interaction between the 3'UTR of *STAT1* and the RNA-binding protein PCBP2. In addition, the T1D-associated SNP region in *Lnc13* (rs917997) is important for this interaction, as the deletion of this region seems to disrupt the binding between PCBP2 and Lnc13, affecting *Lnc13*-induced *STAT1* and chemokine upregulation.

Results 3

Functional characterization of *ARGI* in pancreatic β cells

Atal honetan aurkezten diren emaitzak hurrengo artikuluan izan dira publikatuak:

The results presented in this section have been published in:

González-Moro I., Garcia-Etxebarria K., Mendoza L.M., Fernández-Jimenez N., Mentxaka-Salgado J., Olazagoitia-Garmendia A, Arroyo M.N., Sawatani T., de Beek A.O., Cnop M., Igoillo-Esteve M. & Santin I. (2022) The type 1 diabetes-associated lncRNA *ARGI* participates in virus-induced pancreatic β cell inflammation. bioRxiv; DOI: 10.1101/2022.12.01.518685.

Laburpena:

1 motako diabetesarekin (T1D) loturiko nukleotido bakarreko aldaeren (SNP) gehiengoa giza genomako domeinu ez-kodetzaileetan daude kokatuak. RNA luze ez-kodetzaileetan (lncRNA) kokaturik dauden SNPek bigarren mailako egituraren haustura ekar dezake, molekulen funtzioa aldaraziz. Hemen, infekzio biralen ondorioz T1Dari loturik dagoen lncRNA bat funtzionalki karakterizatu egin da, *ARGI* (Antiviral Response Gene Inducer) deritzona. *ARGI*ren gainadierazpenak gene antibiral eta pro-inflamatorioen transkripzioaren aktibazio dakar pankreako β zeluletan. Infekzio biral baten ostean, *ARGI* pankreako β zeluletako nukleoetan gainadieraztea eragiten du, non CTCF transkripzio faktorearekin elkartzen den *IFNβ* eta interferonez-estimulaturiko geneen domeinu erregulatzaileekin interakzionatzea ahalbidetzen duena. Interakzio honek gene horien transkripzioaren aktibazioa sustatu egiten du era alelo-espezifikoan. T1Derako arrisku aleloa daraman *ARGI*k 1 motatako interferonen erantzuna hiperaktibatzea eragiten du pankreako β zeluletan, gaixotasuna pairatzen duten pertsonen pankrean berezkoa den profila. Datu hauek argitu egiten dute nola T1Darekin loturiko SNP baten agerpenak lncRNA batean gaixotasunaren patogenesian eragin dezaken pankreako β zeluletan.

1. *ARGI* is a nuclear lncRNA upregulated in EndoC-βH1 cells upon a viral infection

As previously described *NONHSAT233405.1* or *ARGI* (Antiviral Response Gene Inducer) is a lncRNA harboring one SNP associated with T1D (rs9585056) located in its third exon. A graphical representation of the genomic localization of *ARGI* and the T1D-associated SNP rs9585056 position is shown in **Figure 34**.

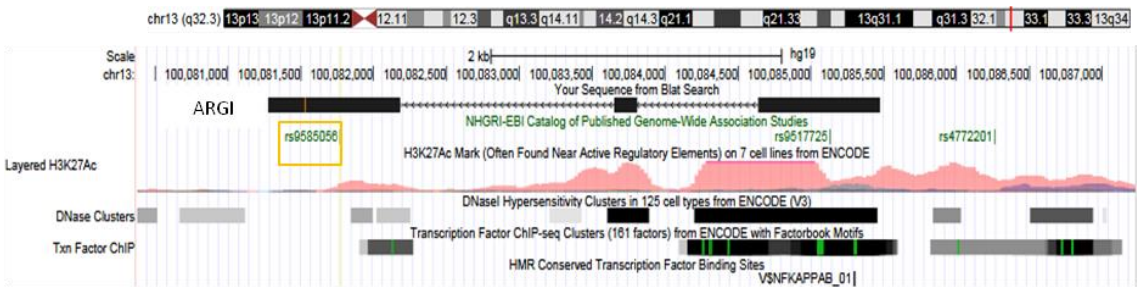


Figure 34. Genomic localization of *ARGI*. *ARGI* is an intergenic lncRNA located in human chromosome 13 (99,429,023-99,433,220; GRCh38/hg38). It has three exons and harbors one type 1 diabetes-associated SNP (rs9585056; chr13: 99,429,262-99) in its third exon (orange box). Epigenetic marks (H3K27Ac and DNase clusters) and a conserved NFkB binding site have been identified close to the transcription-starting site of *ARGI*.

As shown previously (**Figure 18**), *ARGI* is a lncRNA expressed in EndoC-βH1 cells and upregulated by viral dsRNA (e.g. PIC). In line with this observation, infection with two diabetogenic serotypes of Coxsackievirus, CVB1 and CVB4, also upregulated *ARGI* expression in induced pluripotent stem (iPS) cells-derived human β cells **(Figure 35**).

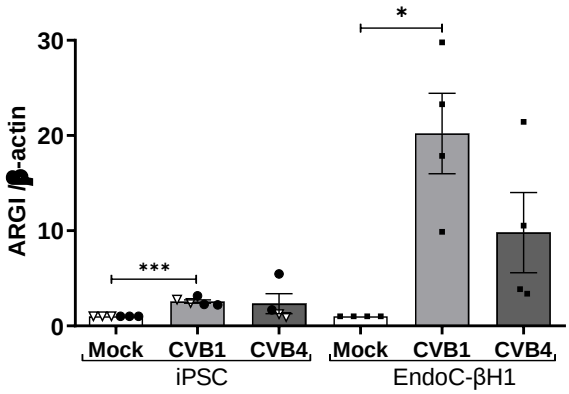


Figure 35. *ARGI* is upregulated in EndoC-βH1 cells and (iPS)-derived pancreatic β cells upon a Coxsackie virus infection. Induced pluripotent stem cell (iPSC)-derived pancreatic islet-like aggregates and EndoC-βH1 were left uninfected (Mock) or infected with CVB1 or CVB4 for 24h. ARGI expression was assessed by qPCR and normalized to the reference gene β-actin. Data are means±SEM of 3 independent 115.6 iPSC differentiations (white triangles), 3 independent 1023A iPSC differentiations (black circles) and 4 independent EndoC-βH1 (black squares); **p<0.001 and *p<0.05; Student's t test.

Considering that cellular location often determines the function of lncRNAs (Bridges et al., 2021; Carlevaro-Fita et al., 2019; Ransohoff et al., 2018), next the subcellular localization of *ARGI* in EndoC-βH1 cells was examined. Determination of subcellular localization was performed in basal condition but also after 24h of PIC transfection or after CVB1 infection. This experiment demonstrated that *ARGI* was expressed preferentially in the nuclei of β cells (**Figure 36A-B**). In addition, both PIC transfection and CVB1 infection increased significantly the presence of this lncRNA in the nuclei of EndoC-βH1 cells.

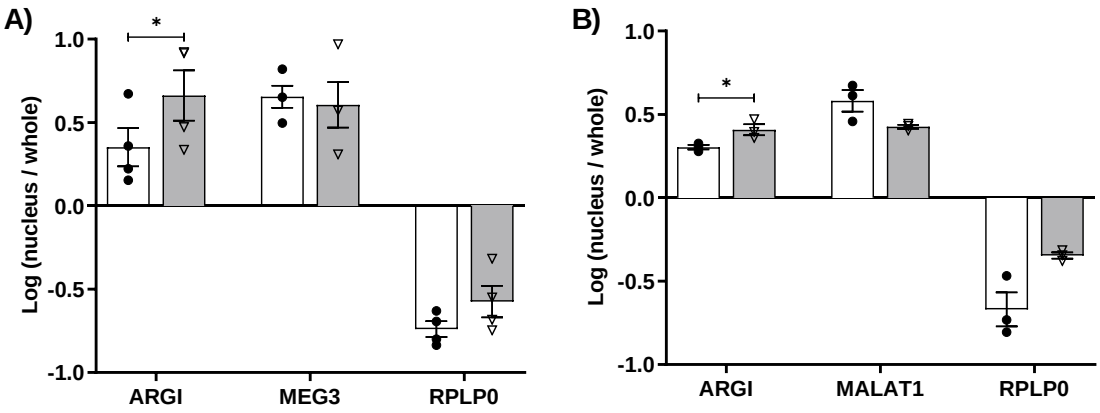


Figure 36. *ARGI* expression is enhanced in the nuclei of EndoC-βH1 cells upon a viral assault. (A) EndoC-βH1 cells were left untransfected (white bars) or exposed to PIC (1 µg/mL) for 24 h (grey bars), and relative ARGI expression was determined in nucleus and whole extracts. Expression of MEG3 and RPLP0 were used as a control for nucleus and whole cell fractions, respectively. Amounts of specific nuclear RNA were measured by qPCR and compared to the total amount of RNA in the whole cell extract. Results are represented as a logarithm (nucleus/whole) and are means ± SEM of four independent experiments; *p< 0.05; Student's t test. (B) EndoC-βH1 cells were left uninfected (white) or infected with CVB1 (MOI= 0.05) (grey). Relative ARGI expression was determined in nucleus and whole extracts. Expression of *MALAT* and *RPLP0* were used as a control for nucleus and whole cell fractions, respectively. Amounts of specific nuclear RNA were measured by qPCR and compared to the total amount of RNA in the whole cell. Results are represented as a logarithm (nucleus/whole) and are means ± SEM of three independent experiments; *p< 0.05; Student's t test.

2. *ARGI* upregulation in pancreatic β cells leads to a hyperactivation of an inflammatory and antiviral gene signature.

The fact that *ARGI* was upregulated upon a viral infection especially in the nuclei of β cells suggested its potential implication in the transcriptional regulation of antiviral and

pro-inflammatory responses. Thus, to identify the potential gene targets of *ARGI*, a RNA sequencing was performed with RNA extracted from control (pCMV6) and *ARGI*-overexpressing (pARGI) EndoC-βH1 cells. As shown in **Figure 37**, pARGI-transfected cells exhibited a 2000-fold increase in *ARGI* upregulation in comparison to pCMV6-transfected control cells.

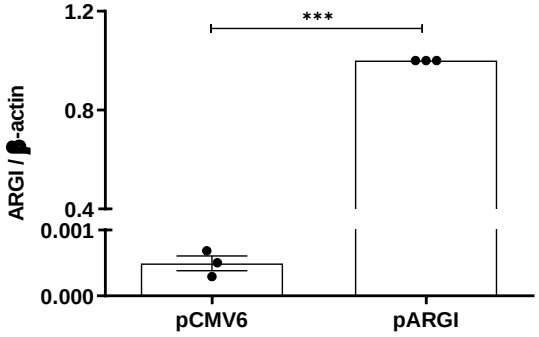


Figure 37. *ARGI* upregulation in pARGI-transfected pancreatic β cells. EndoC-βH1 cells were transfected with a control empty overexpression plasmid (pCMV6) or with a plasmid overexpressing *ARGI* (pARGI) for 24h. *ARGI* expression was determined by qPCR and normalized by the housekeeping gene β-actin. Results are means± SEM of three independent experiments; ***p<0.001; Student's t test.

After RNA sequencing of the samples in the Genomic Platform of CIC bioGUNE, raw data was checked for quality control and analyzed using Trimmomatic 0.39 for read trimming followed by an alignment by means of HIST2 using as reference Human Genome assembly hg38. In order to find genes and gen pathways potentially regulated by *ARGI* in β cells, a Gene Ontology (GO) enrichment analysis of the differentially upregulated genes in *ARGI*-overexpressing β cells was performed using the analysis tools of Enrichr platform (Kuleshov et al., 2016). As shown in **Figure 38**, GO analysis revealed that differentially upregulated genes were enriched in pathways related to antiviral responses and type I IFN signaling pathways, including cellular response to type I IFNs, type I IFN signaling pathway, and defense response to virus, among others.

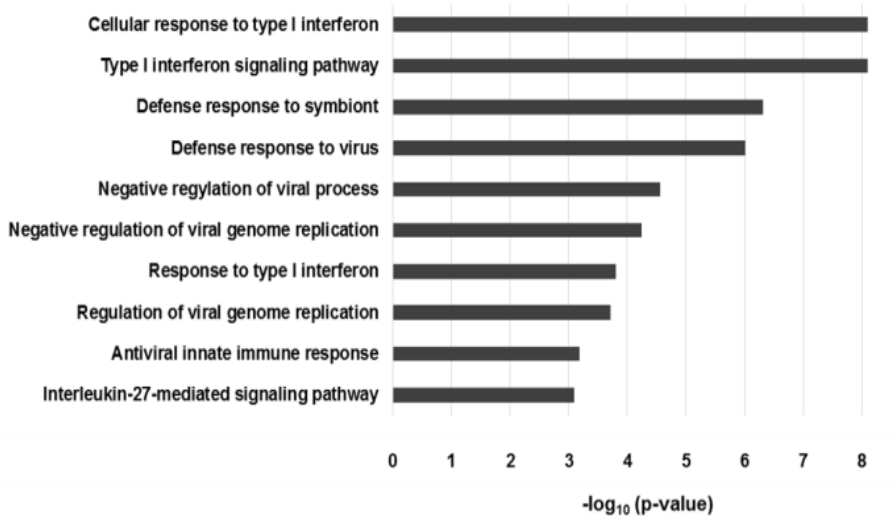


Figure 38. *ARGI* upregulation in pancreatic β cells leads to hyperactivation of an inflammatory and antiviral gene signature. Gene Ontology analysis showed an enrichment of inflammatory and antiviral pathways in *ARGI*-overexpressing pancreatic β cells. RNA sequencing was performed in pCMV6- and pARGI-transfected EndoC-βH1 cells.

Interestingly, among the genes that were differentially upregulated after *ARGI* overexpression, there was a significant overrepresentation of genes belonging to the IDIN network (**Figure 39**). Among the 17.249 transcripts detected in the RNAseq, 430 genes corresponded to the IDIN network, a 2.5% of the total. Besides, 4.32% of the significantly upregulated genes were members of the IDIN pathway, suggesting that *ARGI* controlled the expression of a significant amount of IDIN genes in pancreatic β cells (p-value: 0.0188). Among the top differentially upregulated IDIN genes, there were key antiviral and pro-inflammatory genes, such as the IFN-stimulated genes *MX1, ISG15, IFI6, IFIT1, IFIT3* and *STAT1*. The expression of these top six upregulated genes in *ARGI*-overexpressing β cells was confirmed in an independent sample set by qPCR (**Figure 40**).

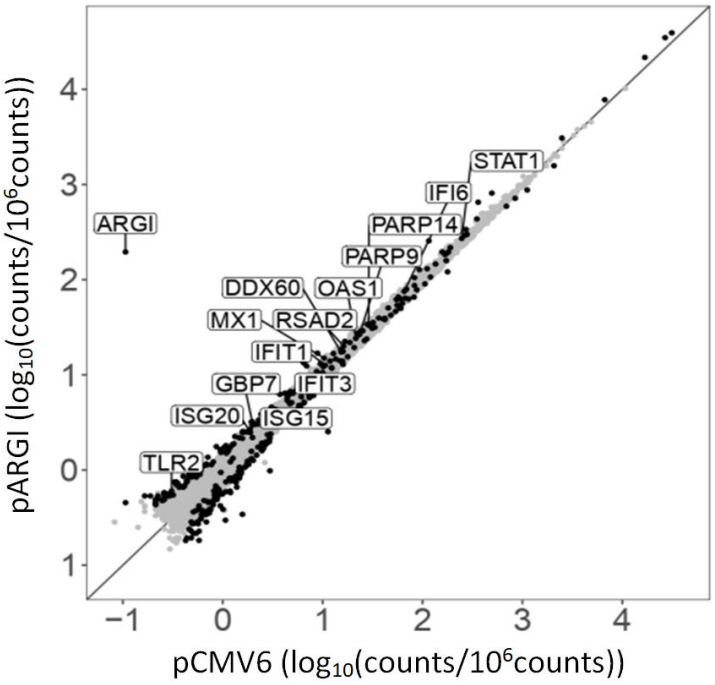


Figure 39. *ARGI* upregulation in pancreatic β cells leads to hyperactivation of genes from the IDIN network. Scatterplot analysis showing differentially expressed genes (black dots) in *ARGI*-overexpressing pancreatic EndoC-βH1 cells compared to pCMV6-transfected control cells. *ARGI* and genes from the IDIN network are highlighted with their corresponding name.

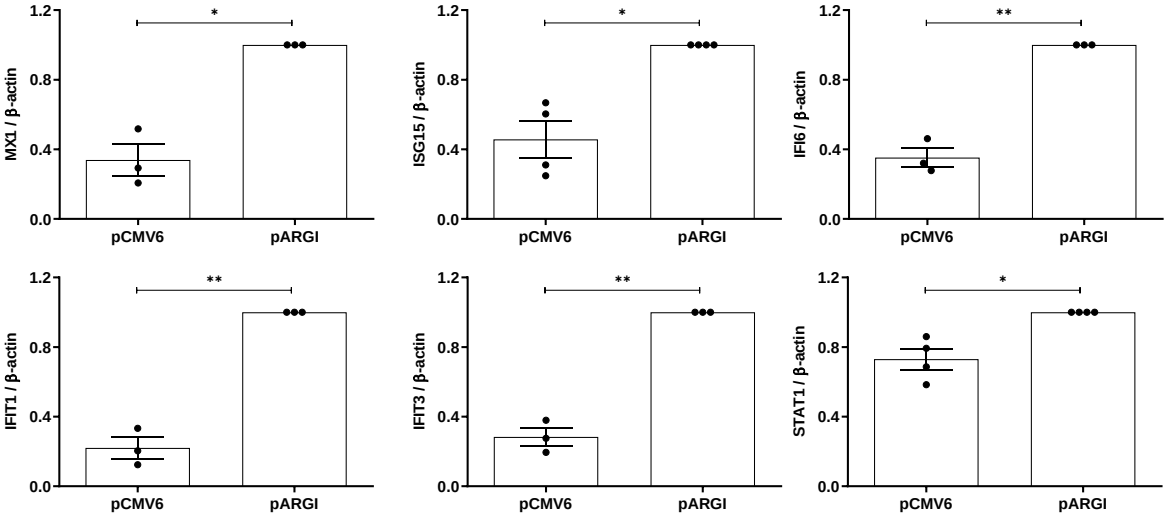


Figure 40. The expression of IDIN genes is increased after overexpressing *ARGI*. The expression of the 6 top IDIN genes differentially upregulated in *ARGI*-overexpressing cells in the RNAseq data was confirmed in an independent sample set. EndoC-βH1 cells were transfected with a control empty overexpression plasmid (pCMV6) or with a plasmid overexpressing *ARGI* and the expression of *MX1*, *ISG15*, *IFI6*, *IFIT1*, *IFIT3* and *STAT1* genes was determined by qPCR and normalized by the housekeeping gene *β-actin*. Results are means ± SEM of three independent experiments; **p < 0.01 and *p < 0.05; Student's t test.

3. *ARGI* participates in the regulation of virus-induced IFNβ and IFN-stimulated gene expression in pancreatic β cells.

Taking into account that most of the IDIN genes upregulated by *ARGI* overexpression are considered as interferon-stimulated genes (ISGs) (Schneider et al., 2014), next I decided to analyze whether *IFNβ* gene expression was also modulated by *ARGI* overexpression in pancreatic β cells. For this purpose, I analyzed *IFNβ* gene expression by qPCR in *ARGI*-overexpressing EndoC-βH1 and observed that *ARGI* upregulation led to a 60% increase in *IFNβ* expression (**Figure 41**). These results suggested that *ARGI* might be regulating the expression of ISGs partially through enhancement of *IFNβ* expression in pancreatic β cells.

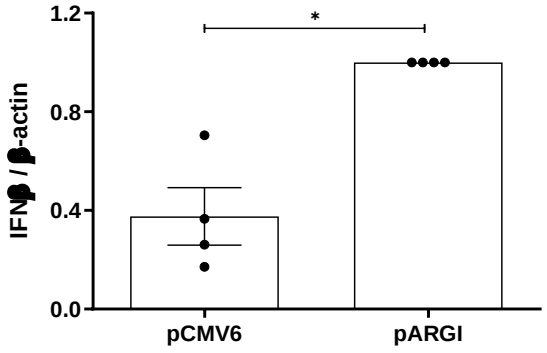


Figure 41. The expression of *IFNβ* is enhanced in *ARGI*-overexpressing pancreatic β cells. EndoC-βH1 cells were transfected with pCMV6 or pARGI and *IFNβ* expression was determined by qPCR and normalized by the housekeeping *β-actin*. Results are means ± SEM of four independent experiments. *p < 0.05; Student's t-test.

To have a mirror image of the overexpression experiments, I decided to perform a complete *ARGI* deletion by CRISPR-Cas9 in pancreatic β cells. To this aim, two different single guide RNA (sgRNA) pairs were designed, and as shown in **Figure 42A**, both guide pairs were designed to target and delete the whole *ARGI* gene. The main limitation of this strategy was to the impossibility to select and clone a complete knockout cell line due to the low duplication rate and sensitivity of the EndoC-βH1 cells. Indeed, as shown in **Figure 42B**, under basal condition, the expression of *ARGI* was similar in CRISPR-Cas9-transfected cells and in control cells. However, in the cells exposed to intracellular PIC, *ARGI* expression was 30-50% lower in CRISPR-Cas9-transfected cells than in control

cells (**Figure 42B**). Under these conditions, I analyzed the expression of *IFNβ* and *ISG15*, and observed that the disruption of *ARGI* expression led to a 60-70% decrease in PIC-induced *IFNβ* (a key regulator of ISG expression) and *ISG15* (a well characterized ISG) expression in pancreatic β cells (**Figure 42C-D**).

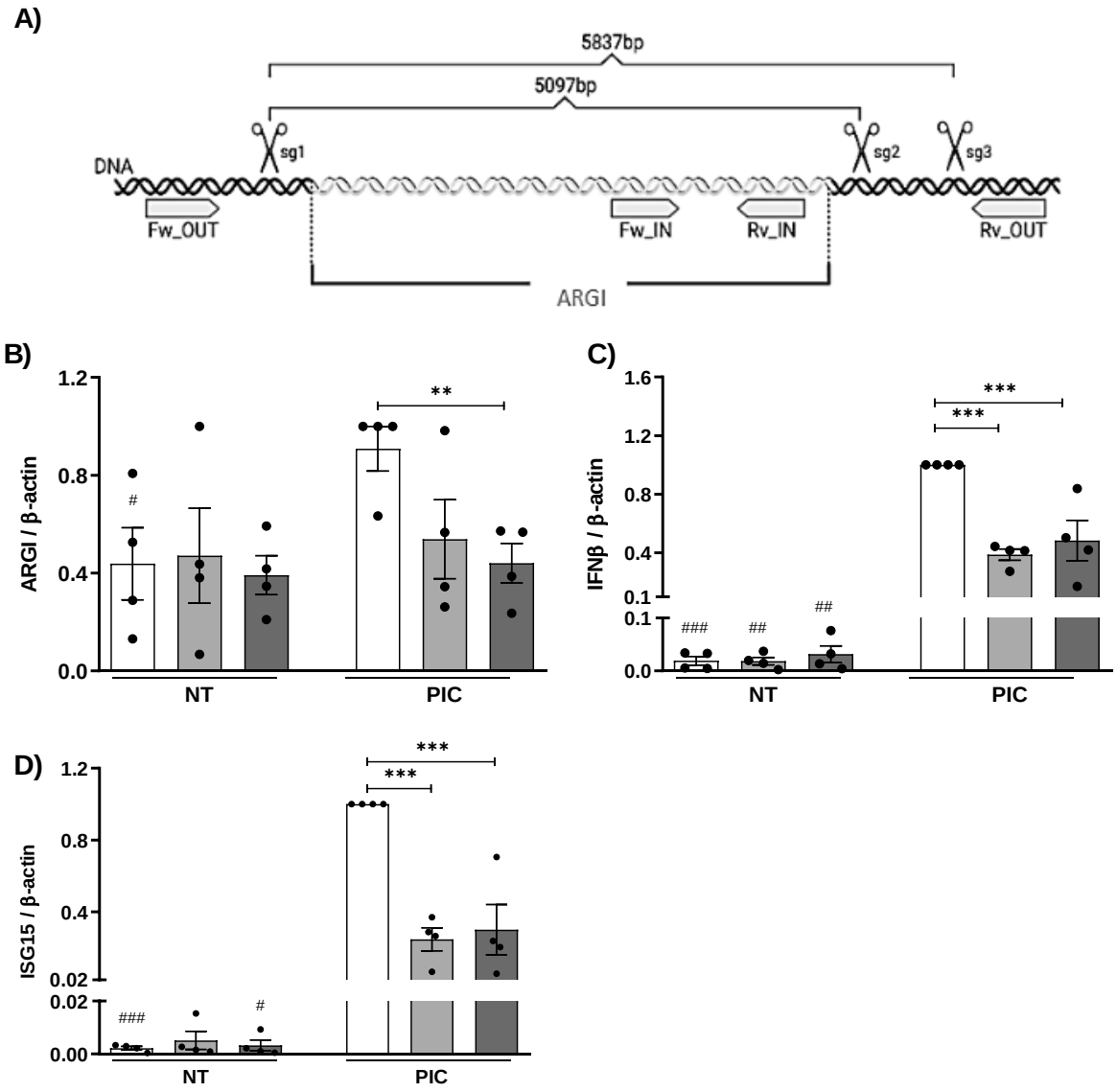


Figure 42. *ARGI* disruption using CRISPR-Cas9 reduces PIC-induced *IFNβ* and *ISG15* expression. (A) *ARGI* disruption was performed by generating a deletion of 5097 bp using sgRNAs 1 and 2, or a deletion of 5837 bp using sgRNAs 1 and 3. The presence of the deletion was confirmed by PCR using a pair of primers located inside the deleted region (for detection of unedited cells; wild type forward (Fw_IN) and wild type reverse (Rv_IN)) and a pair of primers located outside the deleted region (for detection of edited cells; Fw_OUT and Rv_OUT). (B-D) EndoC-βH1 cells were transfected with an empty px459 vector (white bars) or with vectors harboring any of the two combinations of sgRNAs targeting *ARGI* (light and dark grey bars). After 36 h of transfection, cells were left non-transfected (NT) or were transfected with PIC (1µg/ml) for 24h. Expression of *ARGI* (B), *IFNβ* (C) and *ISG15* (D) was determined by

qPCR and normalized by the housekeeping gene β-actin. Results are means ± SEM of four independent experiments; ### p < 0.001, ## p < 0.01 and # p < 0.05 when compared the NT vs the PIC transfected sample with the same vector; ***p < 0.001, **p < 0.01 and *p < 0.05 as indicated; ANOVA followed by Student's t test.

To clarify whether *ARGI* participated in the transcription of *INFβ* and ISGs, an experiment of CRISPRi was performed to inhibit endogenous expression of *ARGI* in β cells. The basis of this methodology resides in the transcriptional inhibition of a gene of interest by directing an inactive Cas9 fused with a transcriptional repressor (e.g. KRAB) to the promoter region of the gene using a complementary sgRNA. In this case, I designed a sgRNA targeting a conserved NFκB binding site in a potential regulatory region located close to the transcription starting site (TSS) of *ARGI.* The decision of targeting this NFκB binding site was taken because in previous experiments it was confirmed that *ARGI* was partially regulated by NFκB. As shown in **Figure 43**, inhibition of NFκB by a specific chemical inhibitor (BAY 11-7082) led to a partial decreased in PIC-induced *ARGI* upregulation. These results suggested that *ARGI* transcription was partially regulated by NFκB activation.

