Carboxyl-ester lipase in human pancreatic disease

A study with focus on genetics, glycosylation and ABO blood groups

Khadija El Jellas

Thesis for the Degree of Philosophiae Doctor (PhD)
University of Bergen, Norway
2018

UNIVERSITY OF BERGEN

Carboxyl-ester lipase in human pancreatic disease

A study with focus on genetics, glycosylation and ABO blood groups

Khadija El Jellas

Thesis for the Degree of Philosophiae Doctor (PhD)
at the University of Bergen

2018

Date of defence: 28.09.2018

Year: 2018

Title: Carboxyl-ester lipase in human pancreatic disease

Name: Khadija El Jellas

Print: Skipnes Kommunikasjon / University of Bergen

Scientific environment

The work presented in this dissertation was funded by the Norwegian Regional Health Authorities (Helse Vest). Additional financial support was obtained from Gades Legat. The thesis studies were carried out at:

Gade Laboratory for Pathology
Department of Clinical Medicine
Faculty of Medicine
University of Bergen
Bergen, Norway

Department of Pathology
Haukeland University Hospital
Bergen, Norway

In collaboration with:

KG Jebsen Center for Diabetes Research
Department of Clinical Science
Faculty of Medicine
University of Bergen
Bergen, Norway

Department of Medical Genetics
Haukeland University Hospital
Bergen, Norway

Biopolymer Mass Spectrometry Lab
Faculty of Natural Sciences
Imperial College London
London, United Kingdom

Acknowledgements

I would like to express my most sincere gratitude to my main supervisor Prof. Anders Molven, you have been a tremendous mentor for me. Thank you for instilling me with your critical thinking and for allowing me to grow as a researcher. I would also like to express my sincere appreciation to my co-supervisors, Researchers Karianne Fjeld and Bente B. Johansson. You have been very supportive during this process; discussions with you have always been exciting and insightful. To my co-supervisor Prof. Pål R. Njølstad, thank you very much for your scientific input and for the positive impact you have had on my work.

This thesis would not have been possible without the efforts of Dr. Dag Hoem. Thanks for supplying the surgical specimens and blood samples, for being so friendly and for inviting me to the operation room. Many thanks to pathologist Heike Immervoll for introducing me to the intricate world of pancreatic pathology. To Prof. Dominique Lombardo, Dr. Eric Mas and the French team; thank you for providing the mAb16D10 antibody, which paved the way for a journey of learning. I also thank all other co-authors.

To Prof. Anne Dell and Dr. Stuart Haslam, thank you very much for hosting me in your lab at Imperial College London and for introducing me to the fascinating world of glycoscience. To Laura Bouche, Aristotelis Antalonopolus, Poh-Choo Pang, Dong Li and Dina Rahman; thank you for teaching me the art of glycomics, for helping me with data acquisition, and for all the fun moments together.

I would like to extend my gratitude to everyone in the MODY group, it has been a pleasure to be part of this "gang". A special thanks to Prof. Stefan Johansson, Prof. Helge Ræader, Dr. Erling Tjora and Janniche Torsvik. To the people at Gades, thank you so much for all the help, technical assistance, coffees and interesting discussions. A special thanks to Solrun Steine, May Brit Kalvenes, Lillian Hallseth, Randi Hope Lavik, Bendik Nordanger, Ingeborg Winge, Kelly Velasco, Åsta Ottesen, Monica Mannelqvist, Benedict Man Hung Choi, Elisabeth Nginamau and Tarig Osman.

Words do not suffice to express my gratitude to my family and my family-like friends overseas and in Bergen, you have been an unlimited source of love and support throughout the entire process; it is great having you.

Bergen, July 2018

Khadija El Jellas

Abstract

Glycosylation is a post-translational modification where carbohydrates are attached to a protein or other organic molecule. Defects in glycosylation are seen in many human pathologies and may even be the underlying cause of some disorders. Here we have studied *CEL* and *ABO*, two genes that play a role in exocrine pancreatic disease, in particular as risk factors for chronic pancreatitis and pancreatic cancer. *CEL* encodes carboxyl-ester lipase, a glycosylated digestive enzyme secreted by the pancreas. The last exon of *CEL* contains a variable number of tandem repeats (VNTR) region, a highly polymorphic sequence translated into a protein C-terminal tail that undergoes mucin-type glycosylation. The *ABO* gene gives rise to a glycosyltransferase that synthesizes the A and B antigens of the ABO blood group system.

We first examined whether genetic variants of *CEL* influenced the risk for the most common and serious type of pancreatic cancer, namely pancreatic ductal adenocarcinoma (PDAC). No association, neither for VNTR length nor for copy number variants (CNVs), was observed. Next, we tested the influence of ABO blood group on the risk for PDAC. Blood group O seemed to have a protective effect as individuals with this blood group were under-represented among cases compared with controls and also exhibited better survival when the tumour was unresectable. We then investigated the expression of *CEL* in PDAC resection specimens and pancreatic cell lines. CEL could not be detected in the neoplastic cells, neither at the mRNA nor at the protein level. When serine-/threonine-linked glycans were released from the mucinous domain of CEL and analysed by high-sensitivity mass spectrometry, we detected core 1- and core 2-based ABH antigen-containing structures in coherence with the genetically determined *ABO* blood type and *FUT2* secretor status of each analyzed individual. The finding of ABH antigens on CEL raises interesting questions of additional biological roles of carboxyl-ester lipase, for example in the gastrointestinal mucosal barrier.

Finally, we performed glycan profiling of the disease-associated CEL variants CEL-MODY and CEL-HYB as well as the normal CEL protein after expression in HEK293 cells. Altered O- and N-linked glycan patters were seen in, respectively, the VNTR and the globular domain of the pathogenic CEL protein variants. In conclusion, although *CEL* VNTR length does not appear to influence susceptibility to exocrine pancreatic disease, aberrant glycosylation of CEL could be involved in pathological processes of the pancreas.

List of publications

Paper I

Monica Dalva, Khadija El Jellas, Solrun J. Steine, Bente B. Johansson, Monika Ringdal, Janniche Torsvik, Heike Immervoll, Dag Hoem, Felix Laemmerhirt, Peter Simon, Markus M. Lerch, Stefan Johansson, Pål R.Njølstad, Frank U. Weiss, Karianne Fjeld and Anders Molven. **Copy number variants and VNTR length polymorphisms of the carboxyl-ester lipase (CEL) gene as risk factors in pancreatic cancer**. *Pancreatology* 2017; 17: 83-88.

Paper II

Khadija El Jellas, Dag Hoem, Kristin G. Hagen, May Britt Kalvenes, Sura Aziz, Solrun J. Steine, Heike Immervoll, Stefan Johansson and Anders Molven. **Associations between ABO blood groups and pancreatic ductal adenocarcinoma: influence on resection status and survival.** *Cancer Medicine* 2017; 6: 1531-1540.

Paper III

Khadija El Jellas, Bente B. Johansson, Karianne Fjeld, Heike Immervoll, Man Hung Choi, Dag Hoem, Mark E. Lowe, Dominique Lombardo, Pål R. Njølstad, Anne Dell, Eric Mas, Stuart M. Haslam and Anders Molven. **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.** Manuscript submitted to *Journal of Biological Chemistry* (2018).

Paper IV

Khadija El Jellas, Stuart M. Haslam, Man Hung Choi, Anne Dell, Pål R. Njølstad, Bente B. Johansson, Karianne Fjeld and Anders Molven. **Altered O- and N-linked glycan profiles in carboxyl ester lipase (CEL) protein variants involved in MODY8 syndrome and chronic pancreatitis.** Manuscript (2018).

Selected abbreviations

CA19.9	Carbohydrate antigen 19
CEA	Carcinoembryonic antigen
CEL/CEL	Carboxyl-ester lipase (gene/protein)
CEL-HYB	*CEL* deletion allele encoding a CEL-CELP hybrid protein
CEL-MODY	MODY8-causing allele encoding the mutant CEL protein
CELP	Carboxyl-ester lipase pseudogene
CHO	Chinese hamster ovary (cell line)
CNV	Copy number variant
CP	Chronic pancreatitis
Del	Deletion
Dup	Duplication
ECM	Extracellular matrix
FAPP	Feto-acinar pancreatic protein
Fuc	Fucose
Gal	Galactose
GalNAc	*N*-acetyl-galactosamine
GlcNAc	*N*-acetyl-glucosamine
GWAS	Genome-wide association study
LacNAc	N-acetyl-lactosamine
MALDI-TOF	Matrix-assisted lased-desorption ionization - Time of flight
Man	Mannose
MS	Mass spectrometry
NeuAC	N-Acetylneuraminic acid
mAb	Monoclonal antibody
MODY	Maturity-onset diabetes of the young
PDAC	Pancreatic ductal adenocarcinoma
PSC	Pancreatic stellate cell
sLe^a	Sialyl-Lewis a
sLe^x	Sialyl-Lewis x
TACA	Tumour-associated carbohydrate antigen
T/sT	T and sialyl-T antigens
Tn/sTn	Tn and sialyl-Tn antigens
UPR	Unfolded protein response
VNTR	Variable number of tandem repeat
WT	Wild type

Contents

1. Introduction

1.1. The healthy pancreas

The mature pancreas is a lobulated organ of approximately 15 cm in length, and weighing, on average, about 80 g in adults. It is located in the peritoneum posterior to the stomach, stretching from the duodenum to the spleen. The gland can be divided into a head, neck, body and the tail region **(Fig. 1)**. The pancreatic head is surrounded by the curved part of the duodenum. The body lies behind the stomach while the protruding tail is in contact with the left kidney and spleen. The main pancreatic duct, averaging 3 mm in diameter, expands across the gland collecting and directing the exocrine secretions into the duodenum. The common bile duct connects the gall bladder with the main pancreatic duct in the head of the pancreas and both open to the duodenum via the major duodenal papilla, also called papilla of Vateri.

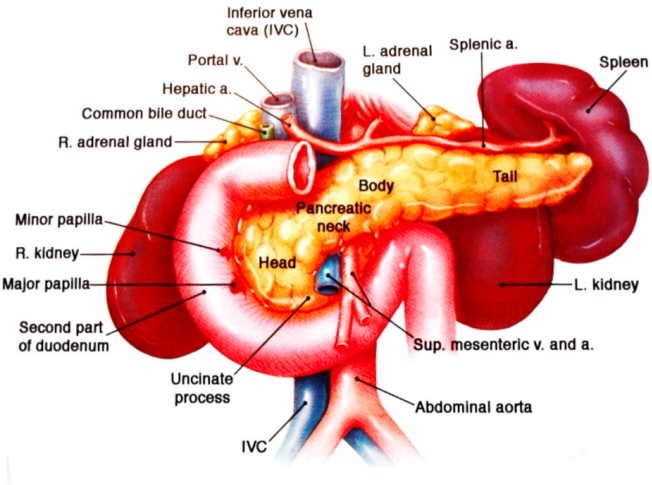


Figure 1. Anatomical relations of the pancreas. The stomach, on the upper front side, and the liver and gall bladder, on the left upper side of the image, are not shown. From the book *Tumors of the pancreas* (1).

The pancreatic gland has a dual physiological role: 1) It stimulates digestion by releasing food-degrading enzymes into the duodenum (exocrine function) and, 2) it produces hormones for maintaining glucose homeostasis (endocrine function).

1.1.1. The exocrine pancreas

The exocrine pancreas consists of acinar cells arranged in small rounded groups (designated acini) and ductal cells **(Fig. 2)**. The acinar cells are highly differentiated and polarized cells with a basally located nucleus and numerous secretory granules in the apical part. The cytoplasm of these cells is strongly basophilic reflecting the specialization of these cells for protein synthesis and secretion of digestive enzymes. In an acinar cell, only 5% of the total cell membrane fraction is plasma membrane whereas up to 60% corresponds to rough endoplasmatic reticulum. In comparison, rough endoplasmatic reticulum constitutes only 35% of total cell membranes in a hepatocyte (2). More than twenty different digestive enzymes, including proteases, amylases, lipases and nucleases (3) are produced and secreted by the acinar cells. Proteases are secreted as inactive precursors, and most of them become activated once they enter the duodenum by a sequential proteolytic cleavage cascade that is initiated by the duodenal enzyme enteropeptidase. On the other hand, amylase, nucleases and lipases are secreted in their partly or fully active form.

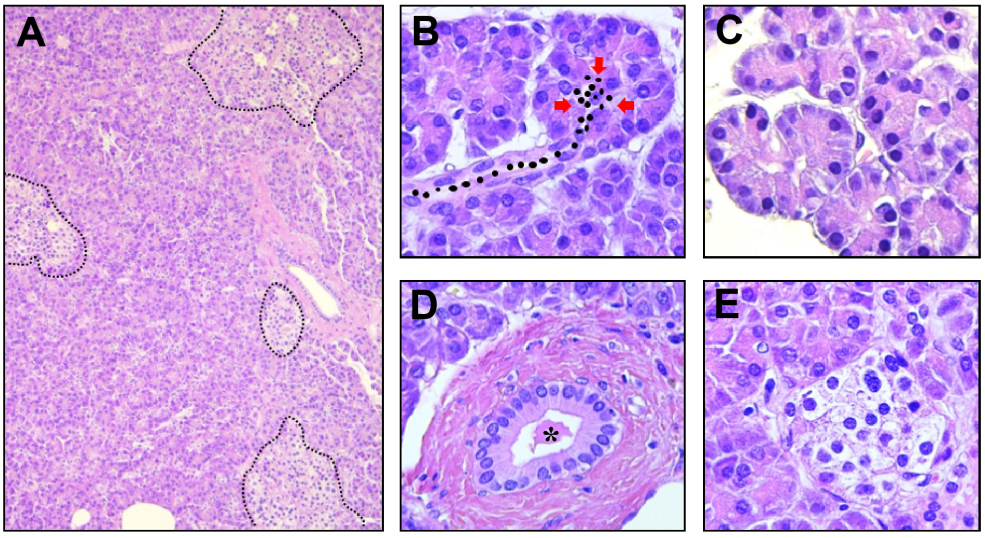


Figure 2. Histology of the normal pancreas. Hematoxylin and eosin staining of normal human pancreas. **(A)** General histological organization of the pancreatic parenchyma. Islets of Langerhans are circumscribed by a dotted line. **(B, C)** Higher magnification of acinar cells arranged in acini. The red arrows indicate how secretion is directed towards an internal lumen created by small interconnecting ducts. The black dots symbolize the transiting digestive enzymes. **(D)** An intralobular duct surrounded by connective tissue. The asterisk indicates amorphic material corresponding to the pancreatic secretion. **(E)** An islet of Langerhans. Magnification in A: 200x, and in B-E: 630x.

The pancreatic ducts are arranged into a branching system that transports the pancreatic secretion towards the duodenum. The ducts vary in diameter and the smallest in size are termed intercalating ducts and are the first that receive the digestive enzymes secreted by the acini (4). The ductal cells secrete water and bicarbonate ($HCO3^-$) in which the digestive enzymes become suspended. This forms the pancreatic secretion, which commonly is referred to as pancreatic juice. From the intercalated ducts the secretion is passed into the intralobular ducts, which then pass it to interlobular ducts, which are larger in size, followed by the main pancreatic duct before its drainage into the duodenum. Around one liter of pancreatic juice is secreted every day, and its alkaline pH serves to neutralize the acidic gastric contents (5).

Another important integrant of the exocrine parenchyma are the so-called pancreatic stellate cells (PSCs). This cell type constitutes roughly 4% of the pancreas cell mass (6). In the healthy pancreas, PSCs appear in a quiescent state where they contribute to maintaining normal tissue structure and architecture by regulating the synthesis and degradation of extracellular matrix (ECM) (7). For this, they do not only secrete ECM proteins but they also produce matrix-degrading enzymes (matrix metalloproteinases), as well as their inhibitors to fine-tune ECM deposition (8). However, in cases of injury, due to a series of phenotypical changes that lead to PSCs activation, the balance between ECM formation and degradation is severely disturbed and excessive fibrogenesis, known as fibrosis, occurs (9). Thus, PSCs play a central role in the desmoplastic reaction in severely damaging pancreatic disorders such as chronic pancreatitis and cancer (10,11).

1.1.2 The endocrine pancreas

The endocrine pancreas consists of a mixed population of cells that are arranged in morphologically recognizable groups called the islets of Langerhans. They are scattered in the exocrine parenchyma and range from 50 to 300 µm in diameter **(Fig. 2A and E)**. The total volume of endocrine cells varies through life, ranging from 15% at the time of birth to only 2-3% in adulthood (12). Each type of endocrine cell produces a single peptide hormone that contributes to the body's glucose homeostasis. α-cells constitute around 15-20% of an islet and secrete glucagon, which stimulates the degradation of glycogen into glucose (glycogenolysis) in liver and skeletal muscle cells and also new synthesis of glucose (gluconeogenesis) in the liver; β-cells make up around 80% of the total islet mass and secrete insulin. This hormone promotes glucose uptake in skeletal muscle, liver and fat cells, and further accumulation of

energy via formation of glycogen (glycogenesis) or fat reserves (lipogenesis). Insulin secretion also blocks glucose production and secretion by the liver. γ-cells or PP-cells are the least frequent cell type in the islets and they produce pancreatic polypeptide, which alters insulin and glucagon secretion, and also has anorexic effects on gastrointestinal functions; δ-cells comprise around 5% of the endocrine tissue and secrete somatostatin that acts in a paracrine manner, exerting inhibitory effects on both α- and β-cells via the somatostatin-receptor expressed in both cell types (13).

1.1.Diseases of the pancreas

The common disorders of the pancreas can be divided into three broad disease groups: diabetes, inflammation (pancreatitis) and neoplasms (cancer). The two latter groups mainly affect the exocrine part of the gland, whereas diabetes is primarily an endocrine disorder.

1.1.1. Diabetes mellitus

The most widespread disease involving the pancreas is diabetes mellitus, often referred to as only diabetes. In this disorder, the hallmark is chronic hyperglycemia due to a deficiency in insulin secretion and/or insulin action. The American Diabetes Association classifies diabetes in four major types: type 1 diabetes, type 2 diabetes, gestational diabetes and other types of diabetes due to specific causes (14).

Type 1 diabetes

Accounting for up to 10% of all diabetes cases, type 1 diabetes is characterized by an autoimmune destruction of the endocrine tissue that leads to severe β-cell loss. Once known as "juvenile-onset diabetes" or "insulin-dependent diabetes", type 1 diabetes is a heterogenous disease most commonly seen in children and youths, and presenting with variable clinical features and disease progression (15,16). The majority of diagnosed type 1 cases are characterized by a T-lymphocyte-mediated autoimmune destruction of the β-cells that leads to insulin deficiency (17). Anti-islet autoantibodies are present in most cases; the most commonly targeted islet components are insulin, glutamic acid decarboxylase protein 65 (GAD65) and zinc transporter 8 (ZnT8) (18). Diabetic ketoacidosis, another hallmark of the disease, is seen in approximately one third of all type 1 patients at the time of diagnosis (19). As the β-cells are permanently destructed, the patients will depend on an exogenous supply of insulin for the rest of their lives.

Type 2 diabetes

The by far most common form of diabetes is type 2, which accounts for around 90% of all diagnosed cases. According to the World Health Organization, the number of people with type 2 diabetes has risen from 108 million in 1980 to 422 million in 2014 (20). The number of affected adults is still expected to rise worldwide, in an epidemic fashion, as this disorder is associated with obesity, a sedentary life style and age (21). Similarly to type 1 diabetes, type 2 diabetes is a heterogeneous disease with a polygenic inheritance pattern and influence of environmental risk factors (22-24).

Type 2 diabetes is characterized by the development of insulin resistance, which occurs in the peripheral tissues such as liver, muscle and adipose tissues when the cells lose their ability to take up glucose from the circulation despite normal or even elevated insulin secretion. As the disease progresses, impaired insulin secretion may also develop. Although there is no autoimmune destruction of β-cells in type 2 diabetes, β-cell dysfunction and finally depletion are linked to the pathophysiology of the disease (25-27). Both type 1 and type 2 diabetes are associated with a high risk for developing diabetes-related complications such as cardiovascular disease, nephropathy, retinopathy and neuropathy (28). This is the reason why diabetes is associated with increased morbidity and mortality.

Monogenic diabetes: MODY

Among the other causes of diabetes recognized by American Diabetes Association are mutations that occur in a single gene. This form of the disease is designated monogenic diabetes. The number of cases probably represents a small fraction of the patients (between 1 and 5%) (29). Up to date, mutations in around 20 different genes have been reported to cause monogenic diabetes (30).

The most common type of monogenic diabetes is called maturity-onset diabetes of the young (MODY) (14). MODY usually occurs in adolescence or early adulthood and an important diagnostic criterion is onset of diabetes before the age of 25 (31). This is, however, the reason why MODY is often misdiagnosed as type 1 diabetes (32). Other criteria for a MODY diagnosis are an autosomal dominant pattern of inheritance and defective insulin production or secretion (31).

Over 90% of all MODY cases are caused by mutations in the genes *GCK, HNF1A* and *HNF4A*. In addition, there are eight much rarer forms of MODY which are caused by mutations in the genes *PDX1*, *HNF1B*, *NEUROD1*, *KLF11*, *CEL*, *PAX4*, *INS* and *BLK* (33,34). Single-base deletions in the *CEL* gene cause MODY8, a syndrome of endocrine and exocrine dysfunction. MODY8 syndrome will be described in more detail in **Section 1.6.**

The different MODY forms are associated with different clinical presentation of the disease. Moreover, the optimal treatment options may vary according to which gene is mutated. Thus, patients with mutations in the genes *HNF1A* and *HNF4A*, with MODY1 and 2, respectively, tend to respond well to sulfonylurea drugs (14,35). On the other hand, MODY2 patients (affecting *GCK*) normally do not require medication as they exhibit mild hyperglycemia that can be controlled mainly by dietary measures. Thus, genetic testing of suspected MODY cases is important for optimal treatment and also to avoid a misdiagnosis as type 1 or type 2 diabetes (30,32).

1.1.2. Inflammatory diseases of the pancreas

Acute pancreatitis

Acute pancreatitis is a sudden and transient inflammation of the pancreas. In most cases, it develops due to passage of biliary tract stones through the main pancreatic duct or as a consequence of prolonged alcohol abuse (36). In the majority of cases, it is a relatively mild form of pancreatitis characterized by interstitial edema and minor foci of necrosis (37). The mild form does usually not involve organ dysfunction and resolves with intravenous fluid rehydration and fasting. However, in 10-20% of all cases, the disease presents in a severe form where abscess, extended necrosis, pseudocysts and hemorrhage are common features. Thus, severe acute pancreatitis can lead to multi-organ failure and be fatal for the patient (38).

Chronic pancreatitis

Chronic pancreatitis (CP) is a long-standing and progressive inflammatory disease characterized by fibrotic destruction of the glandular pancreatic parenchyma (39). The inflammatory process leads to irreversible morphological changes such as loss of the lobular morphology and structure of the pancreas and changes in the arrangement of the islets. Extensive fibrosis, mainly caused by activated PSCs (9), ductal dilation and calcifications are common characteristics of CP (39). The clinical picture includes abdominal pain as well as

pancreatic exocrine and endocrine insufficiency manifesting as maldigestion and diabetes, respectively (39).

CP is caused by a complex combination of environmental and genetic factors. In Western countries, alcohol abuse is the leading cause of CP, accounting for approximately 70% of all cases. In approximately 20% of CP cases, no etiological factors can be identified and these are collectively termed idiopathic CP (40). Hereditary CP, an extremely rare condition with an estimated prevalence of only 3 in one million individuals (41), is diagnosed when three or more family members are affected by chronic pancreatitis and at least two generations are invvolved. The most common cause of hereditary CP are highly penetrant mutations in the cationic trypsinogen gene *PRSS1* (42). This leads to intra-acinar activation of trypsinogen, which is considered pivotal in triggering the pathogenic process. Patients suffering from the hereditary form of CP have a greatly increased risk of developing pancreatic cancer (41,43).

Studies of patient cohorts with familial, early onset and/or or idiopathic CP have revealed several other genes than *PRSS1* to confer susceptibility to the disease. The most classical are the genes coding for the serine protease inhibitor Kazal-type 1 (*SPINK*1) (44), cystic fibrosis transmembrane conductance regulator (*CFTR*) (45), chymotrypsinogen C (*CTRC*) (46) and carboxypeptidase A1 (*CPA1*) (47). In a genome wide association study (GWAS) conducted by Whitcomb *et al.* (48), in addition to the *PRSS1-2* loci, variants of the chromosome X-linked *CLDN2* gene, coding for the cation channel-forming tight junction protein claudin-2, were also seen to influence the risk of pancreatitis. Moreover, Weiss *et al.* identified *ABO* B-blood type and *FUT2* non-secretor status as common population-wide risk factors for developing this disease (49).

