Novel aspects of carboxyl ester lipase (CEL)

Investigations on transcriptional regulation and tissue expression

Hanna Svendsen

This thesis is submitted in partial fulfilment of the requirements for the degree of Master in Biomedical Sciences

Department of Biomedicine, Department of Clinical Medicine
and Department of Clinical Science
University of Bergen

Spring 2024

Table of Contents

ACKNOWLEDGEMENTS **4**

ABBREVIATIONS **5**

ABSTRACT **6**

1. INTRODUCTION **8**

1.1 THE HUMAN PANCREAS 8
1.1.1 The endocrine pancreas *8*
1.1.2 The exocrine pancreas *9*
1.2 DISEASES OF THE PANCREAS 11
1.2.1 Diabetes Mellitus *11*
1.2.2 Pancreatitis *13*
1.3 CARBOXYL ESTER LIPASE 14
1.3.1 The human CEL gene *15*
1.3.2 The CEL protein *15*
1.3.3 Regulation of CEL transcription *17*
1.4 THE POLYMORPHIC NATURE OF *CEL* 18
1.4.1 VNTR lengths *18*
1.4.2 Insertions *18*
1.4.3 Deletions *18*
1.4.4 CNV *18*
1.5 *CEL* IN PANCREATIC DISEASE 19
1.5.1 CEL and monogenic diabetes *19*
1.5.2 CEL and chronic pancreatitis *20*
1.5.3 CEL and type 1 diabetes *22*

2. AIMS OF THE STUDY **23**

3. MATERIALS **24**

4. METHODS **28**

4.1 PREPARATION OF PLASMIDS 28
4.1.1 Transformation of E.coli and plasmid purification *28*
4.1.2 Determination of plasmid concentration and quality *28*
4.2 SITE-DIRECTED MUTAGENESIS 28
4.2.1 Primer design and Polymerase Chain Reaction (PCR) *28*
4.2.2 Transformation and plasmid purification *29*
4.2.3 Sanger Sequencing *30*
4.3 CELL CULTURING AND TRANSFECTION 31
4.3.1 Culturing of cells *31*
4.3.2 Passaging and seeding of cells *31*
4.3.3 Cell freezing and thawing. *31*
4.3.4 Transient transfection of HeLa, HEK293 and ZR-75-1 cells *32*
4.3.5 Tamoxifen and Fulvestrant treatment *32*
4.4 LUCIFERASE-ASSAY 33
4.4.1 Principle *33*
4.4.2 Gaussia Luciferase and Alkaline Phosphatase activity *33*
4.4.3 Statistical analysis *34*
4.5 IMMUNOFLUORESCENCE 34
4.6 SDS-PAGE AND WESTERN BLOT 35
4.6.1 Preparation of cell fractions *35*
4.6.2 Determination of protein concentration *35*
4.6.3 SDS-PAGE *35*
4.6.4 Western Blot *36*
4.7 IMMUNOHISTOCHEMISTRY 36
4.7.1 Chromogenic staining *37*

4.7.2 Fluorescent double-staining ... *37*
4.8 In silico analysis of the *CEL* promoter ... 38

5. RESULTS ... **39**

5.1 Part 1: Analysis of the *CEL* gene promoter ... 39
5.1.1 Preparation of plasmids and introduction of the C>T risk mutation at rs541856133 in the CEL promoter ... *39*
5.1.2 Optimization of the Luciferase assay in HEK293 and HeLa cells ... *40*
5.1.3 The effect of the C>T risk mutation on CEL promoter activity in HEK293 and HeLa cells ... *42*
5.1.4 In silico investigation of the CEL promoter ... *42*
5.1.5 Endogenous expression of ZNF331 in HEK293 cells ... *44*
5.1.6 The effect of ZNF311 on CEL promoter activity in HEK293 cells ... *45*
5.1.7 Transfection efficiency and CEL promoter activity in ZR-75-1 cells ... *47*
5.1.8 The effect of ZNF331 in ZR-75-1 cells ... *49*
5.1.9 The effect of ERα on CEL promoter activity ... *50*
5.1.10 The effect of Tamoxifen and Fulvestrant on CEL promoter activity ... *51*
5.1.11 Endogenous expression of CEL in ZR-75-1 cells ... *53*
5.2 Part 2: CEL protein expression in the pituitary gland ... 53
5.2.1 CEL protein expression in the mouse pituitary gland ... *53*
5.2.2 CEL protein expression in the human pituitary gland ... *55*

6. DISCUSSION ... **57**

6.1 Analysis of the *CEL* gene promoter ... 57
6.1.1 The influence of SNP rs541856133 (T) on CEL promoter activity ... *57*
6.1.2 Impact of transcription factors on CEL promoter activity ... *58*
6.1.3 Hormonal regulation of CEL promoter activity? ... *60*
6.2 CEL's role outside of the digestive tract ... 61
6.2.1 CEL protein expression in the human pituitary gland ... *62*
6.2.2 CEL protein expression in the mouse pituitary gland ... *63*
6.3 Methodological considerations ... 64
6.3.1 Choice of cell model ... *64*
6.3.2 Luciferase-assay ... *65*

7. CONCLUSION ... **67**

8. FUTURE PERSPECTIVES ... **68**

REFERENCES ... **69**

APPENDIX ... **77**

Acknowledgements

I would like to express my deepest gratitude to my main supervisor, Karianne Fjeld, for giving me the opportunity to undertake this exciting project. Karianne, your kindness and support throughout this journey have been incredible. Whether it was lab troubles or day-to-day questions, you were always there to help. Thank you for encouraging me to slowly step out of my comfort zone and helping me grow as a researcher.

I also want to thank my co-supervisors, Jan Inge Bjune and Anders Molven, for your expert advice and helpful feedback on my research and writing. Jan Inge, thank you for sharing your expertise in the lab, especially with the Luciferase-assay. I appreciated your regular insights and thoughts on how to proceed with my results each week.

To the entire CEL research group, thank you for creating such a welcoming and supportive environment. Ranveig Seim Brekke, thank you for all your help in the lab, and for always being there to answer all of my big and small questions. Our small daily chats always brightened my day. Renate Valdersnes Seierstad, I am thankful for your help with immunohistochemistry, and along with Janniche Torsvik, thank you both for all the help with confocal imaging. Solrun Steine, thank you for helping me with sequencing and for managing my countless orders. Diego Iglesias Lopez, our chats and your humor made this master thesis journey much more enjoyable.

Lastly, I owe a huge thanks to my family and friends for their endless encouragement and support. Thank you for being there through every up and down. Lea, our movie nights were the perfect breaks that always lifted my spirits. Finally, a big thank you to all my fellow master students for the friendship and encouragement that helped us all reach the finish line.

Thank you all for your help over this last year. I highly appreciate every one of you.

Bergen, June 2024

Hanna Svendsen

Abbreviations

AP	Acute Pancreatitis
CEL/*CEL*	Carboxyl Ester Lipase protein/gene
CEL-HYB1	Carboxyl Ester Lipase hybrid allele 1
CELP	Carboxyl Ester Lipase pseudogene
CNV	Copy Number Variants
CP	Chronic Pancreatitis
ER	Endoplasmic Reticulum
ERα	Estrogen Receptor Alpha
ERβ	Estrogen Receptor Beta
GLuc	*Gaussia* Luciferase
HEK293	Human Embryonic kidney cell line 293
MODY	Maturity-Onset Diabetes of the Young
MUT	Mutant
SEAP	Secreted Alkaline Phosphatase
SNP	Single Nucleotide Polymorphism
TSH	Thyroid Stimulating Hormone
T1D	Type 1 Diabetes
T2D	Type 2 Diabetes
VNTR	Variable Number of Tandem Repeats
WT	Wild Type

Abstract

Carboxyl ester lipase (CEL) is a digestive enzyme that breaks down dietary fats in the duodenum, and it is mainly expressed by the pancreas and lactating mammary glands. A pathogenic variant of the *CEL* gene, named *CEL-HYB1*, is established as a risk factor for chronic pancreatitis and increases the disease risk by 5-fold in the European population. Interestingly, new data from our research group suggest that *CEL-HYB1* is a risk factor for chronic pancreatitis that predominantly affect women. Furthermore, there are indications that the T-allele of a SNP present in the *CEL* promoter (rs541856133) is associated with *CEL-HYB1*. The same SNP variant may also increase the risk for developing diabetes.

This thesis aimed to understand the effect of the promoter SNP rs541856133 (C>T) on *CEL* gene expression, and to investigate whether CEL could be regulated by female sex hormones. Moreover, we wanted to explore the role of CEL outside of the digestive system, particularly in the pituitary gland.

By cellular models and a dual-reporter luciferase assay, we found that the *CEL* MUT promoter (rs541856133 (T)) exhibited increased activity compared to the *CEL* WT promoter (rs541856133 (C)) in HEK293 cells, but not in ZR-75-1 or HeLa cells. *In silico* analysis of the *CEL* promoter identified the zinc finger protein ZNF331 as a potential repressor predicted to bind to the SNP site, and with reduced binding affinity towards the *CEL* MUT promoter. However, ZNF331 was found to repress the activity of both promoter variants in HEK293 cells, with no effect observed in ZR-75-1 cells.

The *in silico* analysis also revealed putative estrogen receptor (ER) binding sites within the *CEL* promoter. In ZR-75-1 cells, ERα overexpression led to repression of the *CEL* promoter activity, while treatment with the ER modulator Tamoxifen not only reversed this effect but also enhanced the promoter activity. These findings may suggest a regulatory role for estrogen on *CEL* gene expression. Immunohistochemical analysis further confirmed the expression of the CEL protein in the human pituitary gland and that it co-localizes with the thyroid-stimulating hormone (TSH), possibly linking CEL to hormonal pathways.

Based on varying results in HEK293, ZR-75-1 and HeLa cells, we cannot conclude whether the *CEL* promoter SNP rs541856133 (T) significantly impacts *CEL* expression or if it is simply a tag associated with *CEL-HYB1* without affecting gene expression Furthermore, our findings

support a regulatory role for estrogen on CEL expression and highlights the enzyme's potential involvement in endocrine functions through its presence in the pituitary gland.

1. Introduction

1.1 The human pancreas

The human pancreas is a glandular organ located in the upper part of the abdomen (**Figure 1.1**). It possesses an elongated structure, lying transversely behind the stomach, between the spleen and the duodenum. In healthy adults, the pancreas is typically 15-20 cm in length and weights around 80 g (1, 2). The organ can be divided into four anatomical parts: head, neck, body, and tail, with the duodenum curving around its head and the pancreatic tail lying adjacent to the spleen (3). The celiac and superior mesenteric arteries supply the pancreas with blood, while the splenic vein and the superior mesenteric vein provide drainage into the portal vein (1).

The pancreas serves a dual role in the body, having both an endocrine and exocrine function. The endocrine pancreas secretes hormones mainly involved in the regulation of blood glucose homeostasis, including insulin and glucagon. The exocrine part produces digestive enzymes, which are important for nutrient breakdown and absorption in the duodenum (3).

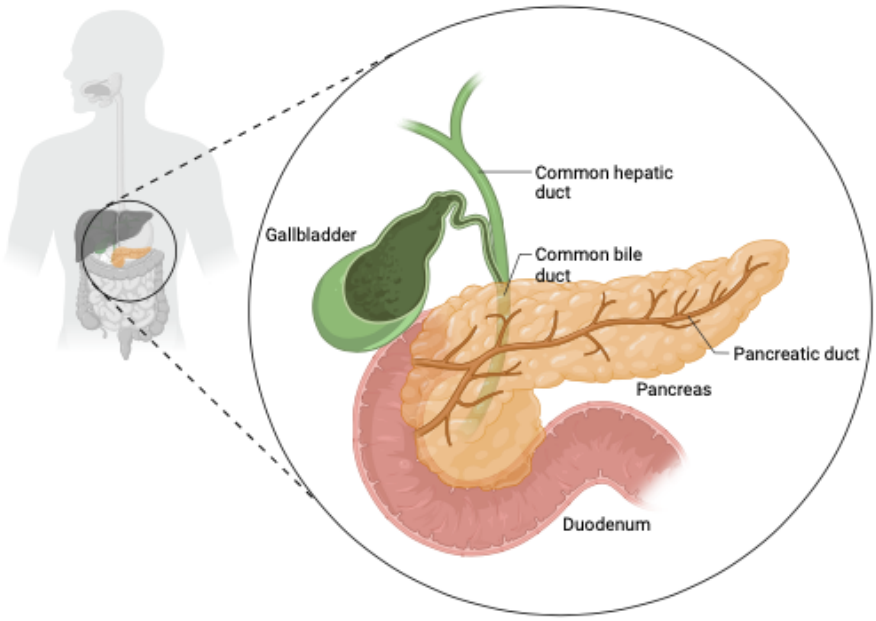


Figure 1.1 Location of the human pancreas. The pancreas is an elongated, glandular organ situated in the upper part of the abdomen, behind the stomach. Adapted from "Pancreas Anatomy", by BioRender.com (2024). Retrieved from https://app.biorender.com/biorender-templates

1.1.1 The endocrine pancreas

The endocrine activity of the pancreas is carried out by clusters of pancreatic cells, called the islets of Langerhans. The islets, which make up about 5% of the total pancreatic volume and are scattered throughout the organ, consist of at least five different hormone producing cells (**Figure 1.2**) (2). The most abundant are the alpha- and beta cells, constituting approximately

20-40% and 50-70% of the cells within human islets, respectively (**4**). Alpha cells are responsible for the production and release of glucagon when blood glucose levels decrease, while beta cells produce and release insulin in response to elevated blood glucose levels. The roles of both alpha- and beta cells are therefore important in maintaining blood glucose homeostasis (2). The remaining cells are delta cells, pancreatic polypeptide cells and epsilon cells that secretes somatostatin, pancreatic polypeptides and ghrelin respectively (**4**). Somatostatin negatively regulates glucagon, insulin, and pancreatic polypeptide secretion. Ghrelin suppresses insulin secretion (5), while pancreatic polypeptide inhibits glucagon and slows gastric emptying (6).

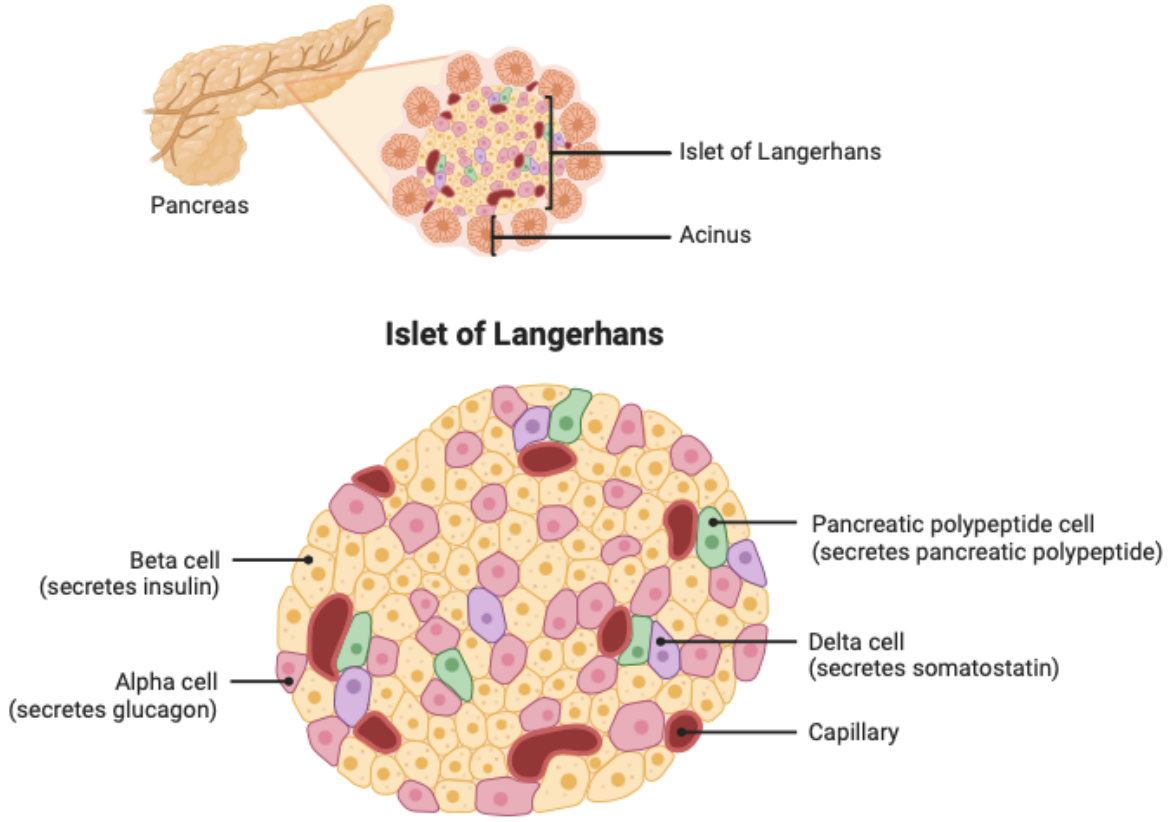


Figure 1.2 Organization of the endocrine pancreas. The endocrine part of the pancreas consists of clusters of pancreatic cells called Islet of Langerhans. Each islet is composed of multiple hormone-producing cells, including beta-cells, alpha-cells, delta-cells, pancreatic polypeptide-cells, and epsilon-cells. Adapted from "Pancreatic Islet of Langerhans" by BioRender.com (2024). Retrieved from https://app.biorender.com/biorender-templates

1.1.2 The exocrine pancreas

Approximately 95% of the pancreatic mass is devoted to its exocrine function (2, 7). The exocrine tissue is structured into lobules, composed of multiple acini. Each acinus comprises a single layer of pyramidal acinar cells, forming a central lumen that connects to a ductal system through centroacinar cells (**Figure 1.3**) (8). Together, the acinus and its draining duct form the

functional exocrine units responsible for producing and releasing enzymes critical for food digestion (9). Among these enzymes are amylase, lipase and protease, which break down carbohydrates, fat, and proteins respectively (10).

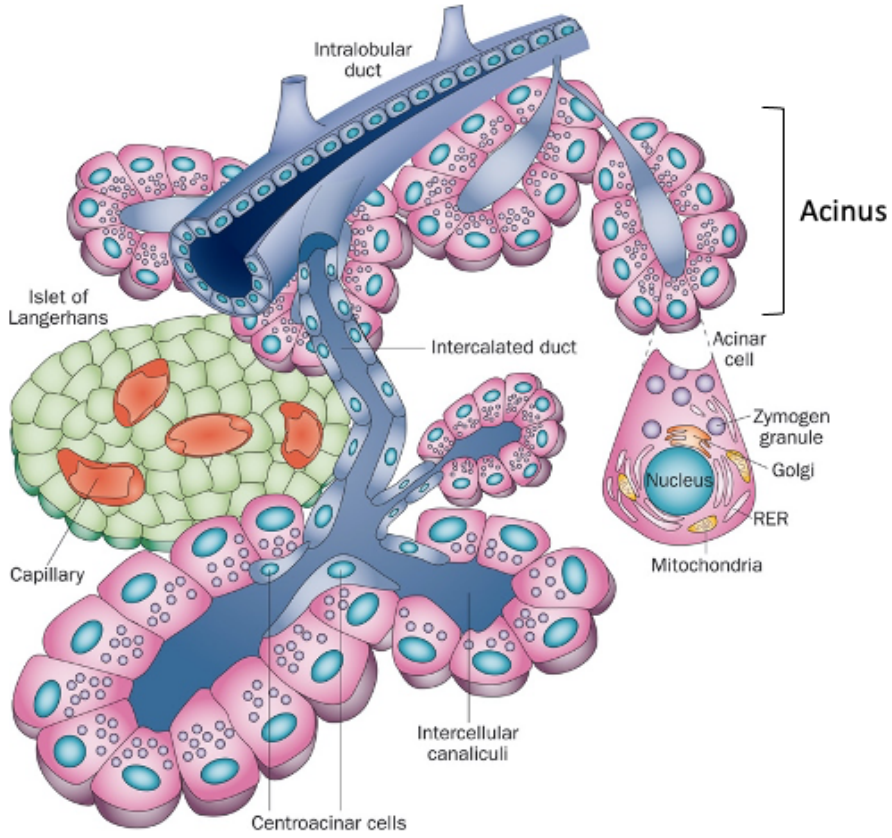


Figure 1.3 Organization of the exocrine pancreas. The functional unit of the exocrine pancreas is the acinus. An acinus consists of acinar cells forming a central lumen, connecting to a ductal system though centroacinar cells. Acinar cells produce, store, and secrete digestive enzymes crucial for the breakdown of food in the duodenum. The enzymes are released into intercalated ducts, where they mix with a bicarbonate-rich fluid produced by the centroacinar and ductal cells. Intercalated ducts connect to larger intralobular and interlobular ducts, which eventually converge into the main pancreatic duct. Retrieved from ref (11).

The digestive enzymes are produced by the acinar cells and stored as inactive proteins (proenzymes) in order to protect the pancreas from self-digestion. The proenzymes are stored in zymogen granules positioned within the acinar cell´s apical region, facing the lumen of the acinus (**Figure 1.3**) (12). These proenzymes are released into small intercalated ducts through exocytosis, a process regulated by neuroendocrine factors such as cholecystokinin (13). The enzyme-rich fluid mixes with a bicarbonate-rich fluid produced by the centroacinar and ductal cells, resulting in the formation of pancreatic juice (8). The secretion of bicarbonate are regulated by the hormone secretin (13). The pancreatic juice travels through the intercalated ducts leading to larger intralobular and interlobular ducts, which eventually converge to form

the main pancreatic duct. The main pancreatic duct, named the duct of Wirsung, spans the entire pancreas and delivers the pancreatic juice into the duodenum, where it combines with bile from the liver (**4**). Here, the bicarbonate neutralizes the gastric acid, thereby creating an optimal pH for digestive enzyme activity.

1.2 Diseases of the pancreas

The pancreas is susceptible to several diseases, including diabetes mellitus, pancreatitis, and cancer (2). Pancreatic cancer is characterized by a very poor prognosis and stands as the 7th leading cause of cancer-related deaths worldwide (14). However, pancreatic cancer is not further discussed in this thesis as the focus will be on diabetes mellitus and pancreatitis. These two diseases affect millions of people worldwide across various age groups and negatively impacts the patients' quality of life (15, 16). Furthermore, these diseases often result in long-term complications that impose a considerable financial burden on healthcare systems (17, 18). A deeper understanding of these conditions is therefore essential for more effective disease prevention and management, and to reduce their economic impact.

1.2.1 Diabetes Mellitus

Diabetes mellitus, often referred to as diabetes, is a group of metabolic diseases characterized by elevated blood glucose levels or hyperglycemia (19). The increase in blood sugar is caused by absolute or relative deprivation in insulin production and/or reduced insulin effectiveness (insulin resistance) (19). Common symptoms of hyperglycemia include thirst, frequent urination, and blurred vision (20). If left untreated, chronic hyperglycemia can lead to long-term microvascular and macrovascular complications, including vision loss, cardiovascular diseases, and nerve damage (21). The disease is typically diagnosed through several blood tests that measure blood glucose levels, including fasting plasma glucose and the oral glucose tolerance test (20).

Diabetes is categorized into several forms, with type 1 diabetes (T1D) and type 2 diabetes (T2D) being the most common, accounting for 5-10% and 90-95% of all diabetes cases, respectively (20). Less common forms include gestational diabetes and monogenic diabetes (20). T1D is characterized by a deficiency or absence of insulin due to autoimmune destruction of the insulin-producing beta cells in the pancreas (22). Management of the disease requires continuous glucose monitoring and insulin therapy (23). Even though the discovery of insulin was over a century ago, there is still no cure of T1D (24). The disease can occur at any stage of

life, but it typically manifests at an early age in children and adolescents (23). One hallmark of T1D is the presence of circulating islet autoantibodies, which serve as diagnostic biomarkers to distinguish T1D from other diabetes subtypes (25). The exact etiology of T1D remains largely unknown, though it likely involves a combination of genetic predisposition and environmental factors (22).

T2D is characterized by relative insulin deficiency combined with insulin resistance, leading to hyperglycemia (20). This condition progresses gradually and tends to manifest later in life compared to T1D (26). Increased weight, age, poor nutrition and lack of physical activity combined with genetic predisposition increases the risk of developing the disease (20). The rising incidence of T2D worldwide is largely due to lifestyle changes and increasing rates of obesity, posing a significant public health concern (26, 27). Unlike T1D, blood glucose homeostasis can be effectively managed through exercise, diet, and weight loss in some patients with T2D (28). Additionally, various oral medications are available that help reverse the mechanisms leading to poor glycemic control in T2D (29). If these options are ineffective, insulin therapy may be necessary (29).

In contrast to T1D and T2D, which have a complex etiology, monogenic diabetes primarily results from variations in a single gene (30). The most prevalent forms of monogenic diabetes, accounting for up to 5% of all diabetes cases, include neonatal diabetes and Maturity-Onset Diabetes of the Young (MODY) (31). Neonatal diabetes appears within the first six months of an infant's life and can be transient, permanent, or syndromic (32). As the disease have been associated with mutations in over 20 different genes, genetic testing to identify the causative mutation is necessary for precise diagnosis and treatment (33). MODY, characterized by an autosomal dominant inheritance pattern, typically manifest before the age of 25 (34). Precise diagnosis and treatment of MODY also require genetic testing, as the disease has been associated with mutations in at least 14 different genes (34).

Gestational diabetes manifests as elevated blood glucose levels that develop or are first recognized during pregnancy (35). Unlike other forms of diabetes that typically require lifelong medical management, gestational diabetes often resolves after childbirth (36). However, individuals with gestational diabetes face an elevated risk of developing T2D later in life (37).

1.2.2 Pancreatitis

Pancreatitis refers to an inflammatory condition affecting the pancreas, which can be divided into two main types: acute and chronic pancreatitis (38). Acute pancreatitis (AP) is characterized by initial swelling and fluid retention in the pancreas (39). The disease has a sudden onset and usually presents as a mild and self-limiting disease. However, approximately 20% of the patients may experience moderate to severe pancreatitis, with different degrees of tissue necrosis and/or organ failure (39, 40). Common symptoms include upper abdominal pain, along with nausea and vomiting (41).

The most common causes of AP are alcohol abuse and gallstones, which together account for over 90% of the cases (41). Less common causes are metabolic disorders such as hypertriglyceridemia and hypercalcemia, genetic factors, and medical procedures like endoscopic retrograde cholangiopancreatography (39). Trauma, certain medications, and infections can also play a role (39). Diagnosis is based on classical abdominal symptoms, elevated serum levels of pancreatic enzymes and imaging of the abdominal area (40). Choice of treatment depends on the severity and the etiology of the disease, but generally involves fluid balance, pain relief, nutritional support, and monitoring of vital signs (39). The overall mortality rate for AP is approximately 3-10%, but for patients suffering from severe acute pancreatitis, the mortality rate increases to 36-50% (42).