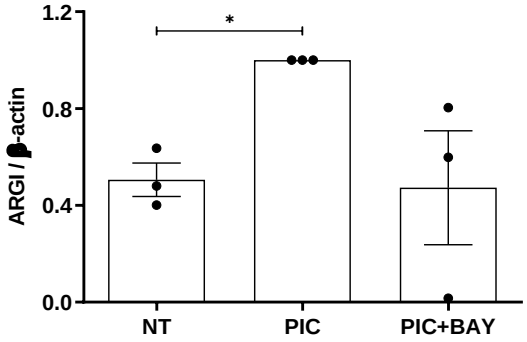


Figure 43. Inhibition of NFκB signaling pathway counteracts PIC-induced *ARGI* upregulation in pancreatic β cells. Human EndoC-βH1 cells were left untreated (NT), treated with intracellular PIC (1 μg/ml) for 24 h (PIC) or treated with PIC and Bay 11-7082 (PIC+BAY). *ARGI* expression was determined by qPCR and normalized by the housekeeping gene *β-actin*. Results are means ± SEM of 3 independent experiments; *p < 0.05; Student's t test.

Regarding CRISPRi experiment, **Figure 44A** shows how the CRISPRi-sgRNA directed to the mentioned NFκB binding site was capable to downregulate *ARGI* expression by approximately a 60%, both in basal and PIC-transfected conditions. These samples

were also employed to analyze the expression of two genes previously seem to be regulated by *ARGI* upregulation, *IFNβ* and *ISG15.* CRISPRi-mediated *ARGI* inhibition resulted in about a 40% decrease in PIC-induced *IFNβ* expression (**Figure 44B**). Results were similar for *ISG15*, in which 60% reduction was observed after CRISPRi technique was applied in PIC transfected EndoC-βH1 (**Figure 44C**). These results confirmed the role of *ARGI* in the regulation of the expression of *IFNβ* and a well-known ISG, named *ISG15*.

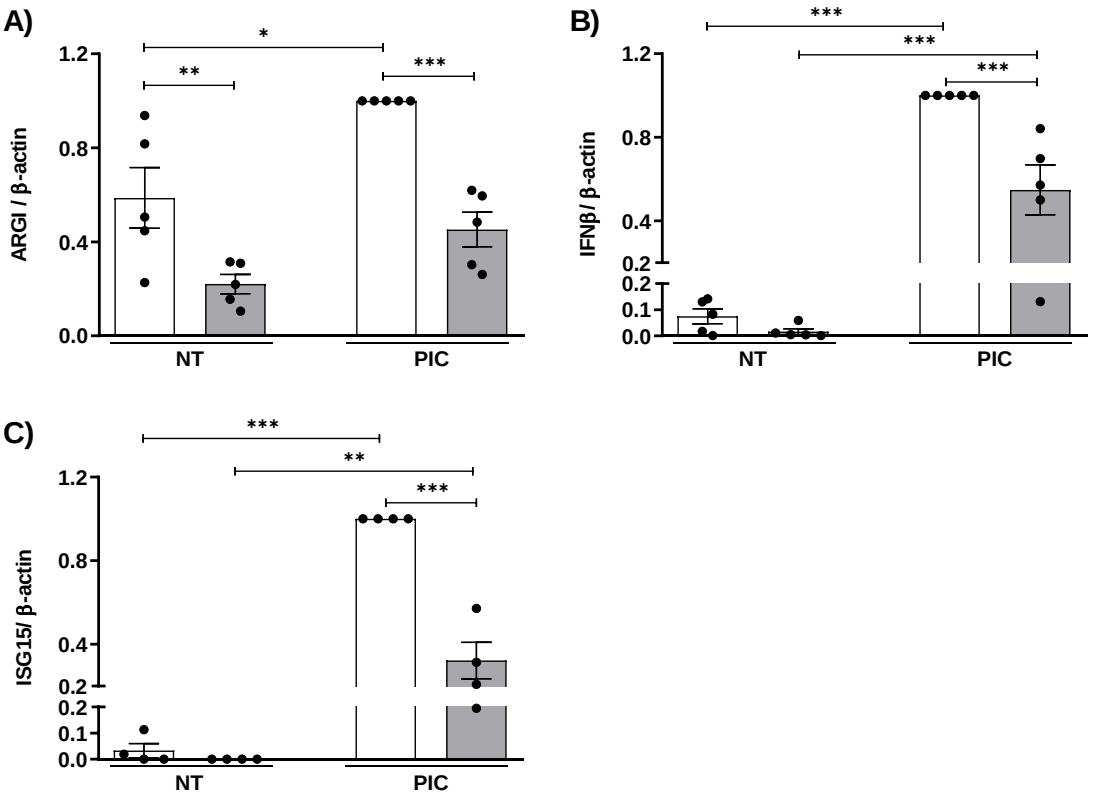


Figure 44. Transcriptional activation of *ARGI* was inhibited using CRISPRi technique. EndoC-βH1 cells were transfected with an empty CRISPRi vector (white bars) or with a CRISPRi vector harboring a sgRNA targeting a conserved NFkB binding site in *ARGI*'s regulatory region (grey bars). After 36 h of transfection, cells were left non-transfected (NT) or were transfected with PIC (1µg/ml) for 24h. Expression of *ARGI* (A), *IFNβ* (B) and *ISG15* (C) were determined by qPCR and normalized by the housekeeping gene *β-actin*. Results are means ± SEM of 4-5 independent experiments; ***p < 0.001, **p < 0.01 and *p < 0.05 as indicated; Student's t test or ANOVA followed by Student's t test.

4. *ARGI* associates with the transcription factor CTCF to bind to the regulatory regions of *IFNβ* and *ISG15* genes in PIC-transfected cells.

The next step on the characterization of *ARGI* was to determine the molecular mechanisms by which this lncRNA was able to regulate the expression of *IFNβ* and *ISG15.* To this aim, a RNA antisense purification (RAP) of *ARGI* was performed in order to determine whether *ARGI* binds to the regulatory regions of *IFNβ* and *ISG15*. This approach allows the purification of proteins, DNA and RNA bound to a target RNA molecule. Using this approach, *ARGI* was purified using biotinylated antisense complementary oligonucleotides, while antisense oligonucleotides complementary to a similar length lncRNA were used as a negative control (**Figure 45**). As shown **Figure 45A**, the purification of *ARGI* was more efficient when the cells where transfected with PIC compared to basal condition, most probably due to the increase in *ARGI* expression induced by PIC. The purification efficiency of the negative control was also analyzed to check the specificity of the d antisense oligonucleotides complementary to *ARGI.* As shown in **Figure 45B**, when the antisense oligonucleotides targeting an irrelevant control RNA were used, *ARGI* was not purified; however, when oligonucleotides complementary to *ARGI* were employed, *ARGI* was efficiently purified, especially in the presence of PIC.

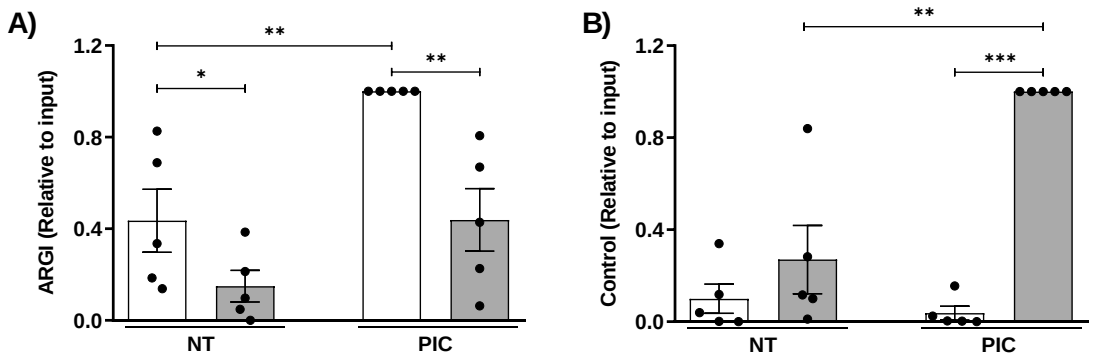
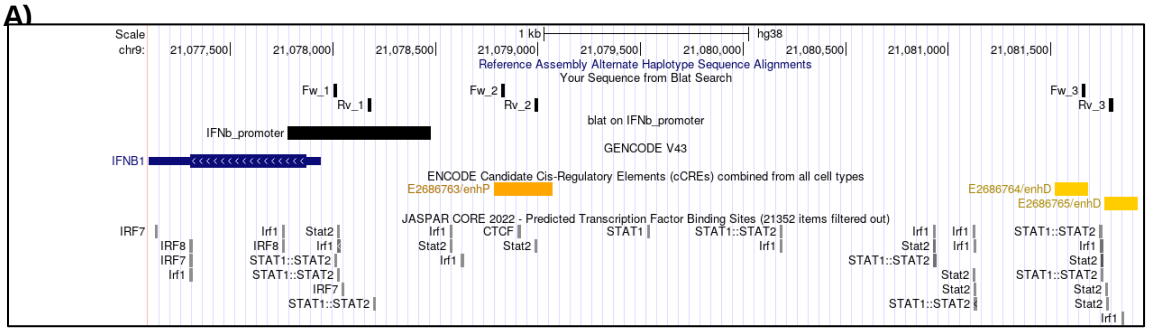


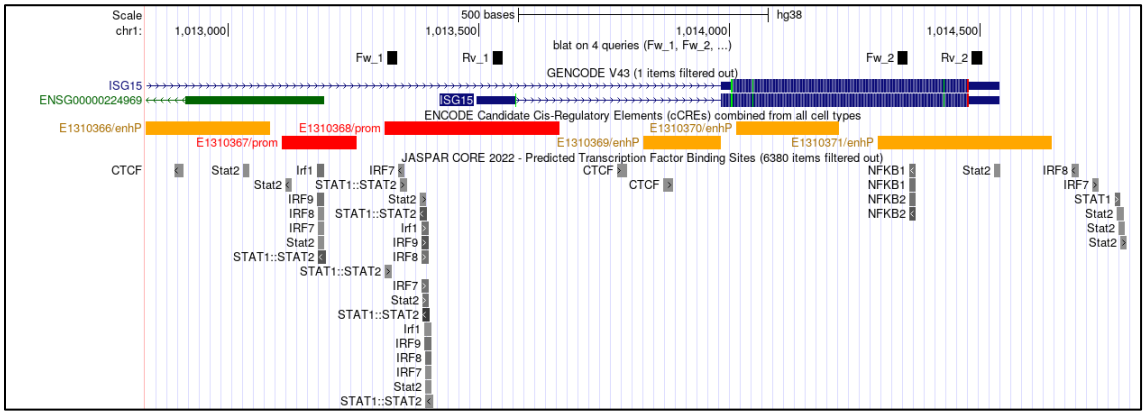
Figure 45. RNA antisense purification of *ARGI* in basal and PIC-transfected EndoC-βH1. RNA antisense purification of *ARGI* was performed in non-transfected (NT) or PIC-transfected EndoC-βH1 cells (PIC). (A) The expression of *ARGI* (A) or the negative control lncRNA (B) was measured in RNA purified using biotinylated antisense oligonucleotides complementary to *ARGI* (white bars) or antisense oligonucleotides complementary to a similar lncRNA used as the negative control (grey bars). Results are expressed as relative to input and are means ± SEM of five independent experiments. ***p<0.001, **p<0.01 and *p < 0.05 as indicated; ANOVA followed by Student's t test.

After *ARGI* purification, the presence of the regulatory regions of *IFNβ* and *ISG15* bound to the lncRNA was determined by qPCR. To this aim, a pair of primers targeting the promoter and enhancer regions of both *IFNβ* and *ISG15* were designed. Localization of these primers are detailed in **Figure 46A-B**. As shown in **Figure 46C-E**, in basal condition (e.g. non-treated cells), *ARGI* was not bound to *ISG15* promoter or enhancer, however it was slightly bound to the *IFNβ* promoter. In contrast, in PIC-transfected EndoC-βH1 cells, *ARGI* was more bound to the promoter region (IFNβ_RAP1) of *IFNβ* gene, but also appeared in the promoter (ISG15_RAP1) and enhancer (ISG15_RAP2) regions of *ISG15* **Figure 46C-E**.

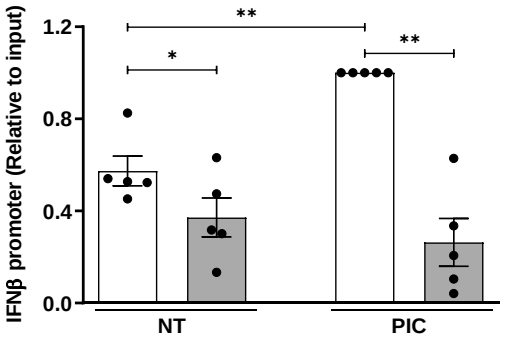
A)


Scale
chr9:
21,077,500
21,078,000
21,078,500
21,079,000
21,079,500
21,080,000
21,080,500
21,081,000
21,081,500
1 kb
hg38
Reference Assembly Alternate Haplotype Sequence Alignments
Your Sequence from Blat Search
Fw_1
Rv_1
Fw_2
Rv_2
Fw_3
Rv_3
blat on IFNb_promoter
IFNb_promoter
GENCODE V43
IFNB1
ENCODE Candidate Cis-Regulatory Elements (cCREs) combined from all cell types
E2686763/enhP
E2686764/enhD
E2686765/enhD
JASPAR CORE 2022 - Predicted Transcription Factor Binding Sites (21352 items filtered out)
IRF7
IRF8
IRF7
Irf1
Irf1
IRF8
Irf1
STAT1::STAT2
STAT1::STAT2
IRF7
STAT1::STAT2
Stat2
Irf1
Stat2
Irf1
CTCF
Stat2
STAT1
STAT1::STAT2
Irf1
Irf1
Irf1
Stat2
STAT1::STAT2
Stat2
Irf1
STAT1::STAT2
Irf1
Stat2
Stat2
STAT1::STAT2
Stat2
Stat2
Irf1

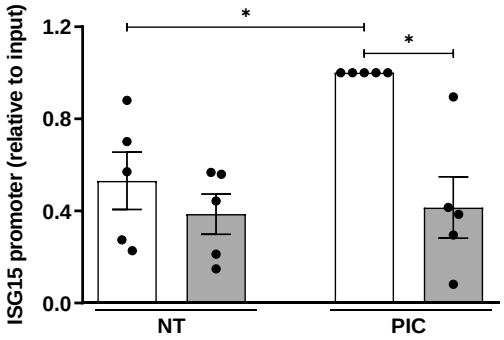
B)


Scale
chr1:
1,013,000
1,013,500
1,014,000
1,014,500
500 bases
hg38
blat on 4 queries (Fw_1, Fw_2, ...)
Fw_1
Rv_1
Fw_2
Rv_2
GENCODE V43 (1 items filtered out)
ISG15
ENSG00000224969
ISG15
ENCODE Candidate Cis-Regulatory Elements (cCREs) combined from all cell types
E1310366/enhP
E1310367/prom
E1310368/prom
E1310369/enhP
E1310370/enhP
E1310371/enhP
JASPAR CORE 2022 - Predicted Transcription Factor Binding Sites (6380 items filtered out)
CTCF
Stat2
Irf1
IRF7
Stat2
STAT1::STAT2
IRF9
IRF8
STAT1::STAT2
IRF7
Stat2
STAT1::STAT2
STAT1::STAT2
IRF7
Stat2
STAT1::STAT2
Irf1
IRF9
IRF8
IRF7
Stat2
STAT1::STAT2
CTCF
CTCF
CTCF
NFKB1
NFKB1
NFKB2
NFKB2
Stat2
IRF8
IRF7
STAT1
Stat2
Stat2
Stat2

C)


IFNβ promoter (Relative to input)
1.2
0.8
0.4
0.0
**
*
**
NT
PIC

D)


ISG15 promoter (relative to input)
1.2
0.8
0.4
0.0
*
*
NT
PIC

E)

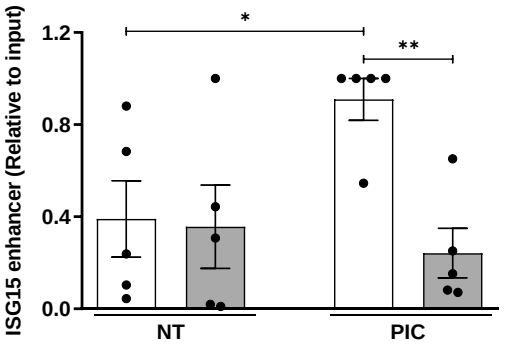

ISG15 enhancer (Relative to input)
1.2
0.8
0.4
0.0
*
**
NT
PIC

Figure 46. *ARGI* binds to the regulatory regions of *IFNβ* and *ISG15* genes upon exposure to a viral insult. RNA antisense purification of *ARGI* was performed in non-transfected (NT) or PIC-transfected EndoC-βH1 cells (PIC). (A-B) *ARGI* presence in *IFNβ* and *ISG15* regulatory regions was measured by qPCR using specific primers: FW1+RV1 were used to amplify *IFNβ* promoter, and FW2+RV2 and FW3+RV3 primer pairs were used to amplify two distal enhancers (A). FW+RV were used to amplify ISG15 promoter and FW2+RV2 were used to amplify a distal enhancer of *ISG15* (B). (C-E) *ARGI*-bound *IFNβ* promoter (C), *ISG15* promoter (D) and *ISG15* enhancer (E) amounts were determined by qPCR. Results are expressed as relative to input and are means ± SEM of five independent experiments. **p<0.01 and *p < 0.05 as indicated; ANOVA followed by Student's t test.

The interaction between *ARGI* and the regulatory regions of *IFNβ* and *ISG15* indicated a potential direct regulatory function of the lncRNA in antiviral gene transcription, however the exact molecular mechanisms underlying this regulation remained known. Therefore, in order to clarify the mechanisms by which *ARGI* regulated the transcription of ISGs, I decided to analyze the transcription factor binding sites present in the regulatory domains of *IFNβ* and *ISG15*. The search revealed that the regulatory regions (promoters and enhancers) of *IFNβ* and *ISG15* contained binding sites for typical pro-inflammatory transcription factors (IRF7, STAT1 and STAT2, among others). Moreover, I detected binding sites for CCCTC-binding factor (CTCF), a conserved zinc finger protein able to act as a transcriptional activator, repressor or insulator protein, blocking the communication between enhancers and promoters (S. M. Kim et al., 2015; Phillips & Corces, 2009). To check the potential of *ARGI* to interact with the mentioned transcription factors, an *in silico* prediction of RNA-protein interactions was performed using the online tool called CatRAPID (Armaos et al., 2021). Results revealed that *ARGI* potentially interacted with CTCF (interaction score: 0.41), but did not show any potential interaction with STAT1, STAT2 or IRF7 proteins.

To confirm the *in silico* prediction showing a potential interaction between CTCF and *ARGI*, a RNA immunoprecipitation (RIP) experiment was performed using an antibody against CTCF. As shown in **Figure 47**, CTCF was efficiently immunoprecipitated in both basal and PIC-transfected EndoC-βH1. Regarding *ARGI* expression in CTCF-bound RNA, as shown in **Figure 47**, in basal condition *ARGI* did not interact with CTCF, but upon exposure to intracellular PIC, *ARGI* was able to bind to CTCF.

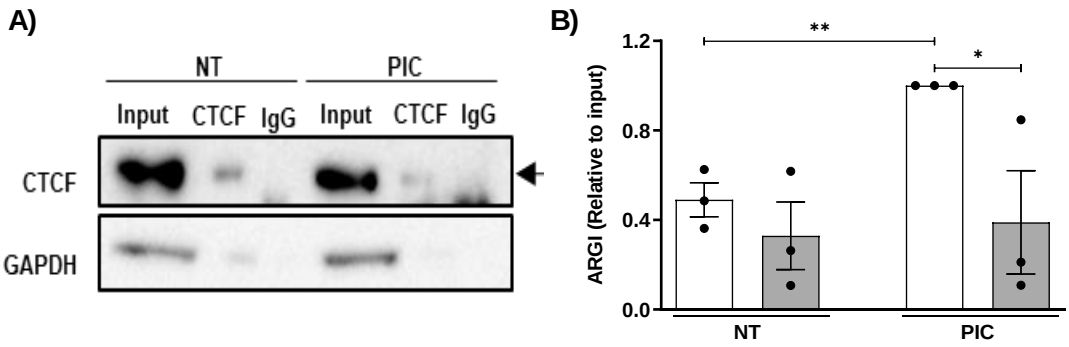


Figure 47. *ARGI* associates with the transcription factor CTCF in PIC-transfected cells. (A) EndoC-βH1 cells were left non-transfected (NT) or transfected with PIC for 24h and RNA immunoprecipitation was performed using a specific antibody for CTCF or an IgG antibody used as a negative control. The image is representative of three independent experiments. (B) *ARGI* expression was determined in CTCF-bound RNA (white bars) and IgG-bound RNA (grey bars) by qPCR. Results are means ± SEM of three independent experiments and the amounts of *ARGI* are expressed as relative to the input. **p<0.01, and *p<0.05 as indicated; ANOVA followed by Student's t test.

Based on the results, the main hypothesis for the mechanism of action of *ARGI* was that after a viral infection *ARGI* interacts with the chromatin-remodeling protein CTCF to regulate the expression of *IFNβ* and *ISG15* genes.

To confirm this possibility, a chromatin-RNA immunoprecipitation was performed to extract RNA and chromatin fragments simultaneously bound to CTCF and check for the presence of *ARGI* and the regulatory regions of *IFNβ* and *ISG15*. To this aim, HEK cells were kept untransfected or transfected with PIC for 24h, and after cell crosslinking, sonication and immunoprecipitation with CTCF, CTCF-bound chromatin and RNA were purified. As shown in **Figure 48A**, *ARGI* was bound to CTCF in PIC-transfected cells, as previously observed in RIP experiments using EndoC-βH1 cells. In addition, in PIC-transfected HEK293FT cells, CTCF was also interacting with the chromatin domains corresponding to *IFNβ* promoter and *ISG15* promoter and enhancer regions **(Figure 48B-D**). These results demonstrated the simultaneous interaction between *ARGI*, CTCF and the regulatory domains of *IFNβ* and *ISG15* genes, especially in the presence of a viral assault (e.g. intracellular PIC).

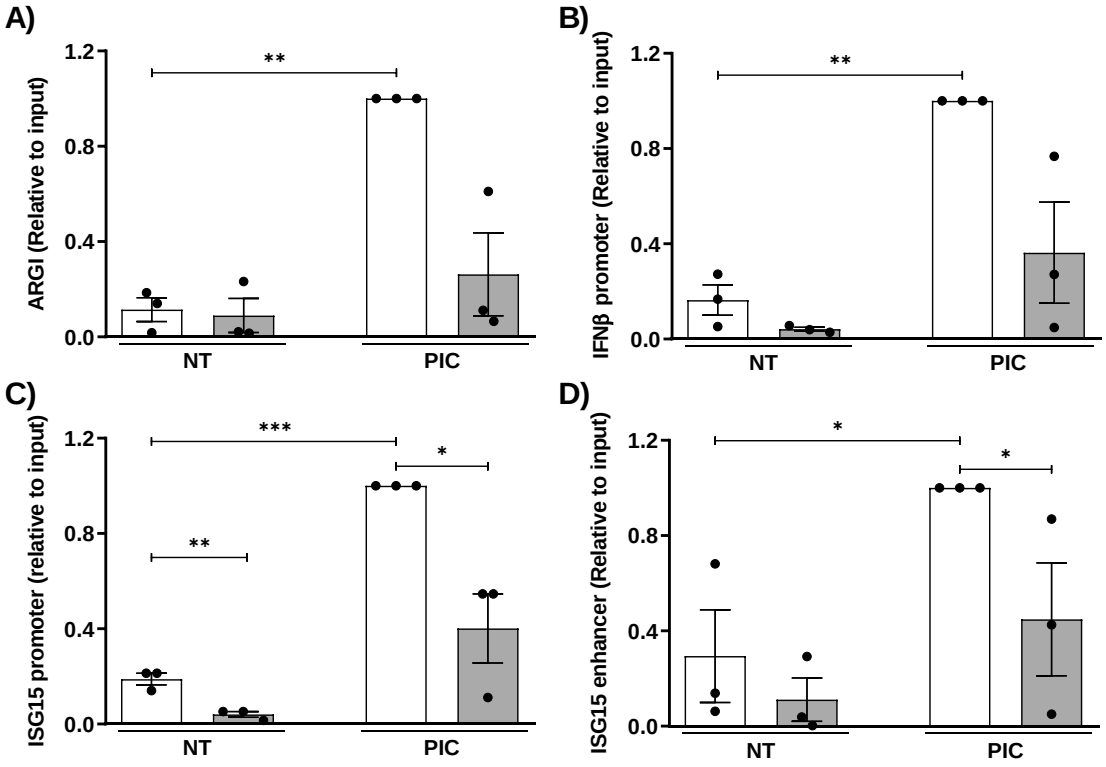


Figure 48. *ARGI* interacts with CTCF to bind with the regulatory regions of *IFNβ* and *ISG15* genes in viral infected cells. Chromatin immunoprecipitation was performed in non-transfected (NT) and PIC-transfected HEK293 cells. CTCF-bound chromatin and RNA was immunoprecipitated using a specific antibody for CTCF (white) and IgG (grey) was used as the negative control. CTCF-bound ARGI expression (A), *IFNβ* promoter (B), *ISG15* promoter (C) and *ISG15* enhancer amounts (D) were determined by qPCR. Results are means ± SEM of three independent experiments and are represented as relative to the input. ***p<0.001, **p<0.01, and *p<0.05 as indicated; ANOVA followed by Student's t test.

5. Allele-specific binding of ARGI to CTCF affects IDIN genes expression level.

As previously explained, the secondary structure of a lncRNA is crucial for its correct function since it may affect the stability and capacity of the molecule to bind to another macromolecules such as DNA, protein and other RNAs (Graf & Kretz, 2020; Robinson et al., 2020; Sanbonmatsu, 2022; Zampetaki et al., 2018). *ARGI* harbors a T1D-associated SNP that was predicted to affect its secondary structure, and potentially also to alter its function. In order to clarify whether the T1D-associated SNP affected *ARGI* function, firstly its potential effect on the interaction between *ARGI* and CTCF was determined.

To this purpose, I performed RIP experiments in cells co-overexpressing *CTCF* with *ARGI* harboring the protective T1D-allele (ARGI-P; rs9585056-T) or the risk T1D-allele (ARGI-R; rs9585056-C). As shown in **Figure 49**, CTCF was efficiently immunoprecipitated in both conditions. In addition, the results showed that both *ARGI* harboring the risk or the protective allele was able to interact with CTCF protein. Nonetheless, the interaction between CTCF protein and *ARGI* harboring the T1D risk allele (ARGI-R) was 6-times stronger than the interaction between CTCF and the lncRNA harboring the protective T1D allele (ARGI-P) (**Figure 49B**).