1.1.3. Neoplastic diseases of the pancreas

From the more than a dozen types of pancreatic neoplasms are listed in the tumour classification system published by the World Health Organization (50), pancreatic ductal adenocarcinoma (PDAC) is the dominating histologic subtype which accounts for more than 90% of all cases. Pancreatic neuroendocrine tumours are very rare in comparison with those arising in the exocrine parenchyma and comprise less than 4% of all cases. Among the neuroendocrine

tumours, nearly half are functioning neoplasms, i.e. those that secrete hormones. Insulinomas are the most frequent type, although with an incidence of only one case per million (51).

There is a variety of uncommon exocrine tumours of the pancreas such as mucinous and serous cystadenocarcinomas, acinar cell carcinomas and solid-pseudopapillary carcinoma (50). However, the next chapter will focus only on PDAC as it encompasses the vast majority of exocrine pancreatic cancer cases.

1.2. Pancreatic ductal adenocarcinoma (PDAC)

1.2.1. Clinical aspects

Among neoplastic malignancies of the gastrointestinal tract, PDAC represents 20% of the total cases, and is the disease that bears the worst prognosis (52,53). With a long-term survival rate as small as 1-3% PDAC is the fourth leading cause of cancer-related deaths in the industrialized countries (54). Surgical resection of the tumour is today's only possible curative treatment, although only offered to 15-20 % of the patients. The inaccessible anatomical location of the pancreatic gland and the asymptomatic course of PDAC aggravates the disease's bad prognosis. Moreover, PDAC is one of the stiffest tumour types, characterized by a very prominent desmoplastic reaction with a dense fibrotic stroma and low cellularity, implying that a typical primary tumour often contains only 5-20% neoplastic cellularity (55). In addition, the low vascular density of PDAC leads to hypoperfusion of the tumours which, together with the dense stroma, could represent a strong barrier for efficient drug delivery (56).

Etiology of pancreatic cancer

The etiology of PDAC is complex, with multiple gene/environmental interactions. CP is among the strongest risk factors for PDAC, especially long-standing inflammation, as found in hereditary pancreatitis (57,58). Smoking and alcohol abuse, which both are risk factors for CP, are known to increase the risk for PDAC 2-fold (59) and 1.5-fold respectively (60).

Other risk factors that contribute to this malignancy are diabetes mellitus, obesity, old age, positive family history and high fat-dietary regimens (61). Individuals infected with the gastric bacteria *Helicobacter pylori* seem also to be at increased risk of pancreatic cancer (62).

Germline mutations as a cause of PDAC

Hereditary pancreatic cancer is defined by at least two PDAC diagnoses among first-degree relatives. It is estimated that up to 10% of the patients with pancreatic cancer have a family history of the disease (63). There are at least six hereditary cancer predisposition syndromes where there is a clearly increased risk for developing PDAC. These are hereditary breast and ovarian cancer syndrome (with germline mutations in *BRCA2*), familial melanoma (*P16/CDKN2A*), Peutz-Jeghers syndrome (*STK11*), Lynch syndrome (*MLH1*, *MSH2*, *PMS2* or *EPCAM*), familial adenomatous polyposis (*APC*) and Li-Fraumeni syndrome (*TP53*) (64). Moreover, germline mutations in hereditary pancreatitis genes such as *PRSS1* are also associated with PDAC (65). Nevertheless, this group of diseases only account for around one tenth of familial pancreatic cancers.

Apart from the rare, high-penetrant mutations in the genes mentioned above there are more common genetic variants associated with a modest increase risk of PDAC (66). One example is polymorphisms at the *ABO* blood group locus (67). In a large GWAS (67) where 558,542 single nucleotide polymorphisms (SNPs) were genotyped in 1,896 individuals with pancreatic cancer and 1,939 controls, the C variant in SNP rs505922 in intron 1 of the *ABO* gene was seen to increase the risk for pancreatic cancer (odds ratio (OR) = 1.20; 95% CI: 1.12-1.28). Since then, this association has been confirmed in several other reports involving different ethnicity groups (68,69). Additional susceptibility alleles with moderate to small effects identified by GWAS are the nuclear receptor *NR5A2* (70) that is predominantly expressed in exocrine pancreas and liver, and involved in pancreatic differentiation and inflammation (71); two Kruppel-like transcription factors, *KLF5* and *KLF12,* that regulate cell growth and transformation (70); and the cleft lip and palate transmembrane 1-like gene (*CLPTM1L*), part of the *CLPTM1L-TERT* locus that includes the telomerase reverse transcriptase gene (*TERT*; (70)).

Somatic mutations in PDAC

With the aim of understanding the somatic DNA changes underlying the development of PDAC the first whole-exome sequencing study on pancreatic cancer was carried out in 2008 (72). In this paper, Jones *et al.* reported the sequencing of 23,219 transcripts representing 20,661 protein-coding genes in PDAC tumours from 24 different patients. The most frequently reported driver mutations were activating *KRAS* mutations (>95% of cases), inactivating

mutations in the tumour suppressor genes *p16*/*CDKN2A* and *TP53* (70-80%, plus 15% hypermethylation-silencing of *p16*/*CDKN2A*), and inactivation of *SMAD4/DPC4* (60%) (72).

Several follow-up studies have validated these results including two studies published by the Cancer Genome Atlas Research Network (73) and Bailey *et al.* (74), where 150 and 456 PDAC cases, respectively, were examined. In the latter report, based on RNA expression profiles, the authors delineated four different subtypes of PDAC. These were: 1) The squamous type, which showed expression of genes related to inflammation, hypoxia response, metabolic reprogramming, TGF-β signaling, MYC pathways activation and autophagy. This type showed the worst prognosis and was enriched in *TP53* and *KDM6A* mutations; 2) the progenitor type which expressed genes involved in early pancreatic development (*FOXA2/3*, *PDX1* and *MNX1*) as well as *MUC5AC* and *MUC1*; 3) the immunogenic tumour type, which shared many of the characteristics of the progenitor class, but was associated with a significant immune infiltrate. Associated gene programs included B cell signaling pathways, antigen presentation, CD4+ T cell, CD8+ T cell and Toll-like receptor signaling pathways; and 4) the aberrantly differentiated endocrine/exocrine type characterized by upregulated networks of *KRAS* activation, exocrine and also endocrine differentiation.

1.2.2. Pancreatic cancer biomarkers

The majority of cancer biomarkers in clinical use are glycoconjugates, i.e. macromolecules containing glycans; a glycan is a carbohydrate molecule formed by different or the same sugar monomer (monosaccharide) assembled in a very specific manner (75-77). Glycans whose expression is related to cancerous states are often termed tumour-associated carbohydrate antigens (TACA). The most widely known serum biomarker for PDAC, Carbohydrate Antigen 19 (CA19.9), was originally isolated in 1979 from a colorectal carcinoma cell line using a mouse monoclonal antibody (78,79). CA19.9 began to attract attention in the pancreatic cancer field because its serum levels were found to be increased in patients suffering from the disease (80).

Although CA19.9 is the only clinically approved biomarker for pancreatic cancer, its use is limited to monitoring therapy response, and not as a diagnostic marker. This is partially because 1) CA19.9 serum levels are elevated in only two-thirds of early stage PDAC cases (81); 2) CA19.9 is frequently elevated in patients with benign disorders (82,83); and 3) CA19.9 is

chemically characterized as a so-called sialyl-Lewis A (sLe^a) glycan, which means that CA19.9 tests always will give negative results in Lewis blood type negative (Le^{a-b-}) individuals. These persons (around 10% of the Caucasian population) are unable to synthesize the CA19.9 structure due to a non-functional α-1,3/4-fucosyl-transferase (*FUT3*) (80,84,85).

Other carbohydrate-based biomarkers that have been employed in pancreatic cancer detection are CEA (86), CA242 (87), sialyl-Lewis x, (sLe^x (88)), and the non-fucosylated sialyl-Lewis C (detected with the monoclonal antibody Dupan-2 (89)). A recently discovered and promising biomarker for early detection of PDAC is glypican-1, a heavily glycosylated protein, which has been detected in the serum of pancreatic patients only (90). Potential PDAC biomarkers not directly involving glycans include the serum protein named C4b-binding protein alfa-chain (C4BPA) (91), and the macrophage inhibitory cytokine 1 (MIC-1), which has been reported to have a diagnostic accuracy for pancreatic cancer similar to CA19.9 (92).

1.3. Protein glycosylation

Glycosylation is a post-translational, covalent modification where carbohydrates are attached to a protein or other organic molecule. Virtually, every secreted and cell surface protein become glycosylated as they pass through the endoplasmatic reticulum (ER) and Golgi apparatus. The glycome, or complete profile of glycan modifications in a specific macromolecule, is assembled by the coordinated action of numerous glycan modifying enzymes including glycosyltransferases and glycosydases. The monosaccharide units, the building blocks of glycans, are usually represented in a uniform symbol nomenclature (93). There are many different types of protein glycosylation. In this thesis, only mucin-type O-linked glycosylation (the glycan is attached through the hydroxyl group of a serine or a threonine residue) and N-linked glycosylation (attachment through the amino group of an asparagine residue) will be addressed.

1.3.1. Mucin-type O-linked glycosylation

A mucin is a large protein where the amino acid sequence contains peptide segments rich in Ser, Thr and Pro that are repeated hundreds of times. Mucins have a very high content of glycans which is responsible for their large molecular weight and for the water-holding properties that the mucins confer in the intestine and other locations where they are produced (94). The extensive O-glycosylation present within the repeating domains serves to keep the

protein backbone in an extended conformation, transforming it from a globular to an extended rod-like structure (95) and also protects against proteolysis (96). O-liked glycans in cell surface proteins and secretions mediate cell adhesion, recognition, and cell-to-cell communication (97,98). These glycan structures also interact with the microbiota (99) as they can be used as energy sources for enteric bacteria (100) or as anchoring receptors (101). Moreover, glycans have been suggested to mediate the cross-talk between the intestinal cells and the immune system (102).

O-glycans are assembled in the Golgi apparatus following a series of steps where different glycosyl-tranferases build different core structures (core 1 to 4; **Fig. 3**). There exist eight different core structures; however, only core 1-4 have been found in human gastrointestinal mucins **(Table 1)**. The different core structures are further extended or "capped" by sialic acid or fucose, forming terminal epitopes with importance for recognition and adhesion.

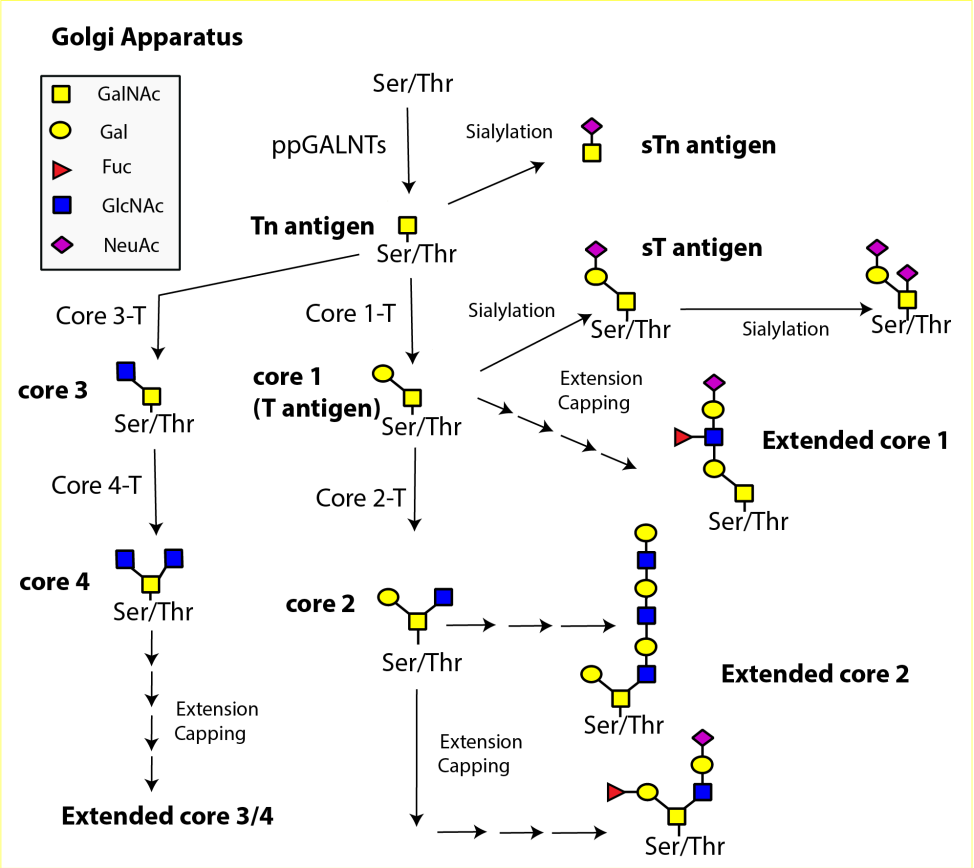


Figure 3: O-linked glycan synthesis pathways. O-glycosylation is initiated by a family of enzymes called polypeptide N-acetylgalactosaminyl-transferases (ppGalNAc-Ts), which is comprised of 20 related enzymes in humans. ppGalNAc-Ts catalyze the addition of GalNAc (via an O-glycosidic linkage) to the hydroxyl groups of Ser/Thr, form an initial structure called Tn. The next step is the addition of galactose (Gal) to form core 1 (or T antigen) structure, catalyzed by the core 1 β1,3-galactosyl-transferase (T-synthase). Core 1 can be further modified to form core 2 structures by core 2 β1,6-N-acetylglucosaminyl-transferase (core 2 synthase). The Tn antigen is also the precursor to core 3 and core 4 structures. All core structures are further modified into extended structures with the addition of other sugars such as Gal, N-acetylgalactosamine (GalNAc), N-acetylglucosamine (GlcNAc), fucose (Fuc), and sialic acid (NeuAc), creating linear or branched glycan structures.

Table 1. O-glycan cores and common antigenic epitopes found in mucins

Core	Glycan
Tn antigen	GalNAcαSer/Thr
Sialyl-Tn antigen	Siaα2-6GalNAcαSer/Thr
Core 1 or T antigen	Galβ1-3GalNAcαSer/Thr
Core 2	GlcNAcβ1-6(Galβ1-3)GalNAcαSer/Thr
Core 3	GlcNAcβ1-3GalNAcαSer/Thr
Core 4	GlcNAcβ1-6(GlcNAcβ1-3)GalNAcαSer/Thr
Terminal antigen*	
Blood group H	Fucα1-2Gal-
Blood group A	GalNAcα1-3(Fucα1-2)Gal-
Blood group B	Galα1-3(Fucα1-2)Gal-
Blood group Lewis a	Galβ1-3(Fucα1-4)GlcNAc-
Blood group Lewis x	Galβ1-4(Fucα1-3)GlcNAc-
Blood group Lewis y	Fucα1-2Galβ1-4(Fucα1-3)GlcNAc-
Sialyl-Lewis x	Siaα2-3Galβ1-4(Fucα1-3)GlcNAc-

*Terminal antigens can also be part of N-linked glycans

Sia, sialic acid or (NeuAc in humans); *Ser/Thr*, serine or threonine residues.

1.3.2. N-linked glycosylation

One of the most common forms of modification for proteins that traverse the ER and secretory pathway is N-linked glycosylation (**Fig. 4**). This attachment occurs early in protein synthesis as the enzyme responsible for the N-glycan transference, the oligosaccharyl-transferase (OST), is part of the ribosome-translocon complex (103,104).

As the newly synthesized peptide emerges in the ER it receives the initial structure (a 14-monosaccharide high mannose glycan) on the amide nitrogen of asparagine in the context of the consensus sequence Asn-X-Ser/Thr (in which X is not a proline). About 70% of proteins contain this recognition sequence. However, it has been estimated that only two-thirds of all N-X-S/T motifs are occupied by N-linked glycans (105). The inital step is followed by a complex process of hydrolytic removal of sugar residues (trimming) and then re-glycosylation with additional sugar residues (processing) such as galactose, fucose, or sialic acid during transit through the endoplasmic ER and Golgi. Contrary to O-glycosylation where biosynthesis occurs by the sequential addition of single monosaccharides, the synthesis of N-glycosylation begins on a lipid and the initial structure is moved *en bloc* to the protein. **Fig. 4** provides a more detailed explanation of N-glycan synthesis. N-linked glycan have a myriad of functions. One example

is that they can assist the glycoprotein maturation process by guiding the protein through a series of quality control steps where the folding is examined by molecular chaperones (106,107).

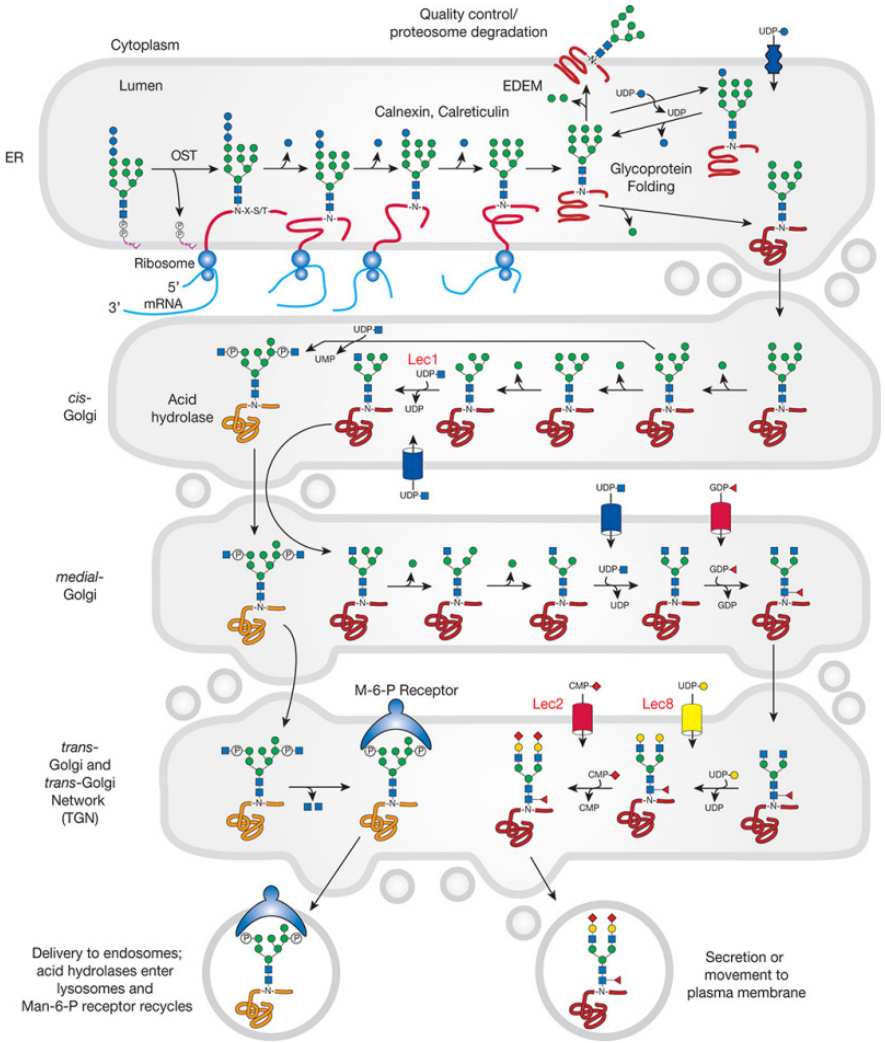


Figure 4: N-glycosylation pathway in the ER and Golgi apparatus. The oligosaccharide GlcNAc2Man9Glc3 is transferred from dolichol-phosphate to the asparagine side chain of the sequence N-X-S/T by the enzyme oligosaccharyl-transferase. Following the transfer of the initial glycan to the protein, glucosidases in the ER remove three glucose (Glc) residues, and ER mannosidase removes a mannose (Man) residue. These reactions are intimately associated with the folding of the glycoprotein assisted by the lectins calnexin and calreticulin, which determine whether the glycoprotein continues to the Golgi or is degraded.

Another lectin, termed ER degradation-enhancing α-mannosidase I-like protein (EDEM), binds to Man residues on misfolded glycoproteins and escorts them via retro-translocation into the cytoplasm for degradation. For correctly folded human proteins, α-mannosidase II removes outer Man residues. The resulting N-glycan is extended by the addition of Fuc, Gal, and NeuAC to generate a complex N-glycan with two or more branches. Image from the book *Essentials of Glycobiology. 2nd edition* (108).

1.3.3. Changes in glycosylation associated with pancreatic disease

Abnormal glycosylation is one of the hallmarks of cancer cells (109,110). Cancer progression and metastasis are characterized by a significant alteration in the glycans signature (111). Such aberrant glycans are functionally important as they can alter cell adhesion, proliferation, survival and metastasis and well as interaction with the immune system (112). Next, a very brief description of the most pronounced changes, with a focus in pancreatic cancer, is provided.

Abnormal O-glycosylation

Abnormal mucin glycosylation is seen in many diseases (113-115). Among the changes seen in pancreatic cancer, perhaps the most predominant is the emergence of the sialylated blood group sLe^a (CA19.9) and sLe^x glycan structures on mucins (116). CA19.9 assays are antibody-based tests (117) that measure the total carbohydrate antigen present in the plasma, although the structure can be attached to different proteins carriers. Yue *et al.* showed that CA19.9 carrier proteins shift between disease states: MUC16 was the preferential carrier in chronic pancreatitis whereas MUC5AC and MUC1 were the most predominant mucin carriers in pancreatic cancer (118).

Another important hallmark of mucins in disease, specially cancer, is their incomplete or truncated glycosylation which leads to the expression of the T and Tn antigens (119) (see also legend to **Fig. 3**). Moreover, a general increase in sialylation accompanies cancerogenesis affecting both O- and N-glycans (120,121). T and Tn, as well as their sialylated variants sialyl-T (sT) and sialyl-Tn (sTn), are small TACAs whose expression at the plasma membrane can enhance tumorigenic and invasive properties (122). This change is often accompanied by up-regulation of sialyl-transferases, e.g. ST3GALI, which convert the T antigen into sialyl-T (sT) antigen, inhibiting the synthesis of core-2 based structures (123).

Abnormal N-glycosylation

N-linked glycans also undergo changes during disease (110). One of the most characteristic changes in neoplastic cells is the increase in β1-6 branching, usually caused by upregulation of the *MGAT5* gene that leads to overexpression of the enzyme N-acetylglucosaminyl-transferase V (121,124). Cells with elevated *MGAT5* expression show an increased frequency of metastasis in mice, and revertants lacking *MGAT5* lose the metastatic phenotype (125).

Increase in fucosylation is also a hallmark of inflammation and cancer, and can affect both N- and O-linked glycans (121). Some proteins with a high degree of fucosylation have been suggested as pancreatic cancer biomarkers, e.g. serum haptoglobin (126) and circulating ribonuclease 1 (127).

1.4. The ABO blood group system

ABO is one of 34 different blood group system known today (128), and together with the Rhesus system, the one with the highest clinical relevance. The ABO antigens, also referred to as ABH, are terminal glycan chain parts of a large soluble or cell surface-anchored glycolipid or glycoprotein. In 1900, Landsteiner observed that the red blood cells of some individuals were agglutinated by the serum of other individuals (129). Looking at the pattern of agglutination he established three different blood groups: A, B and O. Blood group AB was added later, referring to a minority of individuals behaving both A and B. In 1924, Bernstein proposed the "one gene locus-three alleles" model to explain the inheritance of the ABO blood group (130). It was not until 1990 that the elucidation of the molecular genetic basis of the ABO polymorphism was achieved by Yamamoto *et al.* (131).

1.4.1. The ABO locus

The *ABO* locus maps to the long arm of chromosome 9 at q34.2 and codes for different glycosyl-transferases. The coding region comprises 19.5 kilobase (kb) distributed over seven exons. When transcribed, it gives rise to a 1062 base pair (bp) mRNA that is translated into a 41 kDa single pass-transmembrane Golgi-resident protein. Exons 6 and 7 encode over two-thirds of the total protein (354 amino acids) which comprise around 90% of the catalytic domain (131,132). Multiple SNPs have been identified in the *ABO* locus making it one of the most widely characterized polymorphic genes. Today, 335 *ABO* alleles are registered (128). The vast

majority are very uncommon and are categorized as subgroups of the four main alleles: *A*$_1$, *A*$_2$, *B* and *O* (**Fig. 5**). The corresponding phenotypes are present worldwide, but their frequencies differ significantly among ethnicities (133).