Chronic pancreatitis (CP) is a progressive fibroinflammatory condition that evolves gradually over time, characterized by scarring and fibrosis of the pancreatic tissue, which may lead to permanent loss of both exocrine and endocrine function (43, 44). Additional complications such as pseudocysts formation, obstructions in the pancreatic and bile ducts, vascular complications, and duodenal obstruction can also arise (44). Moreover, CP increases the risk of developing pancreatic cancer (45). Patients with CP typically presents with abdominal pain, weight loss due to food malabsorption, and diabetes mellitus, resulting from the development of exocrine and endocrine insufficiency (41). The disease has a huge impact on patient's life, being associated with reduced quality of life as well as increased mortality rate when compared with the general population (44). Diagnosis is based on clinical symptoms and imaging of the abdominal area. While no curative treatment exists, the management of CP involves pain relief, control of complications, and lifestyle changes (44).

The etiology of CP is multifactorial, with excessive alcohol consumption being the most prevalent risk factor (46, 47). The TIGAR-O classification system categorizes other CP risk

factors into Toxic-Metabolic, Idiopathic, Genetic, Autoimmune, Recurrent and Severe Acute Pancreatitis, and Obstructive factors (44, 48). This thesis will focus on the genetic component underlying CP. There are several genes associated with increased disease risk (**Table 1.1**), and they can be grouped into three mechanistic pathways that indicate to their pathological effect, namely the trypsin-dependent, the misfolding dependent and the ductal pathway (49).

Table 1.1 Genetic risk factors in chronic pancreatitis*

Gene	Name	Pathway
CEL	Carboxyl ester lipase	Misfolding-dependent
CTRB1-2	Chymotrypsin B1 and B2	Misfolding-dependent
CPA1	Carboxypeptidase A1	Misfolding-dependent
PRSS1	Cationic trypsinogen	Trypsin-dependent Misfolding-dependent
PRSS2	Anionic trypsinogen	Trypsin-dependent
SPINK1	Trypsin inhibitor	Trypsin-dependent
CTRC	Chymotrypsin C	Trypsin-dependent
CFTR	Cystic fibrosis transmembrane conductance regulator	Ductal physiology
CLDN2	Claudin 2	Ductal physiology
CASR	Calcium-sensing receptor	Ductal physiology

*Based on ref (49).

The focus of this thesis is the digestive enzyme carboxyl ester lipase (*CEL*) gene that is associated with CP through the misfolding-dependent pathway as described in more detail below.

1.3 Carboxyl Ester Lipase

Carboxyl ester lipase (CEL) is a digestive enzyme important for breaking down dietary fat, cholesteryl esters, and fat-soluble vitamins in the duodenum (50). CEL is primarily expressed in the pancreatic acinar cells, accounting for approximately 4% of the proteins in pancreatic juice (51). Being one of the four major lipases secreted from the pancreas into the duodenum, CEL is the only lipase requiring bile salt for activation (50). For this reason, the enzyme is also known as bile-salt stimulated lipase (52) or bile-salt dependent lipase (53). Aside from the pancreas, CEL is also expressed in the lactating mammary glands, where it constitutes around 1-2% of the proteins in mother's milk (54, 55). Moreover, CEL have been identified in fetal liver (56), eosinophils (57), endothelial cells (58), macrophages (59) and in the pituitary gland (60). The latter will be examined in this thesis to explore CEL's potential role outside of the digestive tract.

1.3.1 The human *CEL* gene

The carboxyl ester lipase gene (*CEL*) is approximately 10 kilobases in size and is situated on chromosome band 9q34.13 (61). The gene comprises 11 exons, with the last exon having a variable number of tandem repeat (VNTR) region (**Figure 1.4**). The repeat region is made up of nearly identical 33-bp segments, which are rich in guanine and cytosine (62, 63). Among humans, 16 repeats are the most common VNTR-length, however the number of repeats range from 3-23, contributing to the high polymorphism of *CEL* (64-67). Positioned 11 kilobases downstream from *CEL* is a *CEL*-pseudogene called *CELP* (**Figure 1.4**). *CELP*, which is approximately 5 kilobases in size, lacks exon 2-7 compared to *CEL* but the remaining sequence is highly similar between the two genes (97%) (63, 68). Due to a premature stop codon in exon 8, *CELP* is not anticipated to undergo translation into a functional protein (69).

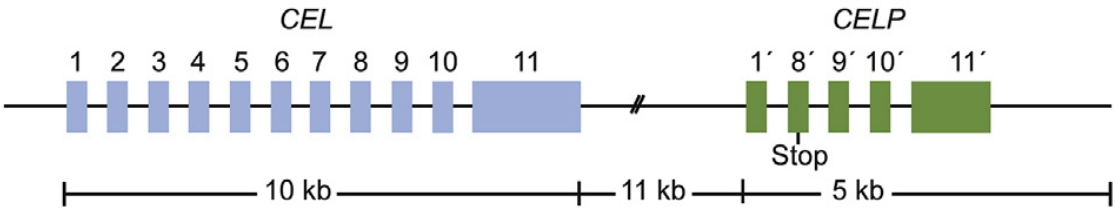


Figure 1.4 The *CEL-CELP* locus. The *CEL* gene, comprising 11 exons, spans a region of 10 kilobases. The *CEL*-like pseudogene *CELP*, located 11 kilobases downstream of *CEL*, is 5 kilobases in length and lacks exon 2-5 compared to *CEL*. Adapted from ref (50).

1.3.2 The CEL protein

The CEL protein comprises a globular N-terminal domain (corresponding to exons 1-10) and a C-terminal with an intrinsically disordered region of varying length (**Figure 1.5**) (50). The globular domain includes a signal peptide, binding sites for bile-salt and the catalytic site (70, 71). The intrinsically disordered region is characterized by a repetitive 11-amino acid sequence, encoded by the 33-bp segments of the *CEL* VNTR region in exon 11 (62). The VNTR-encoded amino acid sequence is enriched in proline, glutamate, serine, and threonine (72), commonly referred to as a "PEST sequence" (73). A VNTR region of 16 repeats results in a protein with 722 amino acids and a theoretical mass of 79 kDa. In addition to variability in the number of VNTR repeats, the molecular weight of the protein fluctuates due to different posttranslational modifications (50).

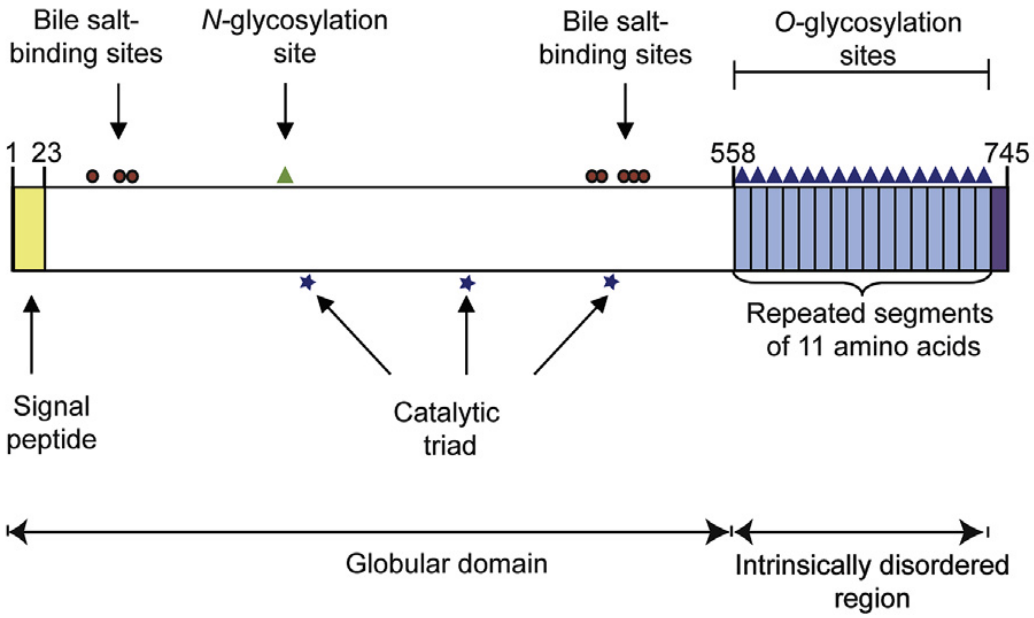


Figure 1.5 The CEL protein. The protein comprises a globular N-terminal and an intrinsically disordered region at the C-terminal. The globular domain includes a signal peptide, binding sites for bile-salt, a *N*-glycosylation site, and the catalytic triad (Ser194, His435, Asp320), all marked with arrows. The intrinsically disordered region consists of a repetitive 11-amino acid sequence, here shown with the most common number of repeats, 16. The repeated amino-acid segments contain multiple *O*-glycosylation sites. Adapted from ref (50).

The secretion of CEL in acinar cells involves a classical secretory pathway, with the protein traveling from the endoplasmic reticulum (ER) to the Golgi-apparatus before its eventual secretion. The process begins with the cleavage of the N-terminal hydrophobic signal peptide, allowing CEL to be transferred to the ER, where it forms a folding complex with several chaperons (50, 74). During translation in the ER, *N*-glycosylation occurs at residue Asn187, a modification important for correct folding and secretion of CEL (75). CEL is then relocated to the *cis*-Golgi with the help of a protein complex, including the chaperone glucose-regulated protein 94 (76). Here, CEL undergoes substantial *O*-glycosylation at the serine and threonine residues within the C-terminal VNTR region (77). The PEST sequence encoded by the VNTR could serve as a signal for rapid protein degradation, and *O*-glycosylation may therefore increase the proteins stability by masking the PEST sequence (73, 78). After being fully *O*-glycosylated, CEL is transferred to the *trans*-Golgi, where it is phosphorylated at Thr340 (79). Finally, CEL is released from the Golgi apparatus and co-stored with other digestive enzymes in zymogen granules until they are collectively secreted into the duodenal lumen (80).

1.3.3 Regulation of *CEL* transcription

Our current understanding of *CEL* gene regulation is based only on a limited number of studies. Classical elements such as the TATA box, crucial for tissue-specific expression, have been identified in the *CEL* gene promoter (**Figure 1.6**) (63, 81-83).

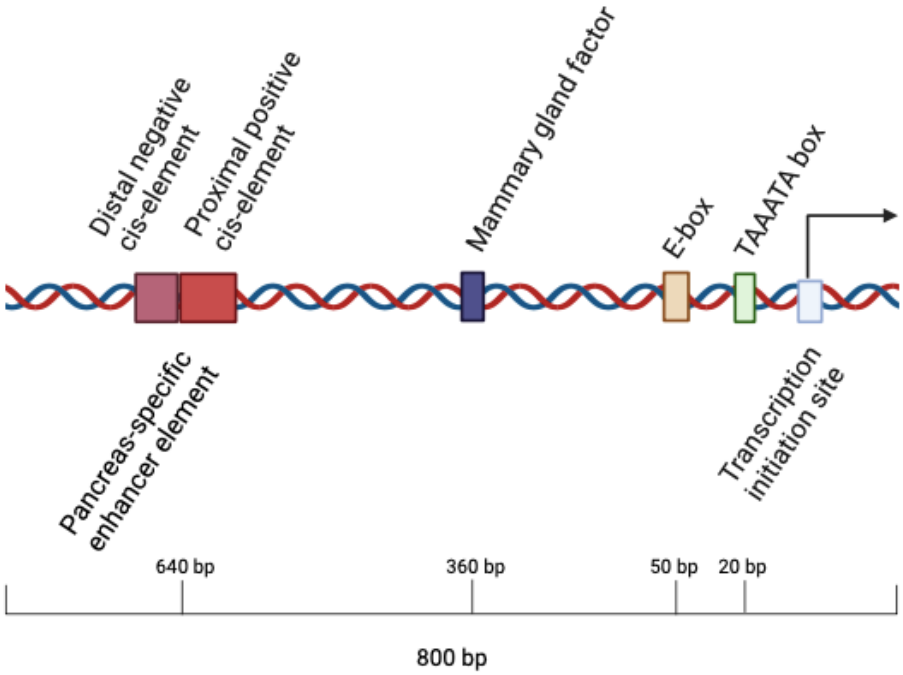


Figure 1.6 Illustration of regulatory elements within the human *CEL* promoter region. The region comprises a consensus sequence recognized by the mammary gland factor, and a pancreas-specific enhancer composed of proximal and distal cis-elements. The promoter also includes an E-box and a TAAATA box. The presented *CEL* promoter region spans approximately 800 base pairs, with the positions of the regulatory elements marked relative to the transcription initiation site. Figure created with BioRender.com, with annotations derived from ref (63, 82-85). The elements are presented schematically and not to scale.

The regulation of *CEL* expression varies across different tissues (84). For instance, high-level expression of *CEL* in the pancreas relies on a pancreas-specific enhancer element, comprising distal negative and proximal positive elements (**Figure 1.6**) (84, 86). None of these cis-elements share homology with other known cis-elements, and the transcription factors involved in the enhancer activity are currently unknown (84). The *CEL* promoter also includes a sequence resembling the consensus sequence of a tissue-specific mammary gland factor (**Figure 1.6**), along with a conserved 11-bp sequence present in other genes encoding milk proteins (63). In human macrophages, *CEL* expression is regulated by an E-box domain (**Figure 1.6**) which interacts with upstream stimulatory factors 1 and 2, in addition to other necessary receptor binding sites (85). The E-box have also been implicated in *CEL* expression in the pancreas and mammary glands, but other unknown tissue-specific elements are also required (85).

1.4 The polymorphic nature of *CEL*

CEL is a highly polymorphic gene, showcasing four categories of genetic variation. In addition to single nucleotide polymorphisms (SNPs) that spans the *CEL* gene locus, VNTR length alterations, single-base pair insertions and deletions within the VNTR and copy number variants (CNV) have been described (50).

1.4.1 VNTR lengths

As previously mentioned, the number of VNTR repeats in exon 11 of the *CEL* gene can vary from 3 to 23 in humans, with 16 repeats being the most prevalent (64). Variation between species has also been observed, with some species such as chicken, cod, and lizards having no repeats, while the gorillas have 39 repeats (87). In humans, shorter VNTR lengths have been associated with diseases such as diabetes and idiopathic chronic pancreatitis (64, 88).

1.4.2 Insertions

Single-base pair insertions located within the VNTR region of *CEL* cause frame-shifts, leading to the production of truncated CEL proteins (50). These insertions have been identified in the 4^{th}, 9^{th}, 10^{th}, 11^{th} and 12^{th} repeat of the *CEL* VNTR, and they generate a stop codon within the same repeat where the mutation occurs (50). The combined allele frequency of these insertions is reported to be 0.07 in normal controls of Northern European descent, suggesting that most of them are benign variants (66, 89). Interestingly, insertions involving a single cytosine residue introduce a novel C-terminal sequence (PRAAHG) in the truncated proteins, distinct from the normal sequence (PAVIRF) (90).

1.4.3 Deletions

Deletions in the 1^{st} (DEL1) or 4^{th} (DEL4) repeat of the VNTR were first discovered in two Norwegian families with suspected monogenic diabetes. The proteins generated by these deletions are predicted to have a different and shorter C-terminal sequence, leading to altered chemical properties and a reduced number of *O*-glycosylation sites (50). Both DEL1 and DEL4 have been identified as causative mutations of CEL-MODY (66), a condition that will be elaborated further in **Section 1.5.1.**

1.4.4 CNV

Three copy number variants involving *CEL* and its pseudogene *CELP* have been reported (64). These variants have most likely arisen through non-allelic homologues recombination between *CEL* and *CELP* (**Figure 1.7**) (91). The initial duplication allele discovered, *CEL-DUP1*, harbors

a premature stop codon in exon 8, and is therefore unlikely to translate into a functional protein (91). In contrast, the reciprocal deletion variant of *CEL-DUP1,* namely *CEL-HYB1*, encodes a functional chimeric protein that is associated with increased risk for developing chronic pancreatitis (91). *CEL-HYB1* is the focus of this thesis and will be further described below.

More recently, another duplication variant (*CEL-DUP2)* was reported, which contains an additional copy of the complete *CEL* gene. The corresponding postulated reciprocal deletion variant has not yet been identified, but is believed to only contain the *CELP* sequence (92).

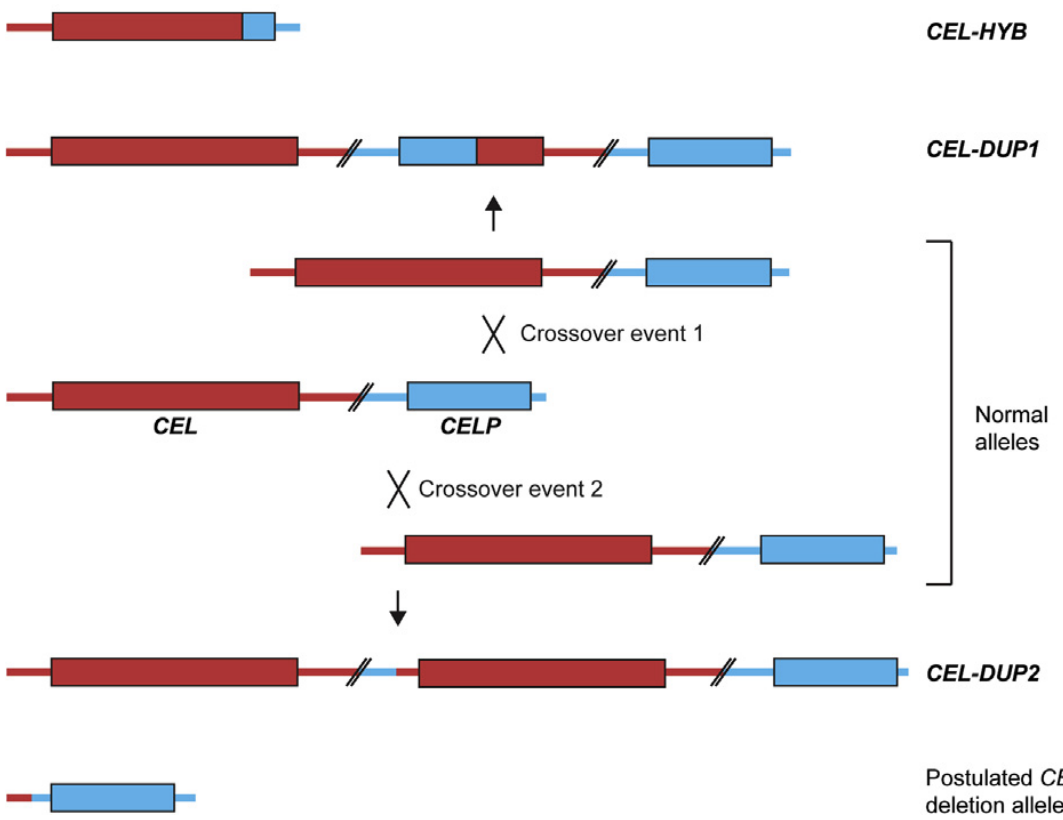


Figure 1.7 Copy number variants identified in the *CEL* locus. The duplication allele *CEL-DUP1* and the deletion allele *CEL-HYB*1 are thought to originate from a crossover event in the intron 9 or exon 10-11 region of *CEL* and *CELP*. Similarly, the duplication allele *CEL-DUP2* and its postulated reciprocal deletion allele presumably result from a crossover event in the promoter regions of *CEL* and *CELP*. Retrieved from ref (92).

1.5 *CEL* in pancreatic disease

1.5.1 *CEL* and monogenic diabetes

CEL-MODY is a rare, dominantly inherited syndrome characterized by monogenic diabetes and exocrine dysfunction (66). The CEL-MODY patients also present with fatty replacement of the pancreatic parenchyma and development of pancreatic cysts (93). The disease was initially discovered in two Norwegian families with single-base deletions in the 1st and 4th CEL

VNTR repeats (66). In recent years, CEL-MODY-like phenotypes have also been identified in Swedish (94), Czech (94), and Italian families (95), with single-base pair deletions in the 1st, 4th, and 5th VNTR repeats, respectively.

Our research group has demonstrated that the CEL-MODY protein tends to form aggregates both intracellularly and extracellularly when overexpressed in cellular systems (96). Additionally, the mutated proteins exhibit reduced secretion compared to normal CEL, and its accumulation within the cells triggers ER stress and apoptosis (97). Moreover, after secretion, CEL-MODY proteins can be internalized via endocytosis in beta cells, which might negatively influence the endocrine function of the pancreas (98, 99).

1.5.2 *CEL* and chronic pancreatitis

In 2015, our research group identified *CEL-HYB1* as a genetic risk factor for CP (91). This variant was overrepresented more than five-fold in individuals with idiopathic pancreatitis compared to healthy controls across three independent European pancreatitis cohorts. *CEL-HYB1* has also been analyzed in three large Asian cohorts. Here, the *CEL-HYB1* allele was not detected suggesting that it is an ethnicity-specific risk factor for CP (100). The *CEL-HYB1* gene encodes a chimeric CEL protein that spans 589 amino acids, comprising the functional catalytic core domain from CEL and a C-terminal with only 3 VNTR repeats from CELP (**Figure 1.8**) (101).

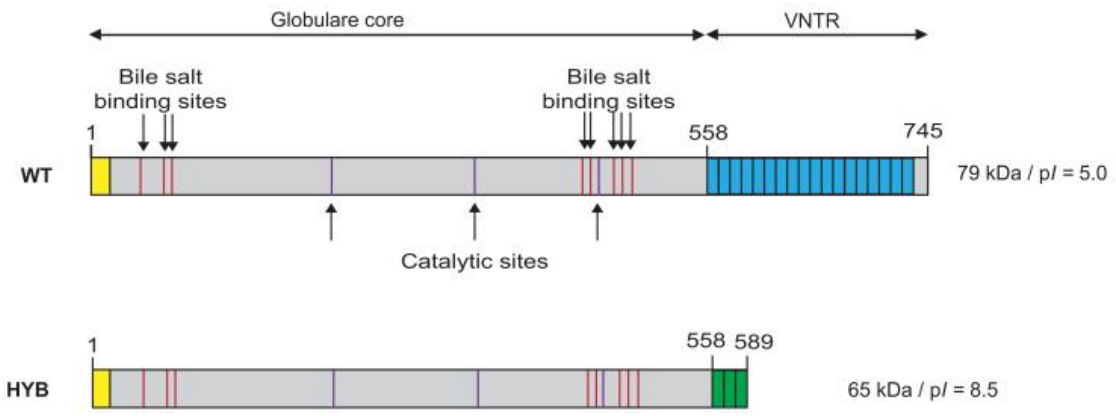


Figure 1.8 The protein structures of CEL and CEL-HYB1. The chimeric CEL-HYB1 protein consists of the globular N-terminal domain from CEL, and the VNTR region at the C-terminal domain from CELP. Thus, CEL-HYB1 has only three VNTR repeats, in contrast to the sixteen repeats found in the most common CEL protein. Adapted from ref (101).

The CEL-HYB1 protein shows similar disease mechanisms as seen with CEL-MODY, but with a less severe phenotype (50). When expressed in cultured cells, CEL-HYB1 exhibits reduced

enzymatic activity and decreased secretion compared to normal CEL (**Figure 1.9**) (102, 103). Additionally, cellular studies demonstrated that CEL-HYB1 forms intracellular aggregates, increases ER-stress, and induces autophagy (103, 104).

To study CEL-HYB1 at the organ level, our research group developed an animal model expressing a humanized version of CEL-HYB1 in the mouse pancreas (105). The model revealed that the *CEL-HYB1* mice spontaneously develop CP, characterized by inflammation, acinar cell atrophy and fatty replacement. The mouse model also confirmed the presence of CEL aggregates and upregulations of ER stress markers, supporting that CEL-HYB1 belongs to the protein-misfolding pathway of CP (103-105).

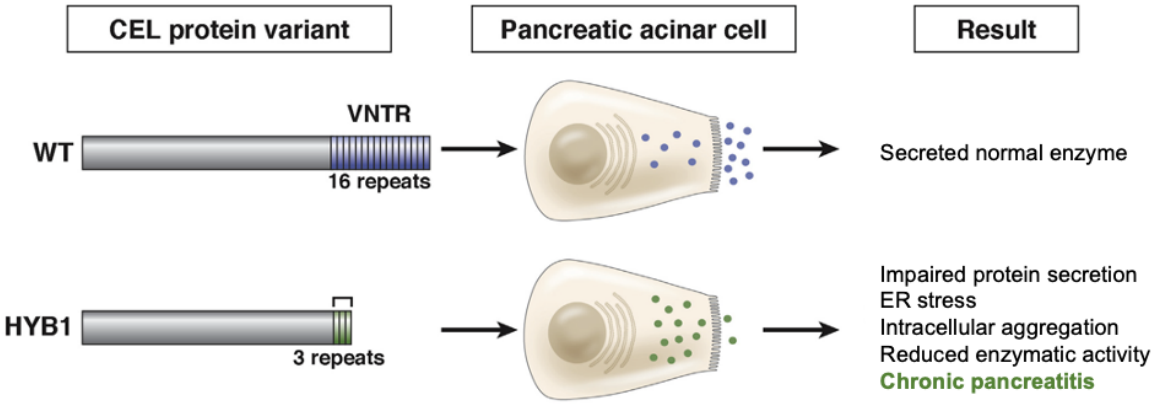


Figure 1.9 CEL-HYB1 secretion. The CEL-HYB1 variant with a truncated VNTR region of only 3 repeats shows impaired protein secretion, reduced enzymatic activity, and is associated with increased risk of developing chronic pancreatitis compared to the wild-type CEL protein. Adapted from ref (106).

Since *CEL-HYB1* is carried by up to 1.0% of the general European population, many carriers do not develop CP, suggesting that additional genetic or environmental factors are necessary to develop the disease (50, 104).

Interestingly, a re-examination of data from the original *CEL-HYB1* discovery paper (101) and revealed that 73% (8/11) of the identified *CEL-HYB1* carriers in the original CP cohort from Leipzig, Germany, were women (n=260, unpublished). Then, in a Norwegian CP material (n=200), we found two *CEL-HYB1* positive, and they were both females (104). To follow up on this gender pattern, we have initiated a collaboration with colleagues at the University of Greifswald, Germany. Their diagnostic lab has been screening CP cases for *CEL-HYB1* since 2017, and they report that 88% (22/25) of all *CEL-HYB1* positive carriers are females (unpublished). These findings strongly suggest that *CEL-HYB1* is a gender specific risk factor

for CP, leading us to hypothesize that *CEL* gene expression may be influenced by female sex hormones.

1.5.3 *CEL* and type 1 diabetes

In 2021, a genome-wide association study (GWAS) involving T1D patients and controls of European ancestry, identified novel genetic risk factors for T1D (107). Among them was the SNP rs541856133 (C>T) located in the *CEL* promoter. The risk variant increases the risk of developing T1D by 3-fold (107).

Given the established link between CEL and diabetes in CEL-MODY (66), and the association between the CEL promoter SNP rs541856133 (C>T) and T1D (107), we sought to also screen for *CEL-HYB1* in individuals diagnosed with T1D. Interestingly, and similar to SNP rs541856133 (C>T), we observed a 3-fold increase in the prevalence of the *CEL-HYB1* gene among T1D patients when compared to healthy controls (**Table 1.2**) (unpublished).