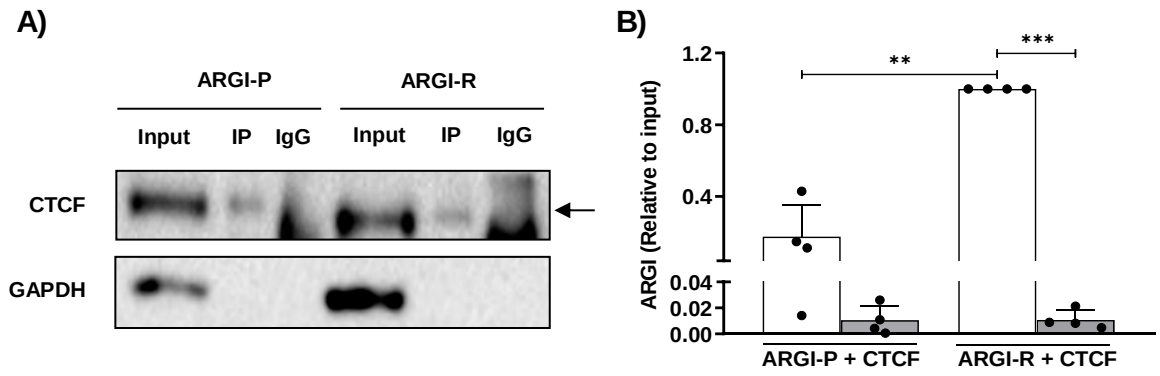


Figure 49. The interaction between *ARGI* and CTCF is allele-specific. (A) Cells were transfected with a plasmid overexpressing ARGI harboring the protective (ARGI-P) or the risk T1D allele (ARGI-R) and co-transfected with a vector overexpressing CTCF. RNA immunoprecipitation was performed using a specific antibody for CTCF and IgG was used as the negative control. (B) CTCF-bound ARGI amounts were determined by qPCR. White bars represent RNA immunoprecipitation performed using a specific antibody for CTCF and grey bars the negative control IgG Results are means ± SEM of four independent experiments and the amounts of *ARGI* are represented as relative to the input. ***p<0.001 and **p<0.01 as indicated; ANOVA followed by Student's t test with Bonferroni correction

Afterwards, I decided to analyze whether the allele-specific structural change of *ARGI* was also altering the effect of *ARGI* in the modulation of IDIN gene upregulation. Allele-specific *ARGI* overexpression plasmids were employed to achieve an allele-specific overexpression of *ARGI* in EndoC-βH1 cells and the expression of several IDIN genes was determined by qPCR. The expression level obtained by both overexpressing vectors led to the same increase in *ARGI* expression, suggesting that the T1D-associated SNP was not affecting *ARGI* stability (**Figure 50**).

As shown in **Figure 50**, the expression of every gene evaluated was significantly increased when *ARGI* harboring the T1D-risk allele was overexpressed. However, when *ARGI* harboring the protective allele for T1D was overexpressed, there was just a slight increment on IDIN gene expression, only statistically significant in the case of *IFNβ* and *MX1* gene. In summary, *ARGI* was able to activate the expression of *IFNβ* and ISGs belonging to the IDIN network in pancreatic β cells, and when it harbored the risk allele for T1D, the effect was exacerbated.

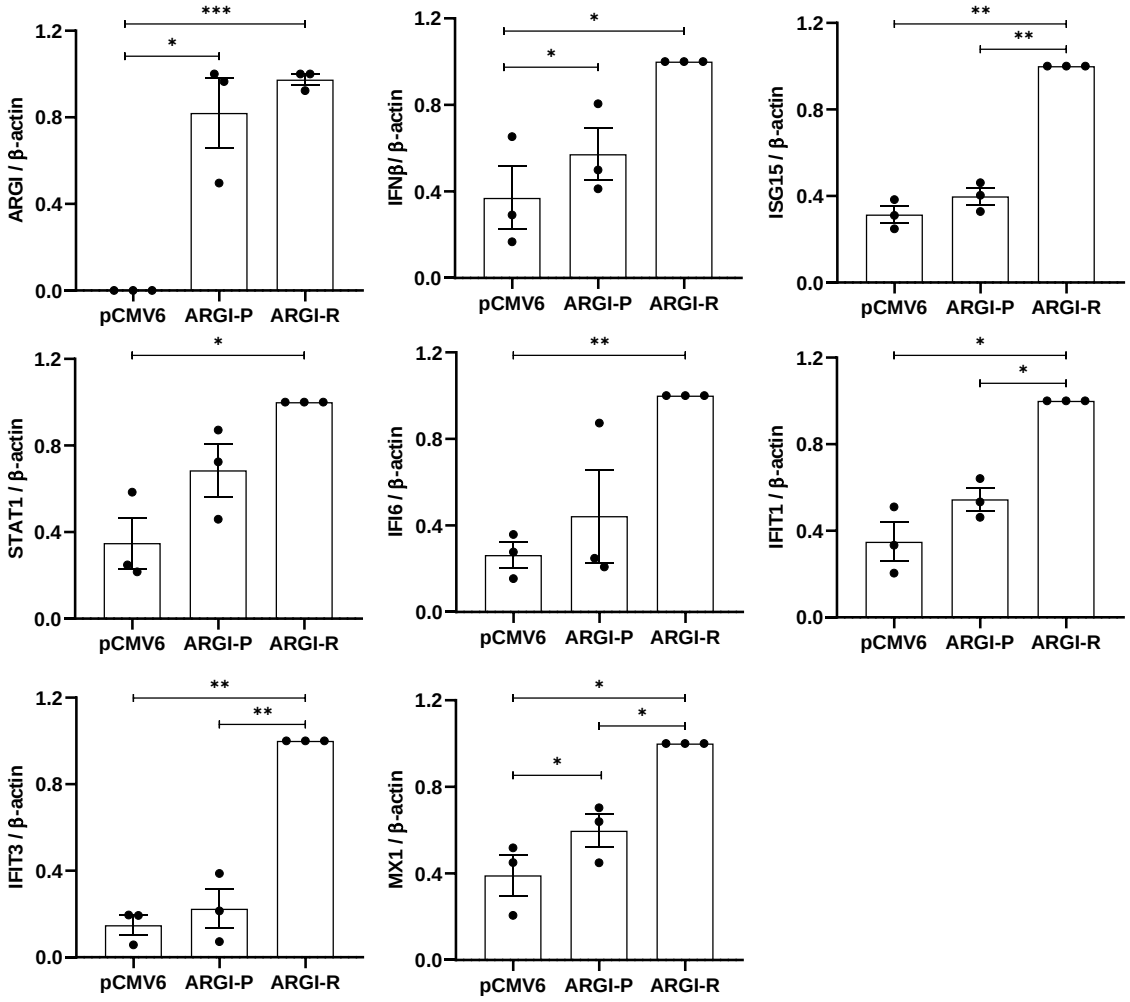


Figure 50. *ARGI* modulates IDIN gene expression in an allele-specific manner. EndoC-βH1 cells were transfected with a control plasmid (pCMV6), a plasmid overexpressing *ARGI* harboring the protective (ARGI-P) or a plasmid overexpressing ARGI harboring the risk T1D allele (ARGI-R). Expression of *ARGI*, *IFNβ*, *ISG15, STAT1, IFI6, IFIT1, IFIT3 and MX1* were determined by qPCR and normalized by the housekeeping gene *β-actin*. Results are means ± SEM of three independent experiments and are represented as relative to the input. ***p<0.001, **p<0.01, and *p<0.05 as indicated; ANOVA followed by Student's t test.

Chapter 6

Discussion

Eztabaida

During the last decades and due to the severity of type 1 diabetes, novel techniques are trying to identify new genomic domains related to the susceptibility of this autoimmune disease. On a first sight, protein-coding genes were suspected to be the main contributors to this pathology. However, this approach is not sufficient for unraveling the genetics behind T1D as the mutations located in protein-coding genes do not explain the etiopathogeny of a considerable proportion of patients. Lately, new efforts have deciphered a collection of new contributors for T1D development, the non-coding genes. In this line, lncRNAs are receiving the greatest attention because of their relevance on multitude biological processes. LncRNAs are very versatile molecules; they can be retained in the nucleus or exported onto the cytoplasm, and consequently adopt different cellular functions, including transcriptional and translational regulation or the modulation of protein and RNA stability, for example (Carlevaro-Fita et al., 2019; Marchese et al., 2017). Quite recently, whole transcriptome studies performed on pancreatic islets revealed the presence of thousands of lncRNAs specifically expressed in pancreatic β cells (Moran et al., 2012).

Regardless genetic susceptibility, T1D is a complex condition in which environmental factor are also key in the development of the disease. Albeit the wide variety of different environmental factors that might trigger this disease, viral infections are in the spotlight (Rewers & Ludvigsson, 2016; Simonsen et al., 2015). The implication of viral infections and the resulting activation of the type I IFN signaling in pancreatic β cells is decisive in the initial stages of T1D (Krogvold, Genoni, et al., 2022). Epidemiological studies have highlighted the implication of enteroviruses, and in particular the Coxsackievirus B subtype, in T1D pathogenesis (Rewers & Ludvigsson, 2016; Simonsen et al., 2015). Although CVB infections are known to contribute to pathogenic processes linked to T1D (Colli et al., 2011, 2019; Laitinen et al., 2014; S. Oikarinen et al., 2014; Op de beeck & Eizirik, 2016; Sioofy-Khojine et al., 2018; Vehik et al., 2019), the mechanisms by which they induce autoimmunity against pancreatic β cells is not yet elucidated. Accumulating evidence propose the presence of T1D-associates genetic variants in genes related to antiviral response as the factors linking viral infections and the activation of autoimmunity against β cells (Marroqui et al.,

2015; Santin et al., 2011). Therefore, the functional characterization of susceptible genes for T1D in virus-infected pancreatic β cells can help in the elucidation of many enigmas that are still undefined in the pathogenesis of the disease.

Considering all these data, this thesis was focused on the identification and posterior characterization of lncRNAs associated with T1D at the pancreatic β cell level. In this sense, our data provide novel information on the molecular mechanisms by which disease-associated SNPs located in lncRNAs might participate in the trigger of inflammation and modulate T1D development.

Thus, the identification of T1D-associated lncRNAs was the first and most decisive step of this work. Although nowadays there is not a standardized protocol for the identification of T1D-linked lncRNAs with a T1D-linked functional role, this work established a workflow to help in this task. Using the data from NONCODE, a database of existing non-coding RNAs, and the GWAS catalog, which contains all the human GWAS data, we were capable of identifying which lncRNAs harbored a T1D-associated SNP. However, the list of candidates was not manageable for a thorough functional characterization, so a selection process was performed to solve this issue and obtain a feasible list of interesting candidates.

One of the main aspects taken into consideration for the prioritization of interesting lncRNAs, was the presence of a T1D-associated SNP in an exonic region predicted to disrupt the secondary structure of the lncRNA. As previously explained, the structure of a lncRNA strongly determines its function (Kino et al., 2010; Mirza et al., 2014; Zampetaki et al., 2018). In this sense, a previous study by Mirza et al. identified a list of 178 T1D-associated SNPs with significant propensity to modify the secondary structure of target lncRNAs. The secondary structure of these molecules is essential for their function, since it is important for their binding capacity with other macromolecules like RNA, DNA or proteins (Marchese et al., 2017; Zampetaki et al., 2018). Among the wide variety of interactions reported, the vast majority of lncRNAs have been described to combine with chromatin-modifying complexes (Engreitz et al., 2013; Guttman & Rinn, 2012; Tsai et al., 2010) and transcription factors (Marchese et al., 2017; Xue et al.,

2016; L. Zhou et al., 2015) mainly resulting in the modulation of gene expression. Interestingly, disease-associated SNPs in these lncRNAs seem to affect their ability to bind a specific protein or even miRNAs, affecting in gene expression regulation (Aznaourova et al., 2020; Fu et al., 2020). This is the case of *LOC146880*, a lncRNA harboring a SNP associated with non-small cell lung cancer (rs140618127) (Feng et al., 2020). The disease-linked variant in *LOC146880* alters its secondary structure creating a binding site for the microRNA miR-539-5p (Feng et al., 2020). A similar effect was described related to the lncRNA named *UCA1* (Fu et al., 2020). The rs12982687 affects the binding capacity of *UCA1* with the microRNA named miR-873-5p, provoking a diminished HIF-1 signaling (Fu et al., 2020).

Many of the T1D candidate protein-coding genes characterized so far have been implicated in the regulation of antiviral and pro-inflammatory responses at the pancreatic β cell level (Santin & Eizirik, 2013). Several genes, such as *PTPN2* and *TYK2*, participate in the regulation of the type I IFN signaling by regulating the type I IFN-induced JAK/STAT signaling pathway (Chandra et al., 2022; Marroqui et al., 2015; Santin et al., 2011), while others (e.g. *MDA5*) encode viral dsRNA cytoplasmic receptors (Colli et al., 2010). However, there is little information about the implication of non-coding genes in virus-induced inflammation and β cell dysfunction. However, some lncRNAs have been associated with antiviral and pro-inflammatory responses in other cell types. For example, a lncRNA named *LUCAT1* was described to interact with STAT1 in the nucleus of THP-1 cells acting as a negative regulator of interferon responses in human (Agarwal et al., 2020). In the case of the lncRNA called *Mirt2,* participates in the inhibition of the NFκB and MAPK pathways to diminish the expression of LPS-induced pro-inflammatory cytokines in tracheal epithelial cells and hepatocytes from C57BL/6 cells (Du et al., 2017).

Because of the potential relevance of viral infections in T1D pathogenesis, this work was focused on the characterization of T1D-associated lncRNAs in virus-induced β cell damage. Thus, this work assumed that a lncRNA which expression is modulated by a viral assault is presumed to be implicated in the modulation of innate immune

response, a crucial pathway in T1D development (Colli et al., 2011, 2019; Laitinen et al., 2014; S. Oikarinen et al., 2014; Op de beeck & Eizirik, 2016; Vehik et al., 2019). In this work the function of two T1D-associated lncRNAs, namely *Lnc13* and *ARGI*, were characterized. The results confirmed that both non-coding molecules modulate virus-induced β cell inflammation, but through a totally different mechanism of action.

Lnc13 modulates virus-induced pancreatic β cell inflammation through stabilization of *STAT1* in an allele-specific manner.

The T1D-associated *Lnc13* is a very ubiquous molecule with an impresively high abundance in the EndoC-βH1 cell line, probably due to the enbrionic nature of this cell line. An important characteristic of lncRNAs is that they expression is very cell- and tissue-specific, suggesting that they might function differently depending on the cell or tissue (L. Chen et al., 2018; Jiang et al., 2016; S. J. Liu et al., 2016). *Lnc13* was previously described to participate in celiac disease (CeD) pathogenesis (Castellanos-Rubio et al., 2016). Interestingly, T1D and CeD are complex autoimmune diseases but althought the SNP present in Lnc13 is linked to both conditions (rs917997), the T1D-associated risk allele is C, whereas the risk alllele for CeD is the opposite allele (rs917997-T). Downregulation of *Lnc13* in intestinal biopsies of celiac patients was linked with increased expression of *STAT1* and other pro-inflammatory genes (Castellanos-Rubio et al., 2016). The functional analysis of *Lnc13* revealed that it interacts with two chromatin-associated proteins (hnRNPD and HDAC1) negatively regulating the expression of CeD-linked pro-inflammatory genes. When the risk allele for celiac disease (allele T) is present on the lncRNA, the affinity to bind hnRNPD and chromatin is decreased resulting in higher pro-inflammatory gene expression, and thus, a higher predisposition to develop the condition. Regarding its implication in T1D, herein, we unraveled how this lncRNA was able to modulate virus-induced inflammation in pancreatic β cells. Differently to the process previously descrived in celiac disease; in pancreatic β cells, *Lnc13* sustains the activation of the STAT1 pro-inflammatory pathway through the stabilization of the STAT1 mRNA molecule.

Interestingly, previous studies have shown that the inflammation caused by viral infections in pancreatic β cells is partially mediated by the activation of type I IFN and the STAT1 signaling pathway (Eizirik et al., 2009; Marroqui et al., 2015; Santin et al., 2011). *STAT1* plays a central role in the regulation of the type I interferon-signaling pathway (Alkanani et al., 2014; Platanias, 2005). The binding of IFNα/β to their receptors results in the auto-phosphorylation and the activation of the Janus activated kinases (JAKs) and the TYK2 (Babon et al., 2014; Quelle et al., 1995). This activation results in the phosphorylation of STAT1 and STAT2 triggering the formation of a complex which translocates to the nucleus and binds the IFN-stimulated response elements in the DNA to initiate the transcription of genes involved in the antiviral response. Activation of the STAT1 signaling pathway is tightly linked with T1D pathogenesis and there are evidences on how STAT1 and IRF1 (one of its downstream transcription factors) are abnormally expressed in pancreatic islets from T1D patients (Colli et al., 2018; Richardson et al., 2016). Additionally, when the STAT1-STAT2 complex interacts with IRF9, this system receives the name of ISGF3. The ISGF3 complex is able to bind to IFN–stimulated response elements (ISRE)s located in ISG promoters activating the essential gene network against viral infections (Au-Yeung et al., 2013; Platanitis et al., 2019).

Considering the biological relevance of STAT1 activation in diabetes development, in this work, the possible link between *Lnc13* and this transcription factor was analyzed. This hypothesis was gaining consistency since the overexpression of the lncRNA in pancreatic β cells provoked the activation of STAT1 and the upregulation of pro-inflammatory chemokines. In the other way round, *Lnc13* disruption in β cells resulted in a partial decrease in viral infection-induced inflammation as it reduced *STAT1* and chemokine expression.

The expression and release of chemokines must be finely regulated to avoid the development of autoimmune diseases like T1D. Indeed, during the first stages of this condition, the immune system is activated and the pro-inflammatory chemokines are released, attracting the immune cells to the islets and generating a local inflammatory

environment that finally could trigger the loss of the pancreatic β cells (Atkinson & Wilson, 2002; Eizirik et al., 2009). Against this background, this work demonstrates that the T1D-associated *Lnc13* activates STAT1 signaling pathway resulting in chemokine production and release. Indeed, the T1D-associated SNP allele enhances the antiviral response, aggravating the local inflammation. These results confirm that T1D risk genetic variants in lncRNAs alter the antiviral innate immune response of β cells, contributing to insulitis and β cell dysfunction.

Regarding the characterization of the molecular mechanisms by which *Lnc13* regulates STAT1 signaling pathway and inflammation of pancreatic β cells, this study demonstrates that *Lnc13* is able to interact with a protein named PCBP2 and stabilize the mRNA of *STAT1*. In a context of a viral infection, *Lnc13* is translocated from the nucleus to the cytoplasm where it interacts with PCBP2 and the 3'UTR of *STAT1* forming a complex that stabilized the mRNA molecule of the *STAT1* gene. The RNA-interacting protein PCBP2 has been previously described to stabilize *STAT1* and *STAT2* mRNA molecules in hepatocytes helping in the antiviral activity of IFNα against Hepatitis C virus (Xin et al., 2011). Another research work described that the stabilization of mRNA molecules through PCBP2 binding relied in its capacity to interact with the 3'UTRs of mRNA molecules (L. Chen et al., 2018; Xin et al., 2011). In contrast, PCBP2 binding to the 5'UTRs of mRNA molecules has been described to inhibit their translation (Smirnova et al., 2019). Our data demonstrates that the presence of the T1D risk allele in *Lnc13* promotes a stronger interaction between PCBP2 and the 3'UTR of *STAT1* when compared to the protective genotype (rs917997-T), inducing a higher stabilization of STAT1 mRNA.

Multiple investigations describe the link between lncRNA secondary structure and their function. Indeed, several studies have described how mutations located inside lncRNAs contribute to disease development through the dysregulation of disease-associated pathways (Aguilo et al., 2016; Castellanos-Rubio et al., 2016; Redis et al., 2016). For example, a lncRNA named *CCAT2* harbors a SNP (rs6983267) associated with cancer risk and regulates the metabolism of cancer by interacting with the CFIm complex in

an allele-specific manner (Redis et al., 2016). Depending on the allele it harbors, *CCTA2* binds the two subunits of CFIm (CFIm25 and CFIm68) with different affinities. Indeed, the G allele for rs6983267, the one linked with greater predisposition for colorectal cancer, creates a structure on the lncRNA which causes higher metastases and cell proliferation compared with the T-allele harboring *CCTA2* (Redis et al., 2016). Another example is *ANRIL.* This lncRNA possesses multiple SNPs modifying its structure and splicing (Aguilo et al., 2016). For instance, rs564389, one of the SNPs most strongly related with *ANRIL* expression, disrupts the Ras Responsive Element Binding Protein 1 (RREB1) and the rs10757278 alters the binding with STAT1, provoking a dysregulated inflammatory response (Aguilo et al., 2016; Harismendy et al., 2011).

As explained before, the rs917997 alters the secondary structure of *Lnc13*. Generating a mutant *Lnc13* lacking a 507bp region (which includes the rs917977 position) we confirmed that this region containing the T1D-associated SNP is fundamental for the conformation of the *Lnc13*-*STAT1*-PCBP2 complex. The absence of this domain alters the configuration of the complex, which, in turn, significantly decreases the ability to activate the signaling pathway of STAT1, and diminished the pro-inflammatory chemokine production. This data highlight the functional effect of the rs917997 SNP in the secondary structure and function of *Lnc13*, and in turn, in T1D pathogenesis at the β cell level. In conclusion, *Lnc13* is implicated in the regulation of virus-induced inflammation in pancreatic β cells via the modulation of STAT1 signaling pathway in an allele-specific manner (**Figure 51**).

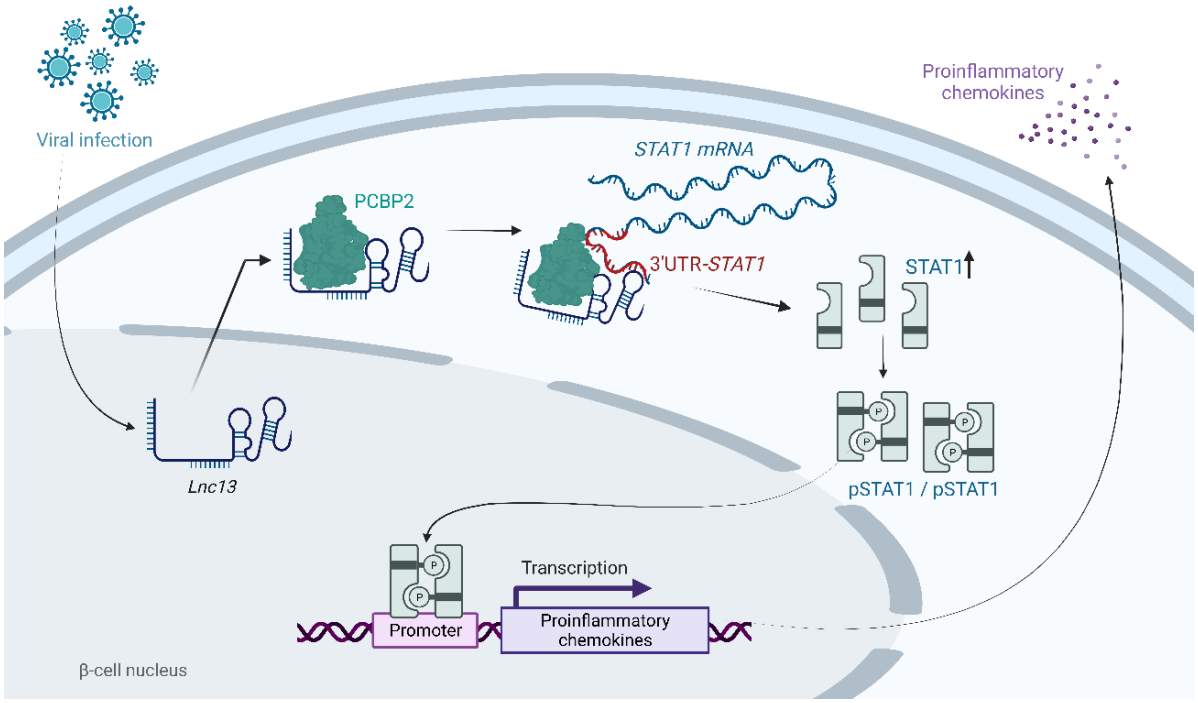


Figure 51. Model of the role of *Lnc13* in the modulation of virus-induced pancreatic β cell inflammation. Viral infections in pancreatic β cells induce an increase in *Lnc13* levels. Additionally, the infection provokes the translocation of this lncRNA from the nucleus of the cells to the cytoplasm. Once in the cytoplasm, *Lnc13* binds the protein PCBP2. The Lnc13-PCBP2 complex interacts with the 3'UTR of STAT1 mRNA and induce its stabilization. The increased stability of the mRNA results in the synthesis of higher quantities of STAT1 protein and stimulate the generation of the active form of STAT1 (phosphorylated STAT1; pSTAT1). The homodimerization of pSTAT1 is translocated onto the nucleus where it promotes the expression of several pro-inflammatory chemokines. Chemokines are releases by the cell, attracting immune cells into the pancreatic islets. In genetically susceptible people harboring the T1D risk genotype for rs917997, the function of *Lnc13* is exharcerbated triggering an excessive antiviral inflammatory response that can contribute to pancreatic β cell destruction.

ARGI regulates the T1D-associated IDIN antiviral network in pancreatic β cells.

ARGI is a newly discovered lncRNA, which has never been characterized. The T1D-associated SNP located in *ARGI* (rs9585056) was previously catalogued as intergenic (Heinig et al., 2010), however the present work has revealed that it is actually located in an exon of *ARGI* lncRNA. In 2010, Heining et al. described that the genotype of the T1D-associated SNP rs9585056 correlated with the expression of an antiviral gene network named IRF7-driven inflammatory gene network (IDIN) in immune cells (Heinig et al., 2010).

The results of this thesis demonstrate that PIC transfection increases *ARGI* expression in β cells, and also Coxsackievirus infection provokes a similar effect. These data confirm that viral infections affect *ARGI* expression, suggesting its implication in antiviral processes. This idea was confirmed in an RNA sequencing experiment in which, we observed an increase in type 1 interferon signaling and antiviral genes in *ARGI*-overexpressing β cells. Interestingly, many of the upregulated genes were part of the IDIN network.

Moreover, we have described how *ARGI* is able to regulate the expression of interferon-stimulated genes (ISGs) by interacting with the transcription factor CTCF. Viral infections mobilize *ARGI* onto the nuclei of pancreatic β cells. Once in the nucleus, *ARGI* interacts with CTCF allowing the binding to regulatory regions of some ISGs and activating their transcription, e.g. *IFNβ* and *ISG15*. Previous investigations have described that the combination between lncRNAs and transcription factors may affect the expression of certain genes (Long et al., 2017; Statello et al., 2021). For example, the lncRNA *Silc1* and the transcription factor SOX11 are co-expressed in mouse dorsal cells and their interaction is essential for the regulation of genes linked with gene regeneration (R. B. Perry et al., 2018). A lncRNA named *PANDA* is induced in a p53-dependent manner and interacts with the transcription factor NF-YA, limiting the expression of pro-apoptotic genes (Hung et al., 2011). In the case of CTCF, it is a transcription factor that is ubiquitously expressed and can either inhibit or activate the transcription of the targeted genes depending on the cell type, stimuli, location of the binding site and presence of additional interacting partners (transcriptional activators, repressors, cohesins and the RNA polymerase II) (Bastiaan Holwerda & de Laat, 2013). Recent studies showed that CTCF is capable of interacting with non-coding RNAs to modulate gene expression; including the regulation of inflammatory genes (Gavrilov et al., 2022; Miyata et al., 2021). The technique called RedCHIP allowed to identify several cis- and trans-acting non-coding RNAs enriched in CTCF-binding sites suggesting a probable implication of CTCF on gene activation and repression (Gavrilov et al., 2022). The versatility and abundance of CTCF allows its implication in a multitude of cellular and biological processes. Some studies have described the interaction of CTCF and non-

coding RNAs in the regulation of inflammatory genes (Saldaña-meyer et al., 2019; Xiang et al., 2014; Yang et al., 2015). In one of these cases, the non-coding RNA named hSATII is described to impair the binding of CTCF to the chromatin, deriving in the alteration of chromatin accessibility and the activation of SASP-like inflammatory genes (Miyata et al., 2021). In our work and thanks to the design of the RNA antisense purification and RNA immunoprecipitation techniques, we demonstrate that in virus-exposed pancreatic β cells, *ARGI* and CTCF are joined and interact with regulatory domains of *IFNβ* and *ISG15* genes. This information evidences that in β cells, viral infections promote the expression of ISGs through the binding of *ARGI* and CTCF to regulatory regions of these antiviral genes.