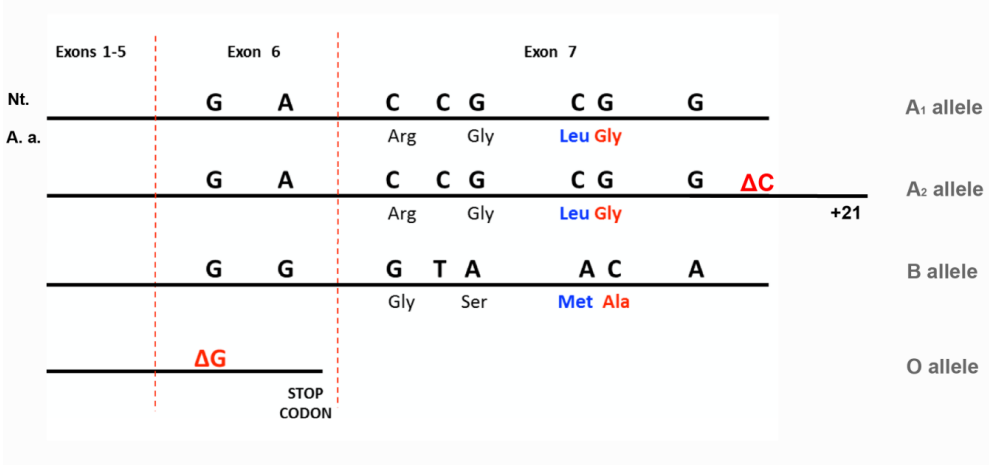


Figure 5. The four main *ABO* alleles. Schematic representation of the nucleotide substitutions in *ABO* gene exons 6 and 7 (letters above the line). The corresponding amino acid substitutions are represented below the line. The A_2 and O allele contain a single base deletion of a C and a G, respectively. The consequence of the former is an extension of the polypeptide by 21 amino acids, whereas the latter results in a truncated protein that lacks the catalytic site due to a premature stop codon. *Nt.*, nucleotide; *A.a.,* amino acid; *Δ*, deletion of a single nucleotide.

1.4.2. The ABO glycosyltransferases

The *A* and *B* alleles encode functional glycosyl-transferases with small differences between them. In contrast, the *O* allele has a single-base deletion, 261delG in exon 6, that shifts the reading frame and results in a truncated protein with no glycosyl-transferase function (131). The *A* and *B* alleles encode two related, but different glycosyl-transferases **(Fig. 6)**. Seven SNPs, one in exon 6 and the rest in exon 7, make transferase A_1 different from the B transferase. Four of these SNPs give rise to amino acid substitutions: Arg176Gly, Gly235Ser, Leu266Met and Gly268Ala. In the catalytic site, Leu266Met and Gly268Ala determine the substrate specificity of the enzymes. Although both enzymes are transmembrane proteins in the Golgi apparatus, it has been noted that upon proteolytical cleavage they can be found as soluble enzymes in plasma, urine and milk (134).

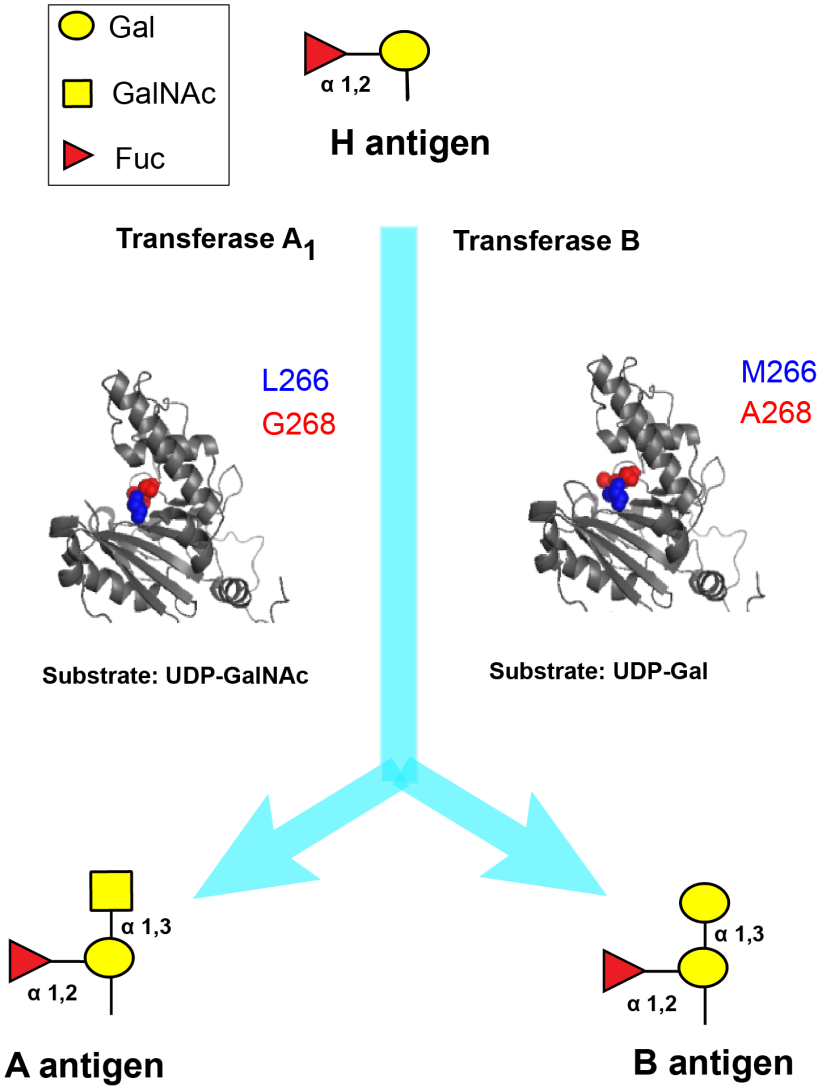


Figure 6. Glycosyl-transferases A_1 and B and their catalytic reaction. Critical residues at the substrate binding pocket (marked in blue and red) determine the substrate specificity of the enzymes. The H antigen is formed by the α-1,2-fucosyl-transferase encoded by the *FUT2* gene, and it is a substrate for A and B transferases to form the A and B antigens, respectively. *L*, leucine; *G*, glycine; *M*, methionine, *A*, alanine; *UDP*, uridine diphosphate.

The A_2 allele is characterized by the deletion 1061delC that disrupts the stop codon. The resulting shift of reading frame yields a protein product with 21 extra amino acid residues at the C-terminus. This slightly longer protein has a diminished A transferase activity characterized by a higher Km. Consequently, there is a qualitative difference in the number of A antigens with $8\text{-}12x10^5$ estimated in an A_1 red blood cell versus $1\text{-}4x10^5$ in an A_2 cell (135). A_1 and A_2 enzymes also differ in biochemical aspects such as differences in optimal catalytic pH and isoelectric point (*pI*).

1.4.3. The ABH antigens

The reaction catalyzed by the *ABO*-coded enzymes is the transference of *N*-acetyl-galactosamine-UDP (GalNAc-UDP) or galactose-UDP (Gal-UDP) to a precursor molecule, the so-called H antigen (Fuc-α1,2-Gal-), to form the A (GalNAc-α1,3(Fuc-α1,2)Gal) and B (Gal-α1,3(Fuc α1,2)Gal) antigens, respectively. Blood group O individuals lack the A or B functional enzymes and therefore express the H antigen without modifications (131).

In addition to red blood cells, the histo-blood group antigens A, B and H (O) are widely present in tissues and also in their soluble form in body fluids. The gene that controls the production and secretion of the H precursor structure in epithelial cells is *FUT2* which determines the "secretor status" of an individual (136), whereas the synthesis of the H antigen in erythrocytes is controlled by a fucosyl-transferase encoded by the *FUT1* gene (137). These genes are homologous and reside on chromosome 19. About 20% of Caucasians are non-secretors, i.e. homozygous for a nonsense mutation in *FUT2*, resulting in a truncated FUT2 fucosyl-transferase (138). Thus, body fluids from A and B non-secretor subjects contain no A or B antigens despite that the subjects have active A and B glycosyl-transferases.

ABH antigens can be present as terminal motifs on mucin-type O-glycans and N-linked glycoproteins that are secreted or membrane-bound. Loss of A and B blood group antigens and the subsequent exposure (or increase) in blood group H (Lew^b) is a common change in glycosylation associated with poor prognosis cancer (139,140).

1.4.4. Anti-ABH natural antibodies

Anti-ABH antibodies are naturally occurring antibodies (NAbs) that belong mainly to the IgM class, but can also be of the IgG3 or IgA type (141). Nabs are part of innate immunity and are present in the blood of mammals at an early age and persist throughout life. They appear from enterobacterial stimulation and are crucially important as they constitute the first line of defense against infections (142). The genesis of Nabs can be understood by employing germ-free mice; In the recent study by Khasbiullina *et al.* (143), it was shown that sterile mice, that were never exposed to any bacteria, had no anti-glycan Abs in their bloodstream. In contrast, animals that were orally inoculated with several bacterial strains as part of the administered food, exhibited significant amounts of antibodies against bacterial polysaccharides, including some reacting towards human ABH blood group antigens.

1.5. Carboxyl-ester lipase: gene and protein

1.5.1. The CEL locus

The human *CEL* gene is located on the long arm of chromosome band 9q34.13 and covers around 10 kb of genomic sequence distributed over 11 exons (144). The last exon contains a very GC-rich variable number of tandem repeats (VNTR) region, constituted by nearly identical 33-bp segments. In all human populations investigated so far, the most frequent *CEL* allele has 16 repeats, although repeat lengths between 3 and 23 have been observed (145,146). Moreover, additional variation is introduced by single-base insertions and deletions within the VNTR (147).

The *CEL* locus also includes a *CEL* pseudogene that is located 11 kb downstream of the *CEL* gene. *CELP* lacks exons 2-7 but the remaining sequence shares 97% homology with *CEL* (148). The pseudogene is present only in man and great apes. It is likely to have originated from a duplication event which occurred relatively late in mammalian evolution (149), and it has been proposed that *CELP* is the original gene because of the sequence similarity between the promoter regions of the mouse *Cel* gene and that of human *CELP*. In humans, *CELP* is reported to be ubiquitously transcribed, but it is not expected to be translated into protein due to a stop codon in exon 8 (148).

1.5.2. CEL glycoprotein: expression, structure and function

The *CEL* gene encodes carboxyl-ester lipase (CEL), also termed bile salt-dependent lipase (BSDL), bile salt-stimulated lipase (BSSL or cholesterol esterase (EC 3.1.1.13). CEL is one of four major lipases secreted by the pancreas to the duodenum, but the only with a requirement for bile salts to be fully active (3,150). CEL hydrolyses a broad spectrum of dietary fats including esters of cholesterol and fat-soluble vitamins (A, D, and E). Moreover, it was recently suggested that CEL has the ability to degrade branched fatty acid esters of hydroxyl fatty acids (FAHFAs), a novel type of fatty acids with anti-diabetic and anti-inflammatory properties (151).

The CEL protein is mainly expressed in the acinar cells of the pancreas where it has been estimated to represent around 4% of the total proteins detected in pancreatic juice (152). However, it is also expressed in human lactating mammary glands, where it is secreted into mother's milk and assists fat digestion in newborns, presumably to compensate for the reduced lipolytic capacity of the newborn's intestine (150). In addition to the high levels in pancreas and mammary glands, *CEL* expression has been documented in the pituitary gland (153), macrophages (154), eosinophils (155), endothelial cells (156), as well as in liver tissue during foetal development (157).

CEL has two major structural domains: an N-terminal globular domain encoded by exons 1-10, and a C-terminal VNTR domain coded by exon 11 **(Fig. 7)**. The globular domain is 535 amino acid long after the signal peptide has been cleaved off, and its three-dimensional structure belongs to the α/β-hydrolase-fold superfamily of hydrolases. The catalytic triad of CEL is composed of Ser194-His435-Asp320, and is part of the catalytic domain that is conserved in all vertebrate species examined to date (158). The asparagine residue at position 210 is modified by a covalently linked N-glycan. This carbohydrate structure has been reported to be required for maximal secretion of CEL (159) and altering it may compromise the enzymatic activity (160). Moreover, two intramolecular disulphide bridges, mediated by Cys64-Cys80 and Cys246-Cys257, stabilize the structure of the CEL globular domain (161).

The COOH-terminus of CEL consists of a VNTR or mucinous domain. This is a structurally disordered tail that adopts a rod-like topological disposition, with a typical architecture of mucins protruding from the globular domain (161). The most common VNTR domain consists

of 16 peptide segments of 11 residues each, all being proline-rich sequences. The repeat numbers 4-9, 11-13, and 15 consist of the basic sequence GAPPVPPTGDS, repeats 1-2 and 14 have EATPVPPTGDS, and the remaining repeats contain minor substitutions of these sequences. The presence of aspartic acid in every repeating unit, and glutamic acid in some, renders this domain highly acidic and contributes to a low isoelectric point (*pI* = 3.3).

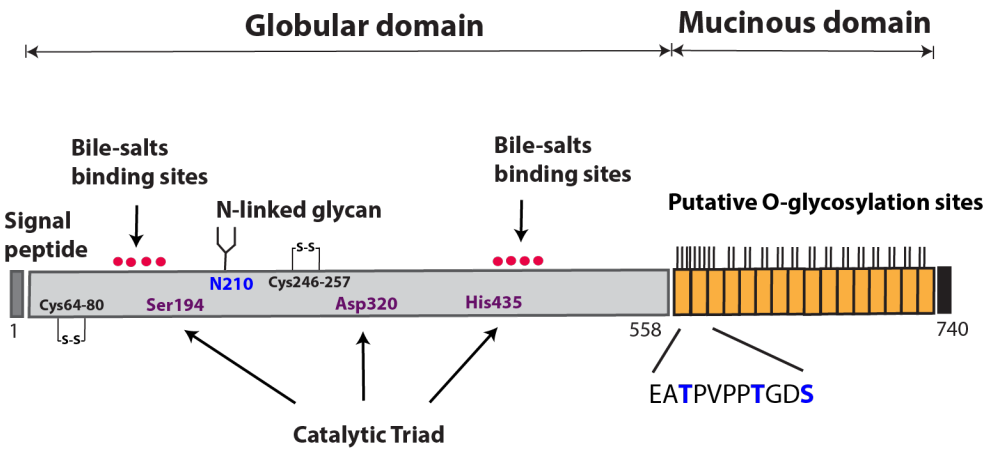


Figure 7. Schematic structure of the CEL protein. The drawing shows the most common CEL variant with 16 tandem repeats in the C-terminal VNTR region. Numbers refer to amino acid positions. An example of the 11 amino acid VNTR repeat sequence (second repeat) is displayed. Vertical lines in the mucin-like domain represent all potential O-glycosylation sites. The structure at residue 210 symbolizes the N-glycan added at the consensus sequence NXT. Two disulfide bridges are also marked. Red dots symbolize bile salt molecules.

1.5.3. CEL protein glycosylation and secretion

The fully processed CEL protein with 16 VNTR repeats is detected by immunoblotting at around 100 kDa in pancreatic secretions (162). This differs from the molecular weight of 79 kDa estimated *in silico* from the 16-repeat protein sequence. This discrepancy is mainly due to O-linked glycosylation of the C-terminus which contains multiple sites susceptible for mucin-type O-glycosylation. According to *in silico* predictions, there are two or three theoretical sites (Ser or Thr) per repeat (163). Studies in BSSL, the milk counterpart of CEL, have found that the sugar content corresponds to Fuc, Gal, GlcNAc, GalNAc, and NeuAc in a molar ratio of 1:3:2:1:0.3 (164). The high content of sialic acid, a negative charged sugar residue, contributes to making the VNTR domain of CEL even more acidic.

The CEL protein begins its journey at the rough ER where the protein is translated. While the nascent polypeptide is still emerging from the ribosome it receives an N-glycan with the structure Glc3Man9GlcNAc2 (165). As shown in **Fig. 4**, this N-glycosylation event marks the beginning of a series of steps where the glucose residues are trimmed or added back allowing the correct folding of the protein through the calnexin/calreticulin cycle (166,167). CEL continues to the Golgi apparatus where the N-glycan is further modified (160). Here mannose residues are trimmed off and the remaining core structure acquires extensions terminating in Fuc and NeuAc residues. In the Golgi apparatus, CEL also undergoes O-glycosylation of the VNTR domain initiated by Golgi-resident ppGalNAcTs **(Fig. 3)**. Once fully N- and O-glycosylated, CEL is phosphorylated at Thr340, which allows final translocation through the secretory pathway, followed by co-storage with other digestive enzymes in zymogen granules (168).

1.5.4. *CEL* in human pancreatic diseases

FAPP – an oncofetal protein form of CEL?

The term oncofetal antigen refers to a substance which is produced by tumours and also by fetal tissues, but is absent or produced in much lower quantities in adult tissues. The term was first used by Gold and Freedman (169) in 1965 to refer to CEA, a cell surface glycoprotein of 180 kDa that was overexpressed in 90% of gastrointestinal malignancies and with 50% of its mass being glycans (170). Soon after, the hepatic glycoprotein α-fetoprotein was also added to this category (171).

To characterize tumour-associated antigens, earlier studies employed xenogeneic antisera. For this, an animal, typically a rabbit or a hamster, was immunized with fetal tissue extract or cancer cell protein extracts, and the antisera used for identifying oncofetal components by posterior immunodetection in diseased or fetal tissues. Using antisera against embryonic pancreas extracts to stain tissue sections of experimentally-induced pancreatic cancer (172) or human pancreatic cancers (173-175), Escribano *et al*. identified an acinar cell-derived molecule which was termed feto-acinar pancreatic protein (FAPP). Later, in the work by Mas *et al.* it was reported that FAPP seemed to have the same amino acid composition as the N-terminal globular domain of CEL (176) and they postulated that it was a CEL variant expressed in embryogenesis that reappears during inflammation and cancer. Since this variant only included six VNTR repeats

in the C-terminus, it was claimed that a deletion of 330 bp in exon 11 of the *CEL* gene was associated with pancreatic cancer.

Subsequently, several papers have been published where FAPP is considered an oncofetal form of CEL (177-179). The monoclonal antibody mAb16D10 was raised against CEL isolated from pancreatic juice of a PDAC patient and was reported to recognize specifically the oncofetal form of CEL (177). In SOJ-6 pancreatic cancer cells, mAb16D10 was found to target an antigen present on the plasma membrane (178), to inhibit the growth of tumour xenografts (178), and to induce cell death by the p53/caspase-dependent apoptotic pathway (177). In human pancreatic tissue sections, mAb16D10 was described to discriminate pancreatic cancer from non-neoplastic tissues and other cancers (179). Although 16D10 glycotope has been reported to be a CEL- and pancreatic cancer-specific glycan in several publications, its exact structure and whether it corresponds to one of the previously described tumour-associated glycan structures is still unknown.

CEL gene variability in pancreatic cancer

Two recent studies have reported that variants within the *CEL* VNTR may contribute to pancreatic cancer risk. Martinez *et al.* (180) claimed that the SNP rs488087 in the second VNTR repeat might serve as a predictive risk factor for pancreatic cancer as carriers of the T variant were overrepresented in a very small set of PDAC cases when compared to different sets of controls. Moreover, from the same group, another report suggested that single-base pair insertions of a cytosine in the VNTR sequence could be early events occurring during pancreatic tumorigenesis (147). This insertion shifts the reading frame and gives rise to a premature stop codon. The result is a slightly truncated protein that the authors propose might be useful for early detection of pancreatic cancer.

The MODY8 syndrome

In 2006, Ræder *et al.* reported that a single-base deletion in the first and fourth repeats of the *CEL* VNTR caused a syndrome of exocrine and endocrine pancreatic dysfunction (33). The disease was discovered in two Norwegian families and fulfilled the MODY criteria (see **Section 1.2.1** above). It was later denoted MODY8 or CEL-MODY (Online Mendelian Inheritance in Man Database # 609812). Both single-base deletions cause a shift in the reading frame and introduce a stop codon after a stretch of around 100 aberrant amino acids (See DEL1 and DEL4

variants in **Fig 8**). The resulting protein contains a truncated COOH-terminus with considerably altered biochemical properties: the CEL-MODY COOH-terminal domain has a very high *pI* value due to an increase in positively charged Arg residues (19 compared to none in normal CEL). As a consequence, this value is shifted from 3.3 in normal CEL to 11.8 in CEL-MODY VNTR domain, and from 5.5 to 9.6 in the overall protein. Moreover, the C-terminal domain of CEL-MODY contains ten Cys residues whose reactive thiol groups can form covalent disulphide bridges (S-S bridges) upon oxidation. The mucinous properties of CEL-MODY are also expected to be altered as it contains a considerably reduced number of putative O-glycosylation sites. Based on the glycosylation-predictor algorithm NetOGlyc 4.0 (163), there are 11 O-glycosylation sites in the VNTR domain of CEL-MODY compared to 36 in the normal protein.

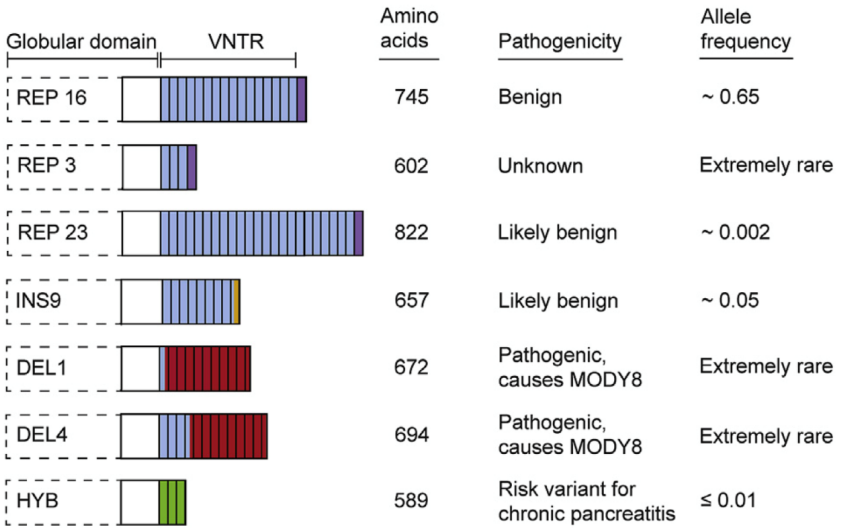


Figure 8. CEL protein variants, their VNTR length, pathogenicity and allele frequency in Northern Europeans. The stippled box represents the CEL globular domain encoded by exons 1-10. The normal VNTR segments are illustrated by blue boxes and the amino acids tailing the VNTR region are marked by a purple box. The most common VNTR variant has 16 repeats (REP16). One extremely short VNTR (REP3) and one very long (REP23) are shown below. INS9 is a normally occurring variant that contains an insertion in the 9th repeat introducing a premature stop codon. A change to red color represents that a frame shift has occurred due to a 1-bp deletion mutation in the first or fourth repeat (DEL1, DEL4), which are the two known MODY8-causing variants. The hybrid CEL protein (HYB) is associated with chronic pancreatitis and

encoded by a recombined allele with three VNTR segments originating from *CELP* (green). Figure published in (181).

Functional studies by Johansson *et al.* and Xiao *et al.* have shown that CEL-MODY adheres to membranes, forms both intra- and extracellular aggregates and introduces ER stress in cell lines (162,182). However, there were discrepancies regarding to which extent the protein was secreted; in the Johansson *et al.* paper, the V5 epitope-tagged mutant protein was constitutively secreted from human embryonic kidney 293 (HEK293) cells, while Xiao *et al.* (182) found that the untagged protein exhibited impaired secretion from rat acinar cells. In the work by Torsvik *et al.* (183), it was demonstrated that CEL-MODY can be endocytosed by HEK293 cells, mouse acinar 266 cells and β-cell-like INS1 cells. Once taken up, the CEL protein seem to undergo lysosomal degradation in the studied cell models (183).

The CEL-HYB allele in chronic pancreatitis

CEL-HYB is a deletion allele and therefore a copy number variant (CNV) of *CEL* that involves the neighbouring *CELP* gene **(Fig. 9)**. *CEL-HYB* was most likely formed by a fusion between the proximal part of *CEL* and the distal part of the pseudogene, resulting in a "hybrid" allele.