Table 1.2 *CEL-HYB1* positive carriers in Norwegian T1D patients and controls

Materials	*CEL-HYB1* positive / sample size	Carrier frequency
Norwegian Childhood Diabetes Registry	28 / 2937	1.0%
Norwegian Healthy Controls	8 / 2516	0.3%

*As of 2022, The Norwegian Childhood Diabetes Registry includes 3206 cases of children diagnosed with T1D (108)

Moreover, preliminary data from our group indicates that there is a link between the SNP rs541856133 (C<T) and the *CEL-HYB1* gene, as all carriers of the *CEL-HYB1* allele also carry the rs541856133 (T) variant (unpublished).

2. Aims of the study

The overall objective of this thesis was to learn more about the role of CEL in health and disease by investigating the promoter region and the expression of the *CEL* gene. We also sought to explore the role CEL outside the digestive system.

The specific aims were:

- To determine the effect of the *CEL* promoter SNP rs541856133 (C>T) on *CEL* gene expression in cellular systems.
- To search for new transcription factor binding sites near rs541856133 in the *CEL* promoter, and to analyze their potential effect on *CEL* promoter activity.
- To investigate whether *CEL* gene expression could be regulated by female sex hormones
- To examine CEL protein expression in the pituitary gland.

3. Materials

Table 3.1 Plasmids

Plasmid (Name in the thesis)	**Encoding**	**Catalog number**	**Supplier**
pCMV6-XL5-ZNF331 untagged (pCMV6-ZNF331)	Human zinc finger protein 331 (ZNF331), transcript variant 1	SC111333	OriGene
pEZX-GA01 (pEZX-empty)	*Gaussia* Luciferase (GLuc) and Secreted Alkaline Phosphatase (SEAP) reporter proteins	ZX103	GeneCopeia Inc.
Human CEL promotor reporter clone in vector pEZX-PG04 (pEZX-hCELp)	*Gaussia* Luciferase (GLuc) and Secreted Alkaline Phosphatase (SEAP) reporter proteins	HPRM43894-PG04	GeneCopeia Inc.
pCMV6-XL4-empty (pCMV6-empty)	Empty vector	PCMV6XL4	OriGene
pSG5-empty	Empty vector	Gift from E. Treuter and Hormonlaboratoriet	
pSG5-ERα	Estrogen Receptor α		
pmaxGFP	Green Fluorescent Protein	V4XC-2024	Amaxa

Table 3.2 Transformation and plasmid purification

Product	**Catalog number**	**Supplier**
ImMedia™ Amp Blue	45-0038	Invitrogen
ImMedia™ Kan Agar	45-0043	Invitrogen
LB Broth (Lennox)	L7275-500TAB	Sigma-Aldrich
Ampicilin sodium salt	A9518-5G	Sigma-Aldrich
Kanamycin sulfate	60615-5G	Sigma-Aldrich
TE buffer (1X), pH 8	A0386, 0500	PanReac, AplliChem
QIAfilter™ Plasmid Maxi Kit (25)	12263	QIAGEN
OneShot TOP10 Chemically Competent *E.coli* cells	C4040-03	Invitrogen
SOC-medium	15544-034	Invitrogen
QIAprep® Spin Miniprep Kit (250)	27106	QIAGEN

Table 3.3 Cell culture and transfection

Product	**Catalog number**	**Supplier**
HEK293 cells	632180	Clontech
HeLa cells	CCL-2	ATCC
ZR-75-1 cells	CRL-1500™	ATCC / prof. Stian Knappskog
Insulin-Transferrin-Selenium-Sodium Pyruvate (ITS-A) (100X)	51300044	Gibco
RPMI 1640 Medium + GlutaMAX™-I (1X)	61870036	Gibco
DMEM (1X) + GlutaMAX™-I Dulbecco’s Modified Eagle Medium	31966-021	Gibco
0.05% Trypsin-EDTA (1X)	25300-054	Gibco
Dulbecco's Phosphate Buffered Saline	D8537-500ML	Sigma-Aldrich
Dimethyl sulfoxide	D8418-50ML	Sigma-Aldrich
Fetal Bovine Serum	F7524-500ML	Sigma-Aldrich
Antibiotic Antimycotic	15240062	ThermoFisher Scientific

4-Hydrotamoxifen	94873	Supelco
Fulvestrant	I4409	Sigma-Aldrich
OPTI-MEM® I (1X) Reduced Serum Medium	31985-062	Gibco
Lipofectamine™ 3000 Transfection Reagent	L3000001	Invitrogen

Table 3.4 Luciferase assay, Mutagenesis, and Sanger Sequencing

Product	Catalog number	Supplier
Secrete-Pair™ Dual Luminescence Assay Kit	LF033	GeneCopoeia
QuikChange II XL Site-Directed Mutagenesis Kit	200522-5	Agilent Technologies
BigDye™ Terminator v3.1 Cycle Sequencing Kit	4337455	Applied Biosystems
BigDye™ Terminator v1.1 & v3.1 5X Sequencing Buffer	4336697	Applied Biosystems
Sephadex® G-50 Superfine	G5050-50G	Sigma-Aldrich
MultiScreen®-HV 96-Well Filter Plate	MAHVN4510	Millipore
MultiScreen Column Loader	MACL09645	Millipore

Table 3.5 SDS-PAGE and Western Blot

Product	Catalog number	Supplier
RIPA Lysis Buffer, 10X	20-188	Millipore
NuPAGE® MOPS SDS Running Buffer (20X)	NP0001-02	Invitrogen
NuPAGE® 10%, Bis-Tris, 1.5 mm, Mini Protein Gel, 10-well	NP0301BOX	Invitrogen
cOmplete Protease inhibitor cocktail tablets	11697498001	Roche
Pierce™ BCA Protein Assay Kit	23225	Thermo Scientific
Nonfat Dry Milk	9999S	Cell Signaling Technology
Microplate, 96 well, PS, F-Bottom, Clear, 10 PCS./BAG	655101	Greiner Bio-One
Magic Mark XP Western Protein Standard	P/N LC5602	Invitrogen
Precision Plus Protein™ Dual Color Standards	161-0374	BioRad
NuPAGE™ Transfer Buffer (20X)	NP00061	Invitrogen
NuPAGE™ Sample Reducing Agent (10X)	NP0009	Invitrogen
NuPAGE™ LDS sample buffer (4X)	NP0007	Invitrogen
PVDF Transfer Membrane	88518	Thermo Scientific
Chromatography Paper, Grade 3MM CHR	3030-866	Whatman
Pierce™ ECL Plus Western Blotting Substrate	32132	Thermo Scientific

Table 3.6 Immunohistochemistry

Product	Catalog number	Supplier
Phosphate-Buffered Saline (PBS) Tablets	18912-014	Gibco
Tween® 20	P1379-250ML	Sigma-Aldrich
Protein Block Serum-Free Ready-To-Use	X0909	Dako
ImmEdge® Pen	H-4000	Vector laboratories
Hydrogen peroxide 30%	1.07298.0500	Suprapur
IHC Antigen Retrieval solution, 10X	00-4956-58	Invitrogen
Liquid DAB+ Substrate Chromogen System	K3468	Dako
Automation Hematoxylin Histological Staining Reagent	S3301	Dako
Mounting Medium	CS703	Dako

Table 3.7 Immunofluorescence

Product	Catalog number	Supplier
Triton® X-100	437002A	VWR
Normal Goat Serum Control	10000C	ThermoFisher Scientific
Poly-L-Lysine 0.01% Solution	P4832	Sigma-Aldrich
Image-iT™ Paraformaldehyde 4% in 0.1 M Phosphate Buffer, pH 7.4	I28800	Invitrogen
ProLong™ Gold Antifade Mountant with DNA Stain DAPI	P36935	Invitrogen
Normal Goat Serum	10189722	Invitrogen

Table 3.8 Antibodies

Product	Catalog number	Supplier	Method
MACH 3 Rabbit HRP Polymer Detection	M3R531H	Biocare Medical	Immunohistochemistry
Anti-ZNF331, Polyclonal antibody produced in rabbit	PA5-66560	Invitrogen	Immunofluorescence
Donkey anti-Rabbit IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Flour™ 488	A21206	Invitrogen	Immunofluorescence
Anti-CEL, Polyclonal antibody produced in rabbit	HPA052701-100ul	Sigma-Aldrich	Immunohistochemistry Immunofluorescence
Anti-CEL**,** Antibody produced in rabbit		Gift from St. Louis lab	Western Blot
Anti-β-actin, Monoclonal antibody produced in mouse	A5441	ThermoFisher	Western Blot
Goat anti-Mouse IgG (H+L) Secondary Antibody, HRP	62-6520	Invitrogen	Western Blot
Goat anti-Rabbit IgG (H+L) Secondary Antibody, HRP	65-6120	Invitrogen	Western Blot
Anti-TSH, Polyclonal antibody produced in rabbit, Alexa Flour® 647	BOSSBS-2676R-A647	VWR	Immunofluorescence
Goat anti-Rabbit IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Flour™ 488	A-11008	Invitrogen	Immunofluorescence

Table 3.9 Buffers and solutions

Buffers and solutions	Composition	Method
PBS-Tween 0.05% (PBS-T)	For 2 L: 4 tablets of PBS dissolved in 2 L MilliQ water + 1 ml Tween® 20	Immunohistochemistry Western Blot
Antibody dilution buffer, pH 7.4	0.05 M Tris + 0.15 M NaCl + 1% BSA + 0.05% Tween® 20	Immunohistochemistry
Wash buffer (PBS-T 0.1%)*	For 1 L: 2 tablets of PBS dissolved in 1 L MilliQ water + 1 ml Tween® 20	Immunofluorescence
Permeabilization solution*	For 15 ml: 15 µl Triton X-100 in 15 ml wash buffer	Immunofluorescence
Blocking solution (5% goat serum in PBS-T 0.1%)*	For 5 ml: 250 µl Normal Goat Serum Control in 5 ml wash buffer	Immunofluorescence
1X RIPA Lysis buffer	For 10 ml: 1 ml 10X RIPA Lysis buffer in 9 ml MilliQ water + 1 cOmplete Protease inhibitor cocktail tablet	Western Blot

1X NuPAGE® MOPS SDS Running Buffer	For 1 L: 50 ml 20X NuPAGE® MOPS SDS Running Buffer in 950 ml MilliQ water	Western Blot
1X NuPAGE® Transfer Buffer	For 1 L: 50 ml 20X NuPAGE® Transfer Buffer in 850 ml MilliQ water + 10% methanol	Western Blot
5% Dry milk	For 100 ml: 5 g Nonfat Dry Milk in 100 ml PBS-Tween 0.05%	Western Blot
PBS	For 1 L: 2 tablets of PBS dissolved in 1 L MilliQ water	Immunofluorescence
Blocking solution (10% goat serum in PBS)*	For 10 ml: 1 ml Normal Goat Serum in 9 ml PBS	Immunofluorescence

*Sterile filtered

Table 3.10 Technical equipment

Product	Supplier
NanoDrop™ One Microvolume UV-Vis Spectrophotometer	Thermo Scientific
ChemiDoc™ MP Imaging System	BioRad
Heraeus Multifuge 3S-R Centrifuge	Thermo Electron Corporation
LUNA-II™ Automated Cell Counter	Logos biosystems
Nikon TMS Inverted Phase Contrast Microscope	Nikon
FLUOstar OPTIMA	BMG LabTech
Veriti™ 96-Well Fast Thermal Cycler	Applied Biosystems
Centrifuge 5425	Eppendorf
3500xL Genetic Analyzer	Applied Biosystems
EVOS M5000 Imaging System	Invitrogen
Dragonfly High Speed Confocal Platform	Andor
Centrifuge 5424 R	Eppendorf
Thermomixer comfort	Eppendorf
PowerPac™ HC	BioRad
Gen5 2.06 software	Biotek
Leica TCS SP8 STED 3X	Leica Microsystems

Table 3.11 Analytical software

Product	Supplier
SnapGene Viewer	Dotmatics
QuikChange® Primer Design	Agilent
GraphPad Prism 10	Dotmatics
BLAST	NCBI
Fiji software	ImageJ
UCSC Genome Browser	UCSC
JASPAR 2024	JASPAR Database, ELIXIR
Human Protein Atlas version 23.0	Human Protein Atlas

4. Methods

4.1 Preparation of plasmids

4.1.1 Transformation of *E.coli* and plasmid purification

For transformation, 1 µl of plasmid was added to 25 µl of OneShot TOP10 Chemically Competent *E.coli* cells and kept on ice for 30 minutes. This was followed by a 30-second heat shock at 42 °C. Subsequently, 125 µl of S.O.C medium was added to the vial, and the mixture was incubated at 37 °C and 225 rpm for 1 hour. Next, 100 µl of the mixture was plated on agar plates containing either Ampicillin (100 µg/ml) or Kanamycin (50 µg/ml) as selection marker and incubated at 37 °C overnight.

The next day, a single colony was selected and inoculated in 4 ml of LB medium containing either Ampicilin (50 µg/ml) or Kanamycin (50 µg/ml). The starter culture was incubated at 37°C with vigorous shaking (300 rpm) for 8 hours before diluted 1:500 in LB medium containing the proper antibiotics. The diluted culture was further incubated as described above for approximately 16 hours. The next day, the bacterial cells were harvested by centrifugation at 4600 rpm and 4°C for 30 minutes. The supernatant was discarded, and plasmid was isolated from the bacterial pellet using the Qiagen Plasmid Maxi Kit protocol according to the manufacturers protocol. The DNA pellet was dissolved in 300 µl TE buffer. The DNA was kept at room temperature overnight, to ensure proper dissolving, before stored at -20°C.

4.1.2 Determination of plasmid concentration and quality

The concentration of the purified plasmids was determined by measuring the absorbance of the sample at 260 nm using the Thermo Scientific™ NanoDrop™ One Microvolume UV-Vis Spectrophotometer. The purity of the DNA sample is indicated by an A_{260}/A_{280} ratio close to 1.80 and an A_{260}/A_{230} ratio ranging from 2.0 to 2.2, according to Thermo Scientific (109).

4.2 Site-Directed Mutagenesis

4.2.1 Primer design and Polymerase Chain Reaction (PCR)

The *CEL*-mutant (rs541856133 (T)) promoter variant was generated using the QuikChange II XL Site-Directed Mutagenesis Kit and the GLuc-On promoter reporter vector containing the *CEL*-wildtype (rs541856133 (C)) promoter as template. Primers were designed using the QuikChange® Primer Design Program from Agilent. The mutagenesis master mix was prepared as presented in **Table 4.1**. The primers used are listed in **Table 4.2**, and the cycling

parameters for the mutagenesis reaction are described in **Table 4.3**. The PCR was performed using a Veriti™ 96-Well Fast Thermal Cycler.

Table 4.1 Master mix for site-directed mutagenesis

Reagent	Volume
10x Reaction buffer	5 µl
dsDNA template (10 ng)	0.5 µl
Primer 1 (125 ng)	1.4 µl
Primer 2 (125 ng)	1.4 µl
dNTP mix	1 µl
QuikSolution	3 µl
ddH_2O	37.7 µl
Pfu Ultra HF DNA polymerase (2.5 U/µl)	1 µl
Total	50 µl

Table 4.2 Primers for site-directed mutagenesis

Primer	Sequence (5' → 3')
SNP C-T (forward)	CAGACGTGGCTCTGTAGGTCCACTCGGTC
SNP C-T (reverse)	GACCGAGTGGACCTACAGAGCCACGTCTG

Table 4.3 Thermal cycling parameters

Temperature (°C)	Time	Cycles
95	1 minute	1
95	50 sec	18
60	50 sec	
68	9 minutes	
68	7 minutes	1
4	∞	1

After the PCR, 1 µl of *Dpn* I restriction enzyme was added to the sample to digest the parental nonmutated DNA. The mixture was centrifuged for 1 minute, incubated at 37 °C for 1 hour and then transformed into competent cell

4.2.2 Transformation and plasmid purification

To prepare the XL10-Gold Ultracompetent Cells for transformation, the cells were thawed gently on ice before 45 µl of the cells were transferred to prechilled 14-ml BD Falcon polypropylene round-bottom tubes. To increase transformation efficiency, 2 µl of β-mercaptoethanol was introduced to the cells, and the mixture was swirled gently and incubated on ice for 10 minutes.

Next, 2 µl of *Dpn* I-treated DNA was added to 45 µl of the XL10-Gold Ultracompetent Cells and incubated on ice for 30 minutes. The cells were then subjected to a 42 °C heat-shock for 30 seconds before incubated on ice for 2 minutes. Next, 0.5 ml of preheated S.O.C medium was

added to the tube, followed by 1 hour incubation at 37 °C with shaking at 225 rpm. The mixture was spread out on an agar plate containing kanamycin and incubated at 37 °C overnight.

Plasmid purification and DNA quality/concentration were performed as described in **Section 4.1.1** and **Section 4.1.2**, respectively.

4.2.3 Sanger Sequencing

Sanger sequencing was performed to verify if the mutagenesis was successful. The sequencing reactions were prepared in PCR-tubes as indicated in **Table 4.4**, with primers listed in **Table 4.5**. Thermal cycling was carried out using a Veriti™ 96-Well Fast Thermal Cycler, and the program used is shown in **Table 4.6.**

Table 4.4 Master mix for sequencing

Reagent	Volume
ddH_2O	3.4 µl
Primer (2 µM)	1.6 µl
Sequencing buffer 5X	2.0 µl
Big Dye	1.0 µl
dsDNA template (200 ng)	2.0 µl
Total	10 µl

Table 4.5 Primers for sequencing

Primer	Sequence (5' → 3')
Forward	GTCATTCTATTCTGGGGG
Reverse	TTGTTCTCGGTGGGCTTGGC
Fw1	GAGTGTGGTTGGTGTGGGAA
Fw2	CTTATTTCCTGTGGCCGAGC
Rv1	AATCCCAGAGCAGCACAAGG
Rv2	CCCTCCACTCATGTGGCTG

Table 4.6 Thermal cycling parameters

Temperature (°C)	Time	Cycles
96	1 minute	1
96	10 sec	35
50	5 sec	
60	4 minutes	
4	∞	1

Following the PCR, the sequencing products were purified using Sephadex G-50 columns. Prior to purification, 10 µl of sequencing products was diluted with 15 µl MilliQ water, then transferred to the Sephadex G-50 columns and centrifuged for 5 minutes at 910 xg. The purified sequencing products were then analyzed using a 3500xL Genetic Analyzer. For reading the DNA sequence, SnapGene Viewer and NCBI BLAST (110) were used.

4.3 Cell culturing and transfection

4.3.1 Culturing of cells

HeLa and HEK293 cells were cultured in high glucose Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% Fetal Bovine Serum (FBS) and 100 U/ml Antibiotic Antimycotic. ZR-75-1 cells were cultured in high glucose RPMI supplemented with 1% Insulin-Transferrin-Selenium-Sodium Pyruvate (ITS-A), 10% FBS and 100 U/ml Antibiotic Antimycotic. If not otherwise stated, the cells were grown in T75 cm^2 flasks and kept in a humidified atmosphere at 37 °C with 5% CO_2.

4.3.2 Passaging and seeding of cells

When passaging HeLa, HEK293, or ZR-75-1 cells, the growth medium was removed, and the cells were washed in 5 ml of pre-warmed Dulbecco's Phosphate Buffered Saline (PBS). Next, 1 ml of pre-warmed 0.05% Trypsin-EDTA (1X) was added to detach the cells from the flask. To halt trypsinization, 3 ml of pre-warmed growth medium was added to the cells. Subsequently, the cells were resuspended, and a preferred amount was transferred to a new T75 cm^2 flask containing pre-warmed cell medium, resulting in a total volume of 13 ml. For experiments requiring a specific number of cells, the cells were quantified using the LUNA-II™ Automated Cell Counter.

4.3.3 Cell freezing and thawing.

For freezing, HeLa, HEK293, and ZR-75-1 cells were cultured in T75 cm^2 flasks until they reached approximately 80% confluency, followed by trypsinization as described in **Section 4.3.2**. The cells were resuspended in 3 ml pre-warmed growth medium and transferred to a 15 ml Falcon tube, followed by centrifugation at 500 xg for 2 minutes. The supernatant was discarded, and the HeLa and HEK293 cells were resuspended in 3 ml freezing medium containing 10% Dimethyl sulfoxide (DMSO) and 90% growth medium (DMEM). The ZR-75-1 cells were resuspended in 3 ml freezing medium containing 5% DMSO and 95% growth medium (RPMI). Next, 1 ml of the cell suspension was transferred to one cryogenic vial and placed in an isopropanol box. The cells were kept at -80 °C overnight before transferred to liquid nitrogen for long-term storage.

Rapid thawing of the cells was performed by rubbing the cryogenic vials in the palm of the hand. The thawed HeLa cells were subsequently transferred to a T75 cm^2 flask, while the HEK293 cells were transferred to a T25 cm^2 flask, both with an adequate amount of pre-warmed growth medium. The next day, the medium was removed, and fresh growth medium was added to remove DMSO residues. The thawed ZR-75-1 cells were transferred to a 15 ml Falcon tube with 9 ml pre-warmed growth medium and centrifuged at 125 xg for 5 minutes. The supernatant was discarded, and the cells were resuspended in 5 ml of pre-warmed growth medium and transferred to a T25 cm^2 flask.

4.3.4 Transient transfection of HeLa, HEK293 and ZR-75-1 cells

Prior to transfection, HeLa and HEK293 cells were seeded in either 96-well plates (7500 cells per well) or 24-well plates (50 000 cells per well), whereas ZR-75-1 cells were seeded in 24-well plates (90 000 cells per well). After 24 hours, the medium was discarded and replaced with DMEM or RPMI without antibiotics.

The cells were transfected with Lipofectamine 3000 following the manufactures protocol. For transfection in 96-well plates, a mixture consisting of 100 ng of plasmid DNA, 10 µl OPTIMEM, 0.2 µl P3000 reagent, and 0.3 µl Lipofectamine 3000 was prepared for each well. To transfect cells in 24-well plates, the transfection mixture was adjusted to 500-600 ng of plasmid DNA, 50 µl OPTIMEM, 1 µl P3000 reagent, and 1.5 µl Lipofectamine 3000 per well.

The transfection mixtures were added to the cells and incubated for 4 hours. The medium was then discarded, and fresh growth medium without antibiotics was added. To measure transfection efficiency, the cells were grown for 24 hours post-transfection and then visualized using the EVOS M5000 Imaging System. For cells undergoing Luciferase assay analysis, the procedure is described below.

4.3.5 Tamoxifen and Fulvestrant treatment

ZR-75-1 cells were seeded and transfected as described in **Section 4.3.4**. Next day, the culture medium was replaced, and half of the cells were treated with either 5 µM of 4-Hydroxy-tamoxifen (Tamoxifen) or 100 nM of Fulvestrant. After 24 hours, Luciferase-assay was performed as described below.

4.4 Luciferase-assay

4.4.1 Principle

For comparative analysis of promoter activity between the wild-type and mutant variants of the *CEL* promoter, the Secrete-Pair™ Dual Luminescence Assay was used (111). This assay quantifies the activity of two secreted reporter proteins, *Gaussia* Luciferase (GLuc) and Secreted Alkaline Phosphatase (SEAP), encoded on the same plasmid, and harvested from the same cell culture medium. The *CEL* promoter is inserted upstream of the GLuc reporter gene in a GLuc-On promoter reporter vector and then transfected into cultured cells (**Figure 4.1**). The *CEL* promoter initiates transcription of the GLuc gene, leading to secretion of the GLuc enzyme into cell culture medium. The activity of GLuc, which is regulated by the *CEL* promoter activity, is measured by adding a luminescent substrate that reacts with GLuc to produce a light signal. This signal is proportional to the GLuc level, which is then indicative of the *CEL* promoter activity. SEAP activity, measured the same way as the GLuc activity, serves as an internal control of transfection efficiency since SEAP expression is constant and independent of the *CEL* promoter (**Figure 4.1**). Thus, the GLuc activity can be normalized to SEAP, to give a more accurate estimate of the *CEL* promoter activity.

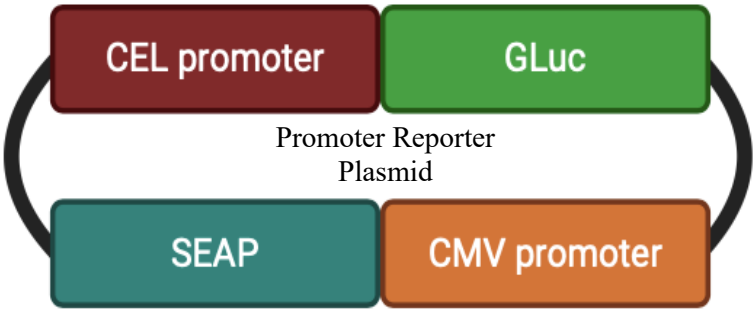


Figure 4.1 Simplified illustration of the GLuc-On promoter reporter plasmid. The *CEL* promoter drives the expression of the GLuc reporter gene, resulting in the secretion of GLuc proteins into the cell culture medium. SEAP expression, controlled by the CMV promoter, serves as an internal control. Figure created with BioRender.com, with inspiration from (112).

4.4.2 *Gaussia* Luciferase and Alkaline Phosphatase activity

Prior to conducting the luciferase assay, cells were seeded and transfected as described in **Section 4.3.4.** The medium was changed after 24 hours, and the next day, the cell medium was collected in PCR-plates before subjected to further analysis. The Luciferase-assay, comprising two steps, was performed according to the Secrete-Pair™ Dual Luminescence Assay Kit protocol. All steps were performed at room temperature unless otherwise stated.

For the *Gaussia* Luciferase (GLuc) assay, a 1X working buffer was prepared by diluting Buffer GL-S (10X) 1:10 in MilliQ water. Subsequently, the working solution for the GLuc-assay was prepared by diluting the Substrate GL (100X) at a 1:100 ratio in the 1X Buffer GL-S. The solution was mixed by inversion and incubated for 25 minutes, shielded from light. Subsequently, 10 µl per well of the harvested cell medium was transferred to a white 96-well plate. Next, 100 µl of the working solution was added to each well and mixed by pipetting. After incubating for 1 minute, the plate was placed in the FLUOstar OPTIMA luminometer to measure *Gaussia* Luciferase activity.

For the Secreted Alkaline Phosphatase (SEAP) assay, a 1X working buffer was prepared by diluting Buffer AP (10X) at a 1:10 ratio with MilliQ water. The remaining cell medium was heated at 65 °C for 15 minutes and subsequently placed on ice. The working solution was prepared by diluting Substrate AP (100X) at a 1:100 ratio in the 1X Buffer AP, followed by a 5-10-minute incubation while being protected from light. The heated medium (10 µl/well) was then transferred into a new white 96-well plate and 100 µl of the working solution was added to each well and mixed by pipetting. The mixture was incubated for 5-10 minutes, before being placed in the luminometer to measure the Alkaline Phosphatase activity.

4.4.3 Statistical analysis

Statistical analysis and visualization of the data were obtained using GraphPad Prism 10. Differences in statistical significances were assessed through Ordinary one-way ANOVA and two-way ANOVA, with Tukey's multiple comparisons test conducted between the individual groups. P-values ≤ 0.05 were considered statistically significant.