It is well reported that in response to a viral infection, interferons play a central role in the regulation of the immune response and, with this aim; they modulate the expression of a network of genes known as interferon stimulated genes (ISGs). The ISGs are essential for the modulation of critical cellular processes, but also are implicated in cell demise as it is the pancreatic β cell apoptosis (Akhbari et al., 2020). Evidences support the importance of the IFN pathway and the ISG modulation in the development of T1D (Akhbari et al., 2020; Ivashkiv. & Donlin, 2014; Lundberg et al., 2016). In a study performed in 2016, it was demonstrated that the infection of human islets with CVB3 provoked the upregulation of type I interferons as well as a strong IFN-driven gene signature characterized by an hyperactivation of several ISGs, such as *ISG15*, *STAT1*, *IRF7*, *MX1* and *CXCL10* (Domsgen et al., 2016). Additionally, ISGs were found to be overrepresented in pancreatic islets from T1D patients in comparison with islets from non-diabetic donors (Lundberg et al., 2016). These data indicate that ISG expression is a typical feature in pancreatic islets of T1D patients. Aberrant activation of the ISGs caused either by genetic mutations/variants or abnormal stimuli can lead to an exacerbated activation of the immune system being key in the progression of T1D (Akhbari et al., 2020; Dias Junior et al., 2019; Ivashkiv. & Donlin, 2014; Lundberg et al., 2016).

We have already reported that the T1D-linked SNP under study is located inside an exon of *ARGI* and, as predicted *in silico,* disrupts its secondary structure. Furthermore, a previously work already linked the genotype of this SNP with regulation differential activation of the IDIN network in immune cells; even if at that time, it was considered as an intergenic SNP (Heinig et al., 2010). In line with this, our data shows that the allele-specific upregulation of *ARGI* triggers an increase in the expression of genes belonging to the IDIN network in an allele-specific manner. Actually, *ARGI* harboring the risk allele for T1D induced higher expression levels of the IDIN genes than ARGI harboring the protective allele. Furthermore, there is a significant stronger interaction between CTCF and *ARGI* when the lncRNA harbors the risk allele, explaining the exacerbated increase in ISG expression in the cells expressing ARGI with the T1D risk allele.

In summary, this work demonstrated that *ARGI* is implicated in the modulation of virus-induced pancreatic β cells inflammation. The molecular mechanism implies an allele-specific binding to the transcription factor CTCF, which in turn regulates the expression of ISG genes (**Figure 52**). These results are in concordance with the expression data from pancreatic islets of T1D donors, in which there is an upregulation of type I IFNs (Foulis et al., 1987; Huang et al., 1995; Pujol-Borrell et al., 1986). Indeed, The Diabetes Virus Detection (DiViD) study has demonstrated that several ISGs, including *STAT1*, *IFI6* or *MX1,* are significantly upregulated in islets from T1D donors (Krogvold, Leete, et al., 2022; Lundberg et al., 2016).

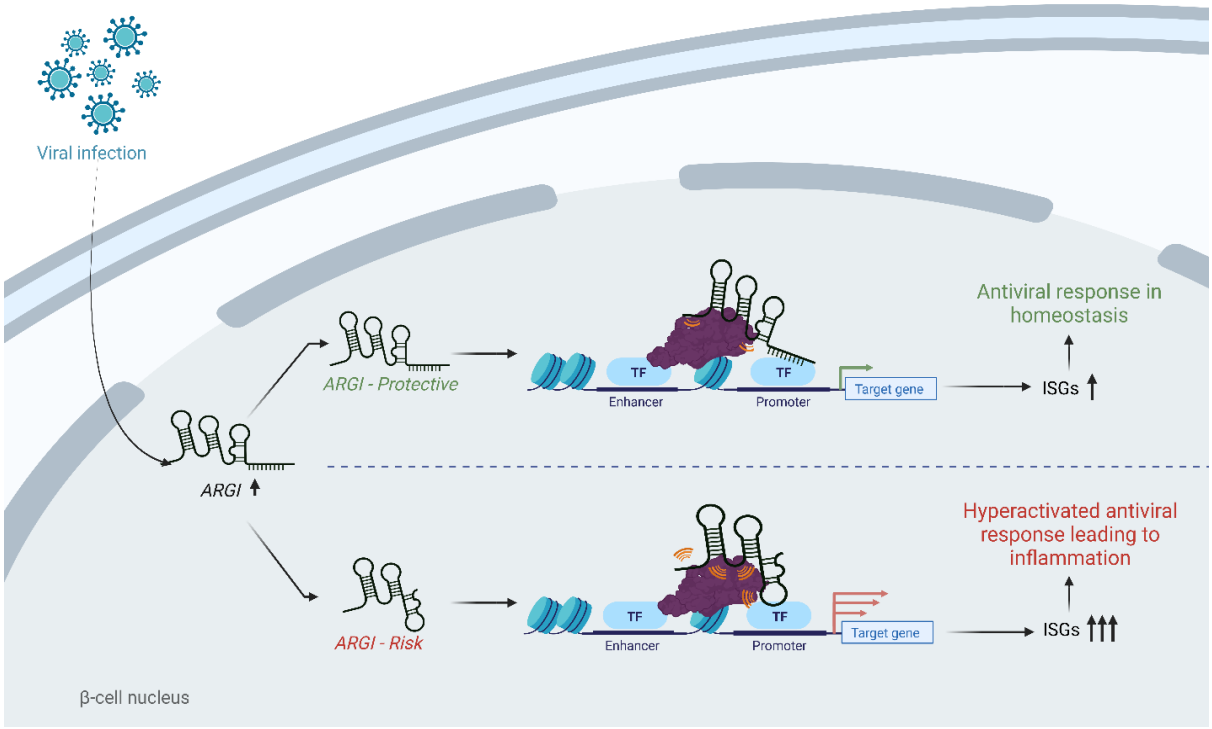


Figure 52. A model of the role of *ARGI* in the allele-specific modulation of virus-induced ISG transcription in pancreatic β cells. Viral infections trigger ARGI upregulation in the nuclei of pancreatic β cells. *ARGI* interacts with the transcription factor CTCF to bind to the regulatory regions of interferon stimulated genes (ISGs), promoting their transcription. When *ARGI* harbors the T1D protective allele, a homeostatic antiviral response is activated; however when the T1D risk allele is present, this process is exacerbated, leading to a hyperactivation of the antiviral response in pancreatic β cells. In the context of T1D pathogenesis, the generation of an excessive pro-inflammatory environment in pancreatic islets, leads to increased insulitis, and eventually,to pancreatic β cell destruction and T1D development.

Chapter 7

Conclusions

Ondorioak

Ondorioak:

1. Gene kodetzaileez gainera, gene ez-kodetzaileek ere, gaixotasun askoren garapenerako ezinbestekoak diren bidezidor molekular anitz modulatzeko ahalmena daukate.

2. Infekzio biralek T1D gaixotasunarekin loturiko lncRNAn adierazpena emendatu dezakete pankreako β zeluletan, hauetan erantzun antibiralaren aktibazioa lagunduz.

3. Infekzio biral baten ostean, *Lnc13* pankreako β zelularen nukleotik zitoplasmara translokatzen da. Behin zitoplasman, *Lnc13* molekulak konplexu bat sortzen du PCBP2 proteina eta *STAT1* molekularen 3'UTRarekin. Konplexu honek *STAT1*aren estabilitatzea emendatzen du, β zeluletan kimiokina pro-inflamatorioen adierazpenea eta askapena handituz.

4. Infekzio biralek *ARGI*ren adierazpena emendatu egiten dute pankreako β zelulen nukleoetan, non CTCF transkripzio faktorearekin erlazionatzen da. Elkartze honek interferoiz-estimulaturiko geneen gune erregulatzaileetara lotzea eragiten du, gene hauen transkripzioa aktibatuz.

5. Infekzio viral baten ostean Lnc13 eta *ARGI*k pankreako β zeluletan daukaten funtzioa alelo-espezifikoa da. LncRNA hauetan dauden T1Dari loturiko SNPek euren bigarren mailako egitura aldatzen dute, beste molekulekin interakzionatzeko duten ahalmenean eraldatzen duena; eta hori dela eta, birusak sorturiko inflamazioan duten eragina aldatzen dute, zeina T1D gaixotasunaren garapenean ere bultzatu dezakeena.

Conclusions:

1. Besides protein coding genes, non-coding genes are also capable of modulating diverse key molecular pathways that are important in the development of multiple pathogenic conditions.

2. Viral infections increase the expression of T1D-associated lncRNAs in pancreatic β cells, facilitating the activation of the antiviral response in these cells.

3. After a viral assault, *Lnc13* translocates from the nucleus to the cytoplasm of pancreatic β cells. Once in the cytoplasm, *Lnc13* forms a complex with PCBP2 and the 3'UTR of *STAT1* increasing *STAT1* stability, which exacerbates the expression and release of pro-inflammatory chemokines in β cells.

4. Viral infections increase the abundance of *ARGI* in the nuclei of pancreatic β cells, where it interacts with the transcription factor CTCF, promoting its binding to the regulatory regions of interferon-stimulated genes and promoting their transcription.

5. The effect of *Lnc13* and *ARGI* in virus-induced pancreatic β cell inflammation is allele-specific. T1D-associated SNPs present in these lncRNAs alter their secondary structure, influencing their ability to interact with other molecules and thus, altering their implication in virus-induced inflammation, and eventually in T1D development.

Chapter 8

References

Erreferentziak

References

Agarwal, S., Vierbuchen, T., Ghosh, S., Chan, J., Jiang, Z., Kandasamy, R. K., Ricci, E., & Fitzgerald, K. A. (2020). The long non-coding RNA LUCAT1 is a negative feedback regulator of interferon responses in humans. *Nature Communications*, *11*(6348). https://doi.org/10.1038/s41467-020-20165-5

Aguilo, F., Di Cecilia, S., & Walsh, M. J. (2016). Long Non-coding RNA ANRIL and Polycomb in Human Cancers and Cardiovascular Disease. *Current Topics in Microbiology and Immunology*, *394*, 29–39. https://doi.org/10.1007/82

Aguilo, F., Zhou, M.-M., & Walsh, M. J. (2011). Long noncoding RNA, polycomb, and the ghosts haunting INK4b-ARK-INK4a expression. *Cancer Research*, *71*(16), 5365–5369. https://doi.org/10.1158/0008-5472.CAN-10-4379.Long

Akerman, I., Tu, Z., Beucher, A., Schadt, E., Ravassard, P., Ferrer, J., Akerman, I., Tu, Z., Beucher, A., Rolando, D. M. Y., Sauty-colace, C., Benazra, M., Nakic, N., Yang, J., Wang, H., Pasquali, L., Moran, I., & Garcia-hurtado, J. (2017). Human Pancreatic b Cell lncRNAs Control Cell- Article Human Pancreatic b Cell lncRNAs Control Cell-Specific Regulatory Networks. *Cell Metabolism*, *25*(2), 400–411. https://doi.org/10.1016/j.cmet.2016.11.016

Akhbari, P., Richardson, S. J., & Morgan, N. G. (2020). Type 1 diabetes: Interferons and the aftermath of pancreatic beta-cell enteroviral infection. *Microorganisms*, *8*(1419), 1–18. https://doi.org/10.3390/microorganisms8091419

Alkanani, A. K., Hara, N., Gianani, R., & Zipris, D. (2014). Kilham rat virus-induced type 1 diabetes involves beta cell infection and intra-islet JAK-STAT activation prior to insulitis. *Virology*, *468*(470), 19–27. https://doi.org/10.1016/j.virol.2014.07.041

Allred, M. G., Chimenti, M. S., Ciecko, A. E., Chen, Y. G., & Lieberman, S. M. (2021). Characterization of type i interferon-associated chemokines and cytokines in lacrimal glands of nonobese diabetic mice. *International Journal of Molecular Sciences*, *22*(7), 1–19. https://doi.org/10.3390/ijms22073767

Aly, T. A., Ide, A., Jahromi, M. M., Barker, J. M., Fernando, M. S., Babu, S. R., Yu, L., Miao, D., Erlich, H. A., Fain, P. R., Barriga, K. I., Norris, J. M., Rewers, M. J., & Eisenbarth, G. S. (2006). Extreme genetic risk for type 1A diabetes. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(38), 14074–14079. https://doi.org/10.1073/pnas.0606349103

Chapter 8: References

American Diabetes Association. (2019). 2. Classification and diagnosis of diabetes: Standards of medical care in diabetes - 2019. *Diabetes Care*, *42*(SUPL. 1), S13–S28. https://doi.org/10.2337/dc19-S002

Armaos, A., Colantoni, A., Proietti, G., Rupert, J., & Tartaglia, G. G. (2021). catRAPID omics v2.0: going deeper and wider in the prediction of protein–RNA interactions. *Nucleic Acids Research*, *49*, W72–W79. https://doi.org/10.1093/nar/gkab393

Assmann, T. S., Brondani, L. de A., Bouças, A. P., Canani, L. H., & Crispim, D. (2015). Toll-like receptor 3 (TLR3) and the development of type 1 diabetes mellitus. *Archives of Endocrinology and Metabolism*, *59*(1), 4–12. https://doi.org/10.1590/2359-3997000000003

Atkinson, M. A., Elisenbarth, G. S., & Michels, A. W. (2014). Type 1 diabetes. *Lancet*, *383*(9911), 69–82. https://doi.org/10.1016/S0140-6736(13)60591-7

Atkinson, M. A., & Maclaren, N. K. (1994). The pathogenesis of insulin-dependent diabetes mellitus. *The New England Journal of Medicine*, *331*(21), 1428–1436.

Atkinson, M. A., & Wilson, S. B. (2002). Fatal attraction : chemokines and type 1 diabetes. *The Journal of Clinical Investigations*, *110*(11), 1611–1613. https://doi.org/10.1172/JCI200217311.Research

Au-Yeung, N., Mandhana, R., & Horvath, C. M. (2013). Transcriptional regulation by STAT1 and STAT2 in the interferon JAK-STAT pathway. *Jak-Stat*, *2*(3), e23931. https://doi.org/10.4161/jkst.23931

Aznaourova, M., Schmerer, N., Schmeck, B., & Schulte, L. N. (2020). Disease-Causing Mutations and Rearrangements in Long Non-coding RNA Gene Loci. *Frontiers in Genetics*, *11*. https://doi.org/10.3389/fgene.2020.527484

Babenko, A. P., Polak, M., Cavé, H., Busiah, K., Czernichow, P., Scharfmann, R., Bryan, J., Aguilar-bryan, L., Vaxillaire, M., & Froguel, P. (2006). Activating mutations in the ABCC8 gene in neonatal diabetes mellitus. *The New England Journal of Medicine*, *355*(5), 456–466.

Babon, J. J., Lucet, I. S., Murphy, J. M., Nicola, N. A., & Varghese, L. N. (2014). The molecular regulation of Janus kinase (JAK) activation. *Biochemical Journal*, *462*(1), 1–13. https://doi.org/10.1042/BJ20140712

Balas, M. M., & Johnson, A. M. (2018). Exploring the mechanisms behind long noncoding RNAs and cancer. *Non-Coding RNA Research*, *3*(3), 108–117. https://doi.org/10.1016/j.ncrna.2018.03.001

Bastiaan Holwerda, S. J., & de Laat, W. (2013). CTCF: The protein, the binding partners, the binding sites and their chromatin loops. In *Philosophical Transactions of the Royal Society B: Biological*

Sciences (Vol. 368, Issue 1620). Royal Society of London. https://doi.org/10.1098/rstb.2012.0369

Batista, T. M., Haider, N., & Kahn, C. R. (2021). Defining the underlying defect in insulin action in type 2 diabetes. *Diabetologia*, *64*(5), 994–1006. https://doi.org/10.1007/s00125-021-05415-5

Beckles, G. L., & Chou, C.-F. (2016). Disparities in the Prevalence of Diagnosed Diabetes — United States, 1999–2002 and 2011–2014. *Morbidity and Mortality Weekly Report*, *65*(45), 1265–1269. https://doi.org/10.15585/mmwr.mm6545a4

Bell, G. I., Selby, M. J., & Rutter, W. J. (1982). The highly polymorphic region near the human insulin gene is composed of simple andemly repeating sequences. *Nature*, *296*(5857), 585. https://doi.org/10.1038/296585c0

Beltrand, J., Busiah, K., Vaivre-Douret, L., Fauret, A. L., Berdugo, M., Cavé, H., & Polak, M. (2020). Neonatal Diabetes Mellitus. *Frontiers in Pediatrics*, *8*, 1–9. https://doi.org/10.3389/fped.2020.540718

Bennett, S. T., Lucassen, A. M., Gough, S. C. L., Powell, E. E., Undlien, D. E., Pritchard, L. E., Merriman, M. E., Kawaguchi, Y., Dronsfield, M. J., Pociot, F., Nerup, J., Bouzekri, N., Cambon-Thomsen, A., Rønningen, K. S., Barnett, A. H., Bain, S. C., & Todd, J. A. (1995). Susceptibility to human type 1 diabetes at IDDM2 is determined by tandem repeat variation at the insulin gene minisatellite locus. *Nature*, *2*(11), 983–989.

Beyerlein, A., Chmiel, R., Hummel, S., Winkler, C., Bonifacio, E., & Ziegler, A. G. (2014). Timing of gluten introduction and islet autoimmunity in young children: Updated results from the Babydiet study. *Diabetes Care*, *37*(9), 194–195. https://doi.org/10.2337/dc14-1208

Bhat, S. A., Ahmad, S. M., Mumtaz, P. T., Malik, A. A., Dar, M. A., Urwat, U., Shah, R. A., & Ganai, N. A. (2016). Long non-coding RNAs: Mechanism of action and functional utility. *Non-Coding RNA Research*, *1*(1), 43–50. https://doi.org/10.1016/j.ncrna.2016.11.002

Boise, L. H., Minn, A. J., Noel, P. J., June, C. H., Accavitti, M. A., Lindsten, T., & Thompson, C. B. (1995). CD28 costimulation can promote T cell survival by enhancing the expression of Bcl-XL. *Immunity*, *3*(1), 87–98. http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=7621080&retmode=ref&cmd=prlinks%5Cnpapers3://publication/uuid/9115F254-4810-4F23-A5F1-579A7102DE75

Bonasio, R., Tu, S., & Reinberg, D. (2010). Molecular signals of epigenetic states. *Science*, *330*(6004), 612–616. https://doi.org/10.1126/science.1191078.Molecular

Chapter 8: References

Bonnefond, A., Lomberk, G., Buttar, N., Busiah, K., Vaillant, E., Lobbens, S., Yengo, L., Dechaume, A., Mignot, B., Simon, A., Scharfmann, R., Neve, B., Tanyolaç, S., Hodoglugil, U., Pattou, F., Cavé, H., Iovanna, J., Stein, R., Polak, M., … Urrutia, R. (2011). Disruption of a novel krüppel-like transcription factor p300-regulated pathway for insulin biosynthesis revealed by studies of the c.-331 INS mutation found in neonatal diabetes mellitus. *Journal of Biological Chemistry*, *286*(32), 28414–28424. https://doi.org/10.1074/jbc.M110.215822

Bridges, M. C., Daulagala, A. C., & Kourtidis, A. (2021). LNCcation: lncRNA localization and function. *Journal of Cell Biology*, *220*(2), 1–17. https://doi.org/10.1083/JCB.202009045

Brock, M., Schuoler, C., Leuenberger, C., Bühlmann, C., Haider, T. J., Vogel, J., Ulrich, S., Gassmann, M., Kohler, M., & Huber, L. C. (2017). Analysis of hypoxia-induced noncoding RNAs reveals metastasis-associated lung adenocarcinoma transcript 1 as an important regulator of vascular smooth muscle cell proliferation. *Experimental Biology and Medicine*, *242*(5), 487–496. https://doi.org/10.1177/1535370216685434

Busiah, K., Drunat, S., Vaivre-Douret, L., Bonnefond, A., Simon, A., Flechtner, I., Gérard, B., Pouvreau, N., Elie, C., Nimri, R., Vries, L. De, Tubiana-Rufi, N., Metz, C., Bertrand, A. M., Nivot-Adamiak, S., de Kerdanet, M., Stuckens, C., Jennane, F., Souchon, P. F., … Cavé, H. (2013). Neuropsychological dysfunction and developmental defects associated with genetic changes in infants with neonatal diabetes mellitus: A prospective cohort study. *The Lancet Diabetes and Endocrinology*, *1*(3), 199–207. https://doi.org/10.1016/S2213-8587(13)70059-7

Butturini, E., de Prati, A. C., & Mariotto, S. (2020). Redox regulation of STAT1 and STAT3 signaling. *International Journal of Molecular Sciences*, *21*(19), 1–18. https://doi.org/10.3390/ijms21197034

Cabrera, O., Berman, D. M., Kenyon, N. S., Ricordi, C., Berggren, P. O., & Caicedo, A. (2006). The unique cytoarchitecture of human pancreatic islets has implications for islet cell function. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(7), 2334–2339. https://doi.org/10.1073/pnas.0510790103

Carlevaro-Fita, J., Polidori, T., Das, M., Navarro, C., Zoller, T. I., & Johnson, R. (2019). Ancient exapted transposable elements promote nuclear enrichment of human long noncoding RNAs. *Genome Research*, *29*, 208–222.

Carrieri, C., Cimatti, L., Biagioli, M., Beugnet, A., Zucchelli, S., Fedele, S., Pesce, E., Ferrer, I., Collavin, L., Santoro, C., Forrest, A. R. R., Carninci, P., Biffo, S., Stupka, E., & Gustincich, S. (2012). Long non-coding antisense RNA controls Uchl1 translation through an embedded SINEB2 repeat.

Nature, *493*, 454–457. https://doi.org/10.1038/nature11508

Castellanos-Rubio, A., Fernandez-Jimenez, N., Kratchmarov, R., Luo, X., Bhagat, G., Green, P. H. R., Schneider, R., Kiledjian, M., Bilbao, J. R., & Ghosh, S. (2016). A long noncoding RNA associated with susceptibility to celiac disease. *Science*, *352*(6281), 91–95. https://doi.org/10.1126/science.aad0467.

Castro-Sanchez, P., Teagle, A. R., Prade, S., & Zamoyska, R. (2020). Modulation of TCR Signaling by Tyrosine Phosphatases: From Autoimmunity to Immunotherapy. *Frontiers in Cell and Developmental Biology*, *8*(December), 1–31. https://doi.org/10.3389/fcell.2020.608747

Chandra, V., Ibrahim, H., Halliez, C., Prasad, R. B., Vecchio, F., Dwivedi, O. P., Kvist, J., Balboa, D., Saarimäki-Vire, J., Montaser, H., Barsby, T., Lithovius, V., Artner, I., Gopalakrishnan, S., Groop, L., Mallone, R., Eizirik, D. L., & Otonkoski, T. (2022). The type 1 diabetes gene TYK2 regulates β-cell development and its responses to interferon-α. *Nature Communications*, *13*(6363), 1–16. https://doi.org/10.1038/s41467-022-34069-z

Chen, L., Zhang, Y. H., Pan, X., Liu, M., Wang, S., Huang, T., & Cai, Y. D. (2018). Tissue expression difference between mRNAs and lncRNAs. *International Journal of Molecular Sciences*, *19*(3416), 1–18. https://doi.org/10.3390/ijms19113416

Chen, M., Xu, R., Ji, H., Greening, D. W., Rai, A., Izumikawa, K., Ishikawa, H., Takahashi, N., & Simpson, R. J. (2016). Transcriptome and long noncoding RNA sequencing of three extracellular vesicle subtypes released from the human colon cancer LIM1863 cell line. *Scientific Reports*, *6*(38397), 1–14. https://doi.org/10.1038/srep38397

Chikuma, S. (2017). CTLA-4, an essential immune-checkpoint for T-Cell activation. In *Current Topics in Microbiology and Immunology* (Vol. 410).

Choo, S. Y. (2007). The HLA system: genetics, immunology, clinical testing and clinical implications. *Yonsei Medical Journal*, *48*(1), 11–23.

Ciężki, S., Kurpiewska, E., Bossowski, A., & Głowińska-Olszewska, B. (2022). Multi-Faceted Influence of Obesity on Type 1 Diabetes in Children – From Disease Pathogenesis to Complications. *Frontiers in Endocrinology*, *13*(June), 1–18. https://doi.org/10.3389/fendo.2022.890833

Cobo-Vuilleumier, N., Lorenzo, P. I., Rodríguez, N. G., Herrera Gómez, I. D. G., Fuente-Martin, E., López-Noriega, L., Mellado-Gil, J. M., Romero-Zerbo, S. Y., Baquié, M., Lachaud, C. C., Stifter, K., Perdomo, G., Bugliani, M., De Tata, V., Bosco, D., Parnaud, G., Pozo, D., Hmadcha, A., Florido, J. P., … Gauthier, B. R. (2018). LRH-1 agonism favours an immune-islet dialogue which protects

against diabetes mellitus. *Nature Communications*, *9*(1), 1–15. https://doi.org/10.1038/s41467-018-03943-0

Colli, M. L., Hill, J. L. E., Marroquí, L., Chaffey, J., Dos Santos, R. S., Leete, P., Coomans de Brachène, A., Paula, F. M. M., Op de Beeck, A., Castela, A., Marselli, L., Krogvold, L., Dahl-Jorgensen, K., Marchetti, P., Morgan, N. G., Richardson, S. J., & Eizirik, D. L. (2018). PDL1 is expressed in the islets of people with type 1 diabetes and is up-regulated by interferons-α and-γ via IRF1 induction. *EBioMedicine*, *36*, 367–375. https://doi.org/10.1016/j.ebiom.2018.09.040

Colli, M. L., Moore, F., Gurzov, E. N., Ortis, F., & Eizirik, D. L. (2010). MDA5 and PTPN2, two candidate genes for type 1 diabetes, modify pancreatic β-cell responses to the viral by-product double-stranded RNA. *Human Molecular Genetics*, *19*(1), 135–146. https://doi.org/10.1093/hmg/ddp474

Colli, M. L., Nogueira, T. C., Allagnat, F., Cunha, D. A., Gurzov, E. N., Cardozo, A. K., Roivainen, M., de beeck, A. O., & Eizirik, D. L. (2011). Exposure to the viral by-product dsrna or coxsackievirus b5 triggers pancreatic beta cell apoptosis via a bim/ mcl-1 imbalance. *PLoS Pathogens*, *7*(9), 1–15. https://doi.org/10.1371/journal.ppat.1002267

Colli, M. L., Paula, F. M., Marselli, L., Marchetti, P., Roivainen, M., Eizirik, D. L., & Op de beeck, A. (2019). Coxsackievirus B tailors the unfolded protein response to favour viral amplification in pancreatic β cells. *Journal of Innate Immunity*, *11*, 375–389.