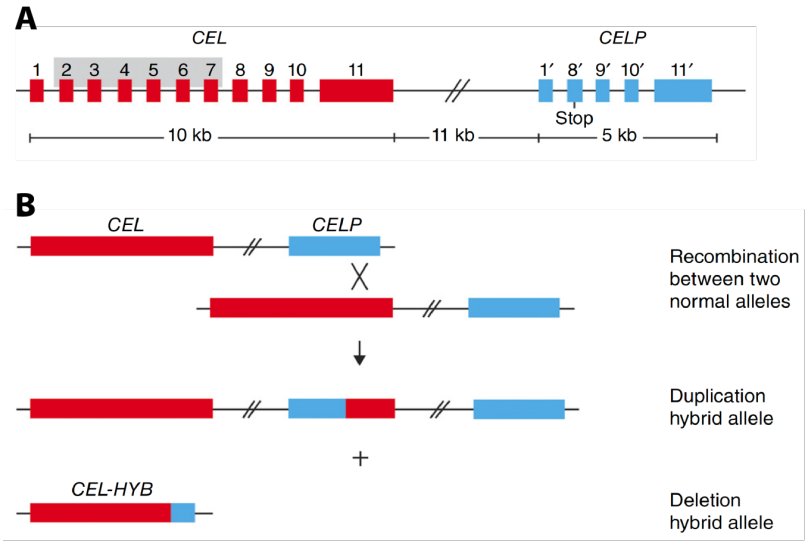


Figure 9. CNVs of the human *CEL* gene. A, structure of the *CEL* locus. The missing region (exon 2–exon 7) in *CELP* is marked in grey in the *CEL* gene. B, proposed mechanism generating *CEL-HYB* by non-allelic homologous recombination between *CEL* and *CELP*. The cross symbolizes a crossover event in the exon 10–exon 11 region. Figure modified from (184).

In 2015, Fjeld *et al.* reported that *CEL-HYB* represents a novel genetic risk factor for chronic pancreatitis (184). In a discovery cohort consisting of 71 German familial chronic pancreatitis cases, *CEL-HYB* was significantly overrepresented among individuals compared to healthy controls (OR = 15.5). This association was replicated in three independent non-alcoholic CP cohorts from France and Germany (n = 1122 cases, OR = 5.2). In an alcoholic CP cohort of 853 cases, there was a smaller, but still significant overrepresentation of the *CEL-HYB* allele (OR = 2.3).

The *CEL-HYB* allele encodes a chimeric CEL protein consisting of the globular domain derived from *CEL* and a COOH-terminal VNTR domain originating from *CELP*. The VNTR domain is shorter, consisting only of 31 amino acids that are arranged in 3 amino acid segments (see HYB in **Fig. 8)** and the total predicted molecular weight of the protein is 65 kDa. When functional consequences of *CEL-HYB* were investigated in HEK293 cells, retention of the protein along the secretory pathway and poor secretion were observed compared to the wild type protein, (184). Moreover, the enzymatic activity of the secreted CEL-HYB fusion protein was reduced to around 40% of the normal CEL protein's activity. Intracellular retention was confirmed by cycloheximide treatment, and an increase of the autophagosomal marker LC3-II suggested that autophagy could be involved in the disease mechanism. Notably, CEL-HYB represents a novel type of CP genetic risk factor as it is not implicated in the well-known trypsin pathway that most CP risk genes are part of.

In an attempt to replicate the *CEL-HYB* association in Asian populations, Zou *et al.* (185) carried out a study using CP patient cohorts from China, Japan and India. They failed to detect the *CEL-HYB* allele in any of the cohorts. Quite surprisingly, though, they discovered a different *CEL-HYB* allele that contained a premature stop-codon in exon 10. This allele was named *CEL-HYB2* and was not associated with increased CP risk, most likely because it resulted in very low expression of the hybrid CEL protein.

2. Aim of the thesis

The general aim of this thesis was to obtain new knowledge about carboxyl-ester lipase and its involvement in human pancreatic diseases. The focus of the work was on genetic variants, glycan patterns and ABO blood groups. Specific aims were as follows:

1) To examine whether *CEL* VNTR length polymorphisms and copy number variants are associated with PDAC risk (Paper I).

2) To investigate whether common variants of the *ABO* blood group gene confer susceptibility to PDAC in a Norwegian patient cohort (Paper II).

3) To elucidate whether glycosylated CEL might be linked to pancreatic cancer and ABO blood groups by examining CEL expression in PDAC tissues, revealing the nature of the 16D10 glycan structure, and identifying O-linked glycans in the mucinous tail region of CEL (Paper III).

4) To determine the glycosylation patterns of pathogenic CEL protein variants expressed in HEK293 cells (Paper IV).

3. Summary of results

Paper I: *Copy number variants and VNTR length polymorphisms of the carboxyl-ester lipase (CEL) gene as risk factors in pancreatic cancer.*

In Paper I, we examined whether CEL CNVs and VNTR length polymorphisms affect the risk for developing pancreatic cancer. CEL CNVs (i.e. duplications and deletions) as well as VNTR lengths were genotyped in 197 Norwegian and 265 German patients diagnosed with pancreatic adenocarcinoma, and in 882 healthy blood donor controls from the same populations. A German family with non-alcoholic chronic pancreatitis and pancreatic cancer was also examined. CNV screening was performed using PCR assays followed by agarose gel electrophoresis, whereas VNTR lengths were determined by DNA fragment analysis and capillary electrophoresis. The investigated family was *CEL-HYB*-positive. However, an association of *CEL-HYB* or a duplication *CEL* allele with pancreatic cancer was not seen in the two patient cohorts examined. The frequency of the 23-repeat VNTR allele was borderline significant in Norwegian cases when compared with controls (1.2% vs. 0.3%; $p = 0.05$). For all other VNTR lengths, no statistically significant difference in frequency was observed. Moreover, no association with pancreatic cancer was detected when CEL VNTR lengths were pooled into groups of short, normal or long alleles. Thus, we were not able to reveal any association between pancreatic cancer and CEL CNVs or *CEL* VNTR lengths. Nevertheless, an effect of very short or long VNTR lengths cannot be ruled out and analyses in larger materials or in other ethnicities are recommended.

Paper II: *Associations between ABO blood groups and pancreatic ductal adenocarcinoma: Influence on resection status and survival.*

Here we investigated the relationship between ABO blood group status and pancreatic ductal adenocarcinoma (PDAC) in a patient cohort from Western Norway ($n = 237$) and two control materials (healthy blood donors, $n = 379$; unselected hospitalized patients, $n = 6149$). When analyzing phenotypes, blood group A was overrepresented among PDAC cases compared to blood donor controls (50.8% vs. 40.6%; OR = 1.51, $p = 0.021$). This enrichment could be explained by the A_1 subgroup distribution as only the A_1 allele (and not A_2) was found to be associated with significantly higher risk for PDAC (23.8% vs. 17.9%; OR = 1.43, $p = 0.018$). When using blood donor controls for comparison, blood group O phenotypic frequency was

lower in cases than in blood donors (33.8% vs. 42.7%; OR = 0.69, $p = 0.039$); this association was also seen when the hospitalized patients group was used for comparison (33.8% vs. 42.9%; OR = 0.68, $p = 0.012$). Results for blood group B were affected by the type of controls selected for statistical comparison. When PDAC patients were divided according to being resected or not, the enrichment of blood group A was most prominent among unresected cases (54.0%), who also had the lowest prevalence of O (28.7%). There was a statistically significant better survival ($p = 0.04$) for blood group O cases than non-O cases among unresected but not among resected patients. The ABO antigens secretor status, assessed by *FUT2* genotyping, did not show an association with PDAC risk in any of the performed comparisons. Thus, we have demonstrated that pancreatic cancer risk is influenced by ABO status in Norwegian patients and, interestingly, that this association might reflect in tumour resectability and survival.

Paper III: *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.*

CEL has been linked to pancreatic cancer through a postulated oncofetal CEL variant termed feto-acinar pancreatic protein (FAPP). This molecule contains a presumably unique glycan structure, 16D10, which is recognized by the monoclonal antibody mAb16D10. However, in this work, we found that CEL was not detectably expressed in neoplastic cells, implying that FAPP is unlikely to be a glyco-isoform of CEL in pancreatic cancer. To investigate whether the 16D10 antigen was an already known cancer-related carbohydrate antigen, we combined characterization by glycan-arrays with matrix-assisted lased-desorption ionization time of flight mass-spectrometry (MALDI-TOF MS) analyses. The mAb16D10 antibody recognized structures containing terminal GalNAc-α1,3(Fuc-α1,2)Gal (= blood group A antigen) and also repeated protein sequences containing GalNAc residues linked to Ser/Thr (= Tn antigen). These findings were supported by immunostainings of human normal and cancerous pancreatic tissue. Further, we used MALDI-TOF mass spectrometry to characterize the released O-glycan pool from isolated CEL from human pancreatic juice. We found that the O-glycome of CEL consisted mainly of core 1/2 structures with a composition depending on *FUT2* and *ABO* gene polymorphisms. When pancreatic juice samples were immunoblotted with anti-Tn antibody, very low levels of Tn antigen were detected at protein bands corresponding to CEL. Moreover, mAb16D10 and anti-A antibodies gave identical immunoblotting patterns suggesting that in pancreatic juice the so-called 16D10 structure corresponds to the A-antigen itself. We conclude that CEL is a glycoprotein with some unique characteristics among digestive enzymes secreted

by the pancreas. This supports the view that CEL could serve biological functions complementary to its cholesteryl esterase activity in the duodenum.

Paper IV: *Altered O- and N-linked glycosylation profiles in carboxyl ester lipase (CEL) protein variants involved in MODY8 syndrome and chronic pancreatitis.*

The aim of this paper was to test whether glycan patterns of the two known pathogenic CEL variants (CEL-HYB, CEL-MODY) would differ from that of the normal CEL protein. To this end, we expressed various CEL constructs in HEK293 cells and determined their O- and N-glycosylation patterns by high-sensitivity MALDI-TOF mass spectrometry. O-glycans attached to the normal CEL protein were found, in almost equal proportions, to be core 1 or core 2 structures terminating with sialic acid but lacking fucose residues. In contrast, over 90% of the O-glycans in CEL-MODY and CEL-HYB were core 1 capped with one or two sialic acids. The N-glycosylation site at N210 carried a mono-sialylated, core-fucosylated bi-antennary structure in normal CEL and CEL-MODY, but contained high levels of fucosylated LacNAc motifs (GalNAc-GlcNAc) in CEL-HYB. In the latter, also the level of unprocessed high mannose structures was increased. When N-glycosylation was disabled by genetically ablating the N210 N-glycosylation site, changes in the subcellular distribution of the CEL variants were noted, indicating that this protein-modification is crucial for correct protein folding, processing and secretion. In conclusion, we observed substantial changes in glycosylation when modelling the two known pathogenic CEL variants in HEK293 cells. To which degree these CEL alterations occur in human acinar tissue and whether they are involved in the initiation/progression of the pancreatic disease process, remains to be investigated.

4. General discussion

4.1. Patient and control selection for association studies

The pancreatic cancer cases from Haukeland University Hospital included in Paper I and Paper II were carefully selected to fulfil diagnostic criteria for PDAC. Histology sections and pathology reports were reviewed by two pathologists. A surgeon collected additional information (i.e. radiology reports, survival data) from the patients' clinical records. In Paper I, 197 Norwegian cases were analysed, of which 156 had a diagnosis of PDAC and 41 were other/unspecified adenocarcinomas anatomically located to the pancreas. This cohort was also studied in Paper II. However, to enlarge the study population we then added a collection of 39 PDAC cases and one other pancreatic adenocarcinoma, for which only formalin-fixed paraffin-embedded samples were available. This cohort had previously been characterized by Immervoll *et al.* (186). As many cases that are referred to as "pancreatic cancer" in the literature are not necessarily of the classical PDAC type, we decided to perform the initial allele frequency analysis of Paper II both for all adenocarcinomas and for the cohort of verified PDAC cases alone. Although our PDAC case number in Paper II is relatively small (n = 195), we consider it a strength that this cohort is very well classified with regard to diagnosis.

We used two control groups, one of 379 blood donors used in both Paper I and II, and another which consisted of 6149 patients admitted Haukeland University Hospital and serotyped during a randomly chosen period of six months. The reason for choosing an additional control group for the *ABO* study was that there could be a bias in ABO blood group distribution among the healthy donors because blood groups is known to associate with some diseases (187). Another point of concern could be that persons with some specific blood types might be more preferred as donors than others. A third problem might be that there are considerable geographical variations in *ABO* frequencies, reflecting human migration movements and possibly also selection pressures from the environment (188,189). Also within Norway, the ABO blood type distribution differs according to geographical localization with the highest blood group O frequency observed in the Western part of the country and the highest B frequency in the northernmost region (190). Blood donors were typically 30 years younger than our patients (Paper II, Table 1), and might be expected to have a wider geographical distribution with regard to origin. The frequency of blood group O was almost identical in the two control groups (42.7 % and 42.9 %), but blood group A and B frequencies varied considerably (A: 40.6 % and 45.8 %; B: 12.1 % and 7.7 % in blood donors and patient controls, respectively). Whether this is due

to people with blood group A more frequently being admitted to hospital than blood group B individuals, whether they have conditions were blood typing is more often needed or whether it reflects a difference in ethnicity of the two control groups, we do not know.

4.2. *CEL* VNTR length polymorphisms in pancreatic disease

Noteworthy, only mammalian *CEL* genes contain the VNTR region as it is absent in lower invertebrates such as fish, reptiles, birds and amphibians (191,192). In mouse and rat, there are 3-4 repeats (192,193), whereas primates generally contain higher repeat numbers (e.g. 39 in gorilla; see refs. (149,192)). The by far most common allele of the human *CEL* gene contains 16 repeats (see e.g. Paper I). How the repetitive tail of *CEL* evolved in mammals is not known. It would be interesting to understand why the tail region expanded in primates and why 16 repeats might be the "optimal" number in humans. We speculate that this new addition to the CEL protein brings along beneficial properties such as better solubility, increased stability and/or enhanced protein secretion. There is also a possibility that the VNTR region confers CEL with advantageous features not related to its role in duodenal hydrolysis of lipids (See Discussion in Paper III). On the other hand, the VNTR did not develop without a cost. A typical sequence of the repeated 33-bp region is GGGGCCCCCCCCGTGCCGCCCACGGG-TGACTCC. The many consecutive guanine and cytosine bases represent a challenge during DNA replication as insertion/deletion mutations tend to arise in such homopolymeric stretches (194). Thus, the MODY8 syndrome mutations are deletions of one base pair within or adjacent to homopolymeric tracts (33).

Due to its polymorphic nature, high GC-content and considerable length variation, sequencing of the VNTR-containing exon 11 of *CEL* is inherently difficult. To capture the full range of variation in this region, DNA sequencing is not sufficient and specific assays such as that developed by Torsvik *et al.* (145) are necessary to perform. This screening method, which consists of a PCR followed by DNA fragment analysis by capillary electrophoresis, was used in Paper I to determine *CEL* VNTR length. We found no association between VNTR length and pancreatic cancer, although an effect from rare alleles could not be excluded and needs to be investigated in very large materials to obtain sufficient statistical power.

Several other studies have linked *CEL* VNTR length to disease susceptibility. Miyasaka *et al.* (195) reported an association between *CEL* VNTR length and alcohol-induced pancreatitis in a Japanese patient cohort. This association, however, was not replicated when German cohorts of idiopathic and alcohol-induced pancreatitis were examined (146). A larger and independent study of German and British patients also failed to replicate the findings of Miyasaka *et al.* but did, on the other hand, observe an association between *CEL* VNTR length and liver cirrhosis (196).

Moreover, a Swedish group (197) reported an association between VNTR length and serum cholesterol profile: carriers of at least one allele with fewer than the common 16 repeats had significantly lower cholesterol levels compared to individuals carrying two normal alleles. A study from the Netherlands, reported that VNTR length polymorphisms are associated with HIV-1 disease progression (198). Heterozygous individuals carrying one short allele (< than 16 repeats) were found to have accelerated disease progression compared to subjects with two normal or longer alleles (≥ 16 repeats). In conclusion, relatively few studies of associations between the highly polymorphic *CEL* VNTR region and human disease have been performed. Although some interesting associations were observed, additional investigations are necessary for verification and for further exploring the role of VNTR-length variation.

4.3. *CEL* CNVs in pancreatic disease

Copy number variation refers to a type of structural genetic variants where a substantial number of base pairs are deleted or duplicated. For the *CEL* gene, both deletions and duplications are known (184,199). However, due to the intrinsic analytical problems regarding the VNTR region (see previous section) and the presence of the neighbouring pseudogene *CELP*, CNVs of *CEL* are not straightforward to identify using high-throughput assays such as GWAS, whole-exome sequencing or whole-genome sequencing (184).

In Paper I, we investigated the association of *CEL* CNVs with pancreatic cancer using a targeted approach. We screened for one deletion (*CEL-HYB*) and one duplication (*CEL-DUP*) variant. CP is a well-established risk factor for pancreatic cancer (200); the discovery of *CEL-HYB* as a *CEL* allele that is associated with CP therefore motivated us to explore whether *CEL* CNVs (both deletions and duplication) could influence the risk of PDAC. Also, we were intrigued by

the finding of *CEL-HYB* carriers in a pedigree with a history of both CP and pancreatic cancer. Nevertheless, we found no association of this *CEL* variant with pancreatic cancer. However, this may not be surprising as there are previous examples where CP genetic risk factors such as *CFTR*, *SPINK*, *CTRC* and *PRSS1* mutations did not associate with increased pancreatic cancer risk (201). In addition, the population carrier frequency of *CEL-HYB* is low, typically less than 1% (184), so that much larger cohorts than ours would be needed if the effect was moderate.

After Paper I was published, a report from Shindo *et al.* also examined *CEL-HYB* in pancreatic cancer (202). The evaluated patient cohort consisted of 850 Americans, of whom 52 were of African descent. At a first glance, their negative result seemed well in line with ours. However, the screening strategy of Shindo *et al.* was probably suboptimal as next-generation sequencing was employed for screening. This most likely implies that a certain number of false positives is expected, which is supported by the relatively high frequency of *CEL-HYB* carriers (2.6% in cases, 1.8% in controls). It is therefore needed to stress the importance of a targeted approach when studying *CEL* genetic diversity. Moreover, when a *CEL-HYB* allele is detected, it should always be verified by Sanger sequencing (184).

Intriguingly, there may exist ethnic-specific variants of *CEL-HYB*. Zou *et al.* (185), when attempting to replicate the association between *CEL-HYB* and chronic pancreatitis in patient cohorts from China, Japan and India, failed to identify any *CEL-HYB* allele carriers. Surprisingly, a different *CEL-HYB* allele that exhibited a premature stop-codon in exon 10, was detected. This variant was named *CEL-HYB2* and did not associate with chronic pancreatitis risk. Based on these findings, it seems that *CEL-HYB* is an ethnic-specific risk allele since it is present in Caucasian populations but not in Asian subjects. Whether *CEL-HYB* is present in other non-Caucasians and, if so, whether it is associated with increased risk for chronic pancreatitis, remains to be investigated.

In Paper I, we also investigated a *CEL* duplication allele (*CEL-DUP;* (184)) with regard to pancreatic cancer risk, again with no significant findings. A limitation to the screening method employed is that it fails to detect duplicated VNTRs of the same length, meaning that a fraction of the "heterozygous samples" (e.g. subjects with VNTR lengths of 15/16) could in fact correspond to CEL-DUP carriers with VNTR lengths of 15/16/16 or 15/15/16. Interestingly, some samples that had a read-out of three different values when assayed for VNTR lengths (e.g. 14/15/17) were not positive for the *CEL-DUP* test. This strongly indicated that other *CEL*

duplication alleles must exist. In fact, a novel duplication allele has recently been discovered in our group (199). In contrast to the duplication allele that we first identified (now denoted *CEL-DUP1*), the newly characterized allele (designated *CEL-DUP2*) contains two complete *CEL* genes which both may give rise to a translated protein. Whether this results in higher levels of CEL protein (e.g. in the duodenum, mother's milk or blood) and whether there are any biological consequences at all of *CEL-DUP2*, remains to be investigated. Fjeld *et al.* also studied *CEL-DUP2* in the context of chronic pancreatitis. The allele frequency was considerably higher in the Chinese than in the French and German materials examined (199). Nevertheless, in none of the cohorts an association was seen with chronic pancreatitis.

4.4. Other variants of the *CEL* gene in pancreatic cancer

One study by Martinez *et al.* (180) has reported an association between a single-nucleotide polymorphism in *CEL* and pancreatic cancer. The authors suggested that C/T at rs488087 in the second VNTR repeat could be a useful marker for defining a population at risk for developing pancreatic cancer as the presented an odds ratio of 4.7 for the T-allele. However, there is a fundamental drawback in this study: an extremely small cohort of pancreatic cancer patients was examined (n = 36); of which 6 cases in fact had a non-PDAC tumour. Moreover, when analysing a small cohort of Norwegian patients (n = 50, of which 31 were PDAC cases and 19 had benign tumours or other pancreatic disease), we found exactly the same T-allele frequency (0.31) as generally seen in European population (203).

In another study by Martinez *et al.* (147), it was claimed that a one-bp insertion in the poly-C tract of the *CEL* VNTR sequence is associated with pancreatic cancer. In human germline DNA, such insertions are commonly seen as 7 of 50 screened samples were positive (203) (see also (33)). The insertions give rise to a premature stop-codon that results in a truncated protein (203). Unfortunately, the patient cohort analysed by Martinez *et al*. was again extremely small (n = 32) and it appears unclear whether the authors refer to germline or somatic insertions variants. The authors present some convincing data with regard to the production of antibodies that are specific for detecting CEL insertion variants at the protein level (147). However, they claim that these antibodies also are positive in pre-neoplastic lesions of the pancreas. In the work carried out in Paper III of this thesis, PanINs as well as malignant infiltrating ducts were negative for CEL expression, both by *in situ* hybridization and immunohistochemistry.

Therefore, the results presented in (147) seem somewhat conflicting with ours in terms of CEL expression.

In a recent study, germline mutations within exons 1-10 in the *CEL* gene have been identified in pancreatic cancer patients. In their patient cohort (n = 1579), Tamura *et al.* (204) described 16 missense mutations in the carboxypeptidase genes *CPA1* and *CPB1* predicted to cause ER-stress as associated with pancreatic cancer risk. No statistical significance was, however, seen with regard to mutations in *CEL*.

4.5. Is there an oncofetal variant of CEL?

In Paper III, contrarily to the work of Benköel *et al.* (179), we did not find evidence that the CEL protein is expressed in PDAC cells. In fact, expression was not detectable at the very early event of acinar-to-ductal metaplasia. Based on the reactivity of the antibodies mAbJ28 and mAb16D10, which recognize glycan structures presumably located in the mucinous domain of CEL, an oncofetal CEL variant has been suggested to exist and denoted FAPP (179). By testing the mAb16D10 antibody in a glycan microarray we demonstrated that it recognized structures containing the blood group A antigen as well as repeated protein sequences containing GalNAc residues linked to serine/threonine (Tn antigen). We did not see a strong Tn band coinciding with CEL on immunoblots (Figure 8 in Paper III). We therefore believe that, when CEL isolated from pancreatic juice was used to produce monoclonal antibodies (178), a blood group A individual may have been selected and that the reactivity of the mAb16D10 antibody then was directed against the blood group A antigen decorating the CEL protein. The findings of Paper III were supported by immunostainings of human pancreatic tissue where anti-A antibodies displayed the same pattern as mAb16D10 in normal or cancerous human pancreatic tissues.

In the study by Pasqualini *et al.* (205), the postulated FAPP was cloned from SOJ-6 cells, and the sequence was reported to contain only six VNTR repeats followed by the normal KEAQMPAVIRF sequence. In light of CEL being a very polymorphic gene due to the VNTR region (Paper I) and knowing the complexity of glycosylation of CEL and its variants (Papers III and IV), we think it is very unlikely that SOJ-6 cells (and other pancreatic cancer cells) contain a specifically glycosylated CEL form with a cancer-specific number of VNTR repeats. Rather, we believe that the SOJ-6 cell line either stems from a subject with only six VNTR

repeats or that it mutated to this repeat number due to the genetic instability of cancer cells. The large majority of pancreatic exocrine cancers exhibits ductal differentiation and are negative for acinar markers. Acinar carcinoma, a very rare subtype of pancreatic cancer, is the only malignant neoplasms where *CEL* expression would be expected to be present (206).

From this and the previous sections, it is clear that CEL has been studied in the context of pancreatic cancer in a number of different ways. The work presented here as Paper III makes it unlikely that CEL is expressed in PDAC neoplastic cells, and our data therefore seriously question the existence of the postulated oncofetal variant FAPP. Taken together, CEL has so far not been convincingly linked to pancreatic cancer risk or progression in any study, neither at the genetic nor at the protein level.