4.5 Immunofluorescence

HEK293 ($3x10^5$ cells/well) or ZR-75-1 cells ($6x10^5$ cells/well) were initially seeded onto coverslips in 12-well plates for immunofluorescence analysis. Since the HEK293 cells easily detach from the coverslips, they were coated with Poly-L-Lysine diluted 1:10 in PBS and incubated for 5 minutes. The coated coverslips were further washed four times with PBS before the HEK293 cells were seeded. The cells were incubated for 24 h, washed with pre-warmed PBS and then fixated in 500 µl of 4% paraformaldehyde for 20 minutes. Following fixation, the cells were subjected to immunostaining, and details of the buffers and solutions used for the staining are described in **Table 3.9.** Unless stated otherwise, all washing steps are carried out on a rocking platform at room temperature, using the washing buffer.

First, the cells were washed for 3x5 minutes in washing buffer. Next, the cells were permeabilized for 20 minutes, and then washed again for 3x5 minutes. Subsequently, the cells were incubated in a 5% blocking solution for 30 minutes, followed by incubation with the primary antibody anti-ZNF331 (1:100) overnight at 4 °C in a humidity chamber. The following day, the cells were washed 3x10 minutes, before incubated with the Alexa Flour™ 488-conjugated secondary antibody (1:200) for 1 hour while shielded from light. The cells were washed 3x10 minutes and then rinsed in PBS before mounted onto glass slides in ProLong Gold Antifade solution with DAPI. The slides were left to dry at room temperature in the dark and overnight before stored at -20 °C. Imaging of the slides were obtained using the Dragonfly High Speed Confocal Platform or Leica TCS STED 3X Confocal. Images were processed using the open-source platform Fiji (113).

4.6 SDS-PAGE and Western Blot

4.6.1 Preparation of cell fractions

ZR-75-1 cells were cultured to 80-90% confluence in T75 cm^2 flasks, as described in **Section 4.3.1**. Subsequently, the cells were washed twice with ice-cold PBS. Cells were harvested by scraping the cells into 2 ml of ice-cold PBS and placed on ice. The cells were centrifuged at 17 000 xg at 4 °C for 15 seconds, and the supernatant was discarded. The cell pellet was then lysed in 150 µl of 1X RIPA buffer, followed by centrifugation at 17 000 xg for 15 minutes at 4 °C. The supernatant, now designated as the cell lysate fraction, was transferred to a new Eppendorf tube and stored at -80 °C until further analysis.

4.6.2 Determination of protein concentration

The concentration of the proteins within the cell lysate fraction were determined using the Pierce BCA Protein Assay kit, following the manufacturer's protocol. The Biotek Gen5 2.06 software was used for detection.

4.6.3 SDS-PAGE

Lysate fractions were prepared for SDS-PAGE as described in **Table 4.7**, with samples containing 60 µg, 80 µg or 100 µg of protein. The samples were denatured at 56 °C for 15 minutes using a heat block, before loaded onto a Bis-Tris gel. The proteins in each sample were separated according to size using the XCell SureLock Mini-Cell Electrophoresis System in 1X MOPS buffer. Magic Mark XP and Precision Plus Protein Dual Color served as protein size markers. The electrophoresis was set to 90 V for 15 minutes, followed by 180 V for 1 hour.

Table 4.7 Sample preparation

Reagent	Volume
Sample/MilliQ water	13.0 µl
LDS Sample Buffer	5.0 µl
Reducing agent	2.0 µl
Total	20 µl

4.6.4 Western Blot

Following SDS-PAGE, proteins were transferred from the Bis-Tris gel to a PVDF membrane using the XCell Blot Module System, according to the manufacturer's instructions. For activation of the PVDF-membrane, the membrane was treated with methanol for 1 minute, then rinsed in MilliQ water for 1 minute before kept in 1X Transfer buffer until use. The protein transfer was conducted at 30 V for 1 hour in 1X Transfer buffer. Next, the PVDF membrane was blocked in 5% dry milk for 1 hour at room temperature, before washed three times in PBS-T (0.05%). For specific protein detection, the membrane was incubated with primary antibodies overnight on a tilting board at 4 °C. The primary antibodies, anti-CEL (1:5000) and anti-β-actin (1:1000), were diluted in 1% dry milk.

Next day, the membrane was washed 3x5 minutes in PBS-T (0.05%), before incubated for 1 hour in room temperature with HRP-conjugated secondary antibodies, diluted 1:5000 in 1% dry milk. The membrane was washed 1x15 minutes and 3x5 minutes in PBS-T (0.05%). Visualization of the protein bands was conducted using the Pierce™ ECL Plus Western Blotting Substrate kit, and imaging were obtained with the ChemiDoc™ MP Imaging System.

4.7 Immunohistochemistry

Formalin Fixed and Paraffin Embedded (FFPE) tissue sections were incubated at 58 °C for 15 minutes, followed by deparaffination in xylene for 2x5 minutes. The slides were rehydrated in decreasing concentrations of ethanol (100%, 96% and 80%) for 2x2 minutes each, before being washed with dH_2O for 1 minute with shaking at 240 rpm. The slides were then placed in PBS-T (0.05%) before being transferred to a high pH retrieval buffer and incubated in a pressure chamber at 120 °C for 1 minute. After the incubation, the slides were cooled down at room temperature for 15 minutes, rinsed with dH_2O and washed in PBS-T (0.05%) for 3x5 minutes with shaking at 240 rpm. The tissue slides were further subjected to chromogen or fluorescence staining as described below. All washing steps were performed in PBS-T (0.05%) with shaking unless stated otherwise. Delineation of the tissue sections were obtained with a ImmEdge pen.

4.7.1 Chromogenic staining

The tissue sections were subjected to DAKO protein blocking for 8 minutes and incubated with the primary antibody anti-CEL (1:200) overnight at 4 °C in a humidity chamber. The following day, the sections were washed for 4x15 minutes, incubated with 3% H_2O_2 for 5 minutes, and washed again for 3x5 minutes. Detection of the primary antibody were obtained using the MACH 3 Rabbit HRP Polymer Detection kit. The sections were incubated with the MACH3 Rabbit Probe, followed by the MACH3 Rabbit Polymer, with each incubation lasting 20 minutes and followed by washing for 3x5 minutes. The 3,3'-Diaminobenzidine (DAB) substrate was used for visualization. The sections were incubated for 10 minutes in hematoxylin for visualization of the nuclei, then rinsed with lukewarm tap water for 4 minutes. Dehydration of the sections were obtained in increasing concentrations of ethanol (80%, 96%, 100%) for 1x1 minute each, before placed in xylene for 2x1 minute. Lastly, the sections were mounted manually in mounting medium and visualized with NanoZoomer XR, Hamamatsu Photonics (Slide Scanner).

4.7.2 Fluorescent double-staining

For fluorescent double-staining, tissue sections were blocked with 10% goat serum for 1 hour. After removal of the protein block, the tissue sections were incubated with the primary antibody anti-CEL (1:100) overnight at 4 °C in a humidity chamber. The following day, the sections were washed for 4x15 minutes and then incubated with the Alexa Flour™ 488-conjugated secondary antibody (1:500) for 1 hour at room temperature while kept in the dark. All further steps were carried out in the dark. The tissue sections were washed 4x15 minutes, blocked with 10% goat serum for 1 hour, then incubated with the Alexa Flour® 647-conjugated antibody anti-TSH (1:50) overnight at 4 °C in a humidity chamber. Next day, the tissue sections were washed 4x15 minutes, then incubated with DAPI (1:2000) for 5 minutes for visualization of the nuclei. The tissue sections were washed 1x2 minutes in PBS before manually mounted with Gold Antifade mounting medium. Imaging of the slides were obtained using the Leica TCS STED 3X Confocal. Images were processed using the open-source platform Fiji (113).

4.8 In silico analysis of the *CEL* promoter

An *in silico* analysis of the human *CEL* promoter was performed using the UCSC Genome Browser on the GRCh38/hg38 human reference genome assembly (114). Putative transcription factors binding to the SNP rs541856133 (C) were identified using the JASPAR Transcription Factor track. Additionally, the UniBind (2021) permissive track hub (115) was used to search for experimentally validated (ChIP-seq) transcription factor binding events.

The binding properties of potential transcription factors towards the SNP rs541856133 (C>T) were analyzed using the JASPAR web page (jaspar.elixir.no) (116). Query sequences of approximately 50 base pairs centering the SNP rs541856133 (C>T) were used to assess the binding properties of the transcription factor ZNF331 (Matrix ID: MA1726.1) (**Table 4.8**).

Table 4.8 Query sequences

***CEL* promoter variant**	**Query sequence**
rs541856133 (C)	ACAGAGGGACAGACGTGGCTCTG**C**AGGTCCACTCGGTCCTGG CACCGGCC
rs541856133 (T)	ACAGAGGGACAGACGTGGCTCTG**T**AGGTCCACTCGGTCCCTG GCACCGGCC

5. Results

5.1 Part 1: Analysis of the *CEL* gene promoter

In this part of the study, we wanted to investigate the role of the promoter SNP rs541856133 (C>T) on *CEL* gene expression. Given the link between the pathogenic variant *CEL-HYB1* and the SNP, our hypothesis was that rs541856133 (T) could upregulate *CEL* gene activity. Moreover, we wanted to explore in more detail regulation of *CEL* gene promoter.

5.1.1 Preparation of plasmids and introduction of the C>T risk mutation at rs541856133 in the *CEL* promoter

Prior to assessing the activity of the *CEL* promoter in cellular systems, all plasmids to be used were amplified by Qiagen Maxi kits. **Table 3.1** provides a detailed description of the plasmids. The concentration and purity of the plasmids were determined by measuring absorbencies at 260 nm and 230-280 nm, respectively. The plasmid yield ranged from 900 ng/µl to 1500 ng/µl in a total of 300 µl, and the ratios $A_{260/280}$ and $A_{260/230}$ were around 1.9 and 2.2, respectively.

To generate the *CEL* promoter variant harboring the rs541856133 (T) risk variant, site-directed mutagenesis was performed using plasmid pEZX-hCELp-WT as template. The mutagenesis was confirmed by Sanger sequencing, and sequencing of nonmutated pEZX-hCELp-WT was included as a negative control. The sequences were evaluated in BLAST (110) and visualized with SnapGene Viewer (**Figure 5.1**). The observed C to T transition shown in **Figure 5.1** indicated a successful mutagenesis experiment. From now on, the plasmids expressing the *CEL* promoter with rs541856133 (C) and rs541856133 (T) will be referred to as pEZX-hCELp-WT and as pEZX-hCELp-MUT, respectively.

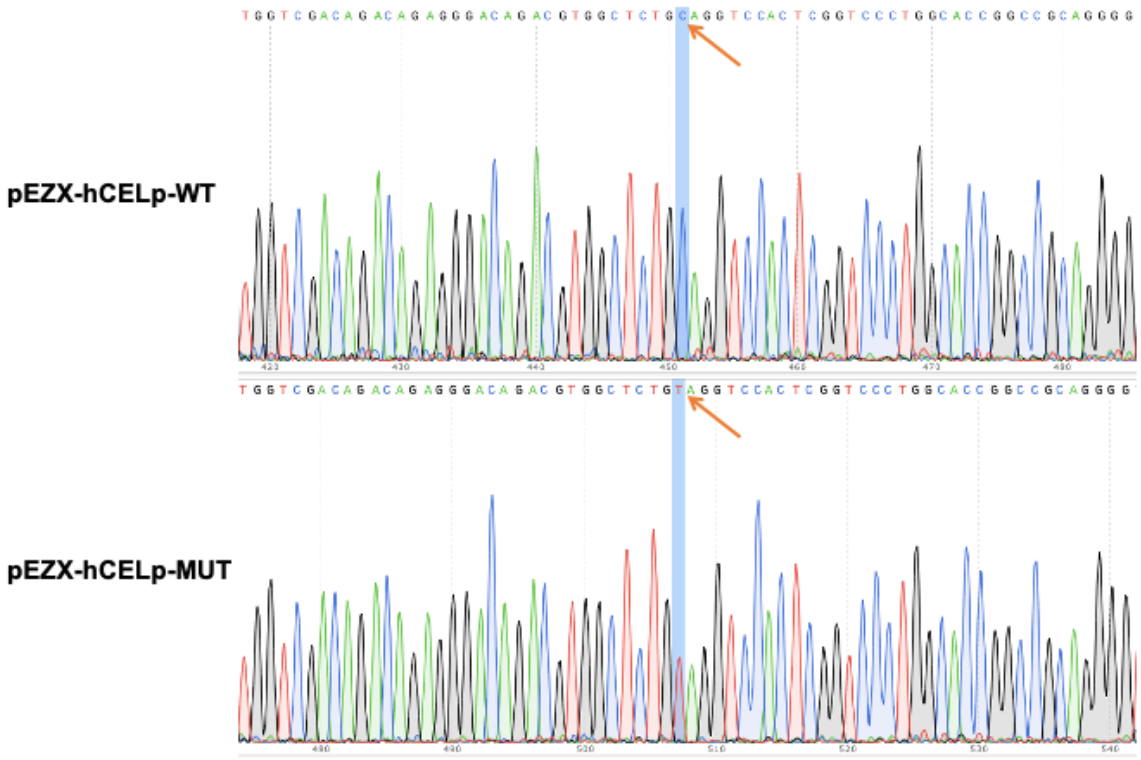


Figure 5.1 *CEL* promoter sequence. Site-directed mutagenesis was used to introduce SNP rs541856133 (T) into the *CEL* promoter of GLuc-On promoter reporter vector (pEZX-hCELp-WT). Sanger sequencing was performed to verify that the mutagenesis was successful. The SNP rs541856133 (C/T) is highlighted with an orange arrow.

5.1.2 Optimization of the Luciferase assay in HEK293 and HeLa cells

The *CEL* promoter activity was measured in both HeLa and HEK293 cells, and the promoter activity was determined using a dual reporter system measuring the activity of *Gaussia* Luciferase (GLuc) and Secreted Alkaline Phosphatase (SEAP) in harvested cell medium. See **Section 4.4.1** for a more detailed description of the assay.

To optimize the Luciferase assay, various experimental setups were tested, including different plate sizes and amounts of plasmid DNA. For instance, HEK293 and HeLa cells were seeded in 96-well plates and transfected with 100 ng of pEZX-hCELp-WT or pEZX-empty. The pEZX-empty vector lacks the *CEL* promoter and served as a negative control. However, the observed activity for the *CEL* WT promoter was not much higher than for the empty reporter, especially in the HeLa cells, and the variability between the biological parallels was high (**Figure 5.2A**). To try to enhance the *CEL* WT promoter activity, HEK293 cells were seeded in 24-well plates and transfected with 500 ng of pEZX-hCELp-WT. This setup resulted in a slight increase in *CEL* WT promoter activity and less variability between the biological parallels (**Figure 5.2B**).

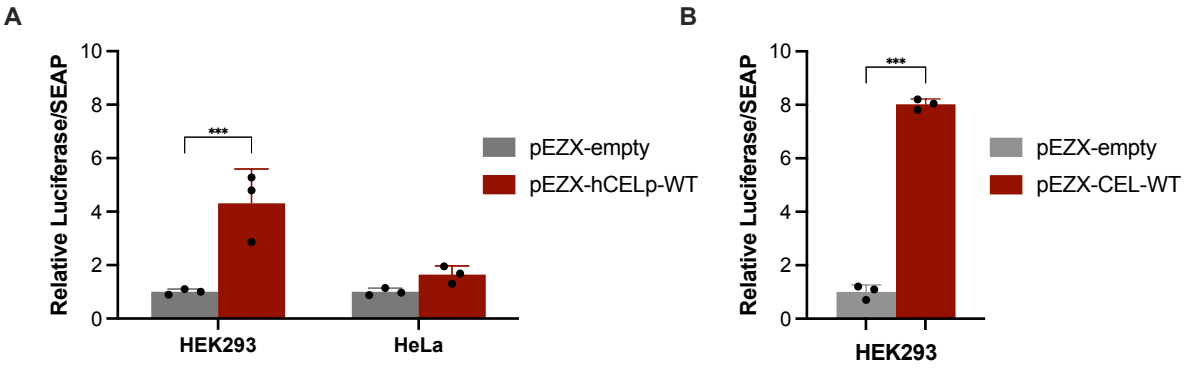


Figure 5.2 Optimization of the Luciferase assay. (A) HeLa and HEK293 cells were seeded in 96-well plates at a density of 7500 cells per well. The cells were transfected with 100 ng of pEZX-hCELp-WT. **(B)** HEK293 cells were seeded in 24-well plates at a density of 50 000 and transfected with 500 ng of pEZX-hCELp-WT. **(A-B)** An empty reporter plasmid served as negative control. *Gaussia* luciferase activity was normalized to secreted alkaline phosphatase activity (SEAP), with values presented relative to the empty reporter. This experimental setup was replicated three times for both cell lines. Statistical significance was determined using one-way ANOVA analysis with significance indicated as *** $P_{adj} <$ 0.001.

To further investigate the transfection response of 500 ng plasmid DNA in HeLa and HEK293 cells, both cell lines were transfected with a GFP-expressing plasmid (pmaxGFP) to evaluate transfection efficiencies. Cells were seeded in 24-well plates and transfected as described in **Section 4.3.4**. The cells were visualized under a fluorescence microscope and the transfection efficiency in both cell lines were estimated to be approximately 70-80% (**Figure 5.3**).

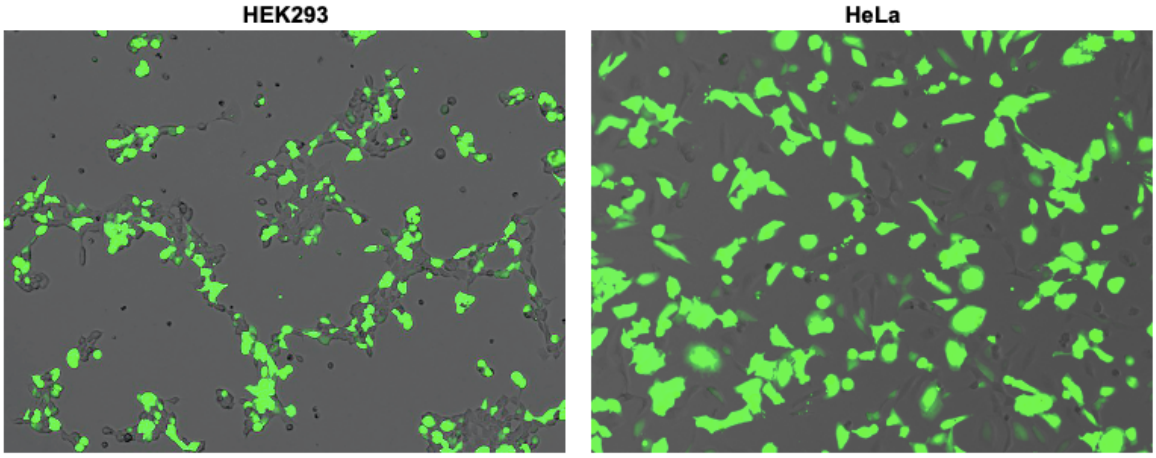


Figure 5.3 Transfection efficiency of HEK293 and HeLa cells. Cells were seeded at a density of 50 000 cells per well in 24-well plates and then transfected with 500 ng of pmaxGFP. Green fluorescence signals mark cells expressing the plasmid. Imaging was performed using a fluorescent microscope at a 10X magnification.

Based on these optimalization experiments, we decided to use 24-well plates and a total of 500 ng plasmid DNA for further analysis.

5.1.3 The effect of the C>T risk mutation on *CEL* promoter activity in HEK293 and HeLa cells

HEK293 and HeLa cells were seeded and transfected with the plasmids pEZX-hCELp-WT (500 ng), pEZX-hCELp-MUT (500 ng) or pEZX-empty (500 ng). The experiment was replicated three times, each having six biological parallels. The result from one representative experiment in each cell line is presented in **Figure 5.4**. Here, the GLuc activity is normalized to SEAP activity and shown relative to the *CEL* WT promoter.

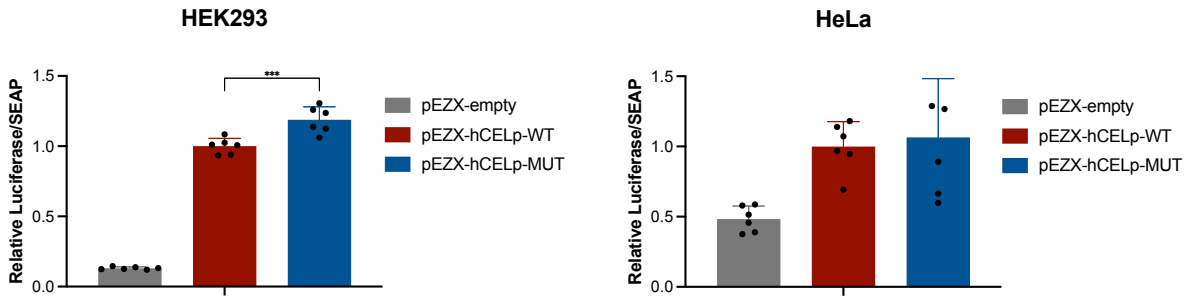


Figure 5.4 *CEL* promoter activity in HEK293 and HeLa cells. Cell medium was collected from both HEK293 cells and HeLa cells transfected with 500 ng of either pEZX-hCELp-WT or pEZX-hCELp-MUT. An empty reporter plasmid served as negative control. *Gaussia* luciferase activity was normalized to SEAP activity, with values presented relative to *CEL* WT promoter activity. This experimental setup was replicated three times for both cell lines. Statistical significance was determined using one-way ANOVA analysis with significance indicated as *** $P_{adj} < 0.001$.

In HEK293 cells, there was a statistically significant increase in GLuc/SEAP activity for the *CEL* MUT promoter compared to the *CEL* WT promoter ($P_{adj} < 0.001$) (**Figure 5.4**). There was no such effect observed in the HeLa cells. Notably, the variability among the biological parallels was higher in HeLa cells compared to HEK293 cells (**Figure 5.4**). Moreover, the relative activity of the *CEL* promoters compared to the empty reporter was still much lower in HeLa cells than in HEK293 cells (**Figure 5.4** and **Appendix 1A-B).** Based on these observations, we decided to exclude the HeLa cells from further experiments.

5.1.4 *In silico* investigation of the *CEL* promoter

As mentioned in the introduction, several regulatory elements upstream of *CEL* have been reported, including an E-box, a TAAATA box, and a pancreas-specific enhancer (63, 82-85). To further explore the regulatory elements upstream of the *CEL* gene, an *in silico* analysis of the human *CEL* promoter was performed using the UCSC Genome Browser on the GRCh38/hg38 human reference genome assembly (114). Here, we identified a zinc finger protein (ZNF331) predicted by the JASPAR Transcription Factor track to bind directly at the

SNP rs541856133, which is situated approximately 1000 base pairs upstream of the transcription initiation site (**Figure 5.5**).

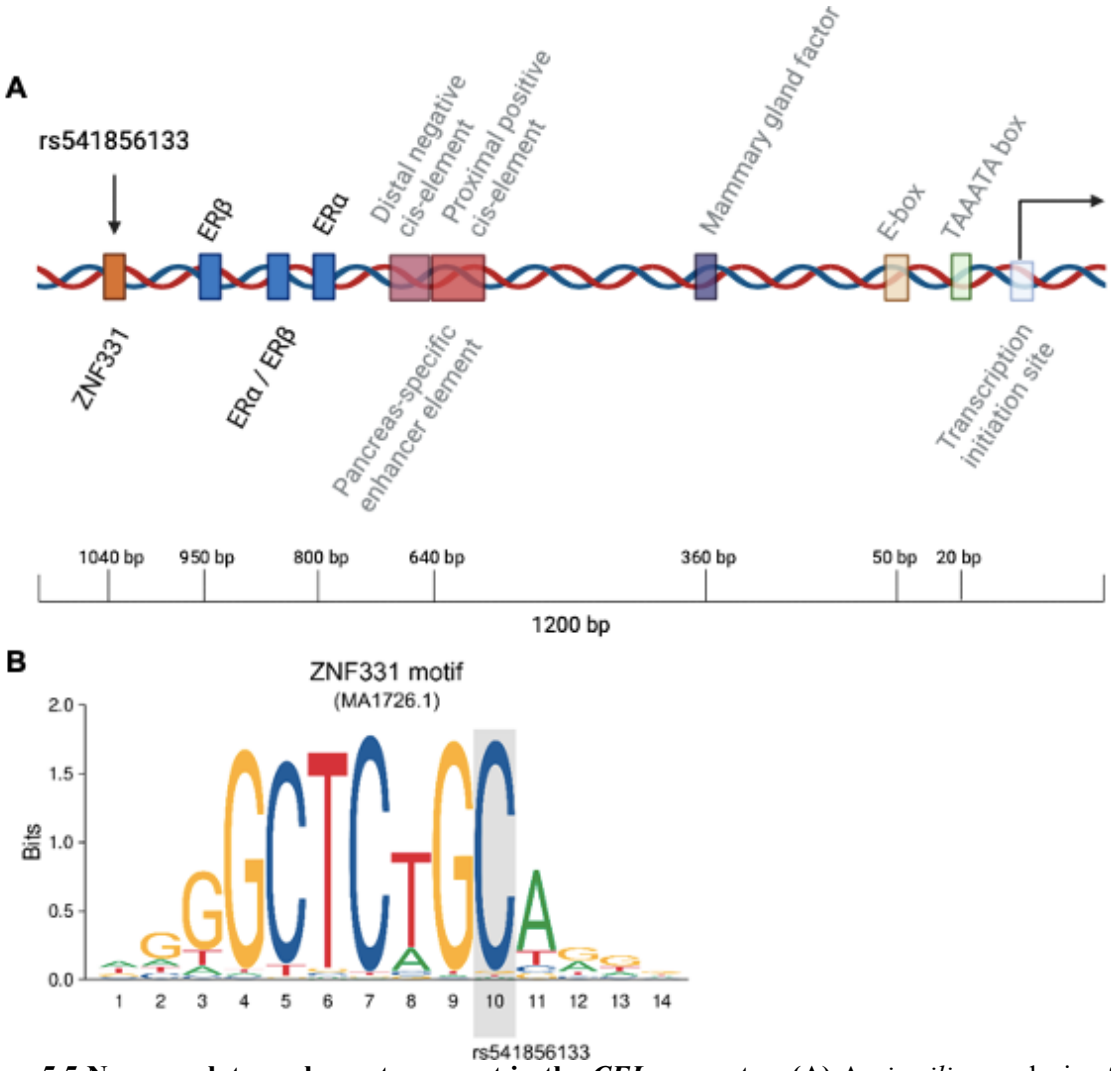


Figure 5.5 New regulatory elements present in the *CEL* promoter. (A) An *in silico* analysis of the *CEL* gene promoter identified three putative binding sites for ERα and ERβ. A potential repressor, ZNF331, binding directly at SNP rs541856133 were also identified. The presented *CEL* promoter region spans approximately 1200 base pairs, with the positions of the regulatory elements marked relative to the transcription initiation site. Figure created with BioRender.com, with annotations derived from ref (63, 82-85, 114). The elements are presented schematically and not to scale. **(B)** The consensus binding motif of ZNF331, shown as a position frequency matrix, retrieved from jaspar.elixir.no. The size of each base reflects its frequency at each position within the binding motif. The SNP rs541856133, highlighted in a grey box, is situated in a conserved position of the motif. The major allele (termed WT in this thesis) is shown.