Colli, M. L., Ramos-Rodríguez, M., Nakayasu, E. S., Alvelos, M. I., Lopes, M., Hill, J. L. E., Turatsinze, J. V., Coomans de Brachène, A., Russell, M. A., Raurell-Vila, H., Castela, A., Juan-Mateu, J., Webb-Robertson, B. J. M., Krogvold, L., Dahl-Jorgensen, K., Marselli, L., Marchetti, P., Richardson, S. J., Morgan, N. G., … Eizirik, D. L. (2020). An integrated multi-omics approach identifies the landscape of interferon-α-mediated responses of human pancreatic beta cells. *Nature Communications*, *11*(2584), 1–17. https://doi.org/10.1038/s41467-020-16327-0

Coomans de Brachène, A., Sousa Dos Santos, R., Marroqui, L., Colli, M. L., Marselli, L., Mirmira, R. G., Marchetti, P., & Eizirik, D. L. (2018). IFNα induces a preferential long-lasting expression of MHC class I in human pancreatic beta cells. *Diabetolo*, *61*(3), 636–640. https://doi.org/10.1007/s00125-017-4536-4.

Coppieters, K. T., Boettler, T., & von Herrath, M. (2012). Virus infections in type 1 diabetes. *Cold Spring Harbor Perspectives in Medicine*, *2*(1), 1–14. https://doi.org/10.1101/cshperspect.a007682

Chapter 8: References

Coppieters, K. T., Wiberg, A., Tracy, S. M., & von Herrath, M. G. (2011). Immunology in the clinic review series: focus on type 1 diabetes and viruses: the role of viruses in type 1 diabetes: a difficult dilemma. *Clinical and Experimental Immunology*, *168*, 5–11.

Cosentino, C., Toivonen, S., Villamil, E. D., Atta, M., Ravanat, J. L., Demine, S., Schiavo, A. A., Pachera, N., Deglasse, J. P., Jonas, J. C., Balboa, D., Otonkoski, T., Pearson, E. R., Marchetti, P., Eizirik, D. L., Cnop, M., & Igoillo-Esteve, M. (2018). Pancreatic β-cell tRNA hypomethylation and fragmentation link TRMT10A deficiency with diabetes. *Nucleic Acids Research*, *46*(19), 10302–10318. https://doi.org/10.1093/nar/gky839

Couture, A., Garnier, A., Docagne, F., Boyer, O., Vivien, D., Le-Mauff, B., Latouche, J. B., & Toutirais, O. (2019). HLA-class II artificial antigen presenting cells in CD4+ T cell-based immunotherapy. *Frontiers in Immunology*, *10*(May), 1–11. https://doi.org/10.3389/fimmu.2019.01081

Cudworth, A. G., & Woodrow, J. C. (1975). Evidence for HL-A-linked Genes in "Juvenile" Diabetes Mellitus. *British Medical Journal*, *3*(5976), 133–135. https://doi.org/10.1136/bmj.3.5976.133

Deakin, J. E., Papenfuss, A. T., Belov, K., Cross, J. G. R., Coggill, P., Palmer, S., Sims, S., Speed, T. P., Beck, S., & Marshall Graves, J. A. (2006). Evolution and comparative analysis of the MHC Class III inflammatory region. *BMC Genomics*, *7*(281), 1–14. https://doi.org/10.1186/1471-2164-7-281

Degos, L., & Dausset, J. (1974). Histocompatibility determinants in multiple sclerosis. *The Lancet*, *10*(3), 307–308. https://doi.org/10.1177/105960118501000301

Demir, S., Nawroth, P. P., Herzig, S., & Ekim Üstünel, B. (2021). Emerging Targets in Type 2 Diabetes and Diabetic Complications. *Advanced Science*, *8*(18), 1–23. https://doi.org/10.1002/advs.202100275

Derrien, T., Johnson, R., Bussotti, G., Tanzer, A., Djebali, S., Tilgner, H., Guernec, G., Martin, D., Merkel, A., Knowles, D. G., Lagarde, J., Veeravalli, L., Ruan, X., Ruan, Y., Lassmann, T., Carninci, P., Brown, J. B., Lipovich, L., Gonzalez, J. M., … Guigó, R. (2012). The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression. *Genome Research*, *22*(9), 1775–1789. https://doi.org/10.1101/gr.132159.111

Dey, B. K., Pfeifer, K., & Dutta, A. (2014). The H19 long noncoding RNA gives rise to microRNAs miR-675-3p and miR-675-5p to promote skeletal muscle differentiation and regeneration. *Genes and Development*, *28*, 491–501. https://doi.org/10.1101/gad.234419.113.3

Dias Junior, A. G., Sampaio, N. G., & Rehwinkel, J. (2019). A Balancing Act : MDA5 in Antiviral Immunity and Autoinflammation. *Trends in Microbiology*, *27*(1), 75–85.

https://doi.org/10.1016/j.tim.2018.08.007

Ding, H., Wang, F., Shi, X., Ma, H., Du, Y., Hou, L., & Xing, N. (2020). LncRNA MALAT1 induces the dysfunction of β cells via reducing the histone acetylation of the PDX-1 promoter in type 1 diabetes. *Experimental and Molecular Pathology*, *114*(12), 104432. https://doi.org/10.1016/j.yexmp.2020.104432

Dixit, E., & Kagan, J. C. (2013). Intracellular Pathogen Detection by RIG-I-Like Receptors. *Advances in Immunology*, *117*, 99–125. https://doi.org/10.1016/B978-0-12-410524-9.00004-9

Docherty, L. E., Kabwama, S., Lehmann, A., Hawke, E., Harrison, L., Flanagan, S. E., Ellard, S., Hattersley, A. T., Shield, J. P. H., Ennis, S., Mackay, D. J. G., & Temple, I. K. (2013). Clinical presentation of 6q24 transient neonatal diabetes mellitus (6q24 TNDM) and genotype-phenotype correlation in an international cohort of patients. *Diabetologia*, *56*(4), 758–762. https://doi.org/10.1007/s00125-013-2832-1

Domsgen, E., Lind, K., Kong, L., Hühn, M. H., Rasool, O., Van Kuppeveld, F., Korsgren, O., Lahesmaa, R., & Flodström-Tullberg, M. (2016). An IFIH1 gene polymorphism associated with risk for autoimmunity regulates canonical antiviral defence pathways in Coxsackievirus infected human pancreatic islets. *Nature Publishing Group*, *6*(39378), 1–14. https://doi.org/10.1038/srep39378

Dorrell, C., Schung, J., Lin, C. F., Canaday, P. S., Fox, A. J., O, S., Bonnah, R., Streeter, P. R., Stoeckert Jr, C. J., Kaestner, K. H., & Grompe, M. (2011). Transcriptomes of the major human pancreatic cell types. *Diabetologia*, *54*(11), 1–23. https://doi.org/10.1007/s00125-011-2283-5

Dotta, F., Censini, S., Van Halteren, A. G. S., Marselli, L., Masini, M., Dionisi, S., Mosca, F., Boggi, U., Muda, A. O., Del Prato, S., Elliott, J. F., Covacci, A., Rappuoli, R., Roep, B. O., Marchetti, P., Covacci, A., Rappuoli, R., Roep, B. O., & Marchetti, P. (2007). Coxsackie B4 virus infection of β cells and natural killer cell insulitis in recent-onset type 1 diabetic patients. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(12), 5115–5120. https://doi.org/10.1073/pnas.0700442104

Du, M., Yuan, L., Tan, X., Huang, D., Wang, X., Zheng, Z., Mao, X., Li, X., Yang, L., Huang, K., Zhang, F., Wang, Y., Luo, X., Huang, D., & Huang, K. (2017). The LPS-inducible lncRNA Mirt2 is a negative regulator of inflammation. *Nature Communications*, *8*(2049). https://doi.org/10.1038/s41467-017-02229-1

Durinovic-Belló, I., Wu, R. P., Gersuk, V. H., Sanda, S., Shilling, H. G., & Nepom, G. T. (2010). Insulin gene VNTR genotype associates with frequency and phenotype of the autoimmune response to

proinsulin. *Genes and Immunity*, *11*(2), 188–193. https://doi.org/10.1038/gene.2009.108

Eizirik, D. L., Colli, M. L., & Ortis, F. (2009). The role of inflammation in insulitis and β-cell loss in type 1 diabetes. In *Nature Reviews Endocrinology* (Vol. 5, Issue 4, pp. 219–226). https://doi.org/10.1038/nrendo.2009.21

Eizirik, D. L., Pasquali, L., & Cnop, M. (2020). Pancreatic β-cells in type 1 and type 2 diabetes mellitus: different pathways to failure. *Nature Reviews Endocrinology*, *16*(7), 349–362. https://doi.org/10.1038/s41574-020-0355-7

Engreitz, J. M., Pandya-Jones, A., McDonel, P., Shishkin, A., Sirokman, K., Surka, C., Kadri, S., Xing, J., Goren, A., Lander, E. S., Plath, K., & Guttman, M. (2013). The Xist lncRNA exploits three-dimensional genome architecture to spread across the X-chromosome. *Science*, *341*(6147). https://doi.org/10.1126/science.1237973

Erlich, H., Valdes, A. M., Noble, J. A., Carlson, J. A., Varney, M., Concannon, P., Mychaleckyj, J. C., Todd, J. A., Bonella, P., Fear, A. L., Lavant, E., Louey, A., & Moonsamy, P. (2008). Effect of cibachron blue F3GA on the polymerization of actin. *Diabetes*, *57*(4), 1084–1092. https://doi.org/doi:10.2337/db07-1331

Esguerra, J. L. S., Ofori, J. K., Nagao, M., Shuto, Y., Karagiannopoulos, A., Fadista, J., Sugihara, H., Groop, L., & Eliasson, L. (2020). Glucocorticoid induces human beta cell dysfunction by involving riborepressor GAS5 LincRNA. *Molecular Metabolism*, *32*(December), 160–167. https://doi.org/10.1016/j.molmet.2019.12.012

Fabbri, M., Frixou, M., Degano, M., & Fousteri, G. (2019). Type 1 diabetes in STAT protein family mutations: Regulating the Th17/Treg equilibrium and beyond. *Diabetes*, *68*(2), 258–265. https://doi.org/10.2337/db18-0627

Fadista, J., Vikman, P., Ottosson, E., Guerra, I., Lou, J., Taneera, J., Storm, P., Osmark, P., Ladenvall, C., Prasad, R. B., Hansson, K. B., Finotello, F., Uvebrant, K., Ofori, J. K., Di Camillo, B., Krus, U., Cilio, C. M., Hansson, O., Eliasson, L., … Groop, L. (2014). Global genomic and transcriptomic analysis of human pancreatic islets reveals novel genes influencing glucose metabolism. *Proceedings of the National Academy of Sciences*, *111*(38), 13924–13929. https://doi.org/10.1073/pnas.1402665111

Faghihi, M. A., Modarresi, F., Khalil, A. M., Wood, D. E., Sahagan, B. G., Morgan, T. E., Finch, C. E., St, G., St. Laurent III, G., Kenny, P. J., & Wahlestedt, C. (2008). Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid fedd-forward regulation of β-secretase

expression. *Nature Medicine*, *14*(7), 723–730. https://doi.org/10.1038/nm1784.Expression

Faghihi, M. A., Zhang, M., Huang, J., Modarresi, F., Van der Brug, M. P., Nalls, M. A., Cookson, M. R., St-Laurent, G., & Wahlestedt, C. (2010). Evidence for natural antisense transcript-mediated inhibition of microRNA function. *Genome Biology*, *11*(5), 1–13. https://doi.org/10.1186/gb-2010-11-5-r56

Falasca, M., Raimondi, C., & Maffucci, T. (2011). Boyden chamber. *Methods in Molecular Biology*, *769*(4), 87–95. https://doi.org/10.1007/978-1-61779-207-6_7

Feng, T., Feng, N., Zhu, T., Li, Q., Zhang, Q., Wang, Y., Gao, M., Zhou, B., Yu, H., Zheng, M., & Qian, B. (2020). A SNP-mediated lncRNA (LOC146880) and microRNA (miR-539-5p) interaction and its potential impact on the NSCLC risk. *Journal of Experimental and Clinical Cancer Research*, *39*(157), 1–12. https://doi.org/10.1186/s13046-020-01652-5

Fernandes, J. C. R., Acuña, S. M., Aoki, J. I., Floeter-Winter, L. M., & Muxel, S. M. (2019). Long non-coding RNAs in the regulation of gene expression: Physiology and disease. *Non-Coding RNA*, *5*(17), 1–25. https://doi.org/10.3390/ncrna5010017

Fève, B., & Scheen, A. J. (2022). When therapeutic drugs lead to diabetes. *Diabetologia*, *65*(5), 751–762. https://doi.org/10.1007/s00125-022-05666-w

Filippi, C. M., & Von Herrath, M. G. (2008). Viral trigger for type 1 diabetes: Pros and cons. *Diabetes*, *57*(11), 2863–2871. https://doi.org/10.2337/db07-1023

Foulis, A. K., Farquharson, M. A., & Meager, A. (1987). Immunoreactive α-interferon in insulin-secreting β cells in type 1 diabetes mellitus. *The Lancet*, 1423–1427.

Frauwirth, K. A., Riley, J. L., Harris, M. H., Parry, R. V, Rathmell, J. C., Plas, D. R., Elstrom, R. L., June, C. H., & Thompson, C. B. (2002). The CD28 Signaling Pathway Regulates Glucose Metabolism ability of resting cells to take up and utilize nutrients at levels sufficient to maintain viability (Rathmell et al in fat and muscle cells insulin induces glucose uptake in excess of that required . *Immunity*, *16*, 769–777.

Fu, Y., Zhang, Y., Cui, J., Yang, G., Peng, S., Mi, W., Yin, X., Yu, Y., Jiang, J., Liu, Q., Qin, Y., & Xu, W. (2020). SNP rs12982687 affects binding capacity of lncRNA UCA1 with miR-873-5p: Involvement in smoking-triggered colorectal cancer progression. *Cell Communication and Signaling*, *18*(37), 1–18. https://doi.org/10.1186/s12964-020-0518-0

Gale, E. A. M. (2002). A missing link in the hygiene hypothesis? *Diabetologia*, *45*(4), 588–594.

https://doi.org/10.1007/s00125-002-0801-1

Gao, N., Li, Y., Li, J., Gao, Z., Yang, Z., Li, Y., Liu, H., & Fan, T. (2020). Long Non-Coding RNAs : The Regulatory Mechanisms, Research Strategies, and Future Directions in Cancers. *Frontiers in Oncology*, *10*(December), 1–13. https://doi.org/10.3389/fonc.2020.598817

Gavrilov, A. A., Sultanov, R. I., Magnitov, M. D., Galitsyna, A. A., Dashinimaev, E. B., Lieberman Aiden, E., & Razin, S. V. (2022). RedChIP identifies noncoding RNAs associated with genomic sites occupied by Polycomb and CTCF proteins. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(1). https://doi.org/10.1073/pnas.2116222119/-/DCSupplemental

Ghafouri-Fard, S., Shoorei, H., Anamag, F. T., & Taheri, M. (2020). The Role of Non-Coding RNAs in Controlling Cell Cycle Related Proteins in Cancer Cells. *Frontiers in Oncology*, *10*(November), 1–19. https://doi.org/10.3389/fonc.2020.608975

Gibadullin, R., Randall, C. J., Sidney, J., Sette, A., & Gellman, S. H. (2021). Backbone Modifications of HLA-A2-Restricted Antigens Induce Diverse Binding and T Cell Activation Outcomes. *Journal of the American Chemical Society*, *143*(17), 6470–6481. https://doi.org/10.1021/jacs.1c00016

Gomez-Lopera, N., Pineda-Trujillo, N., & Diaz-Valencia, P. A. (2019). Correlating the global increase in type 1 diabetes incidence across age groups with national economic prosperity: A systematic review. *World Journal of Diabetes*, *10*(12), 560–580. https://doi.org/10.4239/wjd.v10.i12.560

Goubau, D., Deddouche, S., & Reis e Sousa, C. (2013). Cytosolic Sensing of Viruses. *Immunity*, *38*(5), 855–869. https://doi.org/10.1016/j.immuni.2013.05.007

Graf, J., & Kretz, M. (2020). From structure to function: Route to understanding lncRNA mechanism. *BioEssays*, *42*(12). https://doi.org/10.1002/bies.202000027

Grant, S. F. A., Wells, A. D., & Rich, S. S. (2020). Next steps in the identification of gene targets for type 1 diabetes. *Diabetologia*, *63*(11), 2260–2269. https://doi.org/10.1007/s00125-020-05248-8

Gregory, G. A., Robinson, T. I. G., Linklater, S. E., Wang, F., Colagiuri, S., de Beaufort, C., Donaghue, K. C., Magliano, D. J., Maniam, J., Orchard, T. J., Rai, P., & Ogle, G. D. (2022). Global incidence, prevalence, and mortality of type 1 diabetes in 2021 with projection to 2040: a modelling study. *The Lancet. Diabetes & Endocrinology*, *10*(10), 741–760. https://doi.org/10.1016/S2213-8587(22)00218-2

Chapter 8: References

Grulich-Henn, J., & Klose, D. (2018). Understanding childhood diabetes mellitus: new pathophysiological aspects. *Journal of Inherited Metabolic Disease*, *41*(1), 19–27. https://doi.org/10.1007/s10545-017-0120-9

Guenther, M. G., Levine, S. S., Boyer, L. A., Jaenisch, R., & Young, R. A. (2007). A chromatin landmark and transcription initiation at most promoters in human cells. *Cell*, *130*(1), 77–88. https://doi.org/10.1016/j.cell.2007.05.042.A

Guttman, M., & Rinn, J. L. (2012). Modular regulatory principles of large non-coding RNAs. In *Nature* (Vol. 482, Issue 7385, pp. 339–346). https://doi.org/10.1038/nature10887

Hall, J. R., Messenger, Z. J., Tam, H. W., Phillips, S. L., Recio, L., & Smart, R. C. (2015). Long noncoding RNA lincRNA-p21 is the major mediator of UVB-induced and p53-dependent apoptosis in keratinocytes. *Cell Death and Disease*, *6*(3), 1–9. https://doi.org/10.1038/CDDIS.2015.67

Han, P., & Chang, C.-P. (2015). Long non-coding RNA and chromatin remodeling. *RNA Biology*, *12*(10), 1094–1098. https://doi.org/10.1080/15476286.2015.1063770

Han, S., Donelan, W., Wang, H., Reeves, W., & Yang, L.-J. (2013). Novel autoantigens in type 1 diabetes. *American Journal of Translational Research*, *5*(4), 379–392. www.ajtr.org

Hangauer, M. J., Vaughn, I. W., & McManus, M. T. (2013). Pervasive Transcription of the Human Genome Produces Thousands of Previously Unidentified Long Intergenic Noncoding RNAs. *PLoS Genetics*, *9*(6), 1–13. https://doi.org/10.1371/journal.pgen.1003569

Harder, T., Roepke, K., Diller, N., Stechling, Y., Dudenhausen, J. W., & Plagemann, A. (2009). Birth weight, early weight gain, and subsequent risk of type 1 diabetes: Systematic review and meta-analysis. *American Journal of Epidemiology*, *169*(12), 1428–1436. https://doi.org/10.1093/aje/kwp065

Harismendy, O., Notani, D., Song, X., Rahim, N. G., Tanasa, B., Heintzman, N., Ren, B., Fu, X.-D., Topol, E. J., Rosenfeld, M. G., & Frazer, K. A. (2011). 9p21 DNA variants associated with Coronary Artery Disease impair IFNγ signaling response. *Nature*, *470*(7333), 264–268. https://doi.org/10.1038/nature09753.9p21

He, F., Ge, W., Martinowich, K., Becker-Catania, S., Coskum, V., Zhu, W., Wu, H., Castro, D., Guillemot, F., Fan, G., de Vellis, J., & Sun, Y. E. (2005). A positive autoregulatory loop os Jak-STAT signaling controls the onset of astrogliogenesis. *Nature Neuroscience*, *8*(5), 616–625. https://doi.org/10.1038/nn1440.A

Chapter 8: References

He, S., Su, H., Liu, C., Skogerbø, G., He, H., He, D., Zhu, X., Liu, T., Zhao, Y., & Chen, R. (2008). MicroRNA-encoding long non-coding RNAs. *BMC Genomics*, *11*, 1–11. https://doi.org/10.1186/1471-2164-9-236

Heinig, M., Petretto, E., Wallace, C., Bottolo, L., Rotival, M., Lu, H., Li, Y., Sarwar, R., Langley, S. R., Bauerfeind, A., Hummel, O., Lee, Y. A., Paskas, S., Rintisch, C., Saar, K., Cooper, J., Buchan, R., Gray, E. E., Cyster, J. G., … Cook, S. A. (2010). A trans-acting locus regulates an anti-viral expression network and type 1 diabetes risk. *Nature*, *467*(7314), 460–464. https://doi.org/10.1038/nature09386

Hernández, J. C., Montoya, C. J., & Urcuqui-Inchima, S. (2007). Papel de los receptores tipo toll en las infecciones virales: el VIH-1 como modelo. *Biomédica*, *27*(2), 280. https://doi.org/10.7705/biomedica.v27i2.225

Hickey, M. J., Valenzuela, N. M., & Reed, E. F. (2016). Alloantibody generation and effector function following sensitization to human leukocyte antigen. *Frontiers in Immunology*, *7*(February), 1–13. https://doi.org/10.3389/fimmu.2016.00030

Holling, T. M., Schooten, E., & Van Den Elsen, P. J. (2004). Function and regulation of MHC class II molecules in T-lymphocytes: Of mice and men. *Human Immunology*, *65*(4), 282–290. https://doi.org/10.1016/j.humimm.2004.01.005

Holt, R. I. G., DaVries, J. H., Hess-Fischl, A., Hirsch, I. B., Kirkman, M. S., Klupa, T., Ludwig, B., Nørgaard, K., Pettus, J., Renard, E., Skyler, J. S., Snoek, F. J., Weinstock, R. S., & Peters, A. L. (2021). Interpretation of the management of type 1 diabetes in adults: a consensus report by the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD) in 2021. *Diabetologia*, *64*, 2609–2652. https://doi.org/10.1007/s00125-021-05568-3

Honkanen, H., Oikarinen, S., Nurminen, N., Laitinen, O. H., Huhtala, H., Lehtonen, J., Ruokoranta, T., Hankaniemi, M. M., Lecouturier, V., Almond, J. W., Tauriainen, S., Simell, O., Ilonen, J., Veijola, R., Viskari, H., Knip, M., & Hyöty, H. (2017). Detection of enteroviruses in stools precedes islet autoimmunity by several months: possible evidence for slowly operating mechanisms in virus-induced autoimmunity. *Diabetologia*, *60*(3), 424–431. https://doi.org/10.1007/s00125-016-4177-z

Howard, M. B., Basu, S., Sherwin, E., & Cohen, J. S. (2021). Triple threat: New presentation with diabetic ketoacidosis, COVID-19, and cardiac arrhythmias. *American Journal of Emergency Medicine*, *49*(January).

Huang, X., Xu, Y., Lin, Q., Guo, W., Zhao, D., Wang, C., Wang, L., Zhou, H., Jiang, Y., Cui, W., Qiao, X., Li, Y., Ma, G., & Tang, L. (2020). Determination of antiviral action of long non-coding RNA loc107051710 during infectious bursal disease virus infection due to enhancement of interferon production. *Virulence*, *11*(1), 68–79. https://doi.org/10.1080/21505594.2019.1707957

Huang, X., Yuan, J., Goddard, A., Foulis, A., James, R. F. L., Lernmark, A., Pujol-Borrell, R., Rabinovitch, A., Somoza, N., & Stewart, T. A. (1995). Interferon Expression in the Pancreases of Patients With Type I Diabetes. *Diabetes*, *44*, 658–664.

Hung, T., & Chang, H. Y. (2010). Long noncoding RNA in genome regulation. *RNA Biology*, *7*(5), 582–585.

Hung, T., Wang, Y., Lin, M. F., Koegel, A. K., Koegel, A., Grant, G. D., Horlings, H. M., Shan, N., Umbricht, C., Wang, P., Wang, Y., Kong, B., Langerod, A., Børresen-Dale, A.-L., Kim, S. K., Van de Vijver, M., Sukumar, S., Whitfield, M. L., Kellis, M., … Chang, H. Y. (2011). Extensive and coordinated transcription of noncoding RNAs within cell cycle promoters. *Nature Genetics*, *43*(7), 621–629. https://doi.org/10.1038/ng.848.Extensive

Hunt, K. A., Zhernakova, A., Turner, G., Heap, G. A. R., Franke, L., Bruinenberg, M., Romanos, J., Dinesen, L. C., Ryan, A. W., Panesar, D., Gwilliam, R., Takeuchi, F., Mclaren, W. M., Holmes, G. K. T., Howdle, P. D., Walters, J. R. F., Sanders, D. S., Playford, R. J., Trynka, G., … van Heel, D. A. (2009). Novel celiac disease genetic determinants related to the immune response. *Nature Genetics*, *40*(4), 395–402. https://doi.org/10.1038/ng.102.

Insel, R. A., Dunne, J. L., Atkinson, M. A., Chiang, J. L., Dabelea, D., Gottlieb, P. A., Greenbaum, C. J., Herold, K. C., Krischer, J. P., Lernmark, A., Ratner, R. E., Rewers, M. J., Schatz, D. A., Skyler, J. S., Sosenko, J. M., & Ziegler, A. G. (2015). Staging presymptomatic type 1 diabetes: A scientific statement of jdrf, the endocrine society, and the American diabetes association. *Diabetes Care*, *38*(10), 1964–1974. https://doi.org/10.2337/dc15-1419

International Diabetes Federation. (2021). IDF Diabetes Atlas 2021. In *Diabetes Research and Clinical Practice*. https://doi.org/10.1016/j.diabres.2013.10.013

Isomäki, H., Nissilä, M., Koota, K., & Martio, J. (1974). HL-A 27 and rheumatoid arthritis. *The Lancet*, 1212–1213.