4.6. Glycosylation studies of CEL

4.6.1. Glycosylation of native CEL derived from PDAC patients

In Paper III, the O-glycan profile of CEL was characterized using an MS approach. The protein was purified from pancreatic juice of patients who had been operated for PDAC. The isolated protein could not have stemmed from cancer cells as we found that only the preserved acinar tissue within the morphologically normal-looking pancreatic parenchyma was positive for CEL expression. Notably, the O-glycome of native CEL presented in Paper III varied in composition according to both *ABO* and *FUT2* gene polymorphisms. No classical tumour associated antigens such as $sLew^x$ or $sLew^a$ (CA19.9) were detected, which may support that the protein stemmed from preserved and functionally secreting acini. Moreover, we did not detect the Tn antigen (a structure that results from the truncation of O-glycans to the first monosaccharide only), neither as a molecular ion at m/z 330, nor in immunoblots of isolated CEL protein. However, relatively high levels of sT antigen were seen, but only in the non-secretor sample. These changes in glycosylation among individuals of different *ABO/FUT2* gene makeup reflect an additional level of complexity with regard to CEL protein variability. Whether these differences could accompany a variation in susceptibility risks for certain diseases represents a very interesting question.

The N-glycans of pancreatic CEL are not presented in this thesis. However, our preliminary analyses (El Jellas, unpublished) have revealed that N-linked glycans were mainly of the bi-

antennary complex type with fucose and sialic acid substitutions involving ABH and Lewis blood type terminal moieties. These observations are well in line with previous work on pancreatic CEL (160,207) as well as mother's milk BSSL (208).

4.6.2. Glycosylation studies on recombinant CEL

DNA, RNA or proteins are linear macromolecules that are synthesized by following template structures, and in a similar manner in different cell types. In contrast, glycosylation is a dynamic, not template-driven and strongly cell type-specific process that lead to a heterogeneous pool of branched or linear structures. A group of specific enzymes acts in a sequential manner to build or modify glycans that are generally present in a heterogeneous pool of structures. The same protein expressed in two different cell types can often be modified differently (209-211). Thus, an appropriate selection of cellular model system is pivotal in any glycosylation study.

Although not so widely used as mammalian cells, yeast cells and insect-derived cells have also been exploited with regard to their N- and O-glycosylation capabilities. In the study of Trimble *et al.* (212), CEL was expressed in the yeast *Pichia pastoris*. Not only the N-glycosylation site at N210 of CEL was occupied by high mannose structures, which is the common type of glycosylation in yeast cells, but also the O-linked glycans at the mucinous domain. To avoid this type of structures, which often can be antigenic, human recombinant glycoproteins (e.g. for therapeutic purposes) are preferably produced in mammalian cells (213). Landberg *et al.* (208) compared the O-glycan pattern of native human mammary CEL (i.e. BSSL) to recombinant CEL expressed in CHO cells, the most commonly employed cell line for expressing mammalian glycoproteins. The differences found between the two species of CEL were considerable: native milk CEL contained predominantly large O-linked oligosaccharides, whereas the recombinant form was composed mostly by short O-glycans with a high content of sialic acid. This finding could be reflecting the inability of CHO cells to produce core 2-based O-glycan structures (214,215).

Host selection does not only affect the glycosylation profile, but also the glycosylation site density (210). Consequently, the number of O-glycans attached to native milk CEL estimated by Landberg *et al*. was around nine O-linked glycans and therefore lower than for the

recombinant form produced in CHO cells (between 11 and 16) (208). In another report, Mas *et al.* suggested that pancreatic CEL contained 12-14 O-linked glycans (216). The number of glycosylation sites in CEL will also depend on the number of repeated segments in the mucinous domain, a fact not so often taken into account when studying the protein.

Glycosylation of recombinant CEL-MODY and CEL-HYB

In Paper IV, glycosylation of the pathogenic variants of CEL-MODY and CEL-HYB was examined using HEK293 cells as expression system. With regard to O-linked glycans there were differences between the CEL-WT protein compared to CEL-MODY and CEL-HYB. It has been speculated that the O-glycosylation machinery recognizes the peptide sequence PVPP as consensus motif in CEL (164). The synthesis of core 2-based structures was very low or absent for both pathogenic variants, while nearly half of the CEL-WT O-glycans were core 2-based. Intriguingly, CEL-HYB seemed to be devoid of O-glycans although this finding needs to be verified, for instance, by a glycoproteomics approach or, if possible, by analysing native CEL-HYB protein from carriers of the allele.

The lack of core-2 structures in CEL-MODY glycosylation could be due to the altered peptide backbone biochemical properties, i.e. the high amount of positively charged amino acids and Cys residues, disrupting the glycosyl-transferase-substrate recognition. Whether the lack of proper glycosylation is inducing or contributing to the observed ER stress (182) or other negative intracellular effects, or if the altered glycan pattern is the consequence of the abnormal behavior of the CEL-MODY protein, e. g. by inducing certain glycosyl-transferases such as sialyl-transferases (217), remains to be investigated. Moreover, whether glycosylation of normal CEL (in heterozygous individuals) or other secretory or cell-surface glycoproteins also is affected, is an important question to address.

In the analyses of native pancreatic CEL (Paper III), we did observe fucose residues involved in Lewis and ABO blood group terminal epitopes. However, when analysing O-glycans of recombinant CEL (Paper IV), both core 1- and core 2-based structures were devoid of fucose residues and H antigens were not detected, although *ABO* and *FUT2* genotyping of the HEK293 cells showed that they are derived from a blood group O and a secretor-positive individual. This illustrates that glycosylation patterns obtained from cellular models may not always reflect the genetic constitution of the cells.

The N-glycan pattern of CEL-WT presented in Paper IV, is consistent with reports from others on HEK293 cells with regard to the presence of GalNAc-β-1,4GlcNAc structures (LacdiNAc) (210,218) and of core-fucosylation (219) in mainly bi-antennary complex structures. While CEL-MODY displayed an N-glycan profile similar to that of CEL-WT, the pattern of CEL-HYB was strikingly different. The higher proportion found of oligo- and high-mannose structures could be due to incomplete trimming of the high-mannose glycans, and may reflect a defect in the maturation process of N-glycans. N-glycans are known to be crucial for correct protein folding of glycoproteins (107). In line with this, we observed abnormal CEL-WT/Q210 protein distribution in the cells (Paper IV, **Fig. 5**). Using the rat acinar cells AR42-J cells, Abuakil *et al.* (160) showed that tunicamycin treatment, which inhibits the first step of N-glycan synthesis, leads to a misfolded CEL protein with a dramatic decrease in catalytic activity. Moreover, in the recent report by Dewal *et al.* (220) it was shown that N-glycan architecture is altered upon chemical induction of unfolded-protein response (UPR). Whether the changes in glycosylation seen in CEL-HYB reflect intracellular stress responses to a misfolded CEL-HYB protein requires further investigations.

4.7. The *ABO* gene as a risk factor in pancreatic cancer

The association between ABO blood group and PDAC risk was first reported in 1960 (221), with nearly consistent results in the many studies published since then (for some recent examples, see (67,222-224)). In Paper II, we examined this association in the Norwegian population for the first time (as far as we know). We included the novel aspect of looking into also tumour resectability and survival. We confirmed that the O blood group confers a lower risk for PDAC, as compared with the non-O phenotype. Despite the many published risk studies, the influence of ABO status in cancer prognosis has not been examined extensively. In Paper II we found, in the group of unresected cases, that patients with blood group O survived longer than non-O patients (median 6.7 *vs.* 5.5 months, respectively; $p = 0.04$). In line with our results, Ben *et al.* (225) found that blood group O subjects had better outcome that those with non-O blood group. However, there is a contrary report, by Rahbani *et al.* (226), where the authors observed no statistically significant associations between ABO blood group and PDAC survival. Although the influence of blood groups in PDAC risk is widely accepted, additional, investigations using larger materials are therefore required to clarify the involvement the

common *ABO* genetic polymorphisms in aggressiveness of the disease, as reflected in resectability and survival. In further studies, care should be taken to obtain a correct classification of the PDAC cases, as the term "pancreatic cancer" in many reports may involve cancers originating from the duodenum or bile duct (227) or even endocrine tumours (180).

In Paper II, we also examined the *FUT2* gene variants controlling the ABH blood group secretor status in relation to PDAC risk. *FUT2*/secretor status seems to affect the composition of the intestinal microbiota (228) and the susceptibility to certain gastrointestinal infections (229). Nevertheless, we did not observe a difference in *FUT2* allele distribution or phenotype among pancreatic cancer cases and controls, similar to what a previous report showed (138).

4.7.1. Possible mechanisms behind the ABO and pancreatic cancer link

There is no doubt that *ABO* locus variability influences PDAC risk. However, the exact underlying mechanism by which *ABO* variants interplay with cancer development has not been fully delineated. Molecular mimicry of host glycans by cancer cells and immune system escape is likely to be involved. Upon malignant transformation, truncation of O-linked glycans to small un-extended Tn, or core 1 structure (T antigen) is well documented, especially in mucins (230,231). The structural resemblance of Tn (GalNAc-α) and T (Gal β-1,3GalNAc-α) with the A and B antigens i.e. displaying terminal GalNAc and Gal residues, respectively, could explain the cross-reactivity of naturally occurring anti-A and anti-B antibodies towards such cancer-associated structures.

Having this idea as hypothesis, Hoffman *et al.* (232) showed that sera from healthy blood group O individuals (which contain naturally occurring anti-A and anti-B antibodies), could recognize Tn and T cancer-associated antigens, apart from the A and B antigens. If extensive protein stretches displaying Tn and T antigens, such as seen in cell surface mucins, mimic blood group antigens on epithelial cells of persons of blood group A and B, it is conceivable that cancer cells bearing these antigens more likely would escape the immune system.

Moreover, ABO blood group-containing glycans can modulate the interaction between host and intestinal microbiota by serving as receptors for microorganisms, parasites, and viruses. This is the case of *Helicobacter pylori* which uses Lewis b and blood group O (H) antigens as receptors for colonization of the gastric mucosa, subsequently leading to a peptide ulcer (101,233). Today

the gastrointestinal microbiota is becoming a topic receiving increasing attention. Although no bacterium has been directly linked to PDAC, it cannot be excluded that yet unidentified microbiota interactions that are influenced by ABO blood group, will emerge as involved in PDAC risk.

5. Concluding remarks

The extremely polymorphic nature of *CEL* represents a challenge in genetic studies. In addition to the various SNPs located throughout the gene, *CEL* contains a VNTR region that varies from 3 to 23 repeats. Within the VNTR, there are single-base insertions and deletions that alter the reading frame. Moreover, the presence of an adjacent pseudogene increases the genetic diversity even further, as it predisposes for formation of duplications and deletions involving the whole *CEL-CELP* locus. In this thesis, *CEL* VNTR length and two *CEL* CNVs were examined as potential genetic risk factors for PDAC (Paper I). No associations were revealed. Nevertheless, due to the considerable genetic variability of *CEL* it is still too early to exclude entirely that this gene may influence pancreatic cancer risk.

In contrast, ABO blood group status is a well-established risk factor in pancreatic cancer. We replicated this association in the Norwegian population as we found that both the blood group O allele and the O phenotype was somewhat protective for PDAC and that it also may influence survival (Paper II). Whether naturally occurring antibodies that vary between blood types, contribute to immune protection towards cancer cells by recognizing some tumour-associated carbohydrate antigens, is a very interesting hypothesis that needs to be addressed as a potential PDAC risk modifier.

The VNTR region is an intriguing add-on to the CEL enzyme. The variety of genetic events that alter this region, including some that have been linked to human diseases, suggest that the repetitive C-terminal domain of CEL is relevant for proper functionality of the enzyme. In Paper III, we demonstrated that CEL O-glycans are composed mainly of core 1/core 2 structures with terminal substitutions that depend on the individual's *ABO*/*FUT2* genetic makeup. Thus, the combination of extensive genetic variations in the VNTR and a large number of covalently-coupled carbohydrate branches on the VNTR-encoded protein tail, represents a tremendously large source of CEL protein diversity. Since it is well known that *ABO* and *FUT2* status confer protection/susceptibility to some human disorders, the question of whether specific VNTR lengths in concert with ABO and FUT2 phenotype can be linked to disease is interesting. Moreover, given the very polymorphic nature of the *CEL* gene, it is plausible that new and still uncharacterized *CEL* alleles and their association with disease remain to be unveiled.

When expressing CEL proteins in HEK293 cells (Paper IV) we observed substantial structural changes in glycosylation of the pathogenic CEL-MODY and CEL-HYB variants. Whether similar glycosylation profiles are present in CEL originating from the pancreatic acinar cells of *CEL-MODY* and *CEL-HYB* allele carriers, remains to be investigated. If present, these alterations may play a role in explaining why the pathogenic *CEL* variants predispose for pancreatic disease. Alternatively, the changed glycan profiles may themselves have no pathogenic effect and arise only as a consequence of the altered protein sequence.

6. Future perspectives

The presence of a large and varying set of glycan molecules in the CEL VNTR opens up new possibilities for the roles of this protein, in addition to being a cholesterol esterase in the duodenum. Adhesive properties of the mucinous domain could potentially confer protection against pathogenic viruses or bacteria in the gut and in mother's milk (198,234). Whether changes in glycosylation status, characteristic for certain *CEL* alleles, could be compromising this or other yet uncharacterized functionalities of the CEL glycoprotein remains to be investigated.

The role of the mucinous C-terminal tail of CEL is still enigmatic. As this domain is not needed for enzymatic activity (235,236), it is remarkable that the CEL-HYB protein with only three repeated segments shows a clearly reduced activity in comparison to a variant with no repeats (CEL-TRUNC) (162,184). The latter had an activity similar to CEL-WT when tested *in vitro* (162). Thus, it seems that having no VNTR repeats is less deleterious than having the tail of CEL-HYB with regard to keeping the three-dimensional structure required for a catalytically active enzyme. This may suggest an effect from the two new cysteine residues introduced in the last repeat segment of the CEL-HYB protein (sequence RVCPRPCNG). Potential folding defects compromising the catalytic activity could arise from the presence of reactive thiol groups in these residues, by forming covalent bridges with the four other cysteine residues in the globular domain. Therefore, experiments where the folding of CEL and its protein variants are examined, are highly warranted. This may also shed light on the properties of the CEL-MODY variant, in which there are ten new cysteine residues added, with a great potential for forming intra- and intermolecular cross-linkages that may lead to aggregate formation. Whether the lack of glycosylation sites also contributes to the aggregative behaviour of CEL-MODY is another issue that needs to be addressed.

CEL-HYB increases the risk for chronic pancreatitis in individuals of European ancestry five-fold (184). Still, most carriers of the allele remain healthy, implying that additional genetic and/or environmental factors probably are necessary to precipitate the disease. The dramatic changes that we observed in N-glycosylation of CEL-HYB might be understood in the context of UPR. A very recent study of mice deficient in the UPR/ER stress regulator XBP1 reported that ethanol treatment can cause dimerization of the normal mouse Cel protein, which also contains only three VNTR repeats. Whether ethanol could have an even more dramatic impact

on CEL-HYB due to cysteine-mediated misfolding should now be investigated in order to determine if alcohol abuse could serve as susceptibility factor for chronic pancreatitis development in *CEL-HYB* carriers.

Changes in glycosylation are a common feature in inflammation and cancer, in which an increase in sialylation that particularly affects mucins has been reported (230,231). We propose experiments where stably transfected CEL-MODY-expressing cells are co-transfected with CEL-WT or other secreted mucinous proteins. Such studies might reveal whether the presence of the pathogenic CEL-MODY protein disturbs the glycosylation machinery in general, thereby altering the glycosylation profile of other glycoproteins. Our group has previously shown that the CEL variants, and in particular CEL-MODY, can be endocytosed by non-expressing cells (183). There is also evidence that this process is occurring in the alimentary canal (237). To study whether a possible under-glycosylation of CEL-MODY´s C-terminal tail could stimulate endocytosis, truncation of O-glycosylation should be carried out before performing cellular internalization experiments. To this end, CEL-WT and CEL-MODY Tn antigens could be generated by expression of the protein variants in COSMC chaperon-deficient cells, which are available for CHO (238) or HEK293 cells (163).

In the work done in Paper IV, we employed HEK293 cells because the available acinar cells models, which derive from mouse (266 cells) or rat (AR-42J cells), have proven difficult to transfect with conventional methods. Recently, we have successfully transfected mouse 266 acinar cells with *CEL*-containing plasmids by using electroporation (Johansson *et al.*, unpublished). This opens the possibility to investigate in detail the behaviour of the CEL protein variants in a more relevant model system, although glycosylation differences between rodent and human cells will represent a challenge.

Furthermore, animal studies may serve to give important insights in the pathogenic mechanisms of the CEL-MODY and CEL-HYB variants. A study of whole-body Cel knock-out mice reported no alterations in pancreatic endocrine or exocrine function (239). It was therefore suggested that the MODY8 disease mechanism does not involve a simple loss-of-function and that absent catalytic activity cannot explain the pancreatic phenotype (239). A mouse model is now under construction by our group, and will consist of a humanized knock-in mice where the endogenous mouse *Cel* VNTR is exchanged with the VNTR sequence of CEL-WT, CEL-HYB or CEL-MODY. The latter will hopefully represent the first chronic pancreatitis model that

does not involve disruption of the protease-antiprotease system within the pancreatic acinar cells. Moreover, effects of other risk factors such as smoking and alcohol consumption, can be tested in these mice, with the aim of identifying the precipitating factors of chronic pancreatitis in *CEL-HYB* carriers.

7. References

1. Ralph H. Hruban, M. B. P., David S., M.D. Klimstra. Tumors of the Pancreas (Atlas of Tumor Pathology; 4th Series Fascicle 6) 6th Edition.
2. Alberts B, J. A., Lewis J, et al. Molecular Biology of the Cell. 4th edition. New York: Garland Science; 2002. The Endoplasmic Reticulum. Table 12-12.
Available from: https://www.ncbi.nlm.nih.gov/books/NBK26841/
3. Whitcomb, D. C., and Lowe, M. E. (2007) Human pancreatic digestive enzymes. *Digestive diseases and sciences* **52**, 1-17
4. Stephen J., Pandol. (2010) The Exocrine Pancreas. San Rafael (CA): Morgan & Claypool Life Sciences.
5. Lee, M. G. (2012) Molecular Mechanism of Pancreatic and Salivary Glands Fluid and. **92**, 39-74
6. Apte, M. V., Haber, P. S., Applegate, T. L., Norton, I. D., McCaughan, G. W., Korsten, M. A., Pirola, R. C., and Wilson, J. S. (1998) Periacinar stellate shaped cells in rat pancreas: identification, isolation, and culture. *Gut* **43**, 128-133
7. Means, A. L. (2013) Pancreatic stellate cells: small cells with a big role in tissue homeostasis. *Laboratory investigation; a journal of technical methods and pathology* **93**, 4-7
8. Phillips, P. A., McCarroll, J. A., Park, S., Wu, M. J., Pirola, R., Korsten, M., Wilson, J. S., and Apte, M. V. (2003) Rat pancreatic stellate cells secrete matrix metalloproteinases: implications for extracellular matrix turnover. *Gut* **52**, 275-282
9. Apte, M. V., Pirola, R. C., and Wilson, J. S. (2012) Pancreatic stellate cells: a starring role in normal and diseased pancreas. *Frontiers in physiology* **3**, 344
10. Apte, M. V., Pirola, R. C., and Wilson, J. S. (2015) Pancreatic stellate cell: physiologic role, role in fibrosis and cancer. *Current opinion in gastroenterology* **31**, 416-423
11. Bynigeri, R. R., Jakkampudi, A., Jangala, R., Subramanyam, C., Sasikala, M., Rao, G. V., Reddy, D. N., and Talukdar, R. (2017) Pancreatic stellate cell: Pandora's box for pancreatic disease biology. *World journal of gastroenterology* **23**, 382-405
12. Rahier, J., Wallon, J., and Henquin, J. C. (1981) Cell populations in the endocrine pancreas of human neonates and infants. *Diabetologia* **20**, 540-546
13. Brereton, M. F., Vergari, E., Zhang, Q., and Clark, A. (2015) Alpha-, Delta- and PP-cells: Are They the Architectural Cornerstones of Islet Structure and Co-ordination? *The journal of histochemistry and cytochemistry* **63**, 575-591
14. (2018) 2. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes-2018. *Diabetes care* **41**, S13-S27
15. Eisenbarth, G. S. (2007) Update in type 1 diabetes. *J Clin Endocrinol Metab* **92**, 2403-2407
16. Daneman, D. (2006) Type 1 diabetes. *Lancet* **367**, 847-858
17. Atkinson, M. A., Eisenbarth, G. S., and Michels, A. W. (2014) Type 1 diabetes. *Lancet* **383**, 69-82
18. Pietropaolo, M., Towns, R., and Eisenbarth, G. S. (2012) Humoral autoimmunity in type 1 diabetes: prediction, significance, and detection of distinct disease subtypes. *Cold Spring Harbor perspectives in medicine* **2**, a012831
19. Bluestone, J. A., Herold, K., and Eisenbarth, G. (2010) Genetics, pathogenesis and clinical interventions in type 1 diabetes. *Nature* **464**, 1293-1300

20. Mathers, C. D., and Loncar, D. (2006) Projections of global mortality and burden of disease from 2002 to 2030. *PLoS medicine* **3**, e442

21. Olokoba, A. B., Obateru, O. A., and Olokoba, L. B. (2012) Type 2 Diabetes Mellitus: A Review of Current Trends. *Oman Medical Journal* **27**, 269-273

22. Ali, O. (2013) Genetics of type 2 diabetes. *World Journal of Diabetes* **4**, 114-123

23. Todd, J. A. (1997) Genetics of type 1 diabetes. *Pathologie-biologie* **45**, 219-227

24. Spanakis, E. K., and Golden, S. H. (2013) Race/ethnic difference in diabetes and diabetic complications. *Current diabetes reports* **13**, 814-823

25. Kahn, S. E. (2003) The relative contributions of insulin resistance and beta-cell dysfunction to the pathophysiology of Type 2 diabetes. *Diabetologia* **46**, 3-19

26. Kahn, S. E. (2001) Clinical review 135: The importance of beta-cell failure in the development and progression of type 2 diabetes. *J Clin Endocrinol Metab* **86**, 4047-4058

27. Mitrakou, A., Kelley, D., Mokan, M., Veneman, T., Pangburn, T., Reilly, J., and Gerich, J. (1992) Role of reduced suppression of glucose production and diminished early insulin release in impaired glucose tolerance. *The New England journal of medicine* **326**, 22-29

28. Stolar, M. (2010) Glycemic control and complications in type 2 diabetes mellitus. *The American journal of medicine* **123**, S3-11

29. Steck, A. K., and Winter, W. E. (2011) Review on monogenic diabetes. *Current opinion in endocrinology, diabetes, and obesity* **18**, 252-258

30. Molven, A., and Njolstad, P. R. (2011) Role of molecular genetics in transforming diagnosis of diabetes mellitus. *Expert review of molecular diagnostics* **11**, 313-320

31. Anik, A., Catli, G., Abaci, A., and Bober, E. (2015) Maturity-onset diabetes of the young (MODY): an update. *Journal of pediatric endocrinology & metabolism : JPEM* **28**, 251-263

32. Johansson, B. B., Irgens, H. U., Molnes, J., Sztromwasser, P., Aukrust, I., Juliusson, P. B., Sovik, O., Levy, S., Skrivarhaug, T., Joner, G., Molven, A., Johansson, S., and Njolstad, P. R. (2017) Targeted next-generation sequencing reveals MODY in up to 6.5% of antibody-negative diabetes cases listed in the Norwegian Childhood Diabetes Registry. *Diabetologia* **60**, 625-635

33. Raeder, H., Johansson, S., Holm, P. I., Haldorsen, I. S., Mas, E., Sbarra, V., Nermoen, I., Eide, S. A., Grevle, L., Bjorkhaug, L., Sagen, J. V., Aksnes, L., Sovik, O., Lombardo, D., Molven, A., and Njolstad, P. R. (2006) Mutations in the CEL VNTR cause a syndrome of diabetes and pancreatic exocrine dysfunction. *Nature genetics* **38**, 54-62

34. Vaxillaire, M., Boccio, V., Philippi, A., Vigouroux, C., Terwilliger, J., Passa, P., Beckmann, J. S., Velho, G., Lathrop, G. M., and Froguel, P. (1995) A gene for maturity onset diabetes of the young (MODY) maps to chromosome 12q. *Nature genetics* **9**, 418-423

35. Ellard, S., and Colclough, K. (2006) Mutations in the genes encoding the transcription factors hepatocyte nuclear factor 1 alpha (HNF1A) and 4 alpha (HNF4A) in maturity-onset diabetes of the young. *Human mutation* **27**, 854-869

36. Muddana, V., Whitcomb, D. C., and Papachristou, G. I. (2009) Current management and novel insights in acute pancreatitis. *Expert review of gastroenterology & hepatology* **3**, 435-444