From the literature, ZNF331 has been reported as a potential DNA-binding transcription factor and is thought to suppress the growth of esophageal and gastric cancer (117, 118). Moreover, by analyzing the *CEL* WT and MUT promoter sequences using the JASPAR web page (jaspar.elixir.no) (116), ZNF331 was predicted to bind less efficiently to the *CEL* MUT promoter compared to the *CEL* WT promoter (**Table 5.1**).

Table 5.1 Binding affinity of ZNF331 towards the *CEL* WT and *CEL* MUT promoters predicted by the JASPAR web page*

Promoter variant	Matrix ID /Name	Score	Relative score	Strand	Predicted sequence**
WT (C)	MA1726.1./ZNF331	16.5	0.940	-	GTG**GCTCTGC**AGGT
MUT (T)	MA1726.1./ZNF331	10.7	0.838	-	GTG**GCTCTGT**AGGT

*jaspar.elixir.no
**For simplicity, the reverse complementary sequence, corresponding to the (+) strand is shown. Invariant bases marked with bold letters.

In addition, we found three putative binding sites for the estrogen receptors α (ERα) and β (ERβ) predicted by UniBind (115), located approximately 800-900 base pairs upstream of the transcription initiation site (**Figure 5.5**). Both ERα and ERβ are ligand-dependent transcription factors activated by the female sex hormone estrogen (E2) (119).

Taken together, we identified potential binding sites for three interesting transcription factors in or near rs541856133, and we decided to investigate in more detail the effects of ZNF331 and ER on *CEL* promoter activity *in vitro*.

5.1.5 Endogenous expression of ZNF331 in HEK293 cells

Prior to assessing the possible effect of ZNF331 on *CEL* promotor activity in HEK293 cells, the endogenous expression and intracellular localization of ZNF331 in this cell line were examined. While mRNA levels of ZNF331 in HEK293 cells have been reported in The Human Protein Atlas (version 23.0) (120, 121), we wanted to verify this at the protein level. All further mentioned mRNA levels are presented in **Appendix 2.** HEK293 cells were subjected to immunostaining using an anti-ZNF331 antibody followed by an Alexa-Flour-488 secondary antibody and confocal microscopy. The expression of ZNF331 was observed throughout the whole cell, but with an enhanced signal present within the nucleus (**Figure 5.6**). Based on this finding, we confirmed that ZNF331 is detectable at the protein level in HEK293 cells.

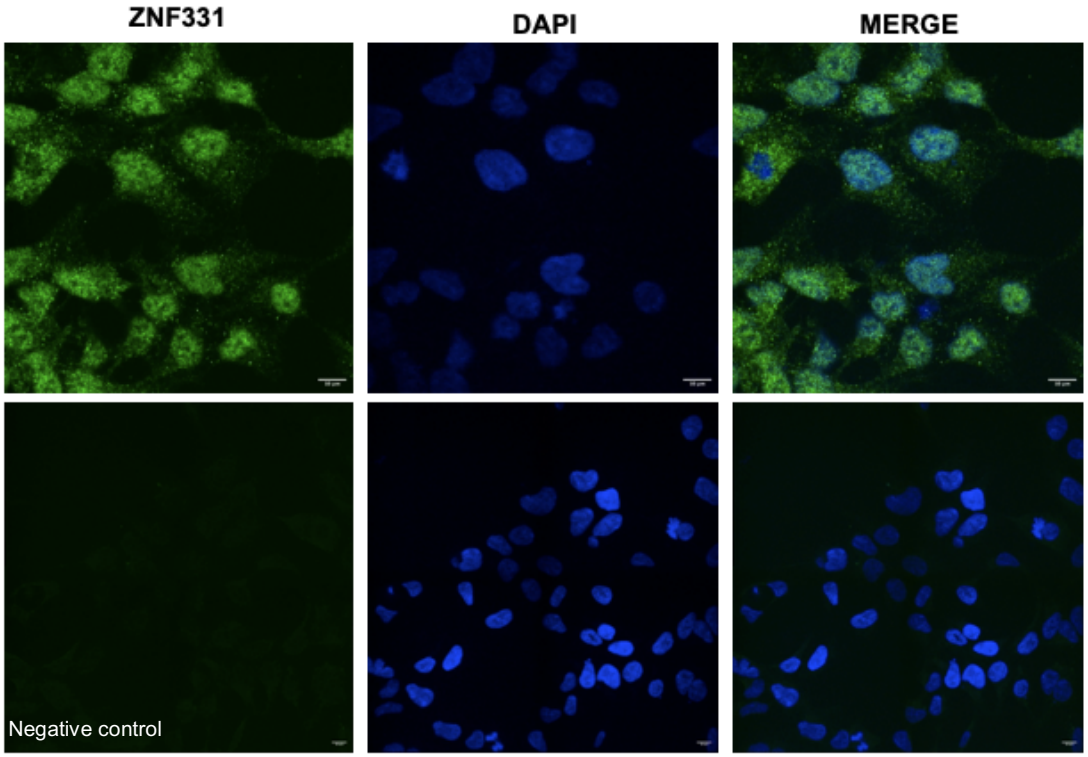


Figure 5.6 Endogenous expression and intracellular localization of ZNF331 in HEK293 cells. Cells were seeded out at a density of 300 000 cells per well in 12-well plates, fixated, and stained using anti-ZNF331 as primary antibody and Alexa-Flour-488 as secondary antibody (upper panel). Cells stained with only secondary antibody were included as negative control (lower panel). ZNF331 protein is seen in green, and the cell nuclei is blue (DAPI). Imaging was performed using a confocal microscope at 100X magnification. Scale bar: 10 µm.

5.1.6 The effect of ZNF311 on *CEL* promoter activity in HEK293 cells

Since ZNF331 is endogenously expressed in HEK293 cells (**Figure 5.6**), we wanted to investigate if increased levels of ZNF331 would influence *CEL* promoter activity. Then, considering the overlap of the ZNF331 binding site with the SNP rs541856133, and the *in silico* data showing differential binding affinity between the two SNP nucleotides (**Table 5.1**), we wanted to examine whether the *CEL* WT and *CEL* MUT promoters would respond differently to ZNF331 overexpression.

Starting with the *CEL* WT promoter, HEK293 cells were initially transfected with 200 ng of pEZX-hCELp-WT together with increasing amounts of pCMV6-ZNF331 (0 ng, 100 ng, 200 ng, and 300 ng). The pCMV6-empty plasmid was also included to ensure that the total amount of transfected DNA was kept constant at 500 ng for all samples. The experiment was performed once with three biological replicates and the result is presented in **Figure 5.7**. We did not observe any effect of ZNF331 on *CEL* WT promoter activity. Notably, the activity of the *CEL*

WT promoter relative to the empty reporter was reduced compared to what was observed when transfecting the cells with only the pEZX-hCELp-WT plasmid (**Figure 5.4**).

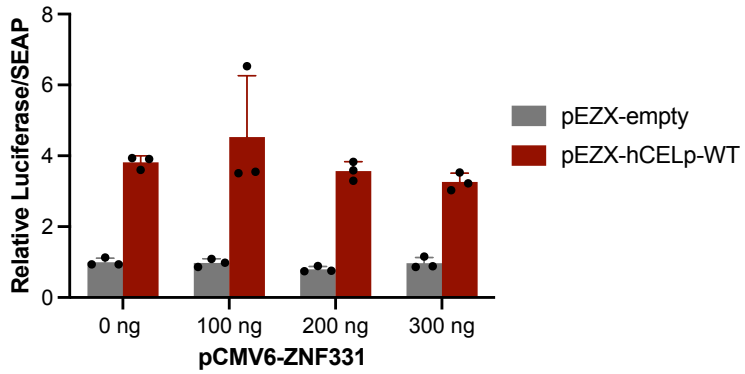


Figure 5.7 ZNF331 and *CEL* WT promoter activity in HEK293 cells. Cell medium was collected from cells transfected with 200 ng of pEZX-hCELp-WT along with increasing amounts of pCMV6-ZNF331. pEZX-empty served as negative control. The *Gaussia* luciferase activity was normalized to the SEAP activity, with values presented relative to the activity of the empty reporter in the absence of pCMV6-ZNF331. Statistical significance was determined using two-way ANOVA analysis.

To investigate if the lack of ZNF331 effect could be due to a low expression of the *CEL* WT promoter, the amount of transfected pEZX-hCELp-WT plasmid was increased from 200 ng to 300 ng. The total amount of transfected plasmid DNA was therefore increased from 500 ng to 600 ng; otherwise, the experimental setup was as previously described. The experiment was repeated four times, with six biological replicates. The result from a representative experiment is presented in **Figure 5.8**.

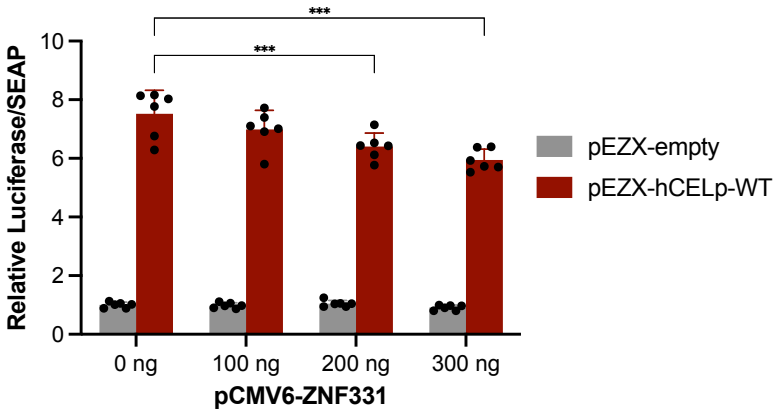


Figure 5.8 The activity of the *CEL* WT promoter in the presence of ZNF331. Cell medium was collected from HEK293 cells transfected with 300 ng of pEZX-hCELp-WT along with increasing amounts of pCMV6-ZNF331 (0 ng, 100 ng, 200 ng and 300 ng). pEZX-empty served as negative control (not shown). The *Gaussia* luciferase activity was normalized to SEAP activity, with values presented relative to the activity of the empty reporter in the absence of pCMV6-ZNF331. The experimental setup was replicated four times. Statistical significance was determined using one-way ANOVA analysis and with significance indicated as *** $P_{adj} < 0.001$.

This time, the presence of ZNF331 lead to a significant reduction in GLuc/SEAP activity for the *CEL* WT promoter. Moreover, there was a dose-response effect, where the highest dose (300 ng pCMV6-ZNF331) resulted in the strongest inhibition ($P_{adj} < 0.001$).

Given the predicted reduced binding of ZNF331 to the *CEL* MUT promoter, we hypothesized that the inhibitory effect of ZNF331 would be abolished or reduced for the *CEL* MUT promoter. To test this hypothesis, HEK293 cells were transfected with 300 ng of either pEZX-hCELp-WT or pEZX-hCELp-MUT, along with pCMV6-ZNF331 (0 ng and 300 ng) and pCMV6-empty. The result is presented in **Figure 5.9**. In the absence of exogenous ZNF331, a significant increase in GLuc/SEAP activity was observed for the *CEL* MUT promoter relative to the *CEL* WT promoter ($P_{adj} < 0.001$), consistent with the results presented in **Figure 5.4**. However, in contrast to the findings in **Figure 5.8**, we here see no repressive effect of ZNF331 on the *CEL* WT promoter activity. Instead, the presence of ZNF331 had a repressive effect on the activity of the *CEL* MUT promoter ($P_{adj} < 0.001$).

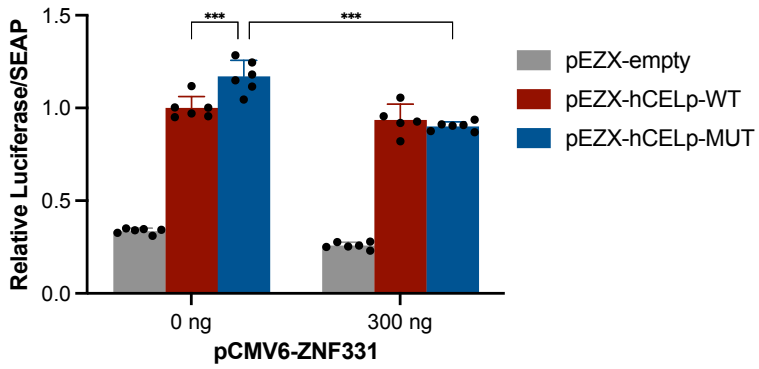


Figure 5.9 Activity of the *CEL* promoter variants in the presence of ZNF331. Cell medium was collected from HEK293 cells transfected with 300 ng of pEZX-hCELp-WT or pEZX-hCELp-MUT together with pCMV6-ZNF331 (0 ng and 300 ng). pEZX-empty served as negative control. The *Gaussia* luciferase activity was normalized to SEAP activity, with values presented relative to the activity of the *CEL* WT promoter in the absence of pCMV6-ZNF331. The experimental setup was replicated three times. Statistical significance was determined using two-way ANOVA analysis with significance indicated as *** $P_{adj} < 0.001$.

5.1.7 Transfection efficiency and *CEL* promoter activity in ZR-75-1 cells

Based on current knowledge, CEL proteins are mainly expressed by the acinar cells of the pancreas (51) and lactating mammary glands (55), and so far no CEL protein expression has been observed in HEK293 cells (98, 122). Therefore, we hypothesized that the *CEL* promoter activity may require additional transcription factors and/or coregulators not present in HEK293 cells. We therefore wanted to test an alternative cell system that endogenously express the CEL protein. We decided to explore the breast cancer cell line ZR-75-1. The choice was driven by the fact that CEL is expressed in lactating mammary gland cells and that the ZR-75-1 cells are positive for CEL mRNA according to the Human Protein Atlas (version 23.0) (120, 123). In addition, this breast cancer cell line is ER positive (124), giving us the opportunity to explore the effect of the ER binding sites detected in the *CEL* promoter (**Section 5.1.4**).

Before analyzing the *CEL* promoter activity in ZR-75-1 cells, the cells were transfected with pmaxGFP to evaluate the transfection efficiency of this cell line. First, we tested a transfection protocol specific for ZR-75-1 cells from Invitrogen, which recommended using 250 ng of plasmid DNA. However, this resulted in a transfection efficiency of only 10-20% (not shown). We therefore decided to test the same transfection protocol as used for HEK293 cells, which lead to an increase in transfection efficiency estimated to be approximately 30-40% in ZR-75-1 cells (**Figure 5.10**). Thus, the latter protocol was used for the following luciferase experiments.

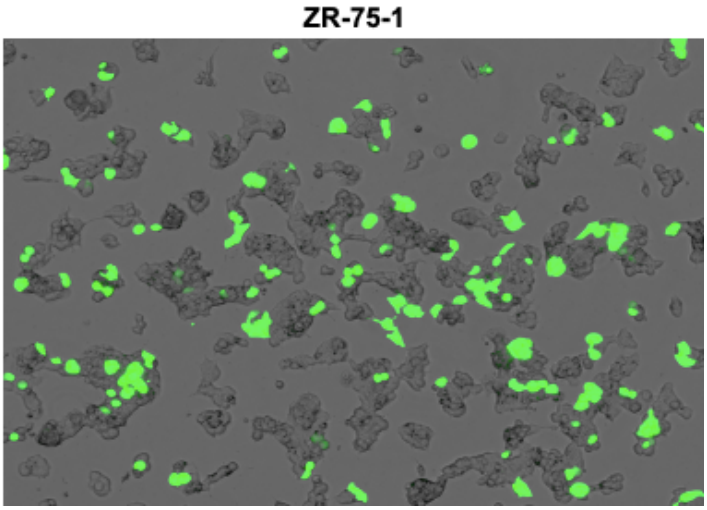


Figure 5.10 Transfection efficiency of ZR-75-1 cells. Cells were seeded at a density of 90 000 cells per well in 24-well plates and transfected with 500 ng of pmaxGFP. Green fluorescence signal indicates cells expressing the plasmid. Imaging was performed using a fluorescent microscope at a 10X magnification.

The *CEL* promoter activity in ZR-75-1 cells were assessed the same way as described in **Section 5.1.3**. The experiment was replicated four times, each with six biological parallels. The result from one representative experiment is presented in **Figure 5.11**. In this cell line, the GLuc/SEAP activity observed for the *CEL* WT promoter compared to the empty vector was higher than in HEK293 cells, suggesting that the breast cancer cell lines do support higher *CEL* promoter activities (**Figure 5.11** and **Appendix 1C**). However, we observed no significant difference in activity for the *CEL* MUT promoter compared to the *CEL* WT promoter.

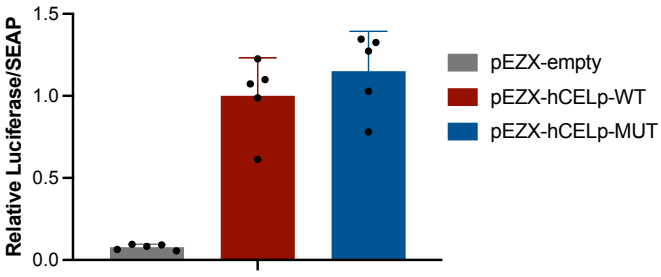


Figure 5.11 Activity of the *CEL* promotor variants in ZR-75-1 cells. Cell medium was harvested from ZR-75-1 cells transfected with 500 ng of pEZX-hCELp-WT, pEZX-hCELp-MUT, or pEZX-empty. The *Gaussia* luciferase activity was normalized to the SEAP activity, with values presented relative to *CEL* WT promoter activity. The experimental setup was replicated four times. Statistical significance was determined using one-way ANOVA analysis.

5.1.8 The effect of ZNF331 in ZR-75-1 cells

We also wanted to analyze the effect of ZNF331 on *CEL* promoter activity in ZR-75-1 cells, but before measuring the promoter activity, we examined endogenous expression and intracellular localization of ZNF331 in the cells. According to the Human Protein Atlas (version 23.0), mRNA levels of ZNF331 have been documented in ZR-75-1 cells (120, 121). We performed immunofluorescence analysis followed by confocal imaging to search for ZNF331 expression at the protein level. We found ZNF331 protein expression throughout the whole cell (**Figure 5.12)** but without the nuclear enrichment observed in the HEK293 cells**.** Even so, this observation confirmed the presence of ZNF331 protein expression in ZR-75-1 cells and we next set out to analyze the effect of ZNF331 on *CEL* promoter activity in these cells.

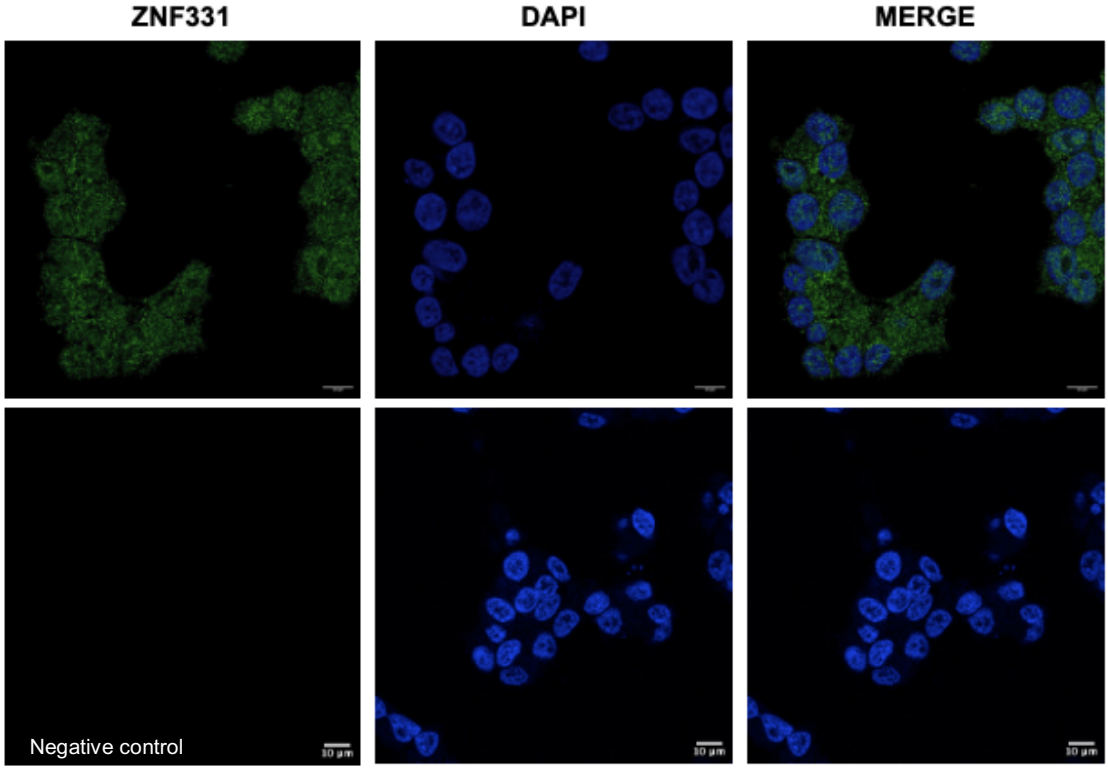


Figure 5.12 Endogenous expression of the ZNF331 protein in ZR-75-1 cells. Cells were seeded out at a density of 600 000 cells per well in 12-well plates, fixated, and stained using anti-ZNF331 as primary antibody and Alexa-Flour-488 as secondary antibody (upper panel). Cells stained with only secondary antibody was included as negative control (lower panel). The ZNF331 protein is seen in green, and the cell nuclei is blue (DAPI). Imaging was performed using a confocal microscope at 100X magnification. Scale bar: 10 µm.

To this end, ZR-75-1 cells were transfected with 300 ng of either pEZX-hCELp-WT or pEZX-hCELp-MUT as before, along with pCMV6-ZNF331 (0 ng and 300 ng). The overexpression of ZNF331 did not lead to any significant change in GLuc/SEAP activity for either variant of the *CEL* promoter (**Figure 5.13**). Furthermore, there was no notable difference in GLuc/SEAP activity between the *CEL* MUT and *CEL* WT promoter, both with and without ZNF331

overexpression. Interestingly, we detected a higher variation in activity among the biological parallels for the *CEL* MUT and *CEL* WT promoters in the absence of ZNF331 overexpression.

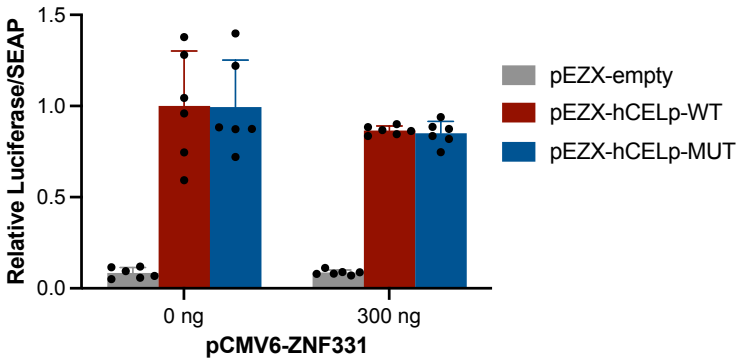


Figure 5.13 Activity of the *CEL* promoter variants in the presence of ZNF331 in ZR-75-1 cells. Cell medium was harvested from ZR-75-1 cells transfected with 300 ng of pEZX-hCELp-WT or pEZX-hCELp-MUT together with pCMV6-ZNF331 (0 ng and 300 ng). pEZX-empty served as negative control. The *Gaussia* luciferase activity was normalized to the SEAP activity, with the presented values relative to *CEL* WT promoter activity in the absence of pCMV6-ZNF331. The experimental setup was replicated two times. Statistical significance was determined using two-way ANOVA analysis.

5.1.9 The effect of ERα on *CEL* promoter activity

Given the identification of three putative ER binding sites within the *CEL* promoter (**Section 5.1.4**) and considering that ZR-75-1 cells are known to be ER-positive (124), we wanted to investigate if overexpressing ERα could have an impact on *CEL* promoter activity in this cell line. ZR-75-1 cells were transfected with 300 ng of pEZX-hCELp-WT together with increasing amounts of pSG5-ERα (0 ng, 100 ng, 200 ng, and 300 ng). To maintain consistent levels of transfected plasmid DNA across different conditions, pSG5-empty was included. One representative experiment is presented in **Figure 5.14,** where the normalized GLuc/SEAP activity is shown relative to the *CEL* WT promoter in the absence of ERα. Strikingly, the *CEL* WT promoter activity significantly decreased in response to ERα co-transfection ($P_{adj} < 0.001$). The effect was observed for all concentrations of ERα with the largest effect at 300 ng of ERα.

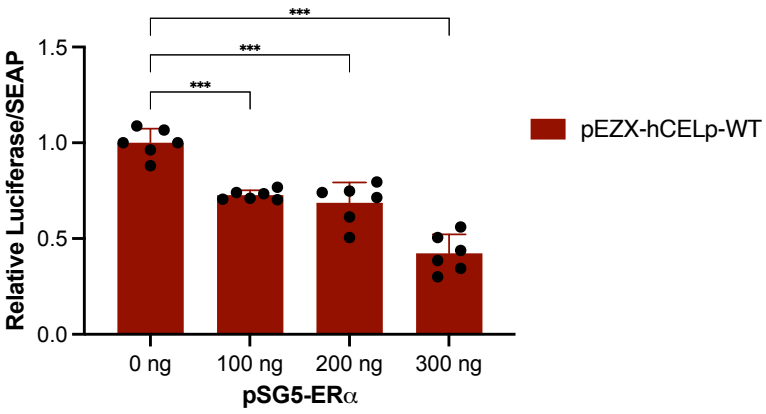


Figure 5.14 The effect of ERα on *CEL* WT promoter activity in ZR-75-1 cells. Cell medium was collected from ZR-75-1 cells transfected with 300 ng of pEZX-hCELp-WT (red column) along with increasing amounts of pSG5-ERα (0 ng, 100 ng, 200 ng and 300 ng). The *Gaussia* luciferase activity was normalized to the SEAP activity, with the presented values relative to *CEL* WT promoter activity in absence of pSG5-ERα. The experimental setup was replicated two times. Statistical significance was determined using one-way ANOVA analysis with significance indicated as *** $P_{adj} < 0.001$.

Next, to determine whether the repressive effect of ERα was affected by the rs541856133 C>T risk mutation, ZR-75-1 cells were transfected with 300 ng of either pEZX-hCELp-WT or pEZX-hCELp-MUT, along with pSG5-ERα (0 ng and 300 ng). The presence of ERα resulted in a significant decrease in GLuc/SEAP activity for both the *CEL* WT ($P_{adj} \leq 0.002$) and the *CEL* MUT promoters ($P_{adj} < 0.001$) (**Figure 5.15**). Notably, the *CEL* MUT promoter activity was significantly lower compared to the *CEL* WT promoter, regardless of ERα overexpression ($P_{adj} < 0.001$ and $P_{adj} \leq 0.033$). This result differs from what observed in **Figure 5.11** where no significant difference in activity was observed between the two *CEL* promoter variants.