Ivashkiv., L. B., & Donlin, L. T. (2014). Regulation of type I interferon responses. *Nature*, *14*(1), 36–49. https://doi.org/10.1038/nri3581

Ize-Ludlow, D., & Sperling, M. A. (2005). The classification of diabetes mellitus: A conceptual

framework. *Pediatric Clinics of North America*, *52*(6), 1533–1552. https://doi.org/10.1016/j.pcl.2005.07.001

Ji, P., Diederichs, S., Wang, W., Böing, S., Metzger, R., Schneider, P. M., Tidow, N., Brandt, B., Buerger, H., Bulk, E., Thomas, M., Berdel, W. E., Serve, H., & Müller-Tidow, C. (2003). MALAT-1, a novel noncoding RNA, and thymosin β4 predict metastasis and survival in early-stage non-small cell lung cancer. *Oncogene*, *22*(39), 8031–8041. https://doi.org/10.1038/sj.onc.1206928

Ji, X., Meng, W., Liu, Z., & Mu, X. (2022). Emerging Roles of lncRNAs Regulating RNA-Mediated Type-I Interferon Signaling Pathway. *Frontiers in Immunology*, *13*(February), 1–14. https://doi.org/10.3389/fimmu.2022.811122

Jiang, C., Li, Y., Zhao, Z., Lu, J., Chen, H., Ding, N., Wang, G., Xu, J., & Li, X. (2016). Identifying and functionally characterizing tissue-specific and ubiquitously expressed human lncRNAs. *Oncotarget*, *7*(6), 7120–7133. https://doi.org/10.18632/oncotarget.6859

Jin, F., Wang, N., Zhu, Y., You, L., Wang, L., De, W., & Tang, W. (2017). Downregulation of Long Noncoding RNA Gas5 Affects Cell Cycle and Insulin Secretion in Mouse Pancreatic β Cells. *Cellular Physiology and Biochemistry*, *43*(5), 2062–2073. https://doi.org/10.1159/000484191

Kambara, H., Niazi, F., Kostadinova, L., Moonka, D. K., Siegel, C. T., Post, A. B., Carnero, E., Barriocanal, M., Fortes, P., Anthony, D. D., & Valadkhan, S. (2014). Negative regulation of the interferon response by an interferon-induced long non-coding RNA. *Nucleic Acids Research*, *42*(16), 10668–10681. https://doi.org/10.1093/nar/gku713

Kanakatti Shankar, R., Pihoker, C., Dolan, L. M., Standiford, D., Badaru, A., Dabelea, D., Rodriguez, B., Black, M. H., Imperatore, G., Hattersley, A., Ellard, S., & Gilliam, L. K. (2013). Permanent neonatal diabetes mellitus: Prevalence and genetic diagnosis in the SEARCH for Diabetes in Youth Study. *Pediatric Diabetes*, *14*(3), 174–180. https://doi.org/10.1111/pedi.12003

Kaur, S., Mirza, A. H., Brorsson, C. A., Fløyel, T., Størling, J., Mortensen, H. B., & Pociot, F. (2016). The genetic and regulatory architecture of ERBB3-type 1 diabetes susceptibility locus. *Molecular and Cellular Endocrinology*, *419*, 83–91. https://doi.org/10.1016/j.mce.2015.10.002

Kazimierczyk, M., Kasprowicz, M. K., Kasprzyk, M. E., & Wrzesinski, J. (2020). Human long noncoding RNA interactome: Detection, characterization and function. *International Journal of Molecular Sciences*, *21*(3), 1–21. https://doi.org/10.3390/ijms21031027

Kern, B. Č., Podkrajšek, K. T., Kovač, J., Šket, R., Bizjan, B. J., Tesovnik, T., Debeljak, M., Battelino, T., & Bratina, N. (2022). The Role of Epigenetic Modifications in Late Complications in Type 1

Diabetes. *Genes*, *13*(705). https://doi.org/10.3390/genes13040705

Kim, K. W., Horton, J. L., Pang, C. N. I., Jain, K., Leung, P., Isaacs, S. R., Bull, R. A., Luciani, F., Wilkins, M. R., Catteau, J., Lipkin, W. I., Rawlinson, W. D., Briese, T., & Craig, M. E. (2019). Higher abundance of enterovirus A species in the gut of children with islet autoimmunity. *Scientific Reports*, *9*(1749), 1–8. https://doi.org/10.1038/s41598-018-38368-8

Kim, S. M., Mayassi, T., & Jabri, B. (2015). Innate immunity: Actuating the gears of celiac disease pathogenesis. *Best Practice and Research: Clinical Gastroenterology*, *29*(3), 425–435. https://doi.org/10.1016/j.bpg.2015.05.001

Kino, T., Hur, D. E., Ichijo, T., Nader, N., & Chrousos, G. P. (2010). Noncoding RNA Gas5 is a growth arrest- and starvation-associated repressor of the glucocorticoid receptor. *Science Signalling*, *3*(107), 1–16. https://doi.org/10.1126/scisignal.2000568.Noncoding

Kitagawa, M., Kitagawa, K., Kotake, Y., Niida, H., & Ohhata, T. (2013). Cell cycle regulation by long non-coding RNAs. *Cellular and Molecular Life Sciences*, *70*(24), 4785–4794. https://doi.org/10.1007/s00018-013-1423-0

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. *Central European Journal of Immunology*, *45*(4), 439–453. https://doi.org/10.5114/CEJI.2020.103386

Kotake, Y., Nakagawa, T., Kitagawa, K., Suzuki, S., Liu, N., Kitagawa, M., & Xiong, Y. (2011). Long non-coding RNA ANRIL is required for the PRC2 recruitment to and silencing of p15INK4B tumor supressor gene. *Oncogene*, *30*(16), 1956–1962. https://doi.org/10.1038/onc.2010.568.Long

Krischer, J. P., Lynch, K. F., Lernmark, A., Hagopian, W. A., Rewers, M. J., She, J. X., Toppari, J., Ziegler, A. G., & Akolkar, B. (2017). Genetic and environmental interactions modify the risk of diabetes-related autoimmunity by 6 years of age: The teddy study. *Diabetes Care*, *40*(9), 1194–1202. https://doi.org/10.2337/dc17-0238

Krogvold, L., Edwin, B., Buanes, T., Frisk, G., Skog, O., Anagandula, M., Korsgren, O., Undlien, D., Eike, M. C., Richardson, S. J., Leete, P., Morgan, N. G., Oikarinen, S., Oikarinen, M., Laiho, J. E., Hyöty, H., Ludvigsson, J., Hanssen, K. F., & Dahl-Jørgensen, K. (2015). Detection of a low-grade enteroviral infection in the islets of langerhans of living patients newly diagnosed with type 1 diabetes. *Diabetes*, *64*(5), 1682–1687. https://doi.org/10.2337/db14-1370

Krogvold, L., Genoni, A., Puggioni, A., Campani, D., Richardson, S. J., Flaxman, C. S., Edwin, B., Buanes, T., Dahl-Jørgensen, K., & Toniolo, A. (2022). Live enteroviruses, but not other viruses, detected

in human pancreas at the onset of type 1 diabetes in the DiViD study. *Diabetologia*, *65*, 2108–2120. https://doi.org/10.1007/s00125-022-05779-2

Krogvold, L., Leete, P., Mynarek, I. M., Russell, M. A., Gerling, I. C., Lenchik, N. I., Mathews, C., Richardson, S. J., Morgan, N. G., & Dahl-Jørgensen, K. (2022). Detection of Antiviral Tissue Responses and Increased Cell Stress in the Pancreatic Islets of Newly Diagnosed Type 1 Diabetes Patients: Results From the DiViD Study. *Frontiers in Endocrinology*, *13*(881997). https://doi.org/10.3389/fendo.2022.881997

Kuleshov, M. V., Jones, M. R., Rouillard, A. D., Fernandez, N. F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S. L., Jagodnik, K. M., Lachmann, A., McDermott, M. G., Monteiro, C. D., Gundersen, G. W., & Maayan, A. (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Research*, *44*(1), W90–W97. https://doi.org/10.1093/nar/gkw377

Kuo, C. C., Hänzelmann, S., Sentürk Cetin, N., Frank, S., Zajzon, B., Derks, J. P., Akhade, V. S., Ahuja, G., Kanduri, C., Grummt, I., Kurian, L., & Costa, I. G. (2019). Detection of RNA-DNA binding sites in long noncoding RNAs. *Nucleic Acids Research*, *47*(6), 1–12. https://doi.org/10.1093/nar/gkz037

Kutty, R. K., Samuel, W., Duncan, T., Postnikova, O., Jaworski, C., Nagineni, C. N., & Redmond, T. M. (2017). Proinflammatory cytokine interferon-γ increases the expression of BANCR, a long non-coding RNA , in retinal pigment epithelial cells. *Cytokine*, *104*(October), 147–150. https://doi.org/10.1016/j.cyto.2017.10.009

Laitinen, O. H., Honkanen, H., Pakkanen, O., Oikarinen, S., Hankaniemi, M. M., Huhtala, H., Ruokoranta, T., Lecouturier, V., André, P., Harju, R., Virtanen, S. M., Lehtonen, J., Almond, J. W., Simell, T., Simell, O., Ilonen, J., Veijola, R., Knip, M., & Hyöty, H. (2014). Coxsackievirus B1 is associated with induction of β-cells autoimmunity that portends type 1 diabetes. *Diabetes*, *63*(2), 446–455. https://doi.org/10.2337/db13-0619

Lam, C. J., Jacobson, D. R., Rankin, M. M., Cox, A. R., & Kushner, J. A. (2017). β Cells persist in T1D pancreata without evidence of ongoing β-Cell turnover or neogenesis. *Journal of Clinical Endocrinology and Metabolism*, *102*(8), 2647–2659. https://doi.org/10.1210/jc.2016-3806

Lanzillotti, C., De Mattei, M., Mazziotta, C., Taraballi, F., Rotonco, J. C., Tognon, M., & Martini, F. (2021). Long Non-coding RNAs and MicroRNAs Interplay in Osteogenic Differentiation of Mesenchymal Stem Cells. *Frontiers in Cell and Developmental Biology*, *9*(April), 1–14. https://doi.org/10.3389/fcell.2021.646032

Chapter 8: References

Lee, J. T. (2009). Lessons from X-chromosome inactivation: Long ncRNA as guides and tethers to the epigenome. *Genes and Development*, *23*(16), 1831–1842. https://doi.org/10.1101/gad.1811209

Lelli, A., Nolan, K., Santambrogio, S., Goncalves, A. F., Schonenberger, M., Guinot, A., Frew, I. J., Marti, H. H., Hoogewijs, D., & Wenger, R. H. (2015). Induction of long noncoding RNA MALAT1 in hypoxic mice. *Hypoxia*, *3*, 45–52. https://doi.org/10.2147/hp.s90555

Lemos, N. E., Dieter, C., Dorfman, L. E., Assmann, T. S., Duarte, G. C. K., Canani, L. H., Bauer, A. C., & Crispim, D. (2018). The rs2292239 polymorphism in ERBB3 gene is associated with risk for type 1 diabetes mellitus in a Brazilian population. *Gene*, *644*(November), 122–128. https://doi.org/10.1016/j.gene.2017.11.009

Lin, W., Zhou, Q., Wang, C. Q., Zhu, L., Bi, C., Zhang, S., Wang, X., & Jin, H. (2020). LncRNAs regulate metabolism in cancer. *International Journal of Biological Sciences*, *16*(7), 1194–1206. https://doi.org/10.7150/ijbs.40769

Lin, X., Xu, Y., Pan, X., Xu, J., Ding, Y., Sun, X., Song, X., Ren, Y., & Shan, P. F. (2020). Global, regional, and national burden and trend of diabetes in 195 countries and territories: an analysis from 1990 to 2025. *Scientific Reports*, *10*(14790). https://doi.org/10.1038/s41598-020-71908-9

Liu, B., Sun, L., Liu, Q., Gong, C., Yao, Y., Lv, X., Lin, L., Yao, H., Su, F., Li, D., Zeng, M., & Song, E. (2015). A Cytoplasmic NF- k B Interacting Long Noncoding RNA Blocks I k B Phosphorylation and Suppresses Breast Cancer Metastasis. *Cancer Cell*, *27*(3), 370–381. https://doi.org/10.1016/j.ccell.2015.02.004

Liu, S. J., Nowakowski, T. J., Pollen, A. A., Lui, J. H., Horlbeck, M. A., Attenello, F. J., He, D., Weissman, J. S., Kriegstein, A. R., Diaz, A. A., & Lim, D. A. (2016). Single-cell analysis of long non-coding RNAs in the developing human neocortex. *Genome Biology*, *17*(67), 1–17. https://doi.org/10.1186/s13059-016-0932-1

Long, Y., Wang, X., Youmans, D. T., & Cech, T. R. (2017). How do lncRNAs regulate transcription ? *Science Advances*, *3*(September), 1–13.

Lu, W., Cao, F., Wang, S., Sheng, X., & Ma, J. (2019). LncRNAs: The Regulator of Glucose and Lipid Metabolism in Tumor Cells. *Frontiers in Oncology*, *9*(November), 1–11. https://doi.org/10.3389/fonc.2019.01099

Lu, W., Xu, Y., Xu, J., Wang, Z., & Ye, G. (2018). Identification of differential expressed lncRNAs in human thyroid cancer by a genome- - wide analyses. *Cancer Medicine*, *7*(April), 3935–3944. https://doi.org/10.1002/cam4.1627

Lundberg, M., Krogvold, L., Kuric, E., Dahl-Jørgensen, K., & Skog, O. (2016). Expression of Interferon-Stimulated Genes in Insulitic Pancreatic Islets of Patients Recently Diagnosed With Type 1 Diabetes. *Diabetes*, *65*, 3104–3110.

Lytrivi, M., Senée, V., Salpea, P., Fantuzzi, F., Philippi, A., Abdulkarim, B., Sawatani, T., Marín-Cañas, S., Pachera, N., Degavre, A., Singh, P., Derbois, C., Lechner, D., Ladrière, L., Igoillo-Esteve, M., Cosentino, C., Marselli, L., Deleuze, J. F., Marchetti, P., … Cnop, M. (2021). DNAJC3 deficiency induces β-cell mitochondrial apoptosis and causes syndromic young-onset diabetes. *European Journal of Endocrinology*, *184*(3), 455–468. https://doi.org/10.1530/EJE-20-0636

Marchese, F. P., Raimondi, I., & Huarte, M. (2017). The multidimensional mechanisms of long noncoding RNA function. *Genome Biology*, *18*(206), 1–13. https://doi.org/10.1186/s13059-017-1348-2

Marchetti, P., Bugliani, M., Lupi, R., Marselli, L., Masini, M., Boggi, U., Filipponi, F., Weir, G. C., Eizirik, D. L., & Cnop, M. (2007). The endoplasmic reticulum in pancreatic beta cells of type 2 diabetes patients. *Diabetologia*, *50*(12), 2486–2494. https://doi.org/10.1007/s00125-007-0816-8

Marroquí, L., Santin, I., Dos Santos, R. S., Marselli, L., Marchetti, P., & Eizirik, D. L. (2014). BACH2, a candidate risk gene for type 1 diabetes, regulates apoptosis in pancreatic β-cells via JNK1 modulation and crosstalk with the candidate gene PTPN2. *Diabetes*, *63*(7), 2516–2527. https://doi.org/10.2337/db13-1443

Marroqui, L., Sousa Dos Santos, R., Fløyel, T., Grieco, F. A., Santin, I., Op De Beeck, A., Marselli, L., Marchetti, P., Pociot, F., & Eizirik, D. L. (2015). TYK2, a candidate gene for type 1 diabetes, modulates apoptosis and the innate immune response in human pancreatic β-cells. *Diabetes*, *64*(11), 3808–3817. https://doi.org/10.2337/db15-0362

Marshall, J. S., Warrington, R., Watson, W., & Kim, H. L. (2018). An introduction to immunology and immunopathology. *Allergy, Asthma and Clinical Immunology*, *14*(s2), 1–10. https://doi.org/10.1186/s13223-018-0278-1

Meier, J. J. (2008). Beta cell mass in diabetes: A realistic therapeutic target? *Diabetologia*, *51*(5), 703–713. https://doi.org/10.1007/s00125-008-0936-9

Melloul, D., Marshak, S., & Cerasi, E. (2002). Regulation of insulin gene transcription. *Diabetologia*, *45*, 309–326.

Mercer, T. R., Dinger, M. E., & Mattick, J. S. (2009). Long non-coding RNAs: insights into function. *Nature Reviews Genetics*, *10*(March), 155–159.

Mercer, T. R., Gerhardt, D. J., Dinger, M. E., Crawford, J., Trapnell, C., Jeddeloh, J. A., Mattick, J. S., & Rinn, J. L. (2012). Targeted RNA sequencing reveals the deep complexity of the human transcriptome. *Nature Biotechnology*, *30*(1), 99–104. https://doi.org/10.1038/nbt.2024

Mirza, A. H., Kaur, S., Brorsson, C. A., & Pociot, F. (2014). Effects of GWAS-associated genetic variants on lncRNAs within IBD and T1D candidate loci. *PLoS ONE*, *9*(8). https://doi.org/10.1371/journal.pone.0105723

Misra, S., & Owen, K. R. (2018). Genetics of Monogenic Diabetes: Present Clinical Challenges. *Current Diabetes Reports*, *18*(141). https://doi.org/10.1007/s11892-018-1111-4

Miyata, K., Imai, Y., Hori, S., Nishio, M., Loo, T. M., Okada, R., Yang, L., Nakadai, T., Maruyama, R., Fujii, R., Ueda, K., Jiang, L., Zheng, H., Toyokuni, S., Sakata, T., Shirahige, K., Kojima, R., Nakayama, M., Oshima, M., … Takahashi, A. (2021). Pericentromeric noncoding RNA changes DNA binding of CTCF and inflammatory gene expression in senescence and cancer. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(35). https://doi.org/10.1073/pnas.2025647118/-/DCSupplemental

Moran, I., Akerman, I., Van de Bunt, M., Xie, R., Benazra, M., Nammo, T., Arnes, L., Nakic, N., Garcia-Hurtado, J., Rodriguez-Seguí, S., Pasquali, L., Sauty-Colace, C., Beucher, A., Scharfmann, R., Van Arensbergen, J., Johnson, P. R., Berry, A., Lee, C., Harkins, T., … Ferrer, J. (2012). Human β Cell Transcriptome Analysis Uncovers lncRNAs That Are Tissue-Specific , Dynamically Regulated , and Abnormally Expressed in Type 2 Diabetes. *Cell Metabolism*, *16*, 435–448. https://doi.org/10.1016/j.cmet.2012.08.010

Morgan, N. G., & Richardson, S. J. (2014). Enteroviruses as causative agents in type 1 diabetes: Loose ends or lost cause? *Trends in Endocrinology and Metabolism*, *25*(12), 611–619. https://doi.org/10.1016/j.tem.2014.08.002

Morran, M. P., Vonberg, A., Khadra, A., & Pietropaolo, M. (2015). Immunogenetics of type 1 diabetes mellitus. *Molecular Aspects of Medicine*, *42*, 42–60. https://doi.org/10.1016/j.mam.2014.12.004

Mostafavi, S., Yoshida, H., Moodley, D., Leboité, H., Rothamel, K., Raj, T., Ye, C. J., Chevrier, N., Zhang, S. Y., Feng, T., Lee, M., Casanova, J. L., Clark, J. D., Hegen, M., Telliez, J. B., Hacohen, N., De Jager, P. L., Regev, A., Mathis, D., & Benoist, C. (2016). Parsing the Interferon Transcriptional Network and Its Disease Associations. *Cell*, *164*(3), 564–578. https://doi.org/10.1016/j.cell.2015.12.032

Chapter 8: References

Motterle, A., Gattesco, S., Peyot, M., Esguerra, J. L. S., Gomez-ruiz, A., Laybutt, D. R., Gilon, P., Burdet, F., Ibberson, M., Eliasson, L., Prentki, M., & Regazzi, R. (2017). Identification of islet-enriched long non-coding RNAs contributing to b -cell failure in type 2 diabetes. *Molecular Metabolism*, *August*, 1–12. https://doi.org/10.1016/j.molmet.2017.08.005

Mulder, D. J. (1974). HL-A antigens and coeliac disease. *The Lancet*, *2*(7878), 727.

Muntoni, S., Cocco, P., Muntoni, S., & Aru, G. (2006). Nitrate in community water supplies and risk of childhood type 1 diabetes in Sardinia, Italy. *European Journal of Epidemiology*, *21*(3), 245–247. https://doi.org/10.1007/s10654-006-0014-x

Murphy, R., Ellard, S., & Hattersley, A. T. (2008). Clinical implications of a molecular genetic classification of monogenic β-cell diabetes. *Nature Clinical Practice: Endocrinology and Metabolism*, *4*(4), 200–213. https://doi.org/10.1038/ncpendmet0778

Nair, G. G., Liu, J. S., Russ, H. A., Tran, S., Saxton, M. S., Chen, R., Juang, C., Li, M. lan, Nguyen, V. Q., Giacometti, S., Puri, S., Xing, Y., Szot, G. L., Oberholzer, J., Bhushan, A., & Hebrok, M. (2019). Recapitulating endocrine cell clustering in culture promotes maturation of human stem-cell-derived β cells. *Nature Cell Biology*, *21*(2), 263–274. https://doi.org/10.1038/s41556-018-0271-4.

Nekoua, M. P., Alidjinou, E. K., & Hober, D. (2022). Persistent coxsackievirus B infection and pathogenesis of type 1 diabetes mellitus. *Nature Reviews Endocrinology*, *18*(8), 503–516. https://doi.org/10.1038/s41574-022-00688-1

Nguyen, C., Varney, M. D., Harrison, L. C., & Morahan, G. (2013). Definition of high-risk type 1 diabetes HLA-DR and HLA-DQ types using only three single nucleotide polymorphisms. *Diabetes*, *62*(6), 2135–2140. https://doi.org/10.2337/db12-1398

Nica, A. C., Ongen, H., Irminger, J., Bosco, D., Berney, T., Antonarakis, S. E., Halban, P. A., & Dermitzakis, E. T. (2013). Cell-type , allelic , and genetic signatures in the human pancreatic beta cell transcriptome. *Genome Research*, *23*, 1554–1562. https://doi.org/10.1101/gr.150706.112.1554

Nigi, L., Brusco, N., Grieco, G. E., Fignani, D., Licata, G., Formichi, C., Aiello, E., Marselli, L., Marchetti, P., Krogvold, L., Jorgensen, K. D., Sebastiani, G., & Dotta, F. (2022). Increased Expression of Viral Sensor MDA5 in Pancreatic Islets and in Hormone-Negative Endocrine Cells in Recent Onset Type 1 Diabetic Donors. *Frontiers in Immunology*, *13*(833141), 1–12. https://doi.org/10.3389/fimmu.2022.833141

Nishitsuji, H., Ujino, S., Yoshio, S., Sugiyama, M., Mizokami, M., Kanto, T., & Shimotohno, K. (2016). Long noncoding RNA #32 contributes to antiviral responses by controlling interferon-stimulated gene expression. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(37), 10388–10393. https://doi.org/10.1073/pnas.1525022113

Noble, J. A., & Erlich, H. A. (2012). Genetics of type 1 diabetes. *Cold Spring Harbor Perspectives in Medicine*, *2*(1), 1–15. https://doi.org/10.1101/cshperspect.a007732

Nogueira, T. C., Paula, F. M., Villate, O., Colli, M. L., Moura, R. F., Cunha, D. A., Marselli, L., Marchetti, P., Cnop, M., Julier, C., & Eizirik, D. L. (2013). GLIS3, a Susceptibility Gene for Type 1 and Type 2 Diabetes, Modulates Pancreatic Beta Cell Apoptosis via Regulation of a Splice Variant of the BH3-Only Protein Bim. *PLoS Genetics*, *9*(5), 1–17. https://doi.org/10.1371/journal.pgen.1003532

Noh, J. H., Kim, K. M., Mcclusky, W., Abdelmohsen, K., & Gorospe, M. (2018). Cytoplasmic functions of lncRNAs. *Wiley Interdisciplinari Review RNA*, *9*(3), 1–24. https://doi.org/10.1002/wrna.1471.Cytoplasmic

Norris, J. M., Yin, X., Lamb, M. M., Barriga, K., Seifert, J., Hoffman, M., Orton, H. D., Barón, A. E., Clare-Salzler, M., Chase, H. P., Szabo, N. J., Erlich, H., Eisenbarth, G. S., & Rewers, M. (2007). Omega-3 polyunsaturated fatty acid intake and islet autoimmunity in children at increased risk for type 1 diabetes. *Jama*, *298*(12), 1420–1428. https://doi.org/10.1001/jama.298.12.1420

Oikarinen, M., Tauriainen, S., Honkanen, T., Vuori, K., Karhunen, P., Vasama-Nolvi, C., Oikarinen, S., Verbeke, C., Blair, G. E., Rantala, I., Ilonen, J., Simell, O., Knip, M., & Hyöty, H. (2008). Analysis of pancreas tissue in a child positive for islet cell antibodies. *Diabetologia*, *51*(10), 1796–1802. https://doi.org/10.1007/s00125-008-1107-8

Oikarinen, S., Krogvold, L., Edwin, B., Buanes, T., Korsgren, O., Laiho, J. E., Oikarinen, M., Ludvigsson, J., Skog, O., Anagandula, M., Frisk, G., Hyöty, H., & Dahl-Jørgensen, K. (2021). Characterisation of enterovirus RNA detected in the pancreas and other specimens of live patients with newly diagnosed type 1 diabetes in the DiViD study. *Diabetologia*, *64*(11), 2491–2501. https://doi.org/10.1007/s00125-021-05525-0

Oikarinen, S., Tauriainen, S., Hober, D., Lucas, B., Vazeou, A., Sioofy-Khojine, A., Bozas, E., Muir, P., Honkanen, H., Ilonen, J., Knip, M., Keskinen, P., Saha, M. T., Huhtala, H., Stanway, G., Bartsocas, C., Ludvigsson, J., Taylor, K., & Hyöty, H. (2014). Virus antibody survey in different european populations indicates risk association between coxsackievirus B1 and type 1 diabetes. *Diabetes*, *63*(2), 655–662. https://doi.org/10.2337/db13-0620

Chapter 8: References

Op de beeck, A., & Eizirik, D. L. (2016). Viral infections in type 1 diabetes mellitus - why the β cells? *Nature Reviews Endocrinology*, *12*(5), 263–273. https://doi.org/10.1038/nrendo.2016.30.

Paquette, J., Giannoukakis, N., Polychronakos, C., Vafiadis, P., & Deal, C. (1998). The INS 5' variable number of tandem repeats is associated with IGF2 expression in humans. *Journal of Biological Chemistry*, *273*(23), 14158–14164. https://doi.org/10.1074/jbc.273.23.14158

Pasquali, L., Gaulton, K. J., Rodríguez-Seguí, S. A., Mularoni, L., Miguel-Escalada, I., Akerman, I., Tena, J. J., Morán, I., Gómez-Marín, C., Van De Bunt, M., Ponsa-Cobas, J., Castro, N., Nammo, T., Cebola, I., García-Hurtado, J., Maestro, M. A., Pattou, F., Piemonti, L., Berney, T., … Ferrer, J. (2014). Pancreatic islet enhancer clusters enriched in type 2 diabetes risk-associated variants. *Nature Genetics*, *46*(2), 136–143. https://doi.org/10.1038/ng.2870

Peng, L., Liu, F., Yang, J., Liu, X., Meng, Y., Deng, X., Peng, C., Tian, G., & Zhou, L. (2020). Probing lncRNA–Protein Interactions: Data Repositories, Models, and Algorithms. *Frontiers in Genetics*, *10*(January), 1–26. https://doi.org/10.3389/fgene.2019.01346

Perry, M. M., Tsitsiou, E., Austin, P. J., Lindsay, M. A., Gibeon, D. S., Adcock, I. M., & Chung, K. F. (2014). Role of non-coding RNAs in maintaining primary airway smooth muscle cells. *Respiratory Research*, *14*(58), 1–12.

Perry, R. B., Hezroni, H., Goldrich, M. J., & Ulitsky, I. (2018). Regulation of neuroregeneration by long noncoding RNAs. *Molecular Cell*, *72*(3), 553–567. https://doi.org/10.1016/j.molcel.2018.09.021.Regulation

Phillips, J. E., & Corces, V. G. (2009). CTCF: Master Weaver of the Genome. In *Cell* (Vol. 137, Issue 7, pp. 1194–1211). Elsevier B.V. https://doi.org/10.1016/j.cell.2009.06.001

Platanias, L. C. (2005). Mechanisms of type-I- and type-II-interferon-mediated signalling. *Nature Reviews Immunology*, *5*(5), 375–386. https://doi.org/10.1038/nri1604

Platanitis, E., Demiroz, D., Schneller, A., Fischer, K., Capelle, C., Hartl, M., Gossenreiter, T., Müller, M., Novatchkova, M., & Decker, T. (2019). A molecular switch from STAT2-IRF9 to ISGF3 underlies interferon-induced gene transcription. *Nature Communications*, *10*(2921), 1–17. https://doi.org/10.1038/s41467-019-10970-y

Pociot, F., Akolkar, B., Concannon, P., Erlich, H. A., Julier, C., Morahan, G., Nierras, C. R., Todd, J. A., Rich, S. S., & Nerup, J. (2010). Genetics of type 1 diabetes: What's next? *Diabetes*, *59*(7), 1561–1571. https://doi.org/10.2337/db10-0076

Chapter 8: References

Porter, J. R., & Barrett, T. G. (2005). Monogenic syndromes of abnormal glucose homeostasis: Clinical review and relevance to the understanding of the pathology of insulin resistance and β cell failure. *Journal of Medical Genetics*, *42*(12), 893–902. https://doi.org/10.1136/jmg.2005.030791

Primavera, M., Giannini, C., & Chiarelli, F. (2020). Prediction and Prevention of Type 1 Diabetes. *Frontiers in Endocrinology*, *11*(June), 1–9. https://doi.org/10.3389/fendo.2020.00248

Pugliese, A. (2017). Autoreactive T cells in type 1 diabetes. *The Journal of Clinical Investigations*, *127*(8), 2881–2891.