37. Di Fabio, F., Abu Hilal, M., and Johnson, C. D. (2011) Acute pancreatitis: mild, severe or potentially fatal. *Pancreatology* **11**, 373-375

38. Zerem, E. (2014) Treatment of severe acute pancreatitis and its complications. *World journal of gastroenterology* **20**, 13879-13892

39. Kleeff, J., Whitcomb, D. C., Shimosegawa, T., Esposito, I., Lerch, M. M., Gress, T., Mayerle, J., Drewes, A. M., Rebours, V., Akisik, F., Munoz, J. E. D., and Neoptolemos, J. P. (2017) Chronic pancreatitis. *Nature reviews. Disease primers* **3**, 17060

40. Masson, E., Chen, J. M., Audrezet, M. P., Cooper, D. N., and Ferec, C. (2013) A conservative assessment of the major genetic causes of idiopathic chronic pancreatitis: data from a comprehensive analysis of PRSS1, SPINK1, CTRC and CFTR genes in 253 young French patients. *PloS one* **8**, e73522

41. Rebours, V., Boutron-Ruault, M. C., Schnee, M., Ferec, C., Le Marechal, C., Hentic, O., Maire, F., Hammel, P., Ruszniewski, P., and Levy, P. (2009) The natural history of hereditary pancreatitis: a national series. *Gut* **58**, 97-103

42. Sahin-Toth, M., and Toth, M. (2000) Gain-of-function mutations associated with hereditary pancreatitis enhance autoactivation of human cationic trypsinogen. *Biochem Biophys Res Commun* **278**, 286-289

43. Howes, N., Lerch, M. M., Greenhalf, W., Stocken, D. D., Ellis, I., Simon, P., Truninger, K., Ammann, R., Cavallini, G., Charnley, R. M., Uomo, G., Delhaye, M., Spicak, J., Drumm, B., Jansen, J., Mountford, R., Whitcomb, D. C., and Neoptolemos, J. P. (2004) Clinical and genetic characteristics of hereditary pancreatitis in Europe. *Clinical gastroenterology and hepatology* **2**, 252-261

44. Sandhu, B., Vitazka, P., Ferreira-Gonzalez, A., Pandya, A., Vachhani, R., Bouhaidar, D., Zfass, A., and Sanyal, A. (2011) Presence of SPINK-1 variant alters the course of chronic pancreatitis. *Journal of gastroenterology and hepatology* **26**, 965-969

45. Weiss, F. U., Simon, P., Bogdanova, N., Mayerle, J., Dworniczak, B., Horst, J., and Lerch, M. M. (2005) Complete cystic fibrosis transmembrane conductance regulator gene sequencing in patients with idiopathic chronic pancreatitis and controls. *Gut* **54**, 1456-1460

46. Zhou, J., and Sahin-Tóth, M. (2011) Chymotrypsin C (CTRC) mutations in chronic pancreatitis. *Journal of gastroenterology and hepatology* **26**, 1238-1246

47. Witt, H., Beer, S., Rosendahl, J., Chen, J. M., Chandak, G. R., Masamune, A., Bence, M., Szmola, R., Oracz, G., Macek, M., Bhatia, E., Steigenberger, S., Lasher, D., Bühler, *et al.* (2013) Variants in CPA1 are strongly associated with early-onset chronic pancreatitis. *Nature genetics* **45**, 1216-1220

48. Whitcomb, D. C., LaRusch, J., Krasinskas, A. M., Klei, L., Smith, J. P., Brand, R. E., Neoptolemos, J. P., Lerch, M. M., Tector, M., Sandhu, B. S., Guda, N. M., Orlichenko, L., *et al.* (2012) Common genetic variants in the CLDN2 and PRSS1-PRSS2 loci alter risk for alcohol-related and sporadic pancreatitis. *Nature genetics* **44**, 1349-1354

49. Weiss, F. U., Schurmann, C., Guenther, A., Ernst, F., Teumer, A., Mayerle, J., Simon, P., Volzke, H., Radke, D., Greinacher, A., Kuehn, J. P., Zenker, M., Volker, U., Homuth, G., and Lerch, M. M. (2015) Fucosyltransferase 2 (FUT2) non-secretor status and blood group B are associated with elevated serum lipase activity in asymptomatic subjects, and an increased risk for chronic pancreatitis: a genetic association study. *Gut* **64**, 646-656

50. Bosman F, C. F., Hruban R, Theise N. (Eds.). WHO Classification of Tumours of the Digestive System. *IARC,* 2010.

51. Hoem, D., Jensen, D., Steine, S., Thorsen, T. E., Viste, A., and Molven, A. (2008) Clinicopathological characteristics and non-adhesive organ culture of insulinomas. *Scandinavian journal of surgery* **97**, 42-49

52. David, M., Lepage, C., Jouve, J. L., Jooste, V., Chauvenet, M., Faivre, J., and Bouvier, A. M. (2009) Management and prognosis of pancreatic cancer over a 30-year period. *British Journal of Cancer* **101**, 215-218

53. Yu, J., Blackford, A. L., Dal Molin, M., Wolfgang, C. L., and Goggins, M. (2015) Time to progression of pancreatic ductal adenocarcinoma from low-to-high tumour stages. *Gut* **64**, 1783-1789

54. Maitra, A., and Hruban, R. H. (2008) Pancreatic cancer. *Annual Reviews of Pathololgy* **3**, 157-188

55. Wood, L. D., and Hruban, R. H. (2015) Genomic Landscapes of Pancreatic Neoplasia. *Journal of Pathology and Translational Medicine* **49**, 13-22

56. Longo, V., Brunetti, O., Gnoni, A., Cascinu, S., Gasparini, G., Lorusso, V., Ribatti, D., and Silvestris, N. (2016) Angiogenesis in pancreatic ductal adenocarcinoma: A controversial issue. *Oncotarget* **7**, 58649-58658

57. Pinho, A. V., Chantrill, L., and Rooman, I. (2014) Chronic pancreatitis: a path to pancreatic cancer. *Cancer letters* **345**, 203-209

58. Weiss, F. U. (2014) Pancreatic cancer risk in hereditary pancreatitis. *Frontiers in physiology* **5**, 70

59. Iodice, S., Gandini, S., Maisonneuve, P., and Lowenfels, A. B. (2008) Tobacco and the risk of pancreatic cancer: a review and meta-analysis. *Langenbeck's archives of surgery* **393**, 535-545

60. Lucenteforte, E., La Vecchia, C., Silverman, D., Petersen, G. M., Bracci, P. M., Ji, B. T., Bosetti, C., Li, D., Gallinger, S., Miller, A. B., Bueno-de-Mesquita, H. B., Talamini, R., *et al.* (2012) Alcohol consumption and pancreatic cancer: a pooled analysis in the International Pancreatic Cancer Case-Control Consortium (PanC4). *Annals of oncology* **23**, 374-382

61. Maisonneuve, P., and Lowenfels, A. B. (2010) Epidemiology of pancreatic cancer: an update. *Digestive Diseases* **28**, 645-656

62. Guo, Y., Liu, W., and Wu, J. (2016) Helicobacter pylori infection and pancreatic cancer risk: A meta-analysis. *Journal of cancer research and therapeutics* **12**, C229-c232

63. Solomon, S., Das, S., Brand, R., and Whitcomb, D. C. (2012) Inherited pancreatic cancer syndromes. *Cancer journal* **18**, 485-491

64. Whitcomb, D. C., Shelton, C. A., and Brand, R. E. (2015) Genetics and Genetic Testing in Pancreatic Cancer. *Gastroenterology* **149**, 1252-1264.e1254

65. Rosendahl, J., Bödeker, H., Mössner, J., and Teich, N. (2007) Hereditary chronic pancreatitis. *Orphanet Journal of Rare Diseases* **2**, 1

66. Manolio, T. A., Collins, F. S., Cox, N. J., Goldstein, D. B., Hindorff, L. A., Hunter, D. J., McCarthy, M. I., Ramos, E. M., Cardon, L. R., Chakravarti, A., Cho, J. H., Guttmacher, A. E., *et al.* (2009) Finding the missing heritability of complex diseases. *Nature* **461**, 747-753

67. Amundadottir, L., Kraft, P., Stolzenberg-Solomon, R. Z., Fuchs, C. S., Petersen, G. M., Arslan, A. A., Bueno-de-Mesquita, H. B., Gross, M., Helzlsouer, K., Jacobs, E. J., *et al.* (2009) Genome-wide association study identifies variants in the ABO locus associated with susceptibility to pancreatic cancer. *Nature genetics* **41**, 986-990

68. Nakao, M., Matsuo, K., Hosono, S., Ogata, S., Ito, H., Watanabe, M., Mizuno, N., Iida, S., Sato, S., Yatabe, Y., Yamao, K., Ueda, R., Tajima, K., and Tanaka, H. (2011) ABO blood group alleles and the risk of pancreatic cancer in a Japanese population. *Cancer science* **102**, 1076-1080

69. Risch, H. A., Lu, L., Wang, J., Zhang, W., Ni, Q., Gao, Y. T., and Yu, H. (2013) ABO Blood Group and Risk of Pancreatic Cancer: A Study in Shanghai and Meta-Analysis. *American Journal of Epidemiology* **177**, 1326-1337

70. Petersen, G. M., Amundadottir, L., Fuchs, C. S., Kraft, P., Stolzenberg-Solomon, R. Z., Jacobs, K. B., Arslan, A. A., Bueno-de-Mesquita, H. B., Gallinger, S., Gross, M., *et al.* (2010) A genome-wide association study identifies pancreatic cancer susceptibility loci on chromosomes 13q22.1, 1q32.1 and 5p15.33. *Nature genetics* **42**, 224-228

71. Cobo, I., Martinelli, P., Flandez, M., Bakiri, L., Zhang, M., Carrillo-de-Santa-Pau, E., Jia, J., Sanchez-Arevalo Lobo, V. J., Megias, D., Felipe, I., Del Pozo, N., Millan, I., *et al.* (2018) Transcriptional regulation by NR5A2 links differentiation and inflammation in the pancreas. *Nature* **554**, 533-537

72. Jones, S., Zhang, X., Parsons, D. W., Lin, J. C., Leary, R. J., Angenendt, P., Mankoo, P., Carter, H., Kamiyama, H., Jimeno, A., Hong, S. M., Fu, B., Lin, M. T., Calhoun, E. S., *et al.* (2008) Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. *Science* **321**, 1801-1806

73. The Cancer Genome Atlas Research Network (2017) Integrated Genomic Characterization of Pancreatic Ductal Adenocarcinoma. *Cancer cell* **32**, 185-203.e113

74. Bailey, P., Chang, D. K., Nones, K., Johns, A. L., Patch, A. M., Gingras, M. C., Miller, D. K., Christ, A. N., Bruxner, T. J., Quinn, M. C., Nourse, C., Murtaugh, L. C., Harliwong, I., *et al.* (2016) Genomic analyses identify molecular subtypes of pancreatic cancer. *Nature* **531**, 47-52

75. Moss, E. L., Hollingworth, J., and Reynolds, T. M. (2005) The role of CA125 in clinical practice. *Journal of clinical pathology* **58**, 308-312

76. Gilgunn, S., Conroy, P. J., Saldova, R., Rudd, P. M., and O'Kennedy, R. J. (2013) Aberrant PSA glycosylation-a sweet predictor of prostate cancer. *Nature reviews. Urology* **10**, 99-107

77. Hammarstrom, S. (1999) The carcinoembryonic antigen (CEA) family: structures, suggested functions and expression in normal and malignant tissues. *Seminars in cancer biology* **9**, 67-81

78. Magnani, J. L., Nilsson, B., Brockhaus, M., Zopf, D., Steplewski, Z., Koprowski, H., and Ginsburg, V. (1982) A monoclonal antibody-defined antigen associated with gastrointestinal cancer is a ganglioside containing sialylated lacto-N-fucopentaose II. *The Journal of biological chemistry* **257**, 14365-14369

79. Koprowski, H., Steplewski, Z., Mitchell, K., Herlyn, M., Herlyn, D., and Fuhrer, P. (1979) Colorectal carcinoma antigens detected by hybridoma antibodies. *Somatic cell genetics* **5**, 957-971

80. Tempero, M. A., Uchida, E., Takasaki, H., Burnett, D. A., Steplewski, Z., and Pour, P. M. (1987) Relationship of carbohydrate antigen 19-9 and Lewis antigens in pancreatic cancer. *Cancer research* **47**, 5501-5503

81. Locker, G. Y., Hamilton, S., Harris, J., Jessup, J. M., Kemeny, N., Macdonald, J. S., Somerfield, M. R., Hayes, D. F., and Bast, R. C., Jr. (2006) ASCO 2006 update of recommendations for the use of tumor markers in gastrointestinal cancer. *Journal of clinical oncology* **24**, 5313-5327

82. Onal, C., Colakoglu, T., Ulusan, S. N., Yapar, A. F., and Kayaselcuk, F. (2012) Biliary obstruction induces extremely elevated serum CA 19-9 levels: case report. *Onkologie* **35**, 780-782

83. Yoshida, E. M., Scudamore, C. H., Erb, S. R., Owen, D. A., and Silver, H. K. (1995) Markedly elevated serum CA 19-9 levels in a case of chronic pancreatitis. *Canadian journal of surgery* **38**, 83-86

84. Nishihara, S., Yazawa, S., Iwasaki, H., Nakazato, M., Kudo, T., Ando, T., and Narimatsu, H. (1993) Alpha (1,3/1,4)fucosyltransferase (FucT-III) gene is inactivated by a single amino acid substitution in Lewis histo-blood type negative individuals. *Biochemical and Biophysical Research Communications* **196**, 624-631

85. Orntoft, T. F., Vestergaard, E. M., Holmes, E., Jakobsen, J. S., Grunnet, N., Mortensen, M., Johnson, P., Bross, P., Gregersen, N., Skorstengaard, K., Jensen, U. B., Bolund, L., and Wolf, H. (1996) Influence of Lewis alpha1-3/4-L-fucosyltransferase (FUT3) gene mutations on enzyme activity, erythrocyte phenotyping, and circulating tumor marker sialyl-Lewis a levels. *The Journal of biological chemistry* **271**, 32260-32268

86. Meng, Q., Shi, S., Liang, C., Liang, D., Xu, W., Ji, S., Zhang, B., Ni, Q., Xu, J., and Yu, X. (2017) Diagnostic and prognostic value of carcinoembryonic antigen in pancreatic cancer: a systematic review and meta-analysis. *OncoTargets and therapy* **10**, 4591-4598

87. Zhang, Y., Yang, J., Li, H., Wu, Y., Zhang, H., and Chen, W. (2015) Tumor markers CA19-9, CA242 and CEA in the diagnosis of pancreatic cancer: a meta-analysis. *International journal of clinical and experimental medicine* **8**, 11683-11691

88. Cazet, A., Julien, S., Bobowski, M., Krzewinski-Recchi, M. A., Harduin-Lepers, A., Groux-Degroote, S., and Delannoy, P. (2010) Consequences of the expression of sialylated antigens in breast cancer. *Carbohydrate research* **345**, 1377-1383

89. Kawa, S., Oguchi, H., Kobayashi, T., Tokoo, M., Furuta, S., Kanai, M., and Homma, T. (1991) Elevated serum levels of Dupan-2 in pancreatic cancer patients negative for Lewis blood group phenotype. *British Journal of Cancer* **64**, 899-902

90. Melo, S. A., Luecke, L. B., Kahlert, C., Fernandez, A. F., Gammon, S. T., Kaye, J., LeBleu, V. S., Mittendorf, E. A., Weitz, J., Rahbari, N., Reissfelder, C., Pilarsky, C., Fraga, M. F., Piwnica-Worms, D., and Kalluri, R. (2015) Glypican-1 identifies cancer exosomes and detects early pancreatic cancer. *Nature* **523**, 177-182

91. Sogawa, K., Takano, S., Iida, F., Satoh, M., Tsuchida, S., Kawashima, Y., Yoshitomi, H., Sanda, A., Kodera, Y., Takizawa, H., Mikata, R., Ohtsuka, M., Shimizu, H., Miyazaki, M., Yokosuka, O., and Nomura, F. (2016) Identification of a novel serum biomarker for pancreatic cancer, C4b-binding protein alpha-chain (C4BPA) by quantitative proteomic analysis using tandem mass tags. *British Journal of Cancer* **115**, 949-956

92. Yang, Y., Yan, S., Tian, H., and Bao, Y. (2018) Macrophage inhibitory cytokine-1 versus carbohydrate antigen 19-9 as a biomarker for diagnosis of pancreatic cancer: A PRISMA-compliant meta-analysis of diagnostic accuracy studies. *Medicine* **97**, e9994

93. Varki, A., Cummings, R. D., Aebi, M., Packer, N. H., Seeberger, P. H., Esko, J. D., Stanley, P., Hart, G., Darvill, A., Kinoshita, T., Prestegard, J. J., Schnaar, R. L., Freeze, *et al.* (2015) Symbol Nomenclature for Graphical Representations of Glycans. *Glycobiology* **25**, 1323-1324

94. Johansson, M. E., and Hansson, G. C. (2016) Immunological aspects of intestinal mucus and mucins. *Nature reviews. Immunology* **16**, 639-649

95. Shogren, R., Gerken, T. A., and Jentoft, N. (1989) Role of glycosylation on the conformation and chain dimensions of O-linked glycoproteins: light-scattering studies of ovine submaxillary mucin. *Biochemistry* **28**, 5525-5536

96. Goettig, P. (2016) Effects of Glycosylation on the Enzymatic Activity and Mechanisms of Proteases. *International Journal of Molecular Sciences* **17**

97. Broide, D. H., Miller, M., Castaneda, D., Nayar, J., Cho, J. Y., Roman, M., Ellies, L. G., and Sriramarao, P. (2002) Core 2 oligosaccharides mediate eosinophil and neutrophil peritoneal but not lung recruitment. *American journal of physiology* **282**, L259-266

98. Ellies, L. G., Tsuboi, S., Petryniak, B., Lowe, J. B., Fukuda, M., and Marth, J. D. (1998) Core 2 oligosaccharide biosynthesis distinguishes between selectin ligands essential for leukocyte homing and inflammation. *Immunity* **9**, 881-890

99. Tailford, L. E., Crost, E. H., Kavanaugh, D., and Juge, N. (2015) Mucin glycan foraging in the human gut microbiome. *Frontiers in Genetics* **6**

100. Hoskins, L. C., and Boulding, E. T. (1976) Degradation of blood group antigens in human colon ecosystems. I. In vitro production of ABH blood group-degrading enzymes by enteric bacteria. *The Journal of clinical investigation* **57**, 63-73

101. Magalhaes, A., and Reis, C. A. (2010) Helicobacter pylori adhesion to gastric epithelial cells is mediated by glycan receptors. *Brazilian journal of medical and biological research* **43**, 611-618

102. Melo-Gonzalez, F., Fenton, T. M., Forss, C., Smedley, C., Goenka, A., MacDonald, A. S., Thornton, D. J., and Travis, M. A. (2018) Intestinal mucin activates human dendritic cells and IL-8 production in a glycan-specific manner. *The Journal of biological chemistry* **293**, 8543-8553

103. Braunger, K., Pfeffer, S., Shrimal, S., Gilmore, R., Berninghausen, O., Mandon, E. C., Becker, T., Förster, F., and Beckmann, R. (2018) Structural basis for coupling protein transport and N-glycosylation at the mammalian endoplasmic reticulum. *Science* **13***, 215-219*

104. Pfeffer, S., Dudek, J., Zimmermann, R., and Forster, F. (2016) Organization of the native ribosome-translocon complex at the mammalian endoplasmic reticulum membrane. *Biochimica et biophysica acta* **1860**, 2122-2129

105. Zielinska, D. F., Gnad, F.,Wisniewski, J. R., Mann, M. (2010) Precision mapping of an in vivo N-glycoproteome reveals rigid topological and sequence constrains. *Cell* **141**, 897-907

106. Hebert, D. N., Garman, S. C., and Molinari, M. (2005) The glycan code of the endoplasmic reticulum: asparagine-linked carbohydrates as protein maturation and quality-control tags. *Trends in cell biology* **15**, 364-370

107. Xu, C., and Ng, D. T. (2015) Glycosylation-directed quality control of protein folding. *Nature reviews. Molecular cell biology* **16**, 742-752

108. Stanley P, S. H., Taniguchi N. N-Glycans. In: Varki A, Cummings RD, Esko JD, et al., editors. Essentials of Glycobiology. 2nd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2009. Chapter 8. Available from: https://www.ncbi.nlm.nih.gov/books/NBK1917/.

109. Vajaria, B. N., and Patel, P. S. (2017) Glycosylation: a hallmark of cancer? *Glycoconjugate journal* **34**, 147-156

110. Nardy, A. F., Freire-de-Lima, L., Freire-de-Lima, C. G., and Morrot, A. (2016) The Sweet Side of Immune Evasion: Role of Glycans in the Mechanisms of Cancer Progression. *Frontiers in Oncology* **6**, 54

111. Munkley, J., and Elliott, D. J. (2016) Hallmarks of glycosylation in cancer. *Oncotarget* **7**, 35478-35489

112. Hakomori, S. (1996) Tumor malignancy defined by aberrant glycosylation and sphingo(glyco)lipid metabolism. *Cancer research* **56**, 5309-5318

113. Ehre, C. (2014) Cystic fibrosis: an inherited disease affecting mucin-producing organs. *The international journal of biochemistry & cell biology* **52**, 136-145

114. Guzman-Aranguez, A., and Argüeso, P. (2010) Structure and Biological Roles of Mucin-type O-glycans at the Ocular Surface. *The ocular surface* **8**, 8-17

115. Cascio, S., and Finn, O. J. (2016) Intra- and Extra-Cellular Events Related to Altered Glycosylation of MUC1 Promote Chronic Inflammation, Tumor Progression, Invasion, and Metastasis. *Biomolecules* **6**

116. Magnani, J. L., Steplewski, Z., Koprowski, H., and Ginsburg, V. (1983) Identification of the gastrointestinal and pancreatic cancer-associated antigen detected by monoclonal antibody 19-9 in the sera of patients as a mucin. *Cancer research* **43**, 5489-5492

117. Partyka, K., Maupin, K. A., Brand, R. E., and Haab, B. B. (2012) Diverse monoclonal antibodies against the CA 19-9 antigen show variation in binding specificity with consequences for clinical interpretation. *Proteomics* **12**, 2212-2220

118. Yue, T., Partyka, K., Maupin, K. A., Hurley, M., Andrews, P., Kaul, K., Moser, A. J., Zeh, H., Brand, R. E., and Haab, B. B. (2011) Identification of blood-protein carriers of the CA 19-9 antigen and characterization of prevalence in pancreatic diseases. *Proteomics* **11**, 3665-3674

119. Radhakrishnan, P., Dabelsteen, S., Madsen, F. B., Francavilla, C., Kopp, K. L., Steentoft, C., Vakhrushev, S. Y., Olsen, J. V., Hansen, L., Bennett, E. P., Woetmann, A., Yin, G., Chen, L., Song,

H., Bak, M., Hlady, R. A., Peters, S. L., Opavsky, R., Thode, C., Qvortrup, K., Schjoldager, K. T., Clausen, H., Hollingsworth, M. A., and Wandall, H. H. (2014) Immature truncated O-glycophenotype of cancer directly induces oncogenic features. *Proceedings of the National Academy of Sciences of the United States of America* **111**, E4066-4075