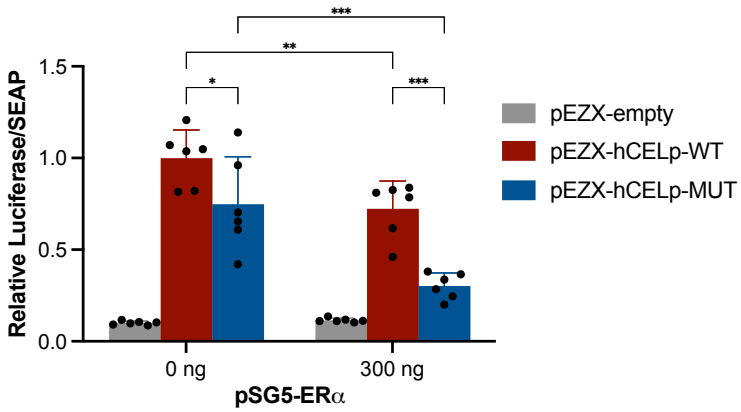


Figure 5.15 The effect of ERα on *CEL* promoter activity in ZR-75-1 cells. Cell medium was collected from ZR-75-1 cells transfected with 300 ng of pEZX-hCELp-WT or pEZX-hCELp-MUT together with pSG5-ERα (0 ng and 300 ng). pEZX-empty served as negative control. The *Gaussia* luciferase activity was normalized to the SEAP activity, with the presented values relative to *CEL* WT promoter activity in the absence of pSG5-ERα. The experimental setup was replicated three times. Statistical significance was determined using two-way ANOVA analysis with significance indicated as * $P_{adj} \leq 0.033$, ** $P_{adj} \leq 0.002$, *** $P_{adj} < 0.001$.

5.1.10 The effect of Tamoxifen and Fulvestrant on *CEL* promoter activity

To investigate if the observed effect of ERα on *CEL* promoter activity could be reversed, the cells were treated with Tamoxifen or Fulvestrant. Tamoxifen is a selective ER modulator that can act both as an estrogenic agonist and antagonist, and it is commonly used in breast cancer treatment (125). Fulvestrant is a selective ER degrader that inhibits ER signaling and is, like Tamoxifen, commonly used in breast cancer treatment (126).

The cells were transfected with 300 ng of either pEZX-hCELp-WT or pEZX-hCELp-MUT, along with pSG5-ERα (0 ng and 300 ng). Three of six biological parallels were exposed to 5 µM of Tamoxifen as suggested by prof. Stian Knappskog (University of Bergen, pers. communication) or 100 nM of Fulvestrant for 24 hours before measuring the GLuc and SEAP activity. The experiment was performed once, and the result is presented in **Figure 5.16**.

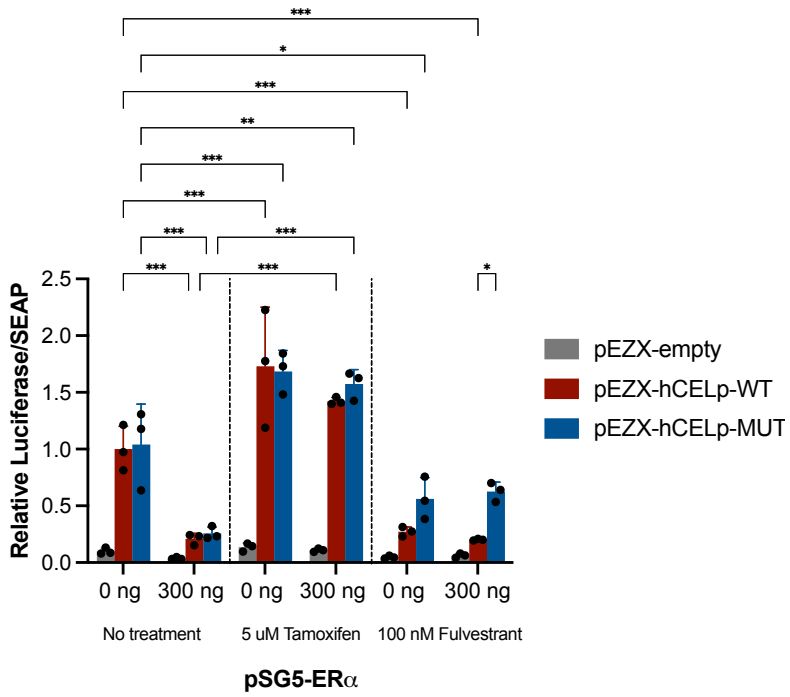


Figure 5.16 Tamoxifen and Fulvestrant treatment and *CEL* promoter activity. Cell medium was collected from ZR-75-1 cells transfected with 300 ng of pEZX-hCELp-WT or pEZX-hCELp-MUT together with pSG5-ERα (0 ng and 300 ng). pEZX-empty served as negative control. The cells were treated with 5 µM of Tamoxifen or 100 nM of Fulvestrant for 24 hours prior to the Luciferase-assay. The *Gaussia* luciferase activity was normalized to the SEAP activity. The presented values are relative to *CEL* WT promoter activity without Tamoxifen or Fulvestrant treatment. Statistical significance was determined using two-way ANOVA analysis with significance indicated as * $P_{adj} \leq 0.033$, ** $P_{adj} \leq 0.002$, *** $P_{adj} < 0.001$.

In the untreated cells (**Figure 5.16**, left side), we observed a strong and significant decrease in GLuc/SEAP activity in response to ERα overexpression for both the *CEL* WT and *CEL* MUT promoters ($P_{adj} < 0.001$), consistent with previous findings (**Figure 5.15)**. Conversely, the treatment with 5 µM of Tamoxifen resulted in significantly higher *CEL* promoter activity for both variants compared to the untreated cells both with and without ERα overexpression ($P_{adj} < 0.001$), thus abolishing the repressive effect of ERα (**Figure 5.16**, middle row).

Conversely, cells treated with 100 nM Fulvestrant (**Figure 5.16**, right side) displayed significantly lower *CEL* promoter activity compared to untreated cells and cells treated with Tamoxifen, and regardless of ERα overexpression ($P_{adj} < 0.001$). Despite this overall reduction in activity, the *CEL* promoter activity in cells treated with Fulvestrant did not decrease further when ERα was overexpressed. Additionally, for the cells treated with Fulvestrant and overexpressing ERα, we observed a significant increase in activity for the *CEL* MUT promoter compared to the *CEL* WT promoter ($P_{adj} \leq 0.033$).

5.1.11 Endogenous expression of CEL in ZR-75-1 cells

As mentioned above, CEL mRNA is detected in ZR-75-1 cells (120, 123). To investigate if the cells also express the CEL protein, we analyzed cell lysates by SDS-PAGE and Western Blotting. Lysates containing 60 µg, 80 µg, and 100 µg of total protein were loaded onto the gel and CEL protein expression was analyzed using an anti-CEL primary antibody and an HRP-conjugated secondary antibody. The theoretical mass of the CEL protein with a 16 repeat VNTR region is 79 kDa, but this can fluctuate due to different posttranslational modifications (50). As displayed in **Figure 5.17**, weak bands were observed just above 75 kDa when loading 80 µg and 100 µg protein, which may represent CEL. The weak bands detected between 50 and 75 kDa are most likely degraded CEL proteins or background staining.

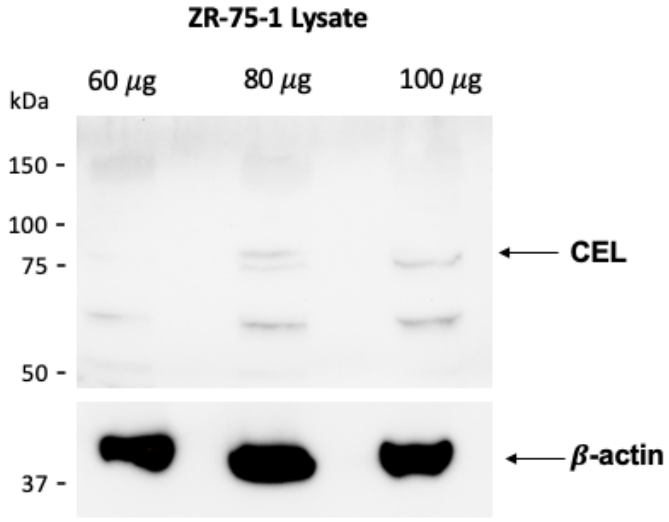


Figure 5.17 Expression of endogenous CEL protein in ZR-75-1 cells. Cell lysates with 60 µg, 80 µg and 100 µg of total protein were subjected to SDS-PAGE and Western Blot analysis. The CEL protein was detected using an anti-CEL primary antibody and an HRP-conjugated secondary antibody. Anti-β-actin was included as loading control. The experiment was performed once.

5.2 Part 2: CEL protein expression in the pituitary gland

Based on our observation that the *CEL-HYB1* variant is a risk factor for CP predominantly in women (unpublished), in combination with the effect of ER on *CEL* promoter activity observed above (**Figur 5.15**), we set out to investigate the expression of CEL in the pituitary gland. There is an older publication showing CEL expression in the this endocrine organ (60), which is critical for production of sex hormones (127). Thus, we wanted to verify this finding to learn more about the potential CEL-hormone interplay.

5.2.1 CEL protein expression in the mouse pituitary gland

Since our research group is working with animal models, we started analyzing CEL protein expression in the pituitary gland of mice. FFPE tissue sections of the mouse pituitary gland

were prepared and subjected to both HE-staining and immunohistochemistry. For the latter, an anti-CEL primary antibody followed by chromogenic staining were used for detection of CEL. Pancreatic tissue from mice was used as both positive and negative controls (not shown). Pituitary tissue sections stained with only secondary antibody was included as negative control. HE-staining was included to assess the histology of the tissue. The results are presented in **Figure 5.18**.

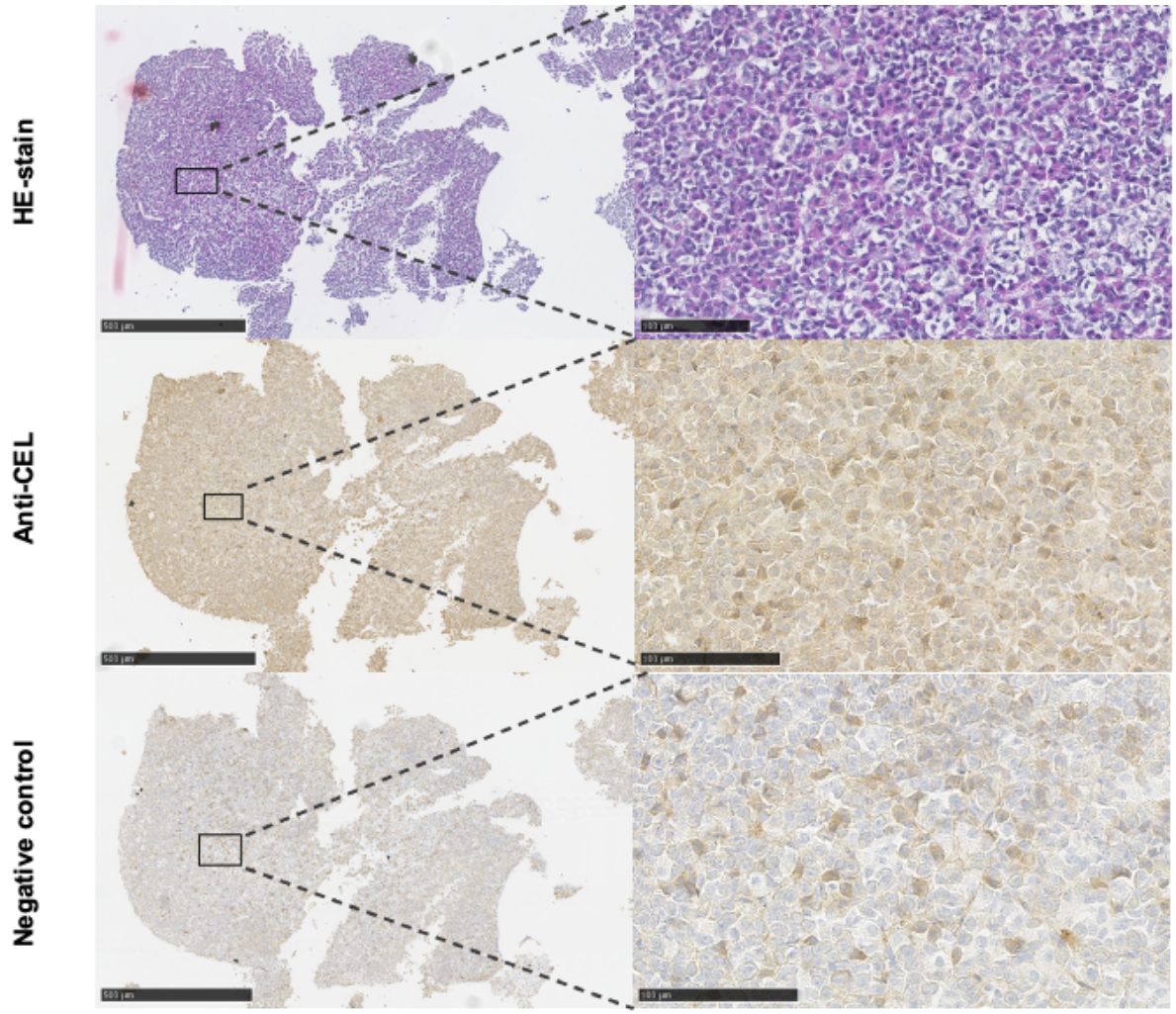


Figure 5.18 CEL protein expression in mouse pituitary gland tissue. HE- and immunohistochemical staining of mouse pituitary tissue using anti-CEL antibody. Brown color indicates positive CEL staining and blue color shows nuclei visualized with hematoxylin. Pituitary tissue without primary antibody incubation served as negative control. Scale bars: 500 µm (left side) and 100 µm (right side).

As seen in the upper panel, the histology of the HE-stained pituitary tissue appeared mostly normal (**Figure 5.18**). In the middle panel, brown color indicates CEL-positive cells, which were observed across the whole tissue section. However, positive staining was also observed in the negative control (**Figure 5.18**, lower panel), indicating false positive CEL staining. No positive staining was detected in the pancreatic tissue used as negative control (not shown).

5.2.2 CEL protein expression in the human pituitary gland

Despite the unexpected false positive CEL-staining in the mouse pituitary gland (**Figure 5.18**), we decided to investigate CEL protein expression in the human pituitary gland. FFPE human pituitary tissue from the adenohypophysis of a 94-year old female with a microadenoma was obtained from the Netherland Brain Bank. Tissue sections were prepared and subjected to both HE-staining and immunohistochemistry as described for the mouse tissue (**Section 5.2.1**). The results are presented in **Figure 5.19**.

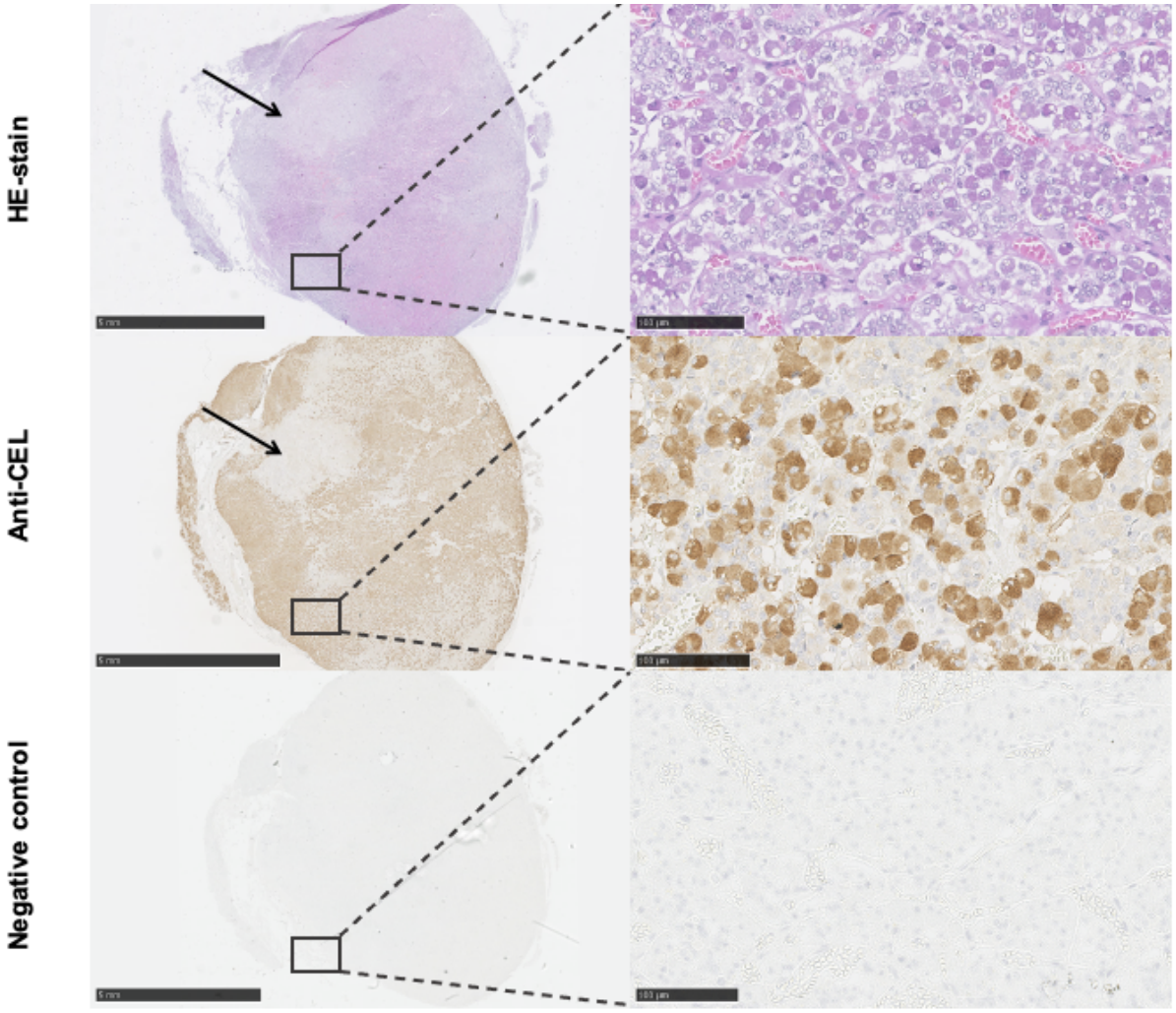


Figure 5.19 CEL protein expression in human pituitary gland tissue. HE- and immunohistochemical staining of human pituitary tissue using anti-CEL antibody. Brown color indicates positive CEL staining and blue color shows nuclei visualized with hematoxylin. Pituitary tissue without primary antibody incubation served as negative control. Scale bar: 5 mm (left side) and 100 µm (right side).

As seen in the upper panel, the histology of the HE-stained pituitary tissue appeared mostly normal, with the exception of a brighter area that most likely represent the microadenoma (**Figure 5.19,** marked by arrow). In the middle panel, CEL-positive cells were detected throughout the tissue. However, the area suspected to include the microadenoma were almost

completely negative for CEL positive cells (**Figure 5.19**, marked by arrow). Notably, no positive staining was detected in the negative control, indicating CEL-specific staining (**Figure 5.19**, lower panel).

To further explore the role of CEL protein in the pituitary gland, we set out to do immunofluorescence staining and confocal microscopy. We wanted to see if CEL co-localizes with any of the six known hormones produced in the adenohypophysis of the pituitary gland. However, due to a lack of time, we were only able to analyze the co-localization of CEL and thyroid-stimulating hormone (TSH). We observed co-localization of CEL and TSH expression in several cells (**Figure 5.20**, upper and middle panel). Additionally, some cells were negative for both CEL and TSH, while others were positive for only one of the proteins.

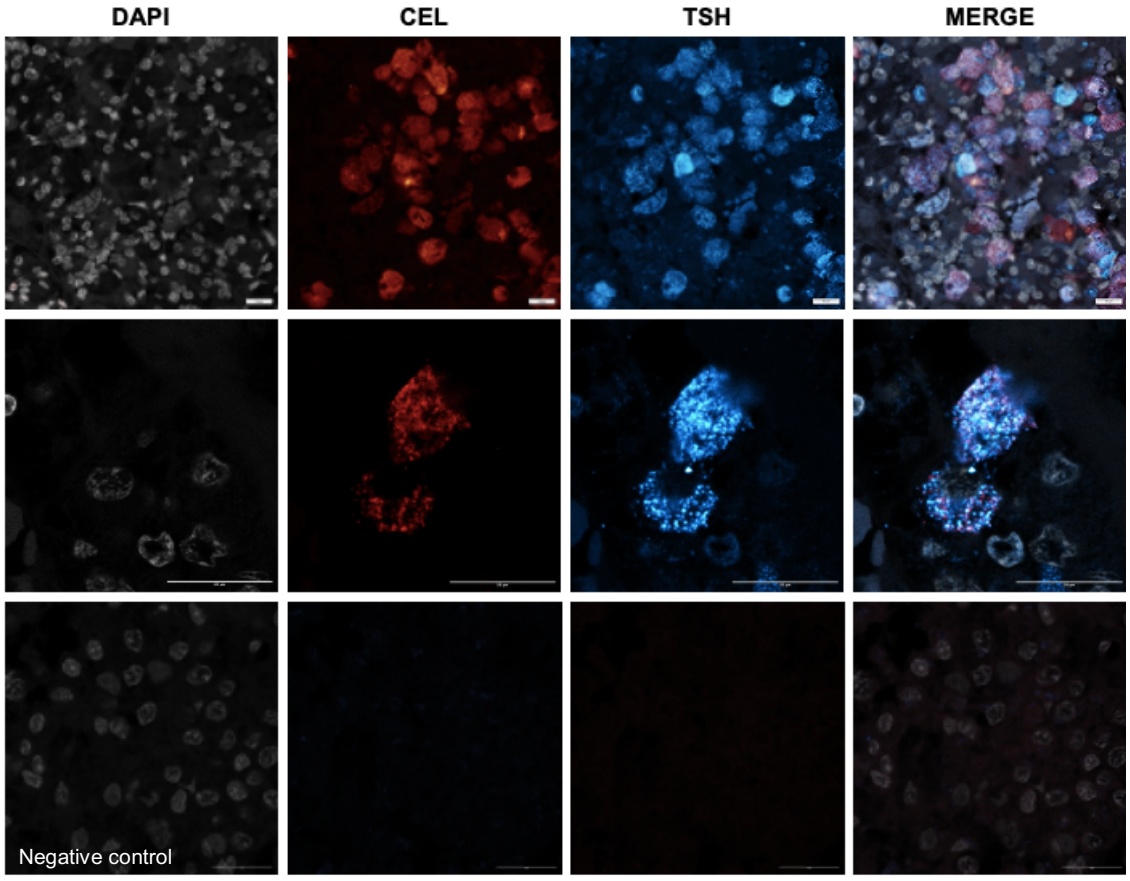


Figure 5.20 CEL and TSH co-localization in human pituitary gland tissue. CEL and TSH was detected in pituitary tissue using immunofluorescent staining and confocal imaging. CEL was identified using an anti-CEL primary antibody followed by a secondary antibody conjugated with Alexa Flour™ 488. TSH was detected using an anti-TSH antibody directly conjugated to Alexa Flour™ 647. CEL protein is seen in red, TSH protein seen in blue, and nuclei seen in grey (DAPI). Tissue stained with only secondary antibody was included as negative control (lower panel). Images were captured at 40X (upper panel) magnification and 100X (two lower panels). Scale bar: 20 µm.

6. Discussion

6.1 Analysis of the *CEL* gene promoter

6.1.1 The influence of SNP rs541856133 (T) on *CEL* promoter activity

Single nucleotide polymorphisms (SNPs) are commonly associated with various traits and diseases, though their individual impact can be subtle and context-dependent (128). SNP rs541856133 (T), located upstream of the *CEL* gene, was identified by GWAS as a new risk factor for T1D (107). We found this association of particular interest, as a pathogenic variant of *CEL* has previously been linked to a monogenic form of diabetes (50). We therefore set out to screen our patient materials for the *CEL* promoter SNP rs541856133 and, to our surprise, we found that rs541856133 (T) was linked to *CEL-HYB1* (unpublished), a variant previously established as a risk factor for CP (91, 105).

Thus, the first task of this thesis was therefore to determine the impact of SNP rs541856133 (T) on *CEL* gene expression. We analyzed the activity of the *CEL* promoter across various cell lines. The *CEL* MUT promoter, containing the rs541856133 (T) risk variant, demonstrated significantly higher activity compared to the *CEL* WT promoter in HEK293 cells, but not in HeLa cells (**Figure 5.4**). Although the observed effect was small, it may suggest that the SNP rs541856133 (T) directly influences *CEL* promoter activity, and that the effect varies between cell lines.

Despite the positive effect of the SNP rs541856133 (T) on *CEL* gene expression in HEK293 cells, the overall activity remained low (**Appendix 1A**). This could be due to the lack of endogenous CEL expression and the necessary machinery for efficient CEL transcription in these cells (98, 122). This led us to explore another cell model that could provide a more representative environment. Given that CEL is expressed in lactating mammary glands, we chose the breast cancer cell line ZR-75-1 cells, which according to the Human Protein Atlas, are positive for CEL mRNA (120, 123). Indeed, the overall *CEL* promoter activity was elevated in ZR-75-1 cells compared to HEK293 cells. However, when comparing the activity between the *CEL* MUT and the *CEL* WT promoters, the results were inconsistent. In different experiments, we observed no significant change (**Figure 5.11**), increased (**Figure 5.15**) and decreased (**Figure 5.16**) activity for the *CEL* MUT promoter relative to the *CEL* WT promoter. Also, when we analyzed the endogenous CEL protein expression in the ZR-75-1 cells, the results were unclear. From the Western blot analysis, when loading a total of 60 ug, 80 ug or

100 ug protein, we did not observe a convincing CEL band (**Figure 5.18**). This result, combined with the inconsistencies observed for the *CEL* promoter activity, may suggest that ZR-75-1 is not the optimal cell line for studying *CEL* promoter activity. Alternative cell models like acinar or ductal-like breast cell lines should therefore be tested. However, such physiological relevant cell models are challenging to work with as further discussed in **Section 6.3.1**.

The variation in *CEL* promoter activity that we observed across and within different cell lines may also suggest that the SNP rs541856133 (T) does not have an effect on CEL expression, and that it is associated with *CEL-HYB1* without affecting the expression of the pathogenic variant. However, non-synonymous SNPs like those found in promoters can have subtle effects and modify transcription factor recognition and DNA binding, potentially leading to notable regulatory effects (128, 129). To address this, we wanted to investigate the regulation of *CEL* expression by exploring potential transcription factors that binds to the *CEL* promoter. This could provide further insights into how SNP rs541856133 (T) and other regulatory elements influence *CEL* promoter activity and gene expression.

6.1.2 Impact of transcription factors on *CEL* promoter activity

We performed *in silico* analysis of the *CEL* promoter region to identify any transcription factors whose binding could be influenced by the C/T transition at rs541856133. Interestingly, the DNA-binding transcription factor ZNF331 (117, 118), was predicted to bind directly at SNP rs541856133 (**Figure 5.5**). In addition, the JASPAR web page (jaspar.elixir.no) (116) suggested reduced binding affinity of ZNF331 towards the *CEL* MUT promoter compared to the *CEL* WT promoter. We therefore set out to examine the regulatory effect of ZNF331 on the *CEL* promoter activity.