Pujol-Borrell, R., Todd, I., Doshi, M., Gray, D., Feldmann, M., & Bottazzo, G. F. (1986). Differential expression and regulation of MHC products in the endocrine and exocrine cells of the human pancreas. *Clinical and Experimental Immunology*, *65*, 128–139.

Quelle, F. W., Thierfelder, W., Witthuhn, B. A., Tang, B., Cohen, S., & Ihle, J. N. (1995). Phosphorylation and activation of the DNA binding activity of purified stat1 by the Janus protein-tyrosine kinases and the epidermal growth factor receptor. *Journal of Biological Chemistry*, *270*(35), 20775–20780. https://doi.org/10.1074/jbc.270.35.20775

Rabi, D. M., Edwards, A. L., Southern, D. A., Svenson, L. W., Sargious, P. M., Norton, P., Larsen, E. T., & Ghali, W. A. (2006). Association of socio-economic status with diabetes prevalence and utilization of diabetes care services. *BMC Health Services Research*, *6*(124). https://doi.org/10.1186/1472-6963-6-124

Ram, R., Mehta, M., Nguyen, Q. T., Larma, I., Boehm, B. O., Pociot, F., Concannon, P., & Morahan, G. (2016). Systematic Evaluation of Genes and Genetic Variants Associated with Type 1 Diabetes Susceptibility. *The Journal of Immunology*, *196*(7), 3043–3053. https://doi.org/10.4049/jimmunol.1502056

Ransohoff, J. D., Wei, Y., & Khavari, P. A. (2018). The functions and unique features of long intergenic non-coding RNA. *Nature Reviews Molecular Cell Biology*, *19*(3), 143–157.

Rashid, F., Shah, A., & Shan, G. (2016). Long Non-coding RNAs in the Cytoplasm. *Genomics, Proteomics & Bioinformatics*, *14*(2), 73–80. https://doi.org/10.1016/j.gpb.2016.03.005

Redis, R. S., Vela, L. E., Lu, W., Ferreira de Oliveira, J., Ivan, C., Rodriguez-Aguayo, C., Adamoski, D., Pasculli, B., Taguchi, A., Chen, Y., Fernandez, A. F., Valledor, L., Van Roosbroeck, K., Chang, S., Shah, M., Kinnebrew, G., Han, L., Atlasi, Y., Cheung, L. H., … Calin, G. A. (2016). Allele-Specific Reprogramming of Cancer Metabolism by the Long Non-coding RNA CCAT2. *Molecular Cell*, *61*(4), 520–534. https://doi.org/10.1016/j.molcel.2016.01.015

Chapter 8: References

Redondo, M. J., Babu, S., Zeidler, A., Orban, T., Yu, L., Greenbaum, C., Palmer, J. P., Cuthbertson, D., Eisenbarth, G. S., Krischer, J. P., & Schatz, D. (2006). Specific human leukocyte antigen DQ influence on expression of antiislet autoantibodies and progression to type 1 diabetes. *Journal of Clinical Endocrinology and Metabolism*, *91*(5), 1705–1713. https://doi.org/10.1210/jc.2005-1695

Redondo, M. J., Steck, A. K., & Pugliese, A. (2018). Genetics of type 1 diabetes. *Pediatric Diabetes*, *19*(3), 346–353. https://doi.org/doi:10.1111/pedi.12597

Rewers, M., & Ludvigsson, J. (2016). Environmental risk factors for type 1 diabetes. *Lancet*, *387*(10035), 2340–2348. https://doi.org/10.1016/S0140-6736(16)30507-4.Environmental

Ricaño-Ponce, I., & Wijmenga, C. (2013). Mapping of immune-mediated disease genes. In *Annual Review of Genomics and Human Genetics* (Vol. 14, pp. 325–353). https://doi.org/10.1146/annurev-genom-091212-153450

Richardson, S. J., Leete, P., Bone, A. J., Foulis, A. K., & Morgan, N. G. (2013). Expression of the enteroviral capsid protein VP1 in the islet cells of patients with type 1 diabetes is associated with induction of protein kinase R and downregulation of Mcl-1. *Diabetologia*, *56*(1), 185–193. https://doi.org/10.1007/s00125-012-2745-4

Richardson, S. J., Rodriguez-Calvo, T., Gerling, I. C., Mathews, C. E., Kaddis, J. S., Russell, M. A., Zeissler, M., Leete, P., Krogvold, L., Dahl-Jørgensen, K., von Herrath, M., Pugliese, A., Atkinson, M. A., & Morgan, N. G. (2016). Islet cell hyperexpression of HLA class I antigens: a defining feature in type 1 diabetes. *Diabetologia*, *59*(11), 2448–2458. https://doi.org/10.1007/s00125-016-4067-4

Richardson, S. J., Willcox, A., Bone, A. J., Foulis, A. K., & Morgan, N. G. (2009). The prevalence of enteroviral capsid protein vp1 immunostaining in pancreatic islets in human type 1 diabetes. *Diabetologia*, *52*(6), 1143–1151. https://doi.org/10.1007/s00125-009-1276-0

Riddle, M. C., Philipson, L. H., Rich, S. S., Carlsson, A., Franks, P. W., Greeley, S. A. W., Nolan, J. J., Pearson, E. R., Zeitler, P. S., & Hattersley, A. T. (2020). Monogenic diabetes: From genetic insights to population-based precision in care. reflections from a diabetes care editors' expert forum. *Diabetes Care*, *43*(12), 3117–3128. https://doi.org/10.2337/dci20-0065

Robinson, E. K., Covarrubias, S., & Carpenter, S. (2020). The how and why of lncRNA function: An innate immune perspective. *BBA-Gene Regulatory Mechanisms*, *1863*(4). https://doi.org/10.1016/j.bbagrm.2019.194419

Roche, P. A., & Furuta, K. (2015). The ins and outs of MHC class II-mediated antigen processing and presentation. *Nature Reviews Immunology*, *15*(4), 203–216. https://doi.org/10.1038/nri3818

Roep, B. O., 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?). *Nature Reviews Endocrinology*, *17*(3), 150–161. https://doi.org/10.1038/s41574-020-00443-4

Rosa, A., & Ballarino, M. (2016). Long Noncoding RNA Regulation of Pluripotency. *Stem Cells International*, *2016*, 1–9. https://doi.org/10.1155/2016/1797692

Ruan, Y., Lin, N., Ma, Q., Chen, R., Zhang, Z., Wen, W., Chen, H., & Sun, J. (2018). Circulating LncRNAs Analysis in Patients with Type 2 Diabetes Reveals Novel Genes Influencing Glucose Metabolism and Islet β-Cell Function. *Cellular Physiology and Biochemistry*, *46*(1), 335–350. https://doi.org/10.1159/000488434

Rubio Cabezas, Ó., & Argente, J. (2012). Diabetes mellitus: formas de presentación clínica y diagnóstico diferencial de la hiperglucemia en la infancia y adolescencia. *Anales de Pediatria*, *77*(5), 344. https://doi.org/10.1016/j.anpedi.2012.06.013

Saldaña-meyer, R., Rodriguez-hernaez, J., Escobar, T., Nishana, M., Jácome-lópez, K., Nora, E. P., Benoit, G., Tsirigos, A., Furlan-magaril, M., Skok, J., & Reinberg, D. (2019). RNA Interactions Are Essential for CTCF-Mediated Genome Organization. *Molecular Cell*, *76*(3), 412–422. https://doi.org/10.1016/j.molcel.2019.08.015.RNA

Salmena, L., Poliseno, L., Tay, Y., Kats, L., & Pandolfi, P. P. (2011). A ceRNA hypothesis: the Rosetta stone of a hidden RNA language? *Cell*, *146*(3), 353–358. https://doi.org/10.1016/j.cell.2011.07.014.A

Salzano, G., Passanisi, S., Mammì, C., Priolo, M., Pintomalli, L., Caminiti, L., Messina, M. F., Pajno, G. B., & Lombardo, F. (2019). Maturity Onset Diabetes of the Young is Not Necessarily Associated with Autosomal Inheritance: Case Description of a De Novo HFN1A Mutation. *Diabetes Therapy*, *10*(4), 1543–1548. https://doi.org/10.1007/s13300-019-0633-3

Sanbonmatsu, K. (2022). Getting to the bottom of lncRNA mechanism: structure–function relationships. *Mammalian Genome*, *33*(2), 343–353. https://doi.org/10.1007/s00335-021-09924-x

Sanjeevi, C. B., Sedimbi, S. K., Landin-Olsson, M., Kockum, I., & Lernmark, Å. (2008). Risk conferred by HLA-DR and DQ for type 1 diabetes in 0-35-year age group in Sweden. *Immunology of Diabetes V*, *1150*, 106–111. https://doi.org/10.1196/annals.1447.061

Chapter 8: References

Santana del Pino, A., Medina-Rodríguez, N., Hernández-García, M., Nóvoa-Mogollón, F. J., & Wägner, A. M. (2017). Is HLA the cause of the high incidence of type 1 diabetes in the Canary Islands? Results from the Type 1 Diabetes Genetics Consortium (T1DGC). *Endocrinologia, Diabetes y Nutricion*, *64*(3), 146–151. https://doi.org/10.1016/j.endinu.2016.12.003

Santin, I., & Eizirik, D. L. (2013). Candidate genes for type 1 diabetes modulate pancreatic islet inflammation and β-cell apoptosis. *Diabetes Obesity and Metabolism*, *15*(Suppl. 3), 71–81.

Santin, I., Moore, F., Colli, M. L., Gurzov, E. N., Marselli, L., Marchetti, P., & Eizirik, D. L. (2011). PTPN2, a candidate gene for type 1 diabetes, modulates pancreatic β-cell apoptosis via regulation of the BH3-only protein Bim. *Diabetes*, *60*(12), 3279–3288. https://doi.org/10.2337/db11-0758

Saxena, A., & Carninci, P. (2011). Long non-coding RNA modifies chromatin: Epigenetic silencing by long non-coding RNAs. *BioEssays*, *33*(11), 830–839. https://doi.org/10.1002/bies.201100084

Scheen, A. J. (2003). Pathophysiology of type 2 diabetes mellitus. *Continuing Medical Education*, *58*(6), 335–341. https://doi.org/10.3390/ijms21176275

Schneider, W. M., Chevillotte, M. D., & Rice, C. M. (2014). Interferon-stimulated genes: A complex web of host defenses. In *Annual Review of Immunology* (Vol. 32, pp. 513–545). Annual Reviews Inc. https://doi.org/10.1146/annurev-immunol-032713-120231

Schofield, C. J., & Sutherland, C. (2012). Disordered insulin secretion in the development of insulin resistance and Type 2 diabetes. *Diabetic Medicine*, *29*(8), 972–979. https://doi.org/10.1111/j.1464-5491.2012.03655.x

Schweiger, D. S., Mendez, A., Kunilo Jamnik, S., Bratanic, N., Bratina, N., Battelino, T., Brecelj, J., & Vidan-Jeras, B. (2016). High-risk genotypes HLA-DR3-DQ2/DR3-DQ2 and DR3-DQ2/DR4-DQ8 in co-occurrence of type 1 diabetes and celiac disease. *Autoimmunity*, *49*(4), 240–247. https://doi.org/10.3109/08916934.2016.1164144

Scoville, D. W., Gruzdev, A., & Jetten, A. M. (2020). Identification of a novel lncrna (G3R1) regulated by GLIS3 in pancreatic β-cells. *Journal of Molecular Endocrinology*, *65*(3), 59–67. https://doi.org/10.1530/JME-20-0082

Sebastian-delacruz, M., Gonzalez-moro, I., Olazagoitia-garmendia, A., Castellanos-rubio, A., & Santin, I. (2021). The Role of lncRNAs in Gene Expression Regulation through. *Non-Coding RNA*, *7*(3), 1–13.

See, D. M., & Tilles, J. G. (1995). Pathogenesis of virus-induced diabetes in mice. *Journal of Infectious Diseases*, *171*(5), 1131–1138. https://doi.org/10.1093/infdis/171.5.1131

Seiler, J., Breinig, M., Caudron-Herger, M., Polycarpou-Schwarz, M., Boutros, M., & Diederichs, S. (2017). The lncRNA VELUCT strongly regulates viability of lung cancer cells despite its extremely low abundance. *Nucleic Acids Research*, *45*(9), 5458–5469. https://doi.org/10.1093/nar/gkx076

Sheen, J. M., Hsieh, C. S., Tain, Y. L., Li, S. W., Yu, H. R., Chen, C. C., Tiao, M. M., Chen, Y. C., & Huang, L. T. (2016). Programming effects of prenatal glucocorticoid exposure with a postnatal high-fat diet in diabetes mellitus. *International Journal of Molecular Sciences*, *17*(4), 1–12. https://doi.org/10.3390/ijms17040533

Shiina, T., Hosomichi, K., Inoko, H., & Kulski, J. K. (2009). The HLA genomic loci map: Expression, interaction, diversity and disease. *Journal of Human Genetics*, *54*(1), 15–39. https://doi.org/10.1038/jhg.2008.5

Simonsen, J. R., Harjutsalo, V., Järvinen, A., Kirveskari, J., Forsblom, C., Groop, P. H., & Lehto, M. (2015). Bacterial infections in patients with type 1 diabetes: A 14-year follow-up study. *BMJ Open Diabetes Research and Care*, *3*(1), 1–9. https://doi.org/10.1136/bmjdrc-2014-000067

Singal, D. P., & Blajchman, M. A. (1973). Histocompatibility (HL A) antigens, lymphocytotoxic antibodies and tissue antibodies in patients with diabetes mellitus. *Diabetes*, *22*(6), 429–432. https://doi.org/10.2337/diab.22.6.429

Singer, R. A., Arnes, L., & Sussel, L. (2015). Noncoding RNAs in Beta Cell Biology Ruth. *Current Opinion in Endodrinology, Diabetes and Obesity*, *22*(2), 77–85. https://doi.org/10.1097/MED.0000000000000141.Noncoding

Sioofy-Khojine, A. B., Lehtonen, J., Nurminen, N., Laitinen, O. H., Oikarinen, S., Huhtala, H., Pakkanen, O., Ruokoranta, T., Hankaniemi, M. M., Toppari, J., Vähä-Mäkilä, M., Ilonen, J., Veijola, R., Knip, M., & Hyöty, H. (2018). Coxsackievirus B1 infections are associated with the initiation of insulin-driven autoimmunity that progresses to type 1 diabetes. *Diabetologia*, *61*(5), 1193–1202. https://doi.org/10.1007/s00125-018-4561-y

Smirnova, V. V, Shestakova, E. D., Bikmetov, D. V, Chugunova, A. A., Osterman, I. A., Serebryakova, M. V, Sergeeva, O. V, Zatsepin, T. S., Shatsky, I. N., & Terenin, I. M. (2019). eIF4G2 balances its own mRNA translation via a PCBP2-based feedback loop. *RNA*, *25*, 757–767. https://doi.org/10.1261/rna.065623.118.

Smyth, D. J., Plagnol, V., Walker, N. M., Cooper, J. D., Downes, K., Yang, J. H. M., Howson, J. M. M.,

Stevens, H., McManus, R., Wijmenga, C., Heap, G. A., Dubois, P. C., Clayton, D. G., Hunt, K. A., van Heel, D. A., & Todd, J. A. (2008). Shared and Distinct Genetic Variants in Type 1 Diabetes and Celiac Disease. *New England Journal of Medicine*, *359*(26), 2767–2777. https://doi.org/10.1056/nejmoa0807917

Song, X., Fan, P.-D., Bantikassegn, A., Guha, U., Threadgill, D. W., Varmus, H., & Politi, K. (2015). ERBB3 Independent Activation of the PI3K Pathway in EGFR Mutant Lung Adenocarcinomas. *Cancer Research*, *75*(6), 1035–1045. https://doi.org/doi:10.1158/0008-5472.CAN-13-1625.

Sosenko, J. M., Skyler, J. S., Krischer, J. P., Greenbaum, C. J., Mahon, J., Rafkin, L. E., Cuthbertson, D., Cowie, C., Herold, K., Eisenbarth, G., & Palmer, J. P. (2010). Glucose excursions between states of glycemia with progression to type 1 diabetes in the Diabetes Prevention Trial-Type 1 (DPT-1). *Diabetes*, *59*(10), 2386–2389. https://doi.org/10.2337/db10-0534

Spalinger, M. R., Kasper, S., Gottier, C., Lang, S., Atrott, K., Vavricka, S. R., Scharl, S., Gutte, P. M., Grütter, M. G., Beer, H. D., Contassot, E., Chan, A. C., Dai, X., Rawlings, D. J., Mair, F., Becher, B., Falk, W., Fried, M., Rogler, G., & Scharl, M. (2016). NLRP3 tyrosine phosphorylation is controlled by protein tyrosine phosphatase PTPN22. *Journal of Clinical Investigation*, *126*(5), 1783–1800. https://doi.org/10.1172/JCI83669

Sperling, M. A., & Garg, A. (2018). Monogenic Forms of Diabetes. *Diabetes in America*, *7*, 1–27. http://www.ncbi.nlm.nih.gov/pubmed/33651552

Spitale, R. C., Tsai, M., & Chang, H. Y. (2011). RNA templating the epigenome. *Epigenetics*, *6*(5), 539–543. https://doi.org/10.4161/epi.6.5.15221

Srivastava, K., Narang, R., Bhatia, J., & Saluja, D. (2016). Expression of heat shock protein 70 gene and its correlation with inflammatory markers in essential hypertension. *PLoS ONE*, *11*(3), 1–15. https://doi.org/10.1371/journal.pone.0151060

Stanford, S. M., Rapini, N., & Bottini, N. (2012). Regulation of TCR signalling by tyrosine phosphatases: From immune homeostasis to autoimmunity. *Immunology*, *137*(1), 1–19. https://doi.org/10.1111/j.1365-2567.2012.03591.x

Statello, L., Guo, C.-J., Chen, L.-L., & Huarte, M. (2021). Gene regulation by long non-coding RNAs and its biological functions. *Nature Reviews Molecular Cell Biology*, *22*(2), 96–118. https://doi.org/10.1038/s41580-020-00315-9

Steck, A. K., Bugawan, T. L., Valdes, A. M., Emery, L. M., Blair, A., Norris, J. M., Redondo, M. J., Babu, S. R., Erlich, H. A., Eisenbarth, G. S., & Rewers, M. J. (2005). Association of non-HLA genes with

type 1 diabetes autoimmunity. *Brief Genetic Report*, *54*(August), 2482–2486.

Steck, A. K., & Rewers, M. J. (2011). Genetics of type 1 diabetes. *Clinical Chemistry*, *57*(2), 176–185. https://doi.org/10.1373/clinchem.2010.148221

Stene, L. C., & Rewers, M. (2011). Immunology in the clinic review series; focus on type 1 diabetes and viruses: the enterovirus link to type 1 diabetes: critical review of human studies. *Clinical and Experimental Immunology*, *168*, 12–23.

Støy, J., Edghill, E. L., Flanagan, S. E., Ye, H., Paz, V. P., Pluzhnikov, A., Below, J. E., Hayes, M. G., Cox, N. J., Lipkind, G. M., Lipton, R. B., Greeley, S. A. W., Patch, A. M., Ellard, S., Steiner, D. F., Hattersley, A. T., Philipson, L. H., & Bell, G. I. (2007). Insulin gene mutations as a cause of permanent neonatal diabetes. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(38), 15040–15044. https://doi.org/10.1073/pnas.0707291104

Stumvoll, M., Goldstein, B. J., & Haeften, T. W. Van. (2005). Type 2 diabetes: principles of pathogenesis and therapy. *The Lancet*, *365*, 1333–1346.

Su, T., Wang, T., Zhang, N., Shen, Y., Li, W., Xing, H., & Yang, M. (2022). Long non-coding RNAs in gastrointestinal cancers : Implications for protein phosphorylation. *Biochemical Pharmacology*, *197*(114907), 1–9. https://doi.org/10.1016/j.bcp.2022.114907

Suciu-foca, N., Buda, J., Almojera, P., & Reemtsma, K. (1974). HL-A antigens and M.L.C. responsiveness in systemic lupus erythematosus. *The Lancet*, *2*(7878), 1972–1973.

Sun, J., Ruan, Y., Wang, M., Chen, R., Yu, N., Sun, L., Liu, T., & Chen, H. (2016). Differentially expressed circulating LncRNAs and mRNA identified by microarray analysis in obese patients. *Scientific Reports*, *6*(35421), 1–10. https://doi.org/10.1038/srep35421

Sundberg, E. J., Deng, L., & Mariuzza, R. A. (2007). TCR recognition of peptide/MHC class II complexes and superantigens. *Seminars in Immunology*, *19*(4), 262–271. https://doi.org/10.1016/j.smim.2007.04.006

Tan, Y., Ju, H., Li, J., Xu, R., & Ju, H. (2021). LncRNA-mediated posttranslational modifications and reprogramming of energy metabolism in cancer. *Cancer Communications*, *41*(July), 109–120. https://doi.org/10.1002/cac2.12108

Tattersall, R. B. (1974). Mild familial diabetes with dominant inheritance. *Quaterly Journal of Medicine*, *43*(179), 339–357. https://doi.org/10.1093/oxfordjournals.qjmed.a067391

Temple, K. I., Gardner, R. J., Mackay, D. J. G., Barber, J. C. K., Robinson, D. O., & Shield, J. P. H. (2000).

Transient Neonatal Diabetes. Widening the understanding of the etiopathogenesis of diabetes. *Diabetes*, *49*, 1359–1566. https://doi.org/10.1177/0973217920030213

Thanabalasingham, G., & Owen, K. R. (2011). Diagnosis and management of maturity onset diabetes of the young (MODY). *BMJ*, *343*(7828), 1–9. https://doi.org/10.1136/bmj.d6044

The ENCODE Project Consortium. (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature*, *447*(7146), 799–816. https://doi.org/10.1038/nature05874

The FANTOM Consortium, RIKEN PMI, & CLST (DGT). (2014). A promoter-level mammalian expression atlas. *Nature*, *507*(7493), 462–470. https://doi.org/10.1038/nature13182.A

Todd, J. A., Bell, J. I., & McDevitt, H. O. (1987). HLA-DQbeta gene contributes to susceptibility and resistance to insulin-dependent diabetes mellitus. *Nature*, *329*, 599–604.

Törn, C., Hadley, D., Lee, H. S., Hagopian, W., Lernmark, Å., Simell, O., Rewers, M., Ziegler, A., Schatz, D., Akolkar, B., Onengut-Gumuscu, S., Chen, W. M., Toppari, J., Mykkänen, J., Ilonen, J., Rich, S. S., She, J. X., Steck, A. K., & Krischer, J. (2015). Role of type 1 diabetes- Associated snps on risk of autoantibody positivity in the TEDDY study. *Diabetes*, *64*(5), 1818–1829. https://doi.org/10.2337/db14-1497

Tsai, M. C., Manor, O., Wan, Y., Mosammaparast, N., Wang, J. K., Lan, F., Shi, Y., Segal, E., & Chang, H. Y. (2010). Long noncoding RNA as modular scaffold of histone modification complexes. *Science*, *329*(5992), 689–693. https://doi.org/10.1126/science.1192002

Udler, M. S. (2019). Type 2 Diabetes: Multiple Genes, Multiple Diseases. *Current Diabetes Reports*, *19*(8), 1–9. https://doi.org/10.1007/s11892-019-1169-7

Undlien, D. E., Bennett, S. T., Todd, J. A., Akselsen, H. E., Ikäheimo, I., Reijonen, H., Knip, M., Thorsby, E., & Rønningen, K. S. (1995). Insulin gene region-encoded susceptibility to IDDM maps upstream of the insulin gene. *Diabetes*, *44*(6), 620–625. https://doi.org/10.2337/diab.44.6.620

Undlien, D. E., Kockum, I., Rønningen, K. S., Lowe, R., Saanjeevi, C. B., Graham, J., Lie, B. A., Akselsen, H. E., Lernmark, Å., & Thorsby, E. (1999). HLA associations in type 1 diabetes among patients not carrying high-risk DR3-DQ2 or DR4-DQ8 haplotypes. *Tissue Antigens*, *54*(6), 543–551. https://doi.org/10.1034/j.1399-0039.1999.540602.x

Urakami, T. (2019). Maturity-onset diabetes of the young (MODY): Current perspectives on diagnosis and treatment. *Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy*, *12*, 1047–

1056. https://doi.org/10.2147/DMSO.S179793

Vehik, K., Lynch, K. F., Wong, M. C., Tian, X., Ross, M. C., Gibbs, R. A., Ajami, N. J., Petrosino, J. F., Rewers, M., Toppari, J., Ziegler, A. G., She, J. X., Lernmark, Å., Akolkar, B., Hagopian, W. A., Schatz, D. A., Krischer, J. P., Hyöty, H., & Richard E, L. (2019). Prospective virome analyses in young children at increased genetic risk for type 1 diabetes. *Physiology & Behavior*, *25*(12), 1865–1872. https://doi.org/10.1038/s41591-019-0667-0.Prospective

Vidan-Jeras, B. (2018). When type 1 diabetes meets celiac disease. *HLA Immune Response Genetics*, *92*(July), 64–66. https://doi.org/10.1111/tan.13441

Virtanen, S. M., Läärä, E., Hyppönen, E., Reijonen, H., Räsänen, L., Aro, A., Knip, M., Ilonen, J., & Åkerblom, H. K. (2000). Cow's milk consumption, HLA-DQB1 genotype, and type 1 diabetes: A nested case-control study of siblings of children with diabetes. *Diabetes*, *49*(6), 912–917. https://doi.org/10.2337/diabetes.49.6.912

Wang, D., & Pan, G. (2019). The Association between rs2292239 Polymorphism in ERBB3 Gene and Type 1 Diabetes: A Meta-Analysis. *BioMed Research International*, *2019*, 1–8. https://doi.org/10.1155/2019/7689642

Wang, H., Jin, Y., Reddy, M. V. P. L., Podolsky, R., Liu, S., Yang, P., Bode, B., Reed, J. C., Steed, R. D., Anderson, S. W., Steed, L., Hopkins, D., Huang, Y., & She, J. X. (2010). Genetically dependent ERBB3 expression modulates antigen presenting cell function and type 1 diabetes risk. *PLoS ONE*, *5*(7), 1–10. https://doi.org/10.1371/journal.pone.0011789

Wang, J. P., Cerny, A., Asher, D. R., Kurt-Jones, E. A., Bronson, R. T., & Finberg, R. W. (2010). MDA5 and MAVS Mediate Type I Interferon Responses to Coxsackie B Virus. *Journal of Virology*, *84*(1), 254–260. https://doi.org/10.1128/jvi.00631-09

Wang, K., Ye, F., Chen, Y., Xu, J., Zhao, Y., Wang, Y., & Lan, T. (2021). Association Between Enterovirus Infection and Type 1 Diabetes Risk: A Meta-Analysis of 38 Case-Control Studies. *Frontiers in Endocrinology*, *12*(September), 1–9. https://doi.org/10.3389/fendo.2021.706964

Wang, Y., Shaked, I., Stanford, S. M., Zhou, W., Julie, M., Mikulski, Z., Shaheen, Z. R., Cheng, G., Sawatzke, K., Campbell, A. M., Auger, J. L., Bilgic, H., Shoyama, F. M., Schmeling, D. O., Balfour Jr, H. H., Hasegawa, K., Chan, A. C., Corbett, J. A., Binstadt, B. A., … Peterson, E. J. (2013). The Autoimmunity-Associated Gene PTPN22 Potentiates Toll- like Receptor-Driven, Type 1 Interferon-Dependent Immunity. *Immunity*, *39*(1), 1–24. https://doi.org/10.1016/j.immuni.2013.06.013.The

Waterhouse, P., Penninger, J. M., Timms, E., Wakeham, A., Shahinian, A., Lee, K. P., Thompson, C. B., Griesser, H., & Mak, T. W. (1995). Lymphoproliferative disorders with early lethality in mice deficient in Ctla-4. *Science*, *270*(5238), 985–988. https://doi.org/10.1126/science.270.5238.985

Wen, L., Ley, R. E., Volchkov, P. V, Stranges, P. B., Avanesyan, L., Stonebraker, A. C., Hu, C., Wong, F. S., Szot, G. L., Bluestone, J. A., Gordon, J. I., & Chervonsky, A. V. (2008). Innate immunity and intestinal microbiota in the development of Type 1 diabetes. *Nature*, *455*(7216), 1109–1113. https://doi.org/10.1038/nature07336.Innate

Wherrett, D. K. (2021). Improving prediction of risk for the development of type 1 diabetes—insights from populations at high risk. *Diabetes Care*, *44*(10), 2189–2191. https://doi.org/10.2337/dci21-0018

Wieczorek, M., Abualrous, E. T., Sticht, J., Álvaro-Benito, M., Stolzenberg, S., Noé, F., & Freund, C. (2017). Major histocompatibility complex (MHC) class I and MHC class II proteins: Conformational plasticity in antigen presentation. *Frontiers in Immunology*, *8*(March), 1–16. https://doi.org/10.3389/fimmu.2017.00292

Wilson, M. E., & Pullen, T. J. (2021). The role of long non-coding RNAs in the regulation of pancreatic beta cell identity. *Biochemical Society Transactions*, *49*(July), 2153–2161.