120. Perez-Garay, M., Arteta, B., Pages, L., de Llorens, R., de Bolos, C., Vidal-Vanaclocha, F., and Peracaula, R. (2010) alpha2,3-sialyltransferase ST3Gal III modulates pancreatic cancer cell motility and adhesion in vitro and enhances its metastatic potential in vivo. *PloS one* **5**
121. Zhao, J., Qiu, W., Simeone, D. M., and Lubman, D. M. (2007) N-linked glycosylation profiling of pancreatic cancer serum using capillary liquid phase separation coupled with mass spectrometric analysis. *Journal of proteome research* **6**, 1126-1138
122. Häuselmann, I., and Borsig, L. (2014) Altered Tumor-Cell Glycosylation Promotes Metastasis. *Frontiers in Oncology* **4**
123. Whitehouse, C., Burchell, J., Gschmeissner, S., Brockhausen, I., Lloyd, K. O., and Taylor-Papadimitriou, J. (1997) A transfected sialyltransferase that is elevated in breast cancer and localizes to the medial/trans-Golgi apparatus inhibits the development of core-2-based O-glycans. *The Journal of cell biology* **137**, 1229-1241
124. Nan, B. C., Shao, D. M., Chen, H. L., Huang, Y., Gu, J. X., Zhang, Y. B., and Wu, Z. G. (1998) Alteration of N-acetylglucosaminyltransferases in pancreatic carcinoma. *Glycoconjugate journal* **15**, 1033-1037
125. Granovsky, M., Fata, J., Pawling, J., Muller, W. J., Khokha, R. and Dennis, J. W. (2000) Suppression of tumor growth and methastasis in Mgat5-deficient mice. *Nature Medicine* **6**, 306-312
126. Okuyama, N., Ide, Y., Nakano, M., Nakagawa, T., Yamanaka, K., Moriwaki, K., Murata, K., Ohigashi, H., Yokoyama, S., Eguchi, H., Ishikawa, O., Ito, T., Kato, M., Kasahara, A., Kawano, S., Gu, J., Taniguchi, N., and Miyoshi, E. (2006) Fucosylated haptoglobin is a novel marker for pancreatic cancer: a detailed analysis of the oligosaccharide structure and a possible mechanism for fucosylation. *International journal of cancer* **118**, 2803-2808
127. Barrabes, S., Pages-Pons, L., Radcliffe, C. M., Tabares, G., Fort, E., Royle, L., Harvey, D. J., Moenner, M., Dwek, R. A., Rudd, P. M., De Llorens, R., and Peracaula, R. (2007) Glycosylation of serum ribonuclease 1 indicates a major endothelial origin and reveals an increase in core fucosylation in pancreatic cancer. *Glycobiology* **17**, 388-400
128. Patnaik, S. K., Helmberg, W., and Blumenfeld, O. O. (2014) BGMUT Database of Allelic Variants of Genes Encoding Human Blood Group Antigens. *Transfusion Medicine and Hemotherapy* **41**, 346-351
129. Landsteiner, K. (1931) Individual differences in human blood. *Science* **73**, 403-409
130. Crow, J. F. (1993) Felix Bernstein and the first human marker locus. *Genetics* **133**, 4-7
131. Yamamoto, F., Clausen, H., White, T., Marken, J., and Hakomori, S. (1990) Molecular genetic basis of the histo-blood group ABO system. *Nature* **345**, 229-233
132. Yamamoto, F., Cid, E., Yamamoto, M., and Blancher, A. (2012) ABO research in the modern era of genomics. *Transfusion Medicine Reviews* **26**, 103-118
133. Garratty, G., Glynn, S. A., and McEntire, R. (2004) ABO and Rh(D) phenotype frequencies of different racial/ethnic groups in the United States. *Transfusion* **44**, 703-706
134. Ginsburg. (1972) Enzymatic basis for blood groups in man. *Advanced Enzymology Related Areas of Molecular Biology.* **36**, 131–149
135. Economidou, J., Hughes-Jones, N. C., and Gardner, B. (1967) Quantitative measurements concerning A and B antigen sites. *Vox sanguinis* **12**, 321-328

136. Kelly, R. J., Rouquier, S., Giorgi, D., Lennon, G. G., and Lowe, J. B. (1995) Sequence and expression of a candidate for the human Secretor blood group alpha(1,2)fucosyltransferase gene (FUT2). Homozygosity for an enzyme-inactivating nonsense mutation commonly correlates with the non-secretor phenotype. *The Journal of biological chemistry* **270**, 4640-4649

137. Larsen, R. D., Ernst, L. K., Nair, R. P., and Lowe, J. B. (1990) Molecular cloning, sequence, and expression of a human GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase cDNA that can form the H blood group antigen. *Proceedings of the National Academy of Sciences of the United States of America* **87**, 6674-6678

138. Wolpin, B. M., Kraft, P., Xu, M., Steplowski, E., Olsson, M. L., Arslan, A. A., Bueno-de-Mesquita, H. B., Gross, M., Helzlsouer, K., Jacobs, E. J., LaCroix, A., Petersen, G., *et al.* (2010) Variant ABO blood group alleles, secretor status, and risk of pancreatic cancer: results from the pancreatic cancer cohort consortium. *Cancer epidemiology, biomarkers & prevention* **19**, 3140-3149

139. Hakomori, S. (2001) Tumor-associated carbohydrate antigens defining tumor malignancy: basis for development of anti-cancer vaccines. *Advances in experimental medicine and biology* **491**, 369-402

140. Varki A, K. R., Toole B, et al. Glycosylation Changes in Cancer. 2017. In: Varki A, Cummings RD, Esko JD, et al., editors. Essentials of Glycobiology [Internet]. 3rd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2015-2017. Chapter 47. Available from: https://www.ncbi.nlm.nih.gov/books/NBK453023/

141. Panda, S., and Ding, J. L. (2015) Natural antibodies bridge innate and adaptive immunity. *Journal of immunology* **194**, 13-20

142. Khasbiullina, N. R., and Bovin, N. V. (2015) Hypotheses of the origin of natural antibodies: a glycobiologist's opinion. *Biochemistry* **80**, 820-835

143. Khasbiullina, N. R., Shilova, N. V., Navakouski, M. E., Nokel, A. Y., Knirel, Y. A., Blixt, O., and Bovin, N. V. (2018) Repertoire of Abs primed by bacteria in gnotobiotic mice. *Innate immunity* **24**, 180-187

144. Taylor, A. K., Zambaux, J. L., Klisak, I., Mohandas, T., Sparkes, R. S., Schotz, M. C., and Lusis, A. J. (1991) Carboxyl ester lipase: a highly polymorphic locus on human chromosome 9qter. *Genomics* **10**, 425-431

145. Torsvik, J., Johansson, S., Johansen, A., Ek, J., Minton, J., Raeder, H., Ellard, S., Hattersley, A., Pedersen, O., Hansen, T., Molven, A., and Njolstad, P. R. (2010) Mutations in the VNTR of the carboxyl-ester lipase gene (CEL) are a rare cause of monogenic diabetes. *Human genetics* **127**, 55-64

146. Ragvin, A., Fjeld, K., Weiss, F. U., Torsvik, J., Aghdassi, A., Mayerle, J., Simon, P., Njolstad, P. R., Lerch, M. M., Johansson, S., and Molven, A. (2013) The number of tandem repeats in the carboxyl-ester lipase (CEL) gene as a risk factor in alcoholic and idiopathic chronic pancreatitis. *Pancreatology* **13**, 29-32

147. Martinez, E., Crenon, I., Silvy, F., Del Grande, J., Mougel, A., Barea, D., Fina, F., Bernard, J. P., Ouaissi, M., Lombardo, D., and Mas, E. (2017) Expression of truncated bile salt-dependent lipase variant in pancreatic pre-neoplastic lesions. *Oncotarget* **8**, 536-551

148. Lidberg, U., Nilsson, J., Stromberg, K., Stenman, G., Sahlin, P., Enerback, S., and Bjursell, G. (1992) Genomic organization, sequence analysis, and chromosomal localization of the human carboxyl ester lipase (CEL) gene and a CEL-like (CELL) gene. *Genomics* **13**, 630-640

149. Madeyski, K., Lidberg, U., Bjursell, G., and Nilsson, J. (1999) Characterization of the gorilla carboxyl ester lipase locus, and the appearance of the carboxyl ester lipase pseudogene during primate evolution. *Gene* **239**, 273-282

150. Blackberg, L., Lombardo, D., Hernell, O., Guy, O., and Olivecrona, T. (1981) Bile salt-stimulated lipase in human milk and carboxyl ester hydrolase in pancreatic juice: are they identical enzymes? *FEBS letters* **136**, 284-288

151. Kolar, M. J., Kamat, S. S., Parsons, W. H., Homan, E. A., Maher, T., Peroni, O. D., Syed, I., Fjeld, K., Molven, A., Kahn, B. B., Cravatt, B. F., and Saghatelian, A. (2016) Branched Fatty Acid Esters of Hydroxy Fatty Acids Are Preferred Substrates of the MODY8 Protein Carboxyl Ester Lipase. *Biochemistry* **55**, 4636-4641

152. Lombardo, D., Guy, O., and Figarella, C. (1978) Purification and characterization of a carboxyl ester hydrolase from human pancreatic juice. *Biochimica et biophysica acta* **527**, 142-149

153. La Rosa, S., Vigetti, D., Placidi, C., Finzi, G., Uccella, S., Clerici, M., Bartolini, B., Carnevali, I., Losa, M., and Capella, C. (2010) Localization of carboxyl ester lipase in human pituitary gland and pituitary adenomas. *The journal of histochemistry and cytochemistry* **58**, 881-889

154. Kodvawala, A., Ghering, A. B., Davidson, W. S., and Hui, D. Y. (2005) Carboxyl ester lipase expression in macrophages increases cholesteryl ester accumulation and promotes atherosclerosis. *The Journal of biological chemistry* **280**, 38592-38598

155. Holtsberg, F. W., Ozgur, L. E., Garsetti, D. E., Myers, J., Egan, R. W., and Clark, M. A. (1995) Presence in human eosinophils of a lysophospholipase similar to that found in the pancreas. *The Biochemical journal* **309 (Pt 1)**, 141-144

156. Li, F., and Hui, D. Y. (1998) Synthesis and secretion of the pancreatic-type carboxyl ester lipase by human endothelial cells. *The Biochemical journal* **329 (Pt 3)**, 675-679

157. Roudani, S., Miralles, F., Margotat, A., Escribano, M. J., and Lombardo, D. (1995) Bile salt-dependent lipase transcripts in human fetal tissues. *Biochimica et biophysica acta* **1264**, 141-150

158. Holmes, R. S., and Cox, L. A. (2011) Comparative Structures and Evolution of Vertebrate Carboxyl Ester Lipase (CEL) Genes and Proteins with a Major Role in Reverse Cholesterol Transport. *Cholesterol* **2011**

159. Morlock-Fitzpatrick, K. R., and Fisher, E. A. (1995) The effects of O- and N-linked glycosylation on the secretion and bile salt-stimulation of pancreatic carboxyl ester lipase activity. *Proceedings of the Society for Experimental Biology and Medicine* **208**, 186-190

160. Abouakil, N., Mas, E., Bruneau, N., Benajiba, A., and Lombardo, D. (1993) 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. *The Journal of biological chemistry* **268**, 25755-25763

161. Baba, T., Downs, D., Jackson, K. W., Tang, J., and Wang, C. S. (1991) Structure of human milk bile salt activated lipase. *Biochemistry* **30**, 500-510

162. Johansson, B. B., Torsvik, J., Bjorkhaug, L., Vesterhus, M., Ragvin, A., Tjora, E., Fjeld, K., Hoem, D., Johansson, S., Raeder, H., Lindquist, S., Hernell, O., Cnop, M., Saraste, J., Flatmark, T., Molven, A., and Njolstad, P. R. (2011) 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. *The Journal of biological chemistry* **286**, 34593-34605

163. Steentoft, C., Vakhrushev, S. Y., Joshi, H. J., Kong, Y., Vester-Christensen, M. B., Schjoldager, K. T., Lavrsen, K., Dabelsteen, S., Pedersen, N. B., Marcos-Silva, L., Gupta, R., Bennett, E. P., Mandel, U., Brunak, S., Wandall, H. H., Levery, S. B., and Clausen, H. (2013) Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. *The EMBO journal* **32**, 1478-1488

164. Wang, C. S., Dashti, A., Jackson, K. W., Yeh, J. C., Cummings, R. D., and Tang, J. (1995) Isolation and characterization of human milk bile salt-activated lipase C-tail fragment. *Biochemistry* **34**, 10639-10644

165. Chen, J. C., Miercke, L. J., Krucinski, J., Starr, J. R., Saenz, G., Wang, X., Spilburg, C. A., Lange, L. G., Ellsworth, J. L., and Stroud, R. M. (1998) Structure of bovine pancreatic cholesterol esterase at 1.6 A: novel structural features involved in lipase activation. *Biochemistry* **37**, 5107-5117

166. Aebi, M., Bernasconi, R., Clerc, S., and Molinari, M. (2010) N-glycan structures: recognition and processing in the ER. *Trends in biochemical sciences* **35**, 74-82

167. Moremen, K. W., and Molinari, M. (2006) N-linked glycan recognition and processing: the molecular basis of endoplasmic reticulum quality control. *Current opinion in structural biology* **16**, 592-599

168. Pasqualini, E., Caillol, N., Valette, A., Lloubes, R., Verine, A., and Lombardo, D. (2000) Phosphorylation of the rat pancreatic bile-salt-dependent lipase by casein kinase II is essential for secretion. *The Biochemical journal* **345 Pt 1**, 121-128

169. Gold, P., Freedman, S. O. (1965) Specific carcinoembryonic antigens of the human digestive system. *Journal of Experimental Medicine* **122**, 467-81

170. Turriziani, M., Fantini, M., Benvenuto, M., Izzi, V., Masuelli, L., Sacchetti, P., Modesti, A., and Bei, R. (2012) Carcinoembryonic antigen (CEA)-based cancer vaccines: recent patents and antitumor effects from experimental models to clinical trials. *Recent patents on anti-cancer drug discovery* **7**, 265-296

171. Wepsic, H. T. (1983) Overview of oncofetal antigens in cancer. *Annals of clinical and laboratory science* **4**, 261-266

172. Escribano, M. J., Carre-Llopis, A., and Loridon-Rosa, B. (1985) Expression of oncofoetal pancreatic antigens in hamster adult pancreas during experimental carcinogenesis. *British Journal of Cancer* **51**, 187-193

173. Escribano, M. J., Cordier, J., Nap, M., Ten Kate, F. J., and Burtin, P. (1986) Differentiation antigens in fetal human pancreas. Reexpression in cancer. *International journal of cancer* **38**, 155-160

174. Albers, G. H., Escribano, M. J., Gonzalez, M., Mulliez, N., and Nap, M. (1987) Fetoacinar pancreatic protein in the developing human pancreas. *Differentiation; research in biological diversity* **34**, 210-215

175. Albers, G. H., Escribano, M. J., Daher, N., and Nap, M. (1990) An immunohistologic study of the feto-acinar pancreatic protein (FAP) in the normal pancreas, chronic pancreatitis, pancreatic adenocarcinoma, and intraabdominal metastases of adenocarcinomas. *American journal of clinical pathology* **93**, 14-19

176. Mas, E., Abouakil, N., Roudani, S., Miralles, F., Guy-Crotte, O., Figarella, C., Escribano, M. J., and Lombardo, D. (1993) Human fetoacinar pancreatic protein: an oncofetal glycoform of the normally secreted pancreatic bile-salt-dependent lipase. *Biochemical Journal* **289**, 609-615

177. Crescence, L., Beraud, E., Sbarra, V., Bernard, J. P., Lombardo, D., and Mas, E. (2012) Targeting a novel onco-glycoprotein antigen at tumoral pancreatic cell surface by mAb16D10 induces cell death. *Journal of immunology* **189**, 3386-3396

178. Panicot-Dubois, L., Aubert, M., Franceschi, C., Mas, E., Silvy, F., Crotte, C., Bernard, J. P., Lombardo, D., and Sadoulet, M. O. (2004) Monoclonal antibody 16D10 to the C-terminal domain of the feto-acinar pancreatic protein binds to membrane of human pancreatic tumoral SOJ-6 cells and inhibits the growth of tumor xenografts. *Neoplasia* **6**, 713-724

179. Benkoel, L., Bernard, J. P., Payan-Defais, M. J., Crescence, L., Franceschi, C., Delmas, M., Ouaissi, M., Sastre, B., Sahel, J., Benoliel, A. M., Bongrand, P., Silvy, F., Gauthier, L., Romagne, F., Lombardo, D., and Mas, E. (2009) Monoclonal antibody 16D10 to the COOH-terminal domain of the feto-acinar pancreatic protein targets pancreatic neoplastic tissues. *Molecular cancer therapeutics* **8**, 282-291

180. Martinez, E., Silvy, F., Fina, F., Bartoli, M., Krahn, M., Barlesi, F., Figarella-Branger, D., Iovanna, J., Laugier, R., Ouaissi, M., Lombardo, D., and Mas, E. (2015) Rs488087 single nucleotide polymorphism as predictive risk factor for pancreatic cancers. *Oncotarget* **6**, 39855-39864

181. Johansson, B. B., Fjeld, K., El Jellas, K., Gravdal, A., Dalva, M., Tjora, E., Raeder, H., Kulkarni, R. N., Johansson, S., Njolstad, P. R., and Molven, A. (2018) The role of the carboxyl ester lipase (CEL) gene in pancreatic disease. *Pancreatology* **18**, 12-19

182. Xiao, X., Jones, G., Sevilla, W. A., Stolz, D. B., Magee, K. E., Haughney, M., Mukherjee, A., Wang, Y., and Lowe, M. E. (2016) A Carboxyl Ester Lipase (CEL) Mutant Causes Chronic Pancreatitis by Forming Intracellular Aggregates That Activate Apoptosis. *The Journal of biological chemistry* **291**, 23224-23236

183. Torsvik, J., Johansson, B. B., Dalva, M., Marie, M., Fjeld, K., Johansson, S., Bjorkoy, G., Saraste, J., Njolstad, P. R., and Molven, A. (2014) Endocytosis of secreted carboxyl ester lipase in a syndrome of diabetes and pancreatic exocrine dysfunction. *The Journal of biological chemistry* **289**, 29097-29111

184. Fjeld, K., Weiss, F. U., Lasher, D., Rosendahl, J., Chen, J. M., Johansson, B. B., Kirsten, H., Ruffert, C., Masson, E., Steine, S. J., Bugert, P., Cnop, M., Grutzmann, R., Mayerle, J., Mossner, *et al.* (2015) A recombined allele of the lipase gene CEL and its pseudogene CELP confers susceptibility to chronic pancreatitis. *Nature genetics* **47**, 518-522

185. Zou, W. B., Boulling, A., Masamune, A., Issarapu, P., Masson, E., Wu, H., Sun, X. T., Hu, L. H., Zhou, D. Z., He, L., Fichou, Y., Nakano, E., Hamada, S., Kakuta, Y., Kume, K., *et al.* (2016) No Association Between CEL-HYB Hybrid Allele and Chronic Pancreatitis in Asian Populations. *Gastroenterology* **150**, 1558-1560.e1555

186. Immervoll, H., Hoem, D., Kugarajh, K., Steine, S. J., and Molven, A. (2006) Molecular analysis of the EGFR-RAS-RAF pathway in pancreatic ductal adenocarcinomas: lack of mutations in the BRAF and EGFR genes. *Virchows Archiv* **448**, 788-796

187. Liumbruno, G. M., and Franchini, M. (2013) Beyond immunohaematology: the role of the ABO blood group in human diseases. *Blood Transfusion* **11**, 491-499

188. Seymour, R. M., Allan, M. J., Pomiankowski, A., and Gustafsson, K. (2004) Evolution of the human ABO polymorphism by two complementary selective pressures. *Proceedings of the Royal Society B* **271**, 1065-1072

189. Cserti, C. M., and Dzik, W. H. (2007) The ABO blood group system and Plasmodium falciparum malaria. *Blood* **110**, 2250-2258

190. Kornstad, L. (1997) Frequency of the blood group antigen K and the A1A2BO groups in the Norwegian counties. *Gene geography* **11**, 37-46

191. Gjellesvik, D. R., Lorens, J. B., and Male, R. (1994) Pancreatic carboxylester lipase from Atlantic salmon (Salmo salar). cDNA sequence and computer-assisted modelling of tertiary structure. *European journal of biochemistry* **226**, 603-612

192. Holmes, R. S., and Cox, L. A. (2011) Comparative Structures and Evolution of Vertebrate Carboxyl Ester Lipase (CEL) Genes and Proteins with a Major Role in Reverse Cholesterol Transport. *Cholesterol* **2011**, 781643

193. Fontaine, R. N., Carter, C. P., and Hui, D. Y. (1991) Structure of the rat pancreatic cholesterol esterase gene. *Biochemistry* **30**, 7008-7014

194. Garcia-Diaz, M., and Kunkel, T. A. (2006) Mechanism of a genetic glissando: structural biology of indel mutations. *Trends in biochemical sciences* **31**, 206-214

195. Miyasaka, K., Ohta, M., Takano, S., Hayashi, H., Higuchi, S., Maruyama, K., Tando, Y., Nakamura, T., Takata, Y., and Funakoshi, A. (2005) Carboxylester lipase gene polymorphism as a risk of alcohol-induced pancreatitis. *Pancreas* **30**, e87-91

196. Fjeld, K., Beer, S., Johnstone, M., Zimmer, C., Mossner, J., Ruffert, C., Krehan, M., Zapf, C., Njolstad, P. R., Johansson, S., Bugert, P., Miyajima, F., Liloglou, T., Brown, L. *et al.* (2016) Length of Variable Numbers of Tandem Repeats in the Carboxyl Ester Lipase (CEL) Gene May Confer Susceptibility to Alcoholic Liver Cirrhosis but Not Alcoholic Chronic Pancreatitis. *PloS one* **11**, e0165567

197. Bengtsson-Ellmark, S. H., Nilsson, J., Orho-Melander, M., Dahlenborg, K., Groop, L., and Bjursell, G. (2004) Association between a polymorphism in the carboxyl ester lipase gene and serum cholesterol profile. *European journal of human genetics* **12**, 627-632

198. Stax, M. J., Kootstra, N. A., van 't Wout, A. B., Tanck, M. W., Bakker, M., Pollakis, G., and Paxton, W. A. (2012) HIV-1 disease progression is associated with bile-salt stimulated lipase (BSSL) gene polymorphism. *PloS one* **7**, e32534

199. Fjeld K, M. E., Michl P, Stokowy T, Lin JH, Steine SJ, Johansson BB, Dalva M, El Jellas K, Ruffert C, Zou WB, Li ZS, Njølstad PR, Chen JM, Liao Z, Johansson S, Rosendahl J, Férec C, Molven A. (2018) The polymorphic carboxyl-ester lipase (CEL) gene: novel copy number variants (CNVs) identified by both classical and high-throughput genetic analysis. *Pancreatology* **4**, S34

200. Lowenfels, A. B., Maisonneuve, P., DiMagno, E. P., Elitsur, Y., Gates, L. K., Jr., Perrault, J., and Whitcomb, D. C. (1997) Hereditary pancreatitis and the risk of pancreatic cancer. International Hereditary Pancreatitis Study Group. *Journal of the National Cancer Institute* **89**, 442-446

201. Schubert, S., Traub, F., Brakensiek, K., von Kopylow, K., Marohn, B., Maelzer, M., Gaedcke, J., Kreipe, H., and Stuhrmann, M. (2014) CFTR, SPINK1, PRSS1, and CTRC mutations are not associated with pancreatic cancer in German patients. *Pancreas* **43**, 1078-1082

202. Shindo, K., Yu, J., Suenaga, M., Fesharakizadeh, S., Tamura, K., Almario, J. A. N., Brant, A., Borges, M., Siddiqui, A., Datta, L., Wolfgang, C. L., Hruban, R. H., Klein, A. P., and Goggins, M. (2017) Lack of association between the pancreatitis risk allele CEL-HYB and pancreatic cancer. *Oncotarget* **8**, 50824-50831

203. Brekke, R. S. (2018) Characterization of insertion variants of the carboxyl-ester lipase (CEL) gene - A role in pancreatic disease? *Master´s thesis submitted to the University of Bergen.*

204. Tamura, K., Yu, J., Hata, T., Suenaga, M., Shindo, K., Abe, T., MacGregor-Das, A., Borges, M., Wolfgang, C. L., Weiss, M. J., He, J., Canto, M. I., Petersen, G. M., et al. (2018) Mutations in the pancreatic secretory enzymes CPA1 and CPB1 are associated with pancreatic cancer. Proceedings of the National Academy of Sciences of the United States of America 115, 4767-4772

205. Pasqualini, E., Caillol, N., Panicot, L., Mas, E., Lloubes, R., and Lombardo, D. (1998) Molecular cloning of the oncofetal isoform of the human pancreatic bile salt-dependent lipase. *The Journal of biological chemistry* **273**, 28208-28218

206. Klimstra, D. S., and Adsay, V. (2016) Acinar neoplasms of the pancreas-A summary of 25 years of research. *Seminars in diagnostic pathology* **33**, 307-318

207. Sugo, T., Mas, E., Abouakil, N., Endo, T., Escribano, M. J., Kobata, A., and Lombardo, D. (1993) The structure of N-linked oligosaccharides of human pancreatic bile-salt-dependent lipase. *European journal of biochemistry* **216**, 799-805

208. Landberg, E., Pahlsson, P., Krotkiewski, H., Stromqvist, M., Hansson, L., and Lundblad, A. (1997) Glycosylation of bile-salt-stimulated lipase from human milk: comparison of native and recombinant forms. *Archives of biochemistry and biophysics* **344**, 94-102

209. Geoghegan, K. F., Song, X., Hoth, L. R., Feng, X., Shanker, S., Quazi, A., Luxenberg, D. P., Wright, J. F., and Griffor, M. C. (2013) Unexpected mucin-type O-glycosylation and host-specific N-glycosylation of human recombinant interleukin-17A expressed in a human kidney cell line. *Protein expression and purification* **87**, 27-34

210. Goh, J. B., and Ng, S. K. (2018) Impact of host cell line choice on glycan profile. *Critical reviews in biotechnology* **38**, 851-867

211. Swiech, K., de Freitas, M. C., Covas, D. T., and Picanco-Castro, V. (2015) Recombinant glycoprotein production in human cell lines. *Methods in molecular biology* **1258**, 223-240

212. Trimble, R. B., Lubowski, C., Hauer, C. R., 3rd, Stack, R., McNaughton, L., Gemmill, T. R., and Kumar, S. A. (2004) Characterization of N- and O-linked glycosylation of recombinant human bile salt-stimulated lipase secreted by Pichia pastoris. *Glycobiology* **14**, 265-274

213. Wurm, F. M. (2004) Production of recombinant protein therapeutics in cultivated mammalian cells. *Nature biotechnology* **22**, 1393-1398

214. Mitoma, J., Petryniak, B., Hiraoka, N., Yeh, J. C., Lowe, J. B., and Fukuda, M. (2003) Extended core 1 and core 2 branched O-glycans differentially modulate sialyl Lewis X-type L-selectin ligand activity. *The Journal of biological chemistry* **278**, 9953-9961

215. Sasaki, H., Bothner, B., Dell, A., and Fukuda, M. (1987) Carbohydrate structure of erythropoietin expressed in Chinese hamster ovary cells by a human erythropoietin cDNA. *The Journal of biological chemistry* **262**, 12059-12076

216. Mas, E., Abouakil, N., Roudani, S., Franc, J. L., Montreuil, J., and Lombardo, D. (1993) Variation of the glycosylation of human pancreatic bile-salt-dependent lipase. *European journal of biochemistry* **216**, 807-812

217. Dalziel, M., Whitehouse, C., McFarlane, I., Brockhausen, I., Gschmeissner, S., Schwientek, T., Clausen, H., Burchell, J. M., and Taylor-Papadimitriou, J. (2001) The relative activities of the C2GnT1 and ST3Gal-I glycosyltransferases determine O-glycan structure and expression of a tumor-associated epitope on MUC1. *The Journal of biological chemistry* **276**, 11007-11015

218. Hamouda, H., Kaup, M., Ullah, M., Berger, M., Sandig, V., Tauber, R., and Blanchard, V. (2014) Rapid analysis of cell surface N-glycosylation from living cells using mass spectrometry. *Journal of proteome research* **13**, 6144-6151

219. Reinke, S. B., Marion & Berger, Markus & Blanchard, Véronique & Hinderlich, Stephan. . (2011) Analysis of Cell Surface N-glycosylation of the Human Embryonic Kidney 293T Cell Line. *Journal of Carbohydrate Chemistry* **5**,218-232.

220. Dewal, M. B., DiChiara, A. S., Antonopoulos, A., Taylor, R. J., Harmon, C. J., Haslam, S. M., Dell, A., and Shoulders, M. D. (2015) XBP1s Links the Unfolded Protein Response to the Molecular Architecture of Mature N-Glycans. *Chemistry & biology* **22**, 1301-1312

221. Aird, I., Lee, D. R., and Roberts, J. A. (1960) ABO blood groups and cancer of oesophagus, cancerof pancreas, and pituitary adenoma. *British medical journal* **1**, 1163-1166

222. Wolpin, B. M., Kraft, P., Gross, M., Helzlsouer, K., Bueno-de-Mesquita, H. B., Steplowski, E., Stolzenberg-Solomon, R. Z., Arslan, A. A., Jacobs, E. J., Lacroix, A., *et al.* (2010) Pancreatic cancer risk and ABO blood group alleles: results from the pancreatic cancer cohort consortium. *Cancer research* **70**, 1015-1023

223. Vasan, S. K., Hwang, J., Rostgaard, K., Nyren, O., Ullum, H., Pedersen, O. B. V., Erikstrup, C., Melbye, M., Hjalgrim, H., Pawitan, Y., and Edgren, G. (2016) ABO blood group and risk of cancer: A register-based cohort study of 1.6 million blood donors. *Cancer epidemiology* **44**, 40-43

224. Antwi, S. O., Bamlet, W. R., Pedersen, K. S., Chaffee, K. G., Risch, H. A., Shivappa, N., Steck, S. E., Anderson, K. E., Bracci, P. M., Polesel, J., Serraino, D., La Vecchia, C., *et al.* (2018) Pancreatic Cancer Risk is Modulated by Inflammatory Potential of Diet and ABO Genotype: A Consortia-based Evaluation and Replication Study. *Carcinogenesis*

225. Ben, Q., Wang, K., Yuan, Y., and Li, Z. (2011) Pancreatic cancer incidence and outcome in relation to ABO blood groups among Han Chinese patients: a case-control study. *International journal of cancer* **128**, 1179-1186

226. Rahbari, N. N., Bork, U., Hinz, U., Leo, A., Kirchberg, J., Koch, M., Buchler, M. W., and Weitz, J. (2012) AB0 blood group and prognosis in patients with pancreatic cancer. *BMC Cancer* **12**, 319

227. Pomianowska, E., Grzyb, K., Westgaard, A., Clausen, O. P., and Gladhaug, I. P. (2012) Reclassification of tumour origin in resected periampullary adenocarcinomas reveals underestimation of distal bile duct cancer. European journal of surgical oncology : the journal of the European Society of Surgical Oncology and the British Association of Surgical Oncology 38, 1043-1050

228. Wacklin, P., Makivuokko, H., Alakulppi, N., Nikkila, J., Tenkanen, H., Rabina, J., Partanen, J., Aranko, K., and Matto, J. (2011) Secretor genotype (FUT2 gene) is strongly associated with the composition of Bifidobacteria in the human intestine. *PloS one* **6**, e20113

229. Carlsson, B., Kindberg, E., Buesa, J., Rydell, G. E., Lidon, M. F., Montava, R., Abu Mallouh, R., Grahn, A., Rodriguez-Diaz, J., Bellido, J., Arnedo, A., Larson, G., and Svensson, L. (2009) The G428A nonsense mutation in FUT2 provides strong but not absolute protection against symptomatic GII.4 Norovirus infection. *PloS one* **4**, e5593

230. Raclawska, D. S., Ttofali, F., Fletcher, A. A., Harper, D. N., Bochner, B. S., Janssen, W. J., and Evans, C. M. (2016) Mucins and Their Sugars. Critical Mediators of Hyperreactivity and Inflammation. *Annals of the American Thoracic Society* **13**, S98-99

231. Wu, Y. M., Nowack, D. D., Omenn, G. S., and Haab, B. B. (2009) Mucin glycosylation is altered by pro-inflammatory signaling in pancreatic-cancer cells. *Journal of proteome research* **8**, 1876-1886

232. Hofmann, B. T., Stehr, A., Dohrmann, T., Gungor, C., Herich, L., Hiller, J., Harder, S., Ewald, F., Gebauer, F., Tachezy, M., Precht, C., Izbicki, J. R., Bockhorn, M., Wagener, C., and Wolters-Eisfeld, G. (2014) ABO blood group IgM isoagglutinins interact with tumor-associated O-glycan structures in pancreatic cancer. *Clinical cancer research* **20**, 6117-6126

233. Falk, P., Roth, K. A., Boren, T., Westblom, T. U., Gordon, J. I., and Normark, S. (1993) An in vitro adherence assay reveals that Helicobacter pylori exhibits cell lineage-specific tropism in the human gastric epithelium. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 2035-2039

234. Ruvoen-Clouet, N., Mas, E., Marionneau, S., Guillon, P., Lombardo, D., and Le Pendu, J. (2006) Bile-salt-stimulated lipase and mucins from milk of 'secretor' mothers inhibit the binding of Norwalk virus capsids to their carbohydrate ligands. *The Biochemical journal* **393**, 627-634

235. Downs, D., Xu, Y. Y., Tang, J., and Wang, C. S. (1994) Proline-rich domain and glycosylation are not essential for the enzymic activity of bile salt-activated lipase. Kinetic studies of T-BAL, a truncated form of the enzyme, expressed in Escherichia coli. *Biochemistry* **33**, 7979-7985

236. Hansson, L., Blackberg, L., Edlund, M., Lundberg, L., Stromqvist, M., and Hernell, O. (1993) Recombinant human milk bile salt-stimulated lipase. Catalytic activity is retained in the absence of glycosylation and the unique proline-rich repeats. *The Journal of biological chemistry* **268**, 26692-26698

237. Lombardo, D., Silvy, F., Crenon, I., Martinez, E., Collignon, A., Beraud, E., and Mas, E. (2018) Pancreatic adenocarcinoma, chronic pancreatitis, and MODY-8 diabetes: is bile salt-dependent lipase (or carboxyl ester lipase) at the crossroads of pancreatic pathologies? *Oncotarget* **9**, 12513-12533

238. Yang, Z., Halim, A., Narimatsu, Y., Jitendra Joshi, H., Steentoft, C., Schjoldager, K. T., Alder Schulz, M., Sealover, N. R., Kayser, K. J., Paul Bennett, E., Levery, S. B., Vakhrushev, S. Y., and Clausen, H. (2014) The GalNAc-type O-Glycoproteome of CHO cells characterized by the SimpleCell strategy. *Molecular & Cellular Proteomics : MCP* **13**, 3224-3235

239. Vesterhus, M., Raeder, H., Kurpad, A. J., Kawamori, D., Molven, A., Kulkarni, R. N., Kahn, C. R., and Njolstad, P. R. (2010) Pancreatic function in carboxyl-ester lipase knockout mice. *Pancreatology* **10**, 467-476

I

Pancreatology 17 (2017) 83–88

Contents lists available at ScienceDirect

Pancreatology

journal homepage: www.elsevier.com/locate/pan

Original article

Copy number variants and VNTR length polymorphisms of the carboxyl-ester lipase (*CEL*) gene as risk factors in pancreatic cancer

Monica Dalva [a, b, c], Khadija El Jellas [a, c, d], Solrun J. Steine [c], Bente B. Johansson [a, b], Monika Ringdal [a, b], Janniche Torsvik [a], Heike Immervoll [d], Dag Hoem [e], Felix Laemmerhirt [f], Peter Simon [f], Markus M. Lerch [f], Stefan Johansson [a, b], Pål R. Njølstad [a, g], Frank U. Weiss [f], Karianne Fjeld [a, b, *, 1], Anders Molven [a, c, d, 1]

[a] KG Jebsen Center for Diabetes Research, Department of Clinical Science, University of Bergen, Bergen, Norway
[b] Center for Medical Genetics and Molecular Medicine, Haukeland University Hospital, Bergen, Norway
[c] Gade Laboratory for Pathology, Department of Clinical Medicine, University of Bergen, Bergen, Norway
[d] Department of Pathology, Haukeland University Hospital, Bergen, Norway
[e] Department of Gastrointestinal Surgery, Haukeland University Hospital, Bergen, Norway
[f] Department of Medicine A, University Medicine Greifswald, Greifswald, Germany
[g] Department of Pediatrics, Haukeland University Hospital, Bergen, Norway

ARTICLE INFO

Article history:
Received 19 June 2016
Received in revised form
5 October 2016
Accepted 9 October 2016
Available online 11 October 2016

Keywords:
Allele frequency
Carboxyl-ester lipase
Copy number variation
Genotyping
Pancreatic cancer
Variable number of tandem repeats

ABSTRACT

Background/Objectives: We have recently described copy number variants (CNVs) of the human carboxyl-ester lipase (*CEL*) gene, including a recombined deletion allele (*CEL-HYB*) that is a genetic risk factor for chronic pancreatitis. Associations with pancreatic disease have also been reported for the variable number of tandem repeat (VNTR) region located in *CEL* exon 11. Here, we examined if *CEL* CNVs and VNTR length polymorphisms affect the risk for developing pancreatic cancer.
Methods: *CEL* CNVs and VNTR were genotyped in a German family with non-alcoholic chronic pancreatitis and pancreatic cancer, in 265 German and 197 Norwegian patients diagnosed with pancreatic adenocarcinoma, and in 882 controls. CNV screening was performed using PCR assays followed by agarose gel electrophoresis whereas VNTR lengths were determined by DNA fragment analysis.
Results: The investigated family was *CEL-HYB*-positive. However, an association of *CEL-HYB* or a duplication *CEL* allele with pancreatic cancer was not seen in our two patient cohorts. The frequency of the 23-repeat VNTR allele was borderline significant in Norwegian cases compared to controls (1.2% vs. 0.3%; $P = 0.05$). For all other VNTR lengths, no statistically significant difference in frequency was observed. Moreover, no association with pancreatic cancer was detected when *CEL* VNTR lengths were pooled into groups of short, normal or long alleles.
Conclusions: We could not demonstrate an association between *CEL* CNVs and pancreatic cancer. An association is also unlikely for *CEL* VNTR lengths, although analyses in larger materials are necessary to completely exclude an effect of rare VNTR alleles.

1. Introduction

The human *carboxyl-ester lipase* (*CEL*) gene is expressed mainly in pancreatic acinar cells [1] and lactating mammary glands [2]. *CEL* encodes a digestive enzyme, which is secreted into the duodenum as a component of pancreatic juice [3]. Upon activation by bile salts, CEL participates in the hydrolysis and absorption of cholesterol and lipid-soluble vitamins [4,5]. The human *CEL* gene spans approximately 10 kb on chromosome 9q34.3 and contains eleven exons [6,7]. The *CEL* locus also includes a *CEL* pseudogene (*CELP*), which lacks the region corresponding to exons 2–7 of *CEL*. Otherwise the two genes show a high degree of sequence homology [7,8].

The last exon of *CEL* contains a variable number of tandem repeat (VNTR) region. Each repeat consists of nearly identical 33-base pair segments. The number of repeated segments has been observed to vary between 3 and 23, with 16 repeats being the by far

* Corresponding author. KG Jebsen Center for Diabetes Research, Department of Clinical Science, University of Bergen, Jonas Lies veg 87, N-5021 Bergen, Norway.
E-mail address: karianne.fjeld@uib.no (K. Fjeld).
[1] Joint senior authorship.

http://dx.doi.org/10.1016/j.pan.2016.10.006
1424-3903/

most frequent number in all cohorts studied [9–13]. We have found that a single-base deletion in the first repeat of the *CEL* VNTR causes maturity-onset diabetes of the young, type 8 (MODY8). This is an autosomal dominantly inherited disease, characterized by diabetes and pancreatic exocrine dysfunction [11] and triggered by altered biochemical and cellular properties of the mutant protein [14,15]. Notably, the number of *CEL* VNTR segments has been reported to influence cholesterol profile of serum [16], rate of HIV-1 disease progression [17] and risk for alcohol-induced pancreatitis [18], although a study from our group could not replicate the latter association [13].

In addition to the VNTR region, copy number variants (CNVs) involving the *CELP* gene contribute to the extensive genetic variability of the *CEL* locus [12,19]. We have recently characterized two such CNVs, one duplication allele (*CEL-DUP*) and one deletion hybrid allele (*CEL-HYB*) [19]. These variants most probably originated from non-allelic homologous recombination between *CEL* and *CELP*. Intriguingly, *CEL-HYB* emerged as a novel genetic risk factor for chronic pancreatitis (CP) where, in contrast to most other known CP genes, the disease mechanism does not directly implicate the protease/antiprotease system of the exocrine pancreas [19].

CP is among the well-established risk factors for pancreatic cancer [20–22]. For patients with hereditary pancreatitis, the lifetime risk of developing pancreatic cancer can be 40% or more [20]. Thus, genetic variants that associate with CP should be considered potential risk factors for pancreatic cancer. Genome-wide association studies (GWAS) have identified multiple loci linked to CP [23,24] and pancreatic cancer [25–27], including a common association with the *ABO* locus [23,26,28]. Also *PRSS1* intron mutations have been reported to confer a risk for both pancreatitis and pancreatic cancer [29], as have mutations in the CP-associated gene *CFTR* [30,31].

Analyses based on GWAS will easily fail to capture associations between the highly polymorphic *CEL* gene and disease [19], and little is known about the gene in the context of malignant pancreatic disease. In light of the well-known links between chronic pancreatitis and pancreas cancer, we therefore investigated whether *CEL* CNVs and VNTR polymorphisms could be genetic risk factors for this form of cancer.

2. Methods

2.1. Patients and DNA samples

This study was performed according to the Helsinki Declaration and all patients gave their written informed consent. The project was approved by the ethical committee of the University of Greifswald and by the Regional Ethical Committee of Western Norway. The patient cohort consisted of 265 German subjects with a diagnosis of pancreatic ductal adenocarcinoma (PDAC) and 197 Norwegian subjects with a diagnosis of PDAC ($n = 156$) or other/unspecified adenocarcinoma anatomically located to the pancreas ($n = 41$). Healthy blood donors (German: $n = 502$, Norwegian: $n = 380$) were used as control groups. A German family with CP and pancreatic cancer was included for *CEL-HYB* screening and DNA samples were available from eight members. The proband of the family had in a previous study been identified as a *CEL-HYB* carrier [19]. Peripheral EDTA blood samples were drawn from all individuals and genomic DNA was extracted according to standard procedures.

2.2. Screening for CNVs of the CEL locus

For screening of the *CEL-HYB* allele, we used the long-range, duplex PCR approach previously described [19]. A similar strategy was used for screening of *CEL-DUP* and both assays are illustrated in Supplementary Fig. 1. Primers used for the *CEL* CNV screenings are listed in Supplementary Table 1. We performed the PCR reactions in a total volume of 10 μl, containing 1× GC buffer, 0.4 mM of each dNTP, 1 M betaine solution, 0.5 U *La Taq* polymerase (TaKaRa Bio Inc.), 0.4 μM of each primer (0.1 μM of Kir6.2 1F and Kir6.2 6R) and 10–50 ng genomic DNA. For *CEL-HYB*, the PCR conditions were as follows: 94 °C for 1 min; 14 cycles of 94 °C for 20 s and 60 °C for 6 min; 16 cycles of 94 °C for 20 s and 62 °C for 6 min; a final elongation step of 72 °C for 10 min and cooling to 4 °C. We used the same conditions for *CEL-DUP*, except that DNA synthesis was extended to 12 min and the number of cycles was 14 followed by 18. Amplified PCR products were analyzed by agarose gel electrophoresis.

2.3. Analysis of CEL VNTR length polymorphisms

CEL VNTR lengths were measured as previously described [12]. In short, PCR amplification of the VNTR was carried out using the unlabeled forward primer Celex11F and the fluorescently NED-labelled reverse primer Celex11R-NED (Supplementary Table 1). The PCR product covered the complete VNTR region as illustrated in Supplementary Fig. 2A. One μl of the PCR product was added to a mixture of 0.2 μl X-Rhodamine MapMarker 1000 (BioVentures Inc.) and 8.8 μl HiDi formaldehyde (Applied Biosystems), and fragment analysis was performed on an ABI3100 capillary sequencer (Applied Biosystems), using Gene Mapper software version 4.0 to analyze the data. Examples of the resulting electrophoretic spectra are given in Supplementary Fig. 2B.

2.4. Statistics

Differences between *CEL* CNVs or *CEL* VNTR length allele frequencies in cases and controls were tested by the two-tailed Fisher's exact test or Pearson's Chi-square test. Odd ratios (OR) and 95% confidence intervals (CI) were calculated by using two-by-two table analysis (http://www.quantitativeskills.com/sisa/). *P*-values <0.05 were considered to be statistically significant.

3. Results

3.1. A CEL-HYB-positive family with pancreatic cancer

During our identification of *CEL-HYB* as a risk factor for CP we analyzed a German discovery cohort of probands with familial idiopathic CP [19]. In one of the *CEL-HYB*-positive families, we noted that two family members also had a diagnosis of pancreatic cancer (Fig. 1). There were reportedly three additional cases of pancreatic cancer in the family (relatives of subject I-2). Although that branch of the pedigree was unavailable for further study, the family history could indicate increased risk for pancreatic cancer among carriers of *CEL-HYB*. Screening for the *CEL-HYB* allele revealed four positive members, including the only cancer case for which DNA was available (Fig. 1). All *CEL-HYB* carriers were affected with CP except for the relatively young subject IV-I. Among the *CEL-HYB* negatives, one subject (III-6) was diagnosed with recurrent acute pancreatitis. However, she also had a history of gallstone, which could explain her disease. Analyzing the VNTR length polymorphism, three of the *CEL-HYB*-positive family members had a single-peak pattern of 14 or 16 repeats, indicating one deletion hybrid allele and one wild type (*CEL-WT*) allele. Subject IV-I was not available for VNTR analysis. Moreover, *CEL-HYB*-positive subjects from a cohort of patients with familial idiopathic CP [19], in which this family's proband originally was included, were analyzed. All

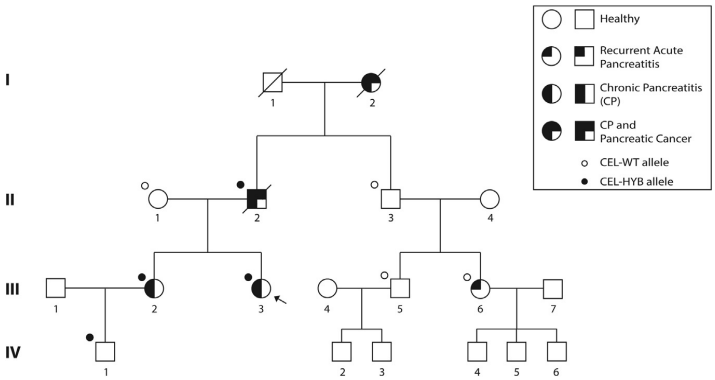


Fig. 1. Pedigree of a *CEL-HYB*-positive family with chronic pancreatitis and pancreatic cancer. Squares represent male family members and circles are females. DNA samples were available for subjects indicated by small circles. Deceased subjects are indicated by a slash through the symbol. The proband is marked with an arrow.

CEL-HYB carriers exhibited a single-peak pattern, with VNTR segments varying from 13 to 17 repeats.

3.2. Analysis of CEL copy number variants

Given the possible link between *CEL-HYB* and pancreatic cancer in the CP family described above, we investigated the prevalence of *CEL-HYB*, as well as the *CEL* duplication allele, in a cohort of German patients with pancreatic adenocarcinoma (Table 1). For *CEL-HYB*, we found a carrier frequency of 2.3% (6/263) in cases and 0.8% (4/495) in controls (OR = 2.87; 95% CI = 0.80–10.25; $P = 0.10$). When screening the same material for *CEL-DUP*, the carrier frequency was 4.2% (11/265) in cases and 5.2% (26/502) in controls (OR = 0.79; 95% CI = 0.39–1.63; $P = 0.60$).

We also screened a Norwegian cohort of subjects with pancreatic adenocarcinoma (Table 1). Here, no carriers of *CEL-HYB* were observed, and there was only one (0.3%) among 380 controls. The *CEL-DUP* allele had a carrier frequency of 5.6% (11/197) in cases and 2.6% (10/380) in controls (OR = 2.19; 95% CI = 0.91–5.25; $P = 0.10$) (Table 1). Taken together, we were therefore unable to detect any association between *CEL* CNVs and pancreatic cancer in our two study materials.

3.3. Analysis of CEL-VNTR length polymorphisms

We determined the number of *CEL* VNTR repeats in all cases and controls. Overall, we observed that the VNTR lengths varied between 4 and 23 repeats, with 16 being clearly the most frequent repeat number in both populations (Tables 2 and 3). Alleles with less than 13 repeats or more than 17 repeats were generally rare, with frequencies below 2%. The electrophoretic spectra obtained in the DNA fragment analysis showed either a single-peak (homozygous) or a double-peak (heterozygous) pattern for the large majority of samples (Supplementary Fig. 2B, panels i-iii). However, as also described previously [12], some samples gave rise to a pattern of three peaks (Supplementary Fig. 2B, panel iv), indicating that they harbor three copies of the complete *CEL* VNTR [12]. We counted the result of these genotypings as three alleles when calculating the allele frequencies. Thus, the number of recorded alleles is slightly higher than twice the number of subjects in Tables 2 and 3.

When comparing VNTR lengths, we noted that the shorter alleles of 4, 6 or 9 repeats were detected in Norwegian controls only (Tables 2 and 3). Moreover, the 23-repeat allele was enriched in Norwegian cases (5/401, 1.2%) when compared with controls (2/768, 0.3%) with a borderline statistical trend (OR = 4.84; 95% CI = 0.93–25.04; $P = 0