We found that ZNF331 had an inhibitory effect on the *CEL* promoter activity in HEK293 cells (**Figure 5.8** and **Figure 5.9**), but in contrast to what might be expected from the JASPAR predictions (**Table 5.1**), the effect of ZNF331 was similar on both *CEL* MUT and *CEL* WT promoters. This repressive effect was not observed in ZR-75-1 cells (**Figure 5.13**), again showing different responses between cell lines. Interestingly, though, we found that HEK293 cells have a higher expression level of the ZNF331 protein than the ZR-75-1 cells, as demonstrated by immunofluorescence analysis (**Figure 5.6** and **Figure 5.12**), respectively. The same trend is reported at the RNA level in the Human Protein Atlas (**Appendix 2**) (120, 121). Thus, the high expression of ZNF331 correlates with a decrease in *CEL* promoter activity in

HEK293 cells, supporting a repressive role for ZNF331. However, further analysis is needed, and one experiment would be to analyze the effect of ZNF331 on the *CEL* promoter activity in a cell line without endogenous ZNF331 expression.

Little is known about ZNF331 from the literature, but it has been found to suppress the growth of esophageal and gastric cancers (117, 118). Our result also suggests an inhibitory effect on CEL expression in HEK293 cells. However, the role of ZNF331 in CEL-expressing tissues such as breast or pancreas remains unclear. According to the Human Protein Atlas, the RNA levels of ZNF331 in these tissues are between those in HEK293 and ZR-75-1 cells (**Appendix 2**) (130, 131), which could suggest a potential regulatory role of ZNF331 in these environments. Notably, as described in **Section 1.3.3**, the *CEL* promoter contains multiple tissue-specific regulatory elements, which might cause ZNF331 to have different influence depending on tissue and cell type.

In a recent update of the JASPAR Transcription Factor Track (March 5, 2024) (114), a new zinc finger protein, ZNF816, has been identified as another transcription factor predicted to bind at the *CEL* promoter SNP rs541856133. Like ZNF331, the role of ZNF816 remains largely unexplored, but its association with the *CEL* promoter is interesting and opens for further investigation to understand its functional effects.

Taken together, our experiments on *CEL* MUT and *CEL* WT promoter activity show that *CEL* MUT has increased activity compared to *CEL* WT in HEK293 cells. Additionally, we observed a repressive effect of ZNF331 on the activity of both *CEL* promoter variants in these cells. However, neither of these findings was replicated in ZR-75-1 cells. The observed increase in activity for the *CEL* MUT promoter in HEK293 cells could be related to functional consequences, such as an increased propensity of the pathogenic variant CEL-HYB1 to form intracellular aggregates and decreased secretion compared to normal CEL (102-105). However, understanding the biological effects of SNPs is challenging, especially in complex disease such as CP. Most SNPs are located in non-coding regions and may not directly impact the nearby genes they are linked with (128). For instance, SNP rs541856133 (T), despite its proximity to the *CEL* gene, may influence other cellular functions or interact with distant loci (132), complicating the interpretation of its functional role in diseases such as CP. SNPs are mainly used as markers to predict genetic susceptibility rather than as a direct causative factor of disease. Even though the risk variant rs541856133 (T) did not seem to have any substantial

impact on *CEL* promoter activity, its association with the *CEL-HYB1* variant presents an opportunity to use this SNP as a biomarker in large-scale studies aiming to identify carries of this condition, even if the SNP does not directly affect CEL protein expression.

6.1.3 Hormonal regulation of *CEL* promoter activity?

Our *in silico* analysis of the *CEL* promoter also revealed binding sites for the ERα and ERβ (**Figure 5.5**). This discovery is particularly interesting as CEL is abundant in mothers milk (55), and that preliminary data from our research group suggests that *CEL-HYB1* is a female-specific risk factor for CP (unpublished). Both ERα and ERβ are ligand-dependent transcription factors activated by the female sex hormone estrogen (119). Understanding how ER might influence *CEL* promoter activity could provide new important insights into how sex hormonal regulation may affect gene expression of a digestive enzyme.

Given that ZR-75-1 is an ER-positive cell line (124), and that these breast cells, according to the Human Protein Atlas, mainly express ERα (not ERβ) (120, 133, 134), we chose to focus on the alpha variant. We observed a significant decrease in activity for both the *CEL* WT and *CEL* MUT promoters in ZR-75-1 cells overexpressing ERα, suggesting a repressive effect on *CEL* promoter activity (**Figure 5.15**). Moreover, treating the same cells with 5 µM Tamoxifen, an ER modulator (125), abolished the repressive effect of ERα on *CEL* promoter activity and also increased the activity compared to the untreated cells (**Figure 5.16**). This effect was not observed with 100 nM of Fulvestrant (**Figure 5.16**), an ER degrader (126). However, this experiment was conducted only once, and needs to be confirmed. Also, the implications of ERα's role in CEL expression *in vivo* or across different cell types are less clear. Another important aspect is that since the ZR-75-1 cell line is ER-positive (124), *CEL* might already be repressed by ERα regardless of ERα overexpression in these cells. It would therefore be interesting to investigate the role of ER on CEL in an ER-negative breast cell line.

Even though it is too early to conclude, our findings may suggest that ERα represses *CEL* promoter activity, which raises questions about the biological significance of this regulation. One interesting aspect is the expression of CEL in lactating mammary glands (55). After childbirth, the hormone level of estrogen drops significantly in the mother, and this loss of estrogen is crucial for initiating and maintaining lactation (135, 136). It can be hypothesized that the reduction in estrogen decreases the repressive effect of ERα on CEL, ensuring that CEL

is only expressed in the mammary gland cells during breastfeeding. However, it has previously been suggested that the increased level of CEL protein expression during breastfeeding is due to the overall increase in differentiated epithelial cells rather than a direct hormonal effect, but this observation was based on mouse studies (86).

Interestingly, CEL has also been suggested as a potential marker for breast cancer (137). In a study by Cui et al, they found a higher CEL protein expression in tumor tissues compared to healthy tissues (137). Additionally, elevated levels of CEL were related to poor survival, particularly in patients with ER-positive tumors (137). The apparent contradiction between our results, which suggest that CEL is repressed by ERα, and the reported high CEL protein expression in ER-positive breast cancer, further shows the complexity of *CEL* gene regulation and its implication in disease.

Finally, it is important to remember that ERα and ERβ exhibit distinct expression patterns in various human tissues, which influence their regulatory roles (**Appendix 2**) (130, 138, 139). For instance, in breast tissue, RNA levels of ERα are notably higher than those of ERβ (130, 138, 139). Conversely, in the pancreas, although overall having low RNA levels, ERβ is more prevalent than ERα (130, 138, 139). This suggests tissue-specific regulatory roles for ERα and ERβ, which might influence CEL expression differently across various tissues. Based on our results, we have gained insights into how CEL is regulated by estrogen in breast cells. Exploring how estrogen affects CEL in acinar cell, where CEL is continuously produced, will be an interesting aspect for further research, potentially uncovering new features of CEL regulation and its implications in different tissues. Additionally, it would be interesting to investigate the effect of overexpressing ERβ on the *CEL* MUT and *CEL* WT promoter activity, given the different expression patterns and regulatory roles of ERα and ERβ.

6.2 CEL's role outside of the digestive tract

CEL expression has been identified in various tissues beyond the pancreatic acinar cells and the lactating mammary gland (51, 55), including the pituitary gland (60), which is an endocrine organ critical for the production and release of hormones such as the thyroid-stimulating hormone (TSH), adrenocorticotropic hormone (ACTH), gonadotropins, prolactin, and growth hormone (GH) (127). Given our preliminary data indicating that the pathogenic *CEL-HYB1* variant is a risk factor for chronic pancreatitis predominantly in women (unpublished), and that

the *CEL* promoter includes several ER binding sites, we wanted to learn more about its potential interactions with hormonal pathways and its role outside of the digestive tract.

6.2.1 CEL protein expression in the human pituitary gland

In this thesis, we detected CEL protein expression in human adenohypophysis tissue using both chromogenic and immunofluorescent staining techniques. CEL-positive cells were observed scattered throughout the whole tissue (**Figure 5.19**), and we also observed co-localization between CEL and the hormone TSH (**Figure 5.20**). With these results we confirm what has been suggested before, that that CEL co-localizes with endocrine cells secreting TSH (60). Further testing is required to assess co-localization with the remaining pituitary hormones: GH, ACTH, gonadotropins, and prolactin.

Thus, despite the expression of CEL in lactating mammary glands (55), CEL was reported not to co-localize with the hormone prolactin in the pituitary gland (60), which is crucial for lactation (140). Such absence of co-localization might be due to limitations of previous staining techniques or the potential repressive effect of CEL expression by ERα, as discussed in **Section 6.1.3**. Moreover, ER is reportedly primarily found in pituitary adenomas producing gonadotropins and prolactin (141), which is the opposite of where CEL is localized (60). This may suggest that CEL is also under ER regulation in the pituitary gland, with the possibility that CEL expression in cells secreting gonadotropins and prolactin is repressed by ERα and therefore not detected. To gain more insights, further studies on the expression patterns of ER, CEL, and other hormones in the pituitary gland are necessary. This is an area we plan to explore further to better understand the role of CEL in this gland.

Interesting, also, data from the Human Protein Atlas show that ZNF331 has higher RNA levels in the pituitary gland compared to the pancreas (**Appendix 2**) (130, 131). Interestingly, the reported RNA levels for ZNF331 in the pituitary gland is comparable with those in HEK293 cells, where we observed a repressive effect of ZNF331 on *CEL* promoter activity (**Figure 5.8** and **Figure 5.9**). This suggests a potential regulatory role for ZNF331 on CEL in the pituitary gland. Investigating the possibility of co-localization of CEL and ZNF331 in pituitary tissue would therefore be interesting for further research.

6.2.2 CEL protein expression in the mouse pituitary gland

As mentioned in the introduction of this thesis, our research group has developed a *CEL-HYB1* mouse model that develops CP spontaneously (105). These animals are therefore excellent models to follow up on various interesting aspects of CEL, including the role of CEL in the pituitary gland. However, it is important to note that this mouse model cannot be used to study the effect of the *CEL* promoter SNP rs541856133, as this area of the promoter region is not conserved between human and mice (63, 82).

However, in contrast to the human pituitary gland, staining of the mouse gland presented with some challenges. We detected positive CEL staining in the negative control, that means in the mouse pituitary tissue that had not been incubated with the primary anti-CEL antibody – only secondary antibody (**Figure 5.18**). This nonspecific staining, which did not occur in pancreatic mouse tissue treated under the same conditions, raises questions about the specificity of CEL staining in the mouse pituitary gland. Staining with a negative control was only performed once and could be due to technical issues like inadequate blocking or the integrity of the tissue. The mouse pituitary gland is a small and delicate organ, and the tissue section used was to some degree distorted and damaged (**Figure 5.18**). Thus, we will follow up on these experiments with good quality tissue sections and hopefully we will be able to use our CEL mice for further studies of the pituitary gland. However, it is important to remember that mouse models do not always perfectly reflect human biology. Although the proximal part of the *CEL* gene promoter is highly conserved between human and rodents (63, 82), the distal part, where SNP rs541856133 is located, is not well conserved. This makes it unsuitable for studying the SNP's effect in mice. Additionally, the regulatory mechanisms involved in pancreatic *CEL* gene expression differ between the two species (83). This suggests that the tissue-specific elements involved in pituitary expression of CEL might also differ between mice and humans, thereby affecting the generalizability of the results.

Altogether, verification of CEL protein expression in the pituitary gland opens up for new and interesting extra-pancreatic roles of CEL. However, the impact of tissue condition and handling on staining results, especially in the mouse tissue, underscores the need for further research and investigations. Staining of CEL in human pituitary tissue was only conducted twice, with tissue derived from only one person, making the significance of these result questionable. Additional research is needed to clarify CEL's function in the pituitary gland and its possible interactions with hormonal pathways. However, the link between CEL and the endocrine system is further

strengthened by our results showing inhibitory effects of ERα on *CEL* promoter activity *in vitro*. Overall, this thesis demonstrates that the role of CEL extends beyond the pancreas and the digestive tract, and that further research on CEL expression is warranted.

6.3 Methodological considerations

6.3.1 Choice of cell model

In this thesis, we used HEK293, HeLa, and ZR-75-1 cells to examine *CEL* promoter activity *in vitro*. Common for all three cell lines is that they are of human origin: HEK293 cells are derived from transformed embryonic kidney cells (142), HeLa cells originate from cervical cancer (143), and ZR-75-1 cells are breast cancer cells (124). These cell lines are commonly used in research as they are easy to culture and transfect, but it is important to be aware of both their advantages and limitations. Given their immortalized and cancerous nature, these cell lines do not represent normal human cell physiology.

When studying CEL, it is important to considering several important characteristics to ensure accurate and meaningful results. Firstly, it would be preferable to use a cell model that endogenously expresses CEL, as it already possesses the necessary machinery for proper expression. Post-translation modifications are crucial for correct folding and function of proteins (144), and CEL is highly *O*-glycosylated at the C-terminal (77). While the use of human cell lines is important to achieve these modifications correctly, it does not always guarantee optimal results. For instance, neither HeLa nor HEK293 cells secrete CEL endogenously (98, 122), and thus may not provide the most effective and correct post-translation modifications. Another important consideration is gender. Here, all three cell lines were derived from female donors, which could impact the experimental results, especially those involving estrogen receptors. It is therefore important to consider how the choice of cell model might affect the interpretation and generalizability of our findings.

The ideal cells for our CEL studies would be acinar cells of the human exocrine pancreas and human lactating mammary gland cells (51, 55). While ZR-75-1 cells were used as a model for the latter, this cell line was proved to not be ideal for studying *CEL* promoter activity, as we were unable to detect CEL protein expression in these cells (**Figure 5.17**). An alternative cell model needs to be considered for further studies. For acinar cells, the only commercially available options include 266-6 and AR42J cell lines, derived from mouse and rat, respectively.

As the distal part of the *CEL* promoter is not well conserved between human and rodents (63, 82), these cell lines cannot be used for analysis of the SNP rs541856133 but could be useful for other CEL related studies. These cell lines are for example suitable for studying acinar cell secretion, as they possess the necessary machinery to secrete CEL. Nevertheless, previous studies by our research group have shown that both 266-6 and AR42J cell lines are challenging to culture and transfect. Also, because they are derived from mice and rat, they may lack optimal post-translation modifications of the CEL protein, and their regulatory mechanism differs from those of human cells (83), making the generalizability of the results difficult and therefore limiting their utility for studying human *CEL* promoter activity.

In conclusion, while using cell lines that closely resemble physiological conditions would be ideal, they are often more challenging to work with. There are advantages and disadvantages to every cell model. Generally, the more physiologically accurate the cells they are, the more difficult they are to work with. Primary cell lines, which gives a closer resembles to the natural state, are even more challenging to keep in culture than commercial cell lines. Despite the challenges, the selection of the correct cell model is important for obtaining accurate and meaningful results when studying CEL.

6.3.2 Luciferase-assay

In this study we chose to use the Secrete-Pair™ Dual Luminescence Assay from GeneCopoeia (111), where we assessed *CEL* promoter activity by measuring the activity of the secreted reporter protein *Gaussia* Luciferase in media from transfected cells. This method offers higher sensitivity compared to traditional luciferases such as firefly and *Renilla* (145), thereby providing a robust way of assessing promoter activity. Despite its advantages, however, we encountered several challenges that might impact the interpretation of our results.

We observed variability in our results both across different cell lines and within the same cell line when performing technical replicates. The observed differences in *CEL* promoter activity across cell lines could be due to variations in secretion capacity. The GLuc and SEAP proteins might accumulate within the cells instead of being completely secreted into the medium, resulting in lower levels of detected reporter proteins even if the *CEL* promoter activity is high. Further investigations, such as performing luciferase assays on cell lysates could give information about whether the accumulations of GLuc and SEAP proteins inside the cells contributes to the low and inconsistent *CEL* promoter activity observed during this thesis. This

aspect could potentially vary across different cell lines and therefore need to be considered when interpretating the results.

Regarding the variability observed within the same cell line, several factors in the luciferase assay setup could explain this. For example, we seeded the cells in 24-well plates instead of 96-well plates. Using 24-well plates has several disadvantages: i) each experiment required multiple plates to cover all parallels and treatments, ii) variations in cell number between plates, and iii) timing inconsistencies. These factors can influence the results due to small variations in cell environment or handling. Also, minor differences in the timing of transfection could affect transcription efficiency and therefore GLuc/SEAP activity. The variability might be less profound using 96-well plates, as these factors would be reduced. However, as previously presented, the *CEL* promoter activity was low in 96-well plates (**Figure 5.2A**), leading us to use 24-well plates despite the potential impact on the results. Additionally, the internal control SEAP, used for transfection normalization, helps account for some of these variabilities.

Additionally, the luminescence reaction of the assay is time sensitive. The duration for which the diluted substrate is mixed with the cell medium and incubated before measuring the luminescent signal can significantly affect the results across the plate. The Secrete-Pair™ Luminescence assay Kit provides two buffer options, GL-S and GL-H, that can impact the results (111). We opted for the GL-S buffer, which provides a more stable and longer-lasting luminescent signal. While using the GL-H buffer could potentially have led to elevated GLuc activity due to its higher sensitivity, its rapidly decaying signal makes it impractical when managing multiple conditions and several biological replicates, as we did in this study.

7. Conclusion

In this study, we investigated several aspects of *CEL* gene expression and its regulation. Conclusions drawn from this thesis are:

- The *CEL* MUT promoter harboring SNP rs541856133 (T) exhibited higher activity in HEK293 cells compared to the *CEL* WT promoter with SNP rs541856133 (C). The effect seems to be cell-specific, as the difference in activity was not observed in HeLa or ZR-75-1 cells.
- We identified potential transcription factor binding sites at or near SNP rs541856133 in the *CEL* promoter region. ZNF331 was predicted to bind directly at the SNP, while ERα and ERβ were predicted to bind close to the SNP.
- ZNF331 was found to repress the activity of both *CEL* promoter variants in HEK293 cells, but it did not affect *CEL* promoter activity in ZR-75-1 cells.
- ERα was shown to repress the activity of the *CEL* promoter variants in ZR-75-1 cells. This repressing effect was abolished with Tamoxifen treatment, which actually increased the *CEL* promoter activity compared to no treatment, further verifying an interaction between CEL expression and estrogen.
- CEL protein was confirmed to be expressed in the human pituitary gland.
- CEL was found to co-localize with TSH within the pituitary gland.

8. Future perspectives

To follow up on the results presented in this study, future research is crucial to gain deeper insights into *CEL* gene expression and regulation. Below are several aspects that would be interesting to explore further:

CEL promoter activity and the effect of SNP rs541856133 (T):

- Perform more studies to determine if the SNP rs541856133 (T) has any specific effect on *CEL* promoter activity.
- Examine the effect of ZNF331 and ERα on *CEL* promoter activity in a cell model that does not endogenously express these transcription factors.
- Explore the impact of ERβon *CEL* promoter activity
- Investigate the impact of the newly discovered transcription factor ZNF816 on *CEL* promoter activity.
- Conduct electrophoretic mobility shift assay (EMSA) to verify if ERα, ERβ, ZNF331 or ZNF816 binds to the *CEL* promoter.
- Study *CEL* promoter activity in an acinar cell model to better replicate physiological conditions, and therefore enhance the generalizability of the findings.

CEL in the pituitary gland:

- Conduct co-staining of human pituitary tissue for CEL with ZNF331, ERα, and the other pituitary hormones (ACTH, GH, prolactin, gonadotrophin).
- Stain additional human pituitary tissues to evaluate if there are gender differences in the expression pattern of CEL.
- Optimize the staining process for CEL in the mouse pituitary gland. Then analyze the differences in CEL expression between male and female mice, and between *CEL-HYB1* carriers versus controls.

References

1. Pandol SJ. The Exocrine Pancreas. Colloquium Series on Integrated Systems Physiology: From Molecule to Function to Disease. San Rafael (CA)2010.
2. Holck P. *Bukspyttkjertelen Store medisinske leksikon*: snl.no; [updated September 29, 2022. Available from: https://sml.snl.no/bukspyttkjertelen.
3. Atkinson MA, Campbell-Thompson M, Kusmartseva I, Kaestner KH. Organisation of the human pancreas in health and in diabetes. Diabetologia. 2020;63(10):1966-73.
4. Dolenšek J, Rupnik MS, Stožer A. Structural similarities and differences between the human and the mouse pancreas. Islets. 2015;7(1):e1024405.
5. Da Silva Xavier G. The Cells of the Islets of Langerhans. J Clin Med. 2018;7(3).
6. Brereton MF, Vergari E, Zhang Q, Clark A. Alpha-, Delta- and PP-cells: Are They the Architectural Cornerstones of Islet Structure and Co-ordination? J Histochem Cytochem. 2015;63(8):575-91.
7. Talathi SS, Zimmerman R, Young M. Anatomy, Abdomen and Pelvis, Pancreas. StatPearls. Treasure Island (FL)2023.
8. Pandiri AR. Overview of exocrine pancreatic pathobiology. Toxicol Pathol. 2014;42(1):207-16.
9. Campbell-Thompson M, Rodriguez-Calvo T, Battaglia M. Abnormalities of the Exocrine Pancreas in Type 1 Diabetes. Curr Diab Rep. 2015;15(10):79.
10. Store norske leksikon (2005 - 2007). *Bukspytt Store medisinske leksikon*: snl.no; [updated August 11, 2023. Available from: https://sml.snl.no/bukspytt.
11. Logsdon CD, Ji B. The role of protein synthesis and digestive enzymes in acinar cell injury. Nat Rev Gastroenterol Hepatol. 2013;10(6):362-70.
12. Low JT, Shukla A, Thorn P. Pancreatic acinar cell: new insights into the control of secretion. Int J Biochem Cell Biol. 2010;42(10):1586-9.
13. Lee MG, Ohana E, Park HW, Yang D, Muallem S. Molecular mechanism of pancreatic and salivary gland fluid and HCO3 secretion. Physiol Rev. 2012;92(1):39-74.
14. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin. 2021;71(3):209-49.
15. Yadav D, Lowenfels AB. The epidemiology of pancreatitis and pancreatic cancer. Gastroenterology. 2013;144(6):1252-61.
16. Ong KL, Stafford LK, McLaughlin SA, Boyko EJ, Vollset SE, Smith AE, et al. Global, regional, and national burden of diabetes from 1990 to 2021, with projections of prevalence to 2050: a systematic analysis for the Global Burden of Disease Study 2021. The Lancet. 2023;402(10397):203-34.
17. Parker ED, Lin J, Mahoney T, Ume N, Yang G, Gabbay RA, et al. Economic Costs of Diabetes in the U.S. in 2022. Diabetes Care. 2023;47(1):26-43.
18. Hall TC, Garcea G, Webb MA, Al-Leswas D, Metcalfe MS, Dennison AR. The socio-economic impact of chronic pancreatitis: a systematic review. J Eval Clin Pract. 2014;20(3):203-7.
19. Ojo OA, Ibrahim HS, Rotimi DE, Ogunlakin AD, Ojo AB. Diabetes mellitus: From molecular mechanism to pathophysiology and pharmacology. Medicine in Novel Technology and Devices. 2023;19:100247.
20. Committee ADAPP. 2. Diagnosis and Classification of Diabetes: Standards of Care in Diabetes—2024. Diabetes Care. 2023;47(Supplement_1):S20-S42.

21. Katwal D, James D, Dagogo-Jack S. Update on Medical Management of Diabetes: Focus on Relevance for Orthopedic Surgeons. Orthopedic Clinics of North America. 2023;54(3):327-40.
22. Quattrin T, Mastrandrea LD, Walker LSK. Type 1 diabetes. The Lancet. 2023;401(10394):2149-62.
23. Holt RIG, DeVries JH, Hess-Fischl A, Hirsch IB, Kirkman MS, Klupa T, et al. 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). Diabetologia. 2021;64(12):2609-52.
24. Mathieu C, Martens P-J, Vangoitsenhoven R. One hundred years of insulin therapy. Nature Reviews Endocrinology. 2021;17(12):715-25.
25. Del Chierico F, Rapini N, Deodati A, Matteoli MC, Cianfarani S, Putignani L. Pathophysiology of Type 1 Diabetes and Gut Microbiota Role. Int J Mol Sci. 2022;23(23).
26. Ahmad E, Lim S, Lamptey R, Webb DR, Davies MJ. Type 2 diabetes. The Lancet. 2022;400(10365):1803-20.
27. Hu FB. Globalization of diabetes: the role of diet, lifestyle, and genes. Diabetes Care. 2011;34(6):1249-57.
28. Committee ADAPP. 8. Obesity and Weight Management for the Prevention and Treatment of Type 2 Diabetes: Standards of Medical Care in Diabetes—2022. Diabetes Care. 2021;45(Supplement_1):S113-S24.
29. DeMarsilis A, Reddy N, Boutari C, Filippaios A, Sternthal E, Katsiki N, et al. Pharmacotherapy of type 2 diabetes: An update and future directions. Metabolism. 2022;137:155332.
30. Deutsch AJ, Ahlqvist E, Udler MS. Phenotypic and genetic classification of diabetes. Diabetologia. 2022;65(11):1758-69.
31. Association AD. 2. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes—2020. Diabetes Care. 2019;43(Supplement_1):S14-S31.
32. Lemelman MB, Letourneau L, Greeley SAW. Neonatal Diabetes Mellitus: An Update on Diagnosis and Management. Clin Perinatol. 2018;45(1):41-59.
33. Greeley SAW, Polak M, Njølstad PR, Barbetti F, Williams R, Castano L, et al. ISPAD Clinical Practice Consensus Guidelines 2022: The diagnosis and management of monogenic diabetes in children and adolescents. Pediatr Diabetes. 2022;23(8):1188-211.
34. Peixoto-Barbosa R, Reis AF, Giuffrida FMA. Update on clinical screening of maturity-onset diabetes of the young (MODY). Diabetology & Metabolic Syndrome. 2020;12(1):50.
35. Fu J, Retnakaran R. The life course perspective of gestational diabetes: An opportunity for the prevention of diabetes and heart disease in women. EClinicalMedicine. 2022;45:101294.
36. Diagnosis and classification of diabetes mellitus. Diabetes Care. 2011;34 Suppl 1(Suppl 1):S62-9.
37. Vounzoulaki E, Khunti K, Abner SC, Tan BK, Davies MJ, Gillies CL. Progression to type 2 diabetes in women with a known history of gestational diabetes: systematic review and meta-analysis. Bmj. 2020;369:m1361.
38. Aabakken L. *Bukspyttkjertelbetennelse Store medisinske leksikon*: snl.no; [updated December 13, 2022. Available from: https://sml.snl.no/bukspyttkjertelbetennelse.
39. Boxhoorn L, Voermans RP, Bouwense SA, Bruno MJ, Verdonk RC, Boermeester MA, et al. Acute pancreatitis. Lancet. 2020;396(10252):726-34.