Wong, W. K. M., Jiang, G., Sørensen, A. E., Chew, Y. V., Lee-Maynard, C., Liuwantara, D., Williams, L., O'Connell, P. J., Dalgaard, L. T., Ma, R. C., Hawthorne, W. J., Joglekar, M. V., & Hardikar, A. A. (2019). The long noncoding RNA MALAT1 predicts human islet isolation quality. *JCI Insight*, *4*(16), 1–13. https://doi.org/10.1172/jci.insight.129299

World Health Organization. (2016). WORLD HEALTH STATISTICS - MONITORING HEALTH FOR THE SDGs. *World Health Organization*, 1.121.

Wu, M., Yang, L. Z., & Chen, L. L. (2021). Long noncoding RNA and protein abundance in lncRNPs. *RNA*, *27*(12), 1427–1440. https://doi.org/10.1261/rna.078971.121

Wu, X., Wang, F., Li, H., Hu, Y., Jiang, L., Zhang, F., Li, M., Wang, X., Jin, Y., Zhang, Y., Lu, W., Wu, W., Shu, Y., Weng, H., Cao, Y., Bao, R., Liang, H., Wang, Z., Zhang, Y.-C., … Liu, Y.-B. (2017). LncRNA-PAGBC acts as a microRNA sponge and promotes gallbladder tumorigenesis. *EMBO Reports*, *18*(10), 1837–1853.

Xiang, J., Yin, Q., Chen, T., Zhang, Y., Zhang, X., Wu, Z., Zhang, S., Wang, H., Ge, J., Lu, X., Yang, L., & Chen, L. (2014). Human colorectal cancer-specific CCAT1-L lncRNA regulates long-range chromatin interactions at the MYC locus. *Cell Research*, *24*(5), 513–531.

https://doi.org/10.1038/cr.2014.35

Xin, Z., Han, W., Zhao, Z., Xia, Q., Yin, B., Yuan, J., & Peng, X. (2011). PCBP2 enhances the antiviral activity of IFN-α against HCV by stabilizing the mRNA of STAT1 and STAT2. *PLoS ONE*, *6*(10), 1–13. https://doi.org/10.1371/journal.pone.0025419

Xue, Z., Hennelly, S., Doyle, B., Gulati, A. A., Novikova, I. V., Sanbonmatsu, K. Y., & Boyer, L. A. (2016). A G-Rich Motif in the lncRNA Braveheart Interacts with a Zinc-Finger Transcription Factor to Specify the Cardiovascular Lineage. *Molecular Cell*, *64*(1), 37–50. https://doi.org/10.1016/j.molcel.2016.08.010

Yang, F., Deng, X., Ma, W., Berletch, J. B., Rabaia, N., Wei, G., Moore, J. M., Filippova, G. N., Xu, J., Liu, Y., Noble, W. S., Shendure, J., & Disteche, C. M. (2015). The lncRNA Firre anchors the inactive X chromosome to the nucleolus by binding CTCF and maintains H3K27me3 methylation. *Genome Biology*, *16*(52), 1–17. https://doi.org/10.1186/s13059-015-0618-0

Yapanis, M., James, S., Craig, M. E., O'Neal, D., & Ekinci, E. I. (2022). Complications of Diabetes and Metrics of Glycemic Management Derived From Continuous Glucose Monitoring. *Journal of Clinical Endocrinology and Metabolism*, *107*(6), e2221–e2236. https://doi.org/10.1210/clinem/dgac034

Yin, D., Zhang, E., You, L., Wang, N., Wang, L., Jin, F., Zhu, Y., Cao, L., Yuan, Q., De, W., & Tang, W. (2015). Downregulation of lncRNA TUG1 Affects Apoptosis and Insulin Secretion in Mouse Pancreatic b Cells. *Cellular Physiology and Biochemistry*, *35*(5), 1892–1904. https://doi.org/10.1159/000373999

Yoon, J. W., Austin, M., Onodera, T., & Notkins, A. L. (1977). Virus-induced diabetes mellitus. *Archives of Virology*, *300*(21), 1173–1179. https://doi.org/10.1007/BF01314374

Yoon, J. W., & Jun, H. S. (2005). Autoimmune destruction of pancreatic β cells. *American Journal of Therapeutics*, *12*(6), 580–591. https://doi.org/10.1097/01.mjt.0000178767.67857.63

You, L., Wang, N., Yin, D., Wang, L., Jin, F., Zhu, Y., Yuan, Q., & De, W. (2015). Downregulation of Long Noncoding RNA Meg3 Affects Insulin Synthesis and Secretion in Mouse Pancreatic Beta Cells. *Journal of Cellular Physiology*, *231*(4), 852–862. https://doi.org/10.1002/jcp.25175

Yu, B., & Wang, S. (2018). Angio-LncRs : LncRNAs that regulate angiogenesis and vascular disease. *Theranostics*, *8*(13), 3654–3674. https://doi.org/10.7150/thno.26024

Yuasa, K., & Hijikata, T. (2016). Distal regulatory element of the STAT1 gene potentially mediates

positive feedback control of STAT1 expression. *Genes to Cells*, *21*(1), 25–40. https://doi.org/10.1111/gtc.12316

Zampetaki, A., Albrecht, A., & Steinhofel, K. (2018). Long non-coding RNA structure and function: Is there a link? In *Frontiers in Physiology* (Vol. 9, Issue 1201, pp. 1–8). Frontiers Media S.A. https://doi.org/10.3389/fphys.2018.01201

Zhang, H., Colclough, K., Gloyn, A. L., & Pollin, T. I. (2021). Monogenic diabetes: A gateway to precision medicine in diabetes. *Journal of Clinical Investigation*, *131*(3), 1–14. https://doi.org/10.1172/JCI142244

Zhang, H., Liang, Y., Han, S., Peng, C., & Li, Y. (2019). Long noncoding RNA and protein interactions: From experimental results to computational models based on network methods. *International Journal of Molecular Sciences*, *20*(6), 1–30. https://doi.org/10.3390/ijms20061284

Zhang, K., Hua, Y. Q., Wang, D., Chen, L. Y., Wu, C. J., Chen, Z., Liu, L. M., & Chen, H. (2019). Systemic immune-inflammation index predicts prognosis of patients with advanced pancreatic cancer. *Journal of Translational Medicine*, *17*(30), 30. https://doi.org/10.1186/s12967-019-1782-x

Zhang, X., Rice, K., Wang, Y., Chen, W., Zhong, Y., Nakayama, Y., Zhou, Y., & Klibanski, A. (2010). Maternally Expressed Gene 3 (MEG3) Noncoding Ribonucleic Acid: Isoform Structure, Expression, and Functions. *Cancer Oncogenes*, *151*(3), 939–947. https://doi.org/10.1210/en.2009-0657

Zhao, J., Sun, B. K., Erwin, J. A., Song, J., & Lee, J. T. (2008). Polycomb proteins targeted by a ashort repeat RNA to the mouse X-chromosome. *Science*, *322*(5902), 750–756. https://doi.org/10.1126/science.1163045.Polycomb

Zheng, Y., Ley, S. H., & Hu, F. B. (2018). Global aetiology and epidemiology of type 2 diabetes mellitus and its complications. *Nature Reviews Endocrinology*, *14*(2), 88–98. https://doi.org/10.1038/nrendo.2017.151

Zhernakova, A., Festen, E. M., Franke, L., Trynka, G., van Diemen, C. C., Monsuur, A. J., Bevova, M., Nijmeijer, R. M., van 't Slot, R., Heijmans, R., Boezen, H. M., van Heel, D. A., van Bodegraven, A. A., Stokkers, P. C. F., Wijmenga, C., Crusius, J. B. A., & Weersma, R. K. (2008). Genetic Analysis of Innate Immunity in Crohn's Disease and Ulcerative Colitis Identifies Two Susceptibility Loci Harboring CARD9 and IL18RAP. *The American Journal of Human Genetics*, *82*(5), 1202–1210. https://doi.org/10.1016/j.ajhg.2008.03.016

Zhou, L., Sun, K., Zhao, Y., Zhang, S., Wang, X., Li, Y., Lu, L., Chen, X., Chen, F., Bao, X., Zhu, X., Wang,

L., Tang, L. Y., Esteban, M. A., Wang, C. C., Jauch, R., Sun, H., & Wang, H. (2015). Linc-YY1 promotes myogenic differentiation and muscle regeneration through an interaction with the transcription factor YY1. *Nature Communications*, *6*(10026), 1–16. https://doi.org/10.1038/ncomms10026

Zhou, Y., Zhang, X., & Klibanski, A. (2012). MEG3 noncoding RNA: a tumor suppressor. *Journal of Molecular Endocrinology*, *48*(3), R45–R53. https://doi.org/10.1530/JME-12-0008.MEG3

Zhu, Y., Liu, Q., Zhou, Z., & Ikeda, Y. (2017). PDX1 , Neurogenin-3 , and MAFA: critical transcription regulators for beta cell development and regeneration. *Stem Cell Research & Therapy*, *8*(240), 1–7. https://doi.org/10.1186/s13287-017-0694-z

Ziegler, A. G., Rewers, M., Simell, O., Simell, T., Lempainen, J., Steck, A., Winkler, C., Ilonen, J., Veijola, R., Knip, M., Bonifacio, E., & Eisenbarth, G. S. (2013). Seroconversion to Multiple Islet Autoantibodies and Risk of Progression to Diabetes in Children. *Journal of the American Medical Association*, *309*(23), 2473–2479. https://doi.org/10.1001/jama.2013.6285

Zipitis, C. S., & Akobeng, A. K. (2008). Vitamin D supplementation in early childhood and risk of type 1 diabetes: a systematic review and meta-analysis. *Archives of Disease in Childhood*, *93*(6), 512–517. https://doi.org/10.1136/adc.2007.128579

Chapter 9

Appendix

Eranskinak

Appendix 1. LncRNAs harboring a T1D-associated SNP

T1D-associated SNPs			Annotated lncRNAs			
Chromosome	Position	SNP code	Chromosome	Start position	End position	lncRNA transcript
chr1	113886839	rs2358994	chr1	113812378	113901238	NONHSAG002475.3 + NONHSAT005341.2 + NONHSAT149626.1 + NONHSAT149627.1
chr1	113834946	rs2476601	chr1	113812378	113901238	NONHSAG002475.3 + NONHSAT149625.1 + NONHSAT149626.1 + NONHSAT225397.1
chr2	100147625	rs13415583	chr2	100109603	100190303	NONHSAG077573.1 + NONHSAT182257.1
chr2	12500615	rs1534422	chr2	11848621	12598651	NONHSAG077122.2 + NONHSAT069214.2 + NONHSAT069223.2 + NONHSAT181280.1 + NONHSAT181281.1 + NONHSAT240366.1 + NONHSAT240368.1 + NONHSAT240369.1 + NONHSAT240370.1
chr2	162254026	rs2111485	chr2	162246093	162259757	NONHSAG029690.3 + NONHSAT241172.1 + NONHSAT241173.1
chr2	28423934	rs6547853	chr2	28423353	28423995	NONHSAG027355.2 + NONHSAT069825.2
chr2	102454108	rs917997	chr2	102452050	102454594	(*LNC13*) NONHSAG069179.1 + NONHSAT348493.1
chr2	100208905	rs9653442	chr2	100208253	100251485	NONHSAG028706.2 + NONHSAT072771.2 + NONHSAT182258.1
chr3	58414687	rs11705721	chr3	58408913	58459793	NONHSAG084679.1 + NONHSAT193943.1
chr3	58414687	rs11705721	chr3	58412947	58426011	NONHSAG084680.1 + NONHSAT193944.1
chr4	26083889	rs10517086	chr4	26074737	26104231	NONHSAT095867.2 + NONHSAT198423.1
chr6	90248512	rs11755527	chr6	90104364	90296742	NONHSAG044354.2 + NONHSAT113982.2
chr6	126431738	rs1538171	chr6	126377654	127119849	NONHSAG044781.2 + NONHSAT114827.2
chr6	126431738	rs1538171	chr6	126339788	126483320	NONHSAG044782.2 + NONHSAT208006.1
chr6	126431738	rs1538171	chr6	126419347	127161542	NONHSAG094480.1 + NONHSAT209660.1
chr6	32395438	rs1980493	chr6	32395325	32396867	NONHSAG112896.1 + NONHSAT251041.1
chr6	31396930	rs2251396	chr6	31394287	31400764	NONHSAT252188.1 + NONHSAT108767.2 + NONHSAT252183.1 + NONHSAT208981.1 + NONHSAT252187.1
chr6	30110498	rs2523989	chr6	30100245	30114725	NONHSAG043397.3 + NONHSAT108601.2 + NONHSAT207067.1
chr6	90247744	rs3757247	chr6	90104364	90296742	NONHSAG044354.2 + NONHSAT113982.2
chr6	90267049	rs72928038	chr6	90104364	90296742	NONHSAG044354.2 + NONHSAT113982.2
chr6	166969779	rs73043122	chr6	166963031	166999911	NONHSAG045405.3 + NONHSAT252915.1 + NONHSAT252916.1 + NONHSAT252917.1
chr6	30814225	rs886424	chr6	30793143	30848302	NONHSAG043436.3 + NONHSAT108682.2 + NONHSAT252173.1 + NONHSAT252176.1 + NONHSAT252177.1 + NONHSAT252178.1 + NONHSAT252174.1 + NONHSAT252171.1 + NONHSAT208963.1 + NONHSAT252172.1 + NONHSAT252175.1

--------- Table continues on the next pages ----------

T1D-associated SNPs			Annotated lncRNAs			
Chromosome	Position	SNP code	Chromosome	Start position	End position	lncRNA transcript
chr6	30814225	rs886424	chr6	30793143	30848302	NONHSAG043436.3 + NONHSAT108682.2 + NONHSAT252173.1 + NONHSAT252176.1 + NONHSAT252177.1 + NONHSAT252178.1 + NONHSAT252174.1 + NONHSAT252171.1 + NONHSAT208963.1 + NONHSAT252172.1 + NONHSAT252175.1
chr6	170063801	rs924043	chr6	170038391	170126890	NONHSAG094711.2 + NONHSAT252969.1
chr6	126377573	rs9388489	chr6	126339788	126483320	NONHSAG044782.2 + NONHSAT208006.1
chr7	50961290	rs10277986	chr7	50866746	51022990	NONHSAG047579.2 + NONHSAT120580.2 + NONHSAT212079.1
chr7	50959497	rs4948088	chr7	50866746	51022990	NONHSAG047579.2 + NONHSAT120580.2 + NONHSAT212079.1
chr7	28189423	rs550448	chr7	28180322	28243991	NONHSAG047220.3 + NONHSAT119753.2 + NONHSAT119754.2 + NONHSAT211856.1 + NONHSAT253229.1 + NONHSAT253230.1
chr7	28189423	rs550448	chr7	28146981	28207902	NONHSAG096236.1 + NONHSAT211855.1
chr7	50398132	rs62447205	chr7	50385894	50400114	NONHSAG097101.1 + NONHSAT213488.1
chr9	4292083	rs10758593	chr9	3824126	4300035	NONHSAG051672.2 + NONHSAT130037.2 + NONHSAT130046.2 + NONHSAT130047.2
chr9	4290823	rs6476839	chr9	3824126	4300035	NONHSAG051672.2 + NONHSAT130037.2 + NONHSAT130046.2 + NONHSAT130047.2
chr9	4291747	rs7020673	chr9	3824126	4300035	NONHSAG051672.2 + NONHSAT130037.2 + NONHSAT130046.2 + NONHSAT130047.2
chr10	122400322	rs113306148	chr10	122375195	122432341	NONHSAG007075.2 + NONHSAT156249.1
chr11	2148544	rs3741208	chr11	2140500	2148666	NONHSAG007412.2 + NONHSAT017485.2 + NONHSAT017486.2 + NONHSAT017487.2
chr11	2163618	rs3842727	chr11	2158719	2168069	NONHSAG007414.2 + NONHSAT017490.2
chr11	69781755	rs4084127	chr11	69781634	69784375	NONHSAG063500.1 + NONHSAT160476.1
chr11	2160994	rs689	chr11	2158719	2168069	NONHSAG007414.2 + NONHSAT017490.2
chr12	9703362	rs11052552	chr12	9702191	9733163	NONHSAG065312.1 + NONHSAT163291.1
chr12	56041720	rs705705	chr12	56019252	56041806	NONHSAG011343.2 + NONHSAT028740.2
chr12	9753255	rs917911	chr12	9752489	9755212	NONHSAG010458.2 + NONHSAT163292.1
chr13	99429512	rs9585056	chr13	99429022	99434056	**(*ARGI*)** NONHSAG107413.1 + NONHSAT233405.1
chr14	98021670	rs1456988	chr14	97925609	98542334	NONHSAG015852.3 + NONHSAT039582.2
chr14	68796882	rs1465788	chr14	68795210	68809156	NONHSAG107517.1 + NONHSAT233762.1
chr14	98032614	rs4900384	chr14	97925609	98542334	NONHSAG015852.3 + NONHSAT039582.2
chr14	100840110	rs56994090	chr14	100834431	100861026	NONHSAG015925.2 + NONHSAT039773.2

--------- Table continues on the next pages ----------

T1D-associated SNPs			Annotated lncRNAs			
Chromosome	Position	SNP code	Chromosome	Start position	End position	lncRNA transcript
chr14	100840110	rs56994090	chr14	100779409	101027415	NONHSAG069030.2 + NONHSAT039736.2 + NONHSAT039737.2 + NONHSAT039738.2 + NONHSAT039739.2 + NONHSAT039740.2 + NONHSAT039741.2 + NONHSAT039742.2 + NONHSAT039743.2 + NONHSAT039744.2 + NONHSAT039745.2 + NONHSAT039746.2 + NONHSAT039747.2 + NONHSAT039748.2 +NONHSAT039749.2 + NONHSAT039750.2 + NONHSAT039751.2 + NONHSAT039753.2 + NONHSAT039754.2 + NONHSAT039755.2 + NONHSAT039756.2 + NONHSAT039757.2 + NONHSAT039758.2 + NONHSAT039759.2 + NONHSAT039763.2 + NONHSAT039764.2 + NONHSAT039765.2 + NONHSAT039767.2 + NONHSAT039771.2 + NONHSAT039772.2 + NONHSAT039773.2 + NONHSAT039774.2 + NONHSAT039775.2 + NONHSAT039776.2 + NONHSAT233952.1 + NONHSAT233953.1 + NONHSAT233954.1 + NONHSAT233955.1 + NONHSAT233956.1 + NONHSAT233957.1 + NONHSAT233958.1 + NONHSAT233959.1 + NONHSAT233960.1 + NONHSAT233961.1 + NONHSAT233962.1 + NONHSAT233963.1 + NONHSAT233965.1
chr14	100839708	rs941576	chr14	100834431	100861026	NONHSAG015925.2 + NONHSAT039773.2
chr14	100839708	rs941576	chr14	100779409	101027415	NONHSAG069030.2 + NONHSAT039736.2 + NONHSAT039737.2 + NONHSAT039738.2 + NONHSAT039739.2 + NONHSAT039740.2 + NONHSAT039741.2 + NONHSAT039742.2 + NONHSAT039743.2 + NONHSAT039744.2 + NONHSAT039745.2 + NONHSAT039746.2 + NONHSAT039747.2 + NONHSAT039748.2 + NONHSAT039749.2 + NONHSAT039750.2 + NONHSAT039751.2 + NONHSAT039753.2 + NONHSAT039754.2 + NONHSAT039755.2 + NONHSAT039756.2 + NONHSAT039757.2 + NONHSAT039758.2 + NONHSAT039759.2 + NONHSAT039763.2 + NONHSAT039764.2 + NONHSAT039765.2 + NONHSAT039767.2 + NONHSAT039771.2 + NONHSAT039772.2 + NONHSAT039774.2 + NONHSAT039775.2 + NONHSAT039776.2 + NONHSAT233952.1 + NONHSAT233953.1 + NONHSAT233954.1 + NONHSAT233955.1 + NONHSAT233956.1 + NONHSAT233957.1 + NONHSAT233958.1 + NONHSAT233959.1 + NONHSAT233960.1 + NONHSAT233961.1 + NONHSAT233962.1 + NONHSAT233963.1 + NONHSAT233965.1
chr15	38614840	rs7171171	chr15	38612281	38618096	NONHSAG070250.1 + NONHSAT170665.1
chr16	11257354	rs193778	chr16	11256306	11266419	NONHSAG108195.1 + NONHSAT235847.1 + NONHSAT235846.1
chr16	75213347	rs7202877	chr16	75136965	75229739	NONHSAG071826.1 + NONHSAT173274.1
chr16	75218429	rs8056814	chr16	75136965	75229739	NONHSAG071826.1 + NONHSAT173274.1
chr16	6086218	rs9934817	chr16	5239801	6776014	NONHSAG018484.2 + NONHSAT140344.2
chr16	6086218	rs9934817	chr16	6056974	6092954	NONHSAG018490.2 + NONHSAT140356.2
chr17	45996523	rs1052553	chr17	45991475	46024114	NONHSAG022015.2 + NONHSAT054194.2
chr17	45996523	rs1052553	chr17	45988501	46002018	NONHSAG073703.1 + NONHSAT176277.1 + NONHSAT176278.1 + NONHSAT176279.1
chr17	39896954	rs12453507	chr17	39895599	39901533	NONHSAG021717.2 + NONHSAT053491.2

--------- Table continues on the next pages ----------

T1D-associated SNPs			Annotated lncRNAs			
Chromosome	Position	SNP code	Chromosome	Start position	End position	lncRNA transcript
chr17	39896954	rs12453507	chr17	39896146	39898684	NONHSAG021718.2 + NONHSAT053492.2
chr17	40614034	rs7221109	chr17	40603857	40638427	NONHSAG073660.1 + NONHSAT176185.1
chr17	40618898	rs757411	chr17	40603857	40638427	NONHSAG073660.1 + NONHSAT176185.1
chr19	46705224	rs425105	chr19	46698538	46706879	NONHSAG076189.1 + NONHSAT179826.1
chr19	48702915	rs516246	chr19	48702516	48705102	NONHSAG109627.1 + NONHSAT240109.1 + NONHSAT240110.1
chr20	1635560	rs6043409	chr20	1633507	1648473	NONHSAG031110.2 + NONHSAT078124.2 + NONHSAT078125.2
chr21	44201934	rs6518350	chr21	44196478	44210467	NONHSAG082969.2 + NONHSAT244336.1 + NONHSAT244337.1
chr21	44201934	rs6518350	chr21	44200254	44207400	NONHSAG083209.2 + NONHSAT244637.1 + NONHSAT082494.2
chr22	37191071	rs229533	chr22	37190379	37199909	NONHSAG033858.3 + NONHSAT245180.1 + NONHSAT245181.1
chr22	37195278	rs229541	chr22	37190379	37199909	NONHSAG033858.3 + NONHSAT245180.1 + NONHSAT245181.1
chr22	21662102	rs428595	chr22	21652204	21670237	NONHSAG033376.3 + NONHSAT083704.2 + NONHSAT083705.2 + NONHSAT083706.2 + NONHSAT244739.1
chr22	30185733	rs5753037	chr22	30184643	30207110	NONHSAG033660.2 + NONHSAT084759.2 + NONHSAT192801.1

Appendix 2. LncRNAs harboring a T1D-associated SNP in an exonic region.

SNP code	lncRNA name	Secondary structure change? (p<0.2)
rs1052553	NONHSAG022015.2 + NONHSAT054194.2	NO
rs1052553	NONHSAG073703.1 + NONHSAT176277.1 + NONHSAT176278.1 + NONHSAT176279.1	NO
rs12453507	NONHSAG021718.2 + NONHSAT053492.2	YES
rs193778	NONHSAG108195.1 + NONHSAT235847.1 + NONHSAT235846.1	NO
rs1980493	NONHSAG112896.1 + NONHSAT251041.1	YES
rs2251396	NONHSAT252188.1 + NONHSAT108767.2 + NONHSAT252183.1 + NONHSAT208981.1 + NONHSAT252187.1	YES
rs3741208	NONHSAG007412.2 + NONHSAT017485.2 + NONHSAT017486.2 + NONHSAT017487.2	NO
rs4084127	NONHSAG063500.1 + NONHSAT160476.1	YES
rs425105	NONHSAG076189.1 + NONHSAT179826.1	NO
rs516246	NONHSAG109627.1 + NONHSAT240109.1 + NONHSAT240110.1	NO
rs6518350	NONHSAG082969.2 +NONHSAT244337.1	NO
rs6518350	NONHSAG083209.2 + NONHSAT244637.1 + NONHSAT082494.2	YES
rs6547853	NONHSAG027355.2 + NONHSAT069825.2	YES
rs705705	NONHSAG011343.2 + NONHSAT028740.2	YES
rs73043122	NONHSAG045405.3 + NONHSAT252915.1 + NONHSAT252916.1 + NONHSAT252917.1	YES
rs886424	NONHSAG043436.3 + NONHSAT108682.2 + NONHSAT252173.1 + NONHSAT252176.1 + NONHSAT252177.1 + NONHSAT252178.1 + NONHSAT252174.1 + NONHSAT252171.1 + NONHSAT208963.1 + NONHSAT252172.1 + NONHSAT252175.1	YES
rs917911	NONHSAG010458.2 + NONHSAT163292.1	NO
rs917997	(***LNC13***) NONHSAG069179.1 + NONHSAT348493.1	YES
rs9585056	**(*ARGI*)** NONHSAG107413.1 + NONHSAT233405.1	YES