40. Banks PA, Bollen TL, Dervenis C, Gooszen HG, Johnson CD, Sarr MG, et al. Classification of acute pancreatitis—2012: revision of the Atlanta classification and definitions by international consensus. Gut. 2013;62(1):102-11.
41. Zerem E, Kurtcehajic A, Kunosić S, Zerem Malkočević D, Zerem O. Current trends in acute pancreatitis: Diagnostic and therapeutic challenges. World J Gastroenterol. 2023;29(18):2747-63.
42. Lee DW, Cho CM. Predicting Severity of Acute Pancreatitis. Medicina (Kaunas). 2022;58(6).
43. Beyer G, Habtezion A, Werner J, Lerch MM, Mayerle J. Chronic pancreatitis. Lancet. 2020;396(10249):499-512.
44. Johnston M, Ravindran R. Chronic pancreatitis. Surgery (Oxford). 2022;40(4):266-73.
45. Vege SS, Chari ST. Chronic Pancreatitis. N Engl J Med. 2022;386(9):869-78.
46. Weiss FU, Laemmerhirt F, Lerch MM. Etiology and Risk Factors of Acute and Chronic Pancreatitis. Visc Med. 2019;35(2):73-81.
47. Hegyi P, Párniczky A, Lerch MM, Sheel ARG, Rebours V, Forsmark CE, et al. International Consensus Guidelines for Risk Factors in Chronic Pancreatitis. Recommendations from the working group for the international consensus guidelines for chronic pancreatitis in collaboration with the International Association of Pancreatology, the American Pancreatic Association, the Japan Pancreas Society, and European Pancreatic Club. Pancreatology. 2020;20(4):579-85.
48. Etemad B, Whitcomb DC. Chronic pancreatitis: Diagnosis, classification, and new genetic developments. Gastroenterology. 2001;120(3):682-707.
49. Strum WB, Boland CR. Advances in acute and chronic pancreatitis. World J Gastroenterol. 2023;29(7):1194-201.
50. Johansson BB, Fjeld K, El Jellas K, Gravdal A, Dalva M, Tjora E, et al. The role of the carboxyl ester lipase (CEL) gene in pancreatic disease. Pancreatology. 2018;18(1):12-9.
51. Lombardo D, Guy O, Figarella C. Purification and characterization of a carboxyl ester hydrolase from human pancreatic juice. Biochim Biophys Acta. 1978;527(1):142-9.
52. Hernell O, Olivecrona T. Human milk lipases. II. Bile salt-stimulated lipase. Biochim Biophys Acta. 1974;369(2):234-44.
53. Abouakil N, Lombardo D. Inhibition of human milk bile-salt-dependent lipase by boronic acids. Implication to the bile salts activator effect. Biochim Biophys Acta. 1989;1004(2):215-20.
54. Ellis LA, Hamosh M. Bile salt stimulated lipase: comparative studies in ferret milk and lactating mammary gland. Lipids. 1992;27(11):917-22.
55. Bläckberg L, Lombardo D, Hernell O, Guy O, Olivecrona T. Bile salt-stimulated lipase in human milk and carboxyl ester hydrolase in pancreatic juice: Are they identical enzymes? FEBS Letters. 1981;136(2):284-8.
56. Roudani S, Miralles F, Margotat A, Escribano MJ, Lombardo D. Bile salt-dependent lipase transcripts in human fetal tissues. Biochim Biophys Acta. 1995;1264(1):141-50.
57. Holtsberg FW, Ozgur LE, Garsetti DE, Myers J, Egan RW, Clark MA. Presence in human eosinophils of a lysophospholipase similar to that found in the pancreas. Biochem J. 1995;309 (Pt 1)(Pt 1):141-4.
58. Li F, Hui DY. Synthesis and secretion of the pancreatic-type carboxyl ester lipase by human endothelial cells. Biochem J. 1998;329 (Pt 3)(Pt 3):675-9.
59. Kodvawala A, Ghering AB, Davidson WS, Hui DY. Carboxyl ester lipase expression in macrophages increases cholesteryl ester accumulation and promotes atherosclerosis. J Biol Chem. 2005;280(46):38592-8.

60. La Rosa S, Vigetti D, Placidi C, Finzi G, Uccella S, Clerici M, et al. Localization of carboxyl ester lipase in human pituitary gland and pituitary adenomas. J Histochem Cytochem. 2010;58(10):881-9.
61. Taylor AK, Zambaux JL, Klisak I, Mohandas T, Sparkes RS, Schotz MC, et al. Carboxyl ester lipase: a highly polymorphic locus on human chromosome 9qter. Genomics. 1991;10(2):425-31.
62. Nilsson J, Blackberg L, Carlsson P, Enerback S, Hernell O, Bjursell G. cDNA cloning of human-milk bile-salt-stimulated lipase and evidence for its identity to pancreatic carboxylic ester hydrolase. Eur J Biochem. 1990;192(2):543-50.
63. Lidberg U, Nilsson J, Stromberg K, Stenman G, Sahlin P, Enerback S, et al. Genomic organization, sequence analysis, and chromosomal localization of the human carboxyl ester lipase (CEL) gene and a CEL-like (CELL) gene. Genomics. 1992;13(3):630-40.
64. Torsvik J, Johansson S, Johansen A, Ek J, Minton J, Ræder H, et al. Mutations in the VNTR of the carboxyl-ester lipase gene (CEL) are a rare cause of monogenic diabetes. Human Genetics. 2010;127(1):55-64.
65. Higuchi S, Nakamura Y, Saito S. Characterization of a VNTR polymorphism in the coding region of the CEL gene. J Hum Genet. 2002;47(4):213-5.
66. Ræder H, Johansson S, Holm PI, Haldorsen IS, Mas E, Sbarra V, et al. Mutations in the CEL VNTR cause a syndrome of diabetes and pancreatic exocrine dysfunction. Nature Genetics. 2006;38(1):54-62.
67. Bengtsson-Ellmark SH, Nilsson J, Orho-Melander M, Dahlenborg K, Groop L, Bjursell G. Association between a polymorphism in the carboxyl ester lipase gene and serum cholesterol profile. European Journal of Human Genetics. 2004;12(8):627-32.
68. Madeyski K, Lidberg U, Bjursell G, Nilsson J. Structure and organization of the human carboxyl ester lipase locus. Mamm Genome. 1998;9(4):334-8.
69. Nilsson J, Hellquist M, Bjursell G. The human carboxyl ester lipase-like (CELL) gene is ubiquitously expressed and contains a hypervariable region. Genomics. 1993;17(2):416-22.
70. Terzyan S, Wang CS, Downs D, Hunter B, Zhang XC. Crystal structure of the catalytic domain of human bile salt activated lipase. Protein Sci. 2000;9(9):1783-90.
71. Hui DY, Howles PN. Carboxyl ester lipase: structure-function relationship and physiological role in lipoprotein metabolism and atherosclerosis. J Lipid Res. 2002;43(12):2017-30.
72. Reue K, Zambaux J, Wong H, Lee G, Leete TH, Ronk M, et al. cDNA cloning of carboxyl ester lipase from human pancreas reveals a unique proline-rich repeat unit. J Lipid Res. 1991;32(2):267-76.
73. Rogers S, Wells R, Rechsteiner M. Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. Science. 1986;234(4774):364-8.
74. Lombardo D. Bile salt-dependent lipase: its pathophysiological implications. Biochim Biophys Acta. 2001;1533(1):1-28.
75. Abouakil N, Mas E, Bruneau N, Benajiba A, Lombardo D. Bile salt-dependent lipase biosynthesis in rat pancreatic AR 4-2 J cells. Essential requirement of N-linked oligosaccharide for secretion and expression of a fully active enzyme. J Biol Chem. 1993;268(34):25755-63.
76. Bruneau N, Lombardo D. Chaperone function of a Grp 94-related protein for folding and transport of the pancreatic bile salt-dependent lipase. J Biol Chem. 1995;270(22):13524-33.
77. Bruneau N, Nganga A, Fisher EA, Lombardo D. O-Glycosylation of C-terminal tandem-repeated sequences regulates the secretion of rat pancreatic bile salt-dependent lipase. J Biol Chem. 1997;272(43):27353-61.

78. Loomes KM, Senior HE, West PM, Roberton AM. Functional protective role for mucin glycosylated repetitive domains. Eur J Biochem. 1999;266(1):105-11.
79. Pasqualini E, Caillol N, Valette A, Lloubes R, Verine A, Lombardo D. Phosphorylation of the rat pancreatic bile-salt-dependent lipase by casein kinase II is essential for secretion. Biochem J. 2000;345 Pt 1(Pt 1):121-8.
80. Bruneau N, Lombardo D, Bendayan M. Participation of GRP94-related protein in secretion of pancreatic bile salt-dependent lipase and in its internalization by the intestinal epithelium. J Cell Sci. 1998;111 (Pt 17):2665-79.
81. Kumar VB, Sasser T, Mandava JB, Sadi HA, Spilburg C. Identification of 5' flanking sequences that affect human pancreatic cholesterol esterase gene expression. Biochemistry and Cell Biology. 1997;75(3):247-54.
82. Kannius-Janson M, Lidberg U, Bjursell G, Nilsson J. The tissue-specific regulation of the carboxyl ester lipase gene in exocrine pancreas differs significantly between mouse and human. Biochem J. 2000;351 Pt 2(Pt 2):367-76.
83. Lidmer AS, Kannius M, Lundberg L, Bjursell G, Nilsson J. Molecular cloning and characterization of the mouse carboxyl ester lipase gene and evidence for expression in the lactating mammary gland. Genomics. 1995;29(1):115-22.
84. Lidberg U, Kannius-Janson M, Nilsson J, Bjursell G. Transcriptional regulation of the human carboxyl ester lipase gene in exocrine pancreas. Evidence for a unique tissue-specific enhancer. J Biol Chem. 1998;273(47):31417-26.
85. Bengtsson SH, Madeyski-Bengtson K, Nilsson J, Bjursell G. Transcriptional regulation of the human carboxyl ester lipase gene in THP-1 monocytes: an E-box required for activation binds upstream stimulatory factors 1 and 2. Biochem J. 2002;365(Pt 2):481-8.
86. Kannius-Janson M, Lidberg U, Hulten K, Gritli-Linde A, Bjursell G, Nilsson J. Studies of the regulation of the mouse carboxyl ester lipase gene in mammary gland. Biochem J. 1998;336 (Pt 3)(Pt 3):577-85.
87. Holmes RS, Cox LA. Comparative Structures and Evolution of Vertebrate Carboxyl Ester Lipase (CEL) Genes and Proteins with a Major Role in Reverse Cholesterol Transport. Cholesterol. 2011;2011:781643.
88. Mao X-T, Deng S-J, Kang R-L, Wang Y-C, Li Z-S, Zou W-B, et al. Homozygosity of short VNTR lengths in the CEL gene may confer susceptibility to idiopathic chronic pancreatitis. Pancreatology. 2021;21(7):1311-6.
89. Brekke RS, Gravdal A, El Jellas K, Curry GE, Lin J, Wilhelm SJ, et al. Common single-base insertions in the VNTR of the carboxyl ester lipase (CEL) gene are benign and also likely to arise somatically in the exocrine pancreas. Hum Mol Genet. 2024.
90. Martinez E, Crenon I, Silvy F, Del Grande J, Mougel A, Barea D, et al. Expression of truncated bile salt-dependent lipase variant in pancreatic pre-neoplastic lesions. Oncotarget. 2016;8(1).
91. Fjeld K, Weiss FU, Lasher D, Rosendahl J, Chen JM, Johansson BB, et al. A recombined allele of the lipase gene CEL and its pseudogene CELP confers susceptibility to chronic pancreatitis. Nat Genet. 2015;47(5):518-22.
92. Fjeld K, Masson E, Lin JH, Michl P, Stokowy T, Gravdal A, et al. Characterization of CEL-DUP2: Complete duplication of the carboxyl ester lipase gene is unlikely to influence risk of chronic pancreatitis. Pancreatology. 2020;20(3):377-84.
93. Ræder H, McAllister FE, Tjora E, Bhatt S, Haldorsen I, Hu J, et al. Carboxyl-ester lipase maturity-onset diabetes of the young is associated with development of pancreatic cysts and upregulated MAPK signaling in secretin-stimulated duodenal fluid. Diabetes. 2014;63(1):259-69.

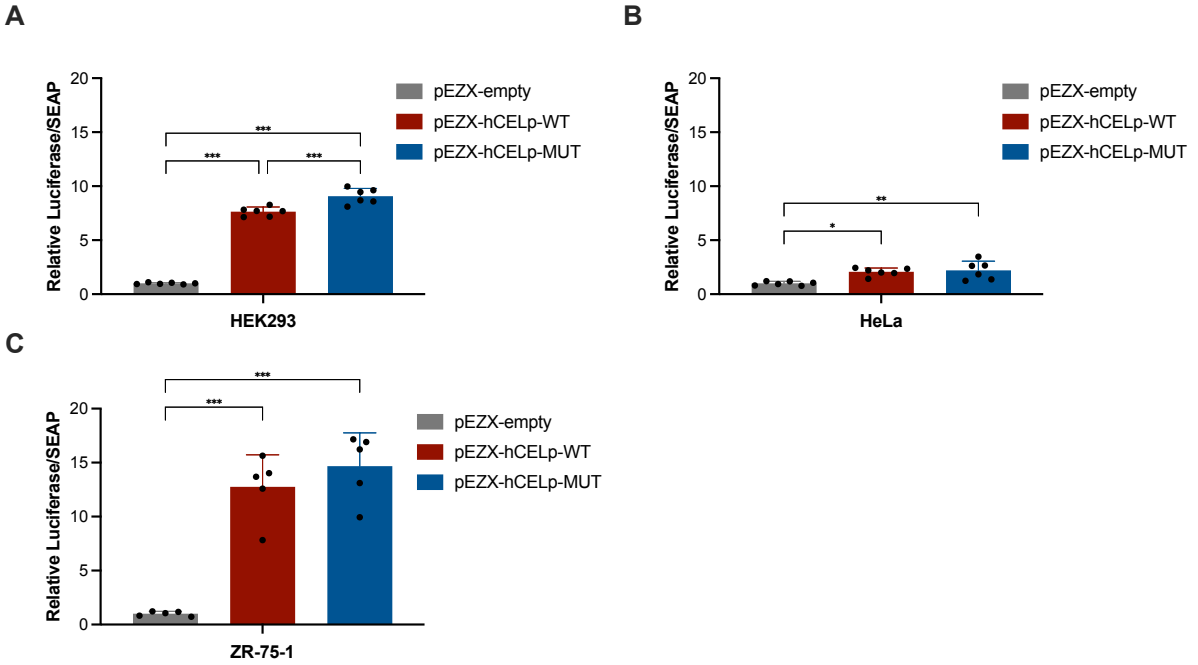
94. El Jellas K, Dušátková P, Haldorsen IS, Molnes J, Tjora E, Johansson BB, et al. Two New Mutations in the CEL Gene Causing Diabetes and Hereditary Pancreatitis: How to Correctly Identify MODY8 Cases. J Clin Endocrinol Metab. 2022;107(4):e1455-e66.
95. Pellegrini S, Pipitone GB, Cospito A, Manenti F, Poggi G, Lombardo MT, et al. Generation of β Cells from iPSC of a MODY8 Patient with a Novel Mutation in the Carboxyl Ester Lipase (CEL) Gene. J Clin Endocrinol Metab. 2021;106(5):e2322-e33.
96. Johansson BB, Torsvik J, Bjørkhaug L, Vesterhus M, Ragvin A, Tjora E, et al. Diabetes and pancreatic exocrine dysfunction due to mutations in the carboxyl ester lipase gene-maturity onset diabetes of the young (CEL-MODY): a protein misfolding disease. J Biol Chem. 2011;286(40):34593-605.
97. Xiao X, Jones G, Sevilla WA, Stolz DB, Magee KE, Haughney M, et al. A Carboxyl Ester Lipase (CEL) Mutant Causes Chronic Pancreatitis by Forming Intracellular Aggregates That Activate Apoptosis. J Biol Chem. 2016;291(44):23224-36.
98. Torsvik J, Johansson BB, Dalva M, Marie M, Fjeld K, Johansson S, et al. Endocytosis of secreted carboxyl ester lipase in a syndrome of diabetes and pancreatic exocrine dysfunction. J Biol Chem. 2014;289(42):29097-111.
99. Kahraman S, Dirice E, Basile G, Diegisser D, Alam J, Johansson BB, et al. Abnormal exocrine-endocrine cell cross-talk promotes β-cell dysfunction and loss in MODY8. Nat Metab. 2022;4(1):76-89.
100. Zou WB, Boulling A, Masamune A, Issarapu P, Masson E, Wu H, et al. No Association Between CEL-HYB Hybrid Allele and Chronic Pancreatitis in Asian Populations. Gastroenterology. 2016;150(7):1558-60 e5.
101. Fjeld K, Weiss FU, Lasher D, Rosendahl J, Chen JM, Johansson BB, et al. A recombined allele of the lipase gene CEL and its pseudogene CELP confers susceptibility to chronic pancreatitis. Nat Genet. 2015;47(5):518-22.
102. Dalva M, Lavik IK, El Jellas K, Gravdal A, Lugea A, Pandol SJ, et al. Pathogenic Carboxyl Ester Lipase (CEL) Variants Interact with the Normal CEL Protein in Pancreatic Cells. Cells. 2020;9(1).
103. Cassidy BM, Zino S, Fjeld K, Molven A, Lowe ME, Xiao X. Single nucleotide polymorphisms in CEL-HYB1 increase risk for chronic pancreatitis through proteotoxic misfolding. Hum Mutat. 2020;41(11):1967-78.
104. Tjora E, Gravdal A, Engjom T, Cnop M, Johansson BB, Dimcevski GG, et al. Protein misfolding in combination with other risk factors in CEL-HYB1-mediated chronic pancreatitis. Eur J Gastroenterol Hepatol. 2021;33(6):839-43.
105. Fjeld K, Gravdal A, Brekke RS, Alam J, Wilhelm SJ, El Jellas K, et al. The genetic risk factor CEL-HYB1 causes proteotoxicity and chronic pancreatitis in mice. Pancreatology. 2022;22(8):1099-111.
106. Molven A, Fjeld K, Lowe ME. Lipase Genetic Variants in Chronic Pancreatitis: When the End Is Wrong, All's Not Well. Gastroenterology. 2016;150(7):1515-8.
107. Chiou J, Geusz RJ, Okino ML, Han JY, Miller M, Melton R, et al. Interpreting type 1 diabetes risk with genetics and single-cell epigenomics. Nature. 2021;594(7863):398-402.
108. Torild Skrivarhaug SJK, Nicolai Andre Lund-Blix. The Norwegian Childhood Diabetes Registry (NCDR) Annual Report 2022. Barnediabetesregisteret, Barne- og ungdomsklinikken, Oslo universitetssykehus HF; 2023.
109. Scientific TF. Interpretation of Nucleic Acid 260/280 Ratios 2012 [Available from: https://assets.thermofisher.com/TFS-Assets/CAD/Product-Bulletins/T123-NanoDrop-Lite-Interpretation-of-Nucleic-Acid-260-280-Ratios.pdf.

110. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. Journal of Molecular Biology. 1990;215(3):403-10.
111. GeneCopoeia I. Secrete-Pair™ Dual Luminescence Assay Kit [User Manual]. 2021 [Available from: https://www.genecopoeia.com/wp-content/uploads/2017/02/Secrete-Pair-%E2%84%A2-Luminescence-Assay-Kit-protocol-20160516.pdf.
112. GeneCopoeia. GLuc-ON Promoter Reporter Clones 2011 [Available from: https://www.genecopoeia.com/wp-content/uploads/2012/08/GLuc-ON_Promoter_Reporter_Clones.pdf.
113. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. Nature Methods. 2012;9(7):676-82.
114. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, et al. The Human Genome Browser at UCSC. Genome Research. 2002;12(6):996-1006.
115. Puig RR, Boddie P, Khan A, Castro-Mondragon JA, Mathelier A. UniBind: maps of high-confidence direct TF-DNA interactions across nine species. bioRxiv. 2021:2020.11.17.384578.
116. Rauluseviciute I, Riudavets-Puig R, Blanc-Mathieu R, Castro-Mondragon Jaime A, Ferenc K, Kumar V, et al. JASPAR 2024: 20th anniversary of the open-access database of transcription factor binding profiles. Nucleic Acids Research. 2023;52(D1):D174-D82.
117. Jiang S, Linghu E, Zhan Q, Han W, Guo M. Methylation of ZNF331 Promotes Cell Invasion and Migration in Human Esophageal Cancer. Current Protein & Peptide Science. 2015;16(4):322-8.
118. Yu J, Liang QY, Wang J, Cheng Y, Wang S, Poon TC, et al. Zinc-finger protein 331, a novel putative tumor suppressor, suppresses growth and invasiveness of gastric cancer. Oncogene. 2013;32(3):307-17.
119. Kuiper GG, Enmark E, Pelto-Huikko M, Nilsson S, Gustafsson JA. Cloning of a novel receptor expressed in rat prostate and ovary. Proc Natl Acad Sci U S A. 1996;93(12):5925-30.
120. Thul PJ, Åkesson L, Wiking M, Mahdessian D, Geladaki A, Ait Blal H, et al. A subcellular map of the human proteome. Science. 2017;356(6340).
121. Zink Finger Protein 331 (Cell Line) Human Protein Atlas v23.0 [Available from: https://www.proteinatlas.org/ENSG00000130844-ZNF331/cell+line.
122. El Jellas K, Johansson BB, Fjeld K, Antonopoulos A, Immervoll H, Choi MH, et al. The mucinous domain of pancreatic carboxyl-ester lipase (CEL) contains core 1/core 2 O-glycans that can be modified by ABO blood group determinants. J Biol Chem. 2018;293(50):19476-91.
123. Carboxyl Ester Lipase (Cell Line) Human Protein Atlas v23.0 [Available from: https://www.proteinatlas.org/ENSG00000170835-CEL/cell+line.
124. Engel LW, Young NA, Tralka TS, Lippman ME, O'Brien SJ, Joyce MJ. Establishment and characterization of three new continuous cell lines derived from human breast carcinomas. Cancer Res. 1978;38(10):3352-64.
125. Quirke VM. Tamoxifen from Failed Contraceptive Pill to Best-Selling Breast Cancer Medicine: A Case-Study in Pharmaceutical Innovation. Front Pharmacol. 2017;8:620.
126. Nathan MR, Schmid P. A Review of Fulvestrant in Breast Cancer. Oncol Ther. 2017;5(1):17-29.
127. Hiller-Sturmhöfel S, Bartke A. The endocrine system: an overview. Alcohol Health Res World. 1998;22(3):153-64.

128. Fadason T, Farrow S, Gokuladhas S, Golovina E, Nyaga D, O'Sullivan JM, et al. Assigning function to SNPs: Considerations when interpreting genetic variation. Seminars in Cell & Developmental Biology. 2022;121:135-42.
129. Degtyareva AO, Antontseva EV, Merkulova TI. Regulatory SNPs: Altered Transcription Factor Binding Sites Implicated in Complex Traits and Diseases. Int J Mol Sci. 2021;22(12).
130. Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Proteomics. Tissue-based map of the human proteome. Science. 2015;347(6220):1260419.
131. Zink Finger Protein 331 (Tissue) Human Protein Atlas v23.0 [Available from: https://www.proteinatlas.org/ENSG00000130844-ZNF331/tissue.
132. Brodie A, Azaria JR, Ofran Y. How far from the SNP may the causative genes be? Nucleic Acids Res. 2016;44(13):6046-54.
133. Estrogen Receptor Alpha (Cell Line) Human Protein Atlas v23.0 [Available from: https://www.proteinatlas.org/ENSG00000091831-ESR1/cell+line.
134. Estrogen Receptor Beta (Cell Line) Human Protein Atlas v23.0 [Available from: https://www.proteinatlas.org/ENSG00000140009-ESR2/cell+line.
135. Jin X, Perrella SL, Lai CT, Taylor NL, Geddes DT. Causes of Low Milk Supply: The Roles of Estrogens, Progesterone, and Related External Factors. Advances in Nutrition. 2024;15(1):100129.
136. Canul-Medina G, Fernandez-Mejia C. Morphological, hormonal, and molecular changes in different maternal tissues during lactation and post-lactation. The Journal of Physiological Sciences. 2019;69(6):825-35.
137. Cui Y, Jiao Y, Wang K, He M, Yang Z. A new prognostic factor of breast cancer: High carboxyl ester lipase expression related to poor survival. Cancer Genet. 2019;239:54-61.
138. Estrogen Receptor Alpha (Tissue) Human Protein Atlas v23.0 [Available from: https://www.proteinatlas.org/ENSG00000091831-ESR1/tissue.
139. Estrogen Receptor Beta (Tissue) Human Protein Atlas v23.0 [Available from: https://www.proteinatlas.org/ENSG00000140009-ESR2/tissue.
140. Bernard V, Young J, Binart N. Prolactin — a pleiotropic factor in health and disease. Nature Reviews Endocrinology. 2019;15(6):356-65.
141. Friend KE, Chiou YK, Lopes MB, Laws ER, Jr., Hughes KM, Shupnik MA. Estrogen receptor expression in human pituitary: correlation with immunohistochemistry in normal tissue, and immunohistochemistry and morphology in macroadenomas. J Clin Endocrinol Metab. 1994;78(6):1497-504.
142. Graham FL, Smiley J, Russell WC, Nairn R. Characteristics of a Human Cell Line Transformed by DNA from Human Adenovirus Type 5. Journal of General Virology. 1977;36(1):59-72.
143. Gey GO. Tissue Culture Studies of the proliferative capacity of cervical carcinoma and normal epithelium. Cancer Research. 1952;12:264-5.
144. Dumont J, Euwart D, Mei B, Estes S, Kshirsagar R. Human cell lines for biopharmaceutical manufacturing: history, status, and future perspectives. Crit Rev Biotechnol. 2016;36(6):1110-22.
145. Tannous BA, Kim DE, Fernandez JL, Weissleder R, Breakefield XO. Codon-optimized Gaussia luciferase cDNA for mammalian gene expression in culture and in vivo. Mol Ther. 2005;11(3):435-43.
146. Carboxyl Ester Lipase (Tissue) Human Protein Atlas v23.0 [Available from: https://www.proteinatlas.org/ENSG00000170835-CEL/tissue.

Appendix

Appendix 1. The Gluc/SEAP activity for the *CEL* promoter variants from **(A-B)** Figure 5.4, **(C)** Figure 5.11, here presented relative to the activity of the empty reporter pEZX-empty.



Appendix 2. RNA levels, shown in normalized transcripts per million (nTPM), of CEL, ZNF331, ERα, and ERβ in various cell lines and tissues, derived from the Human Protein Atlas.

Protein	**Cell lines (120)**		
	ZR-75-1 (nTPM)	**HEK293 (nTPM)**	**HeLa (nTPM)**
CEL (123)	3.1	1.1	0.2
ZNF331 (121)	7.8	74.9	13.7
ERα (133)	31.2	0.0	0.2
ERβ (134)	1.5	0.3	0.6
Protein	**Tissues (130)**		
	Breast (nTPM)	**Pancreas (nTPM)**	**Pituitary gland (nTPM)**
CEL (146)	10.1	107112.3	196.4
ZNF331(131)	18.8	20.3	74.1
ERα (138)	28.3	1	6.4
ERβ (139)	0.9	2	0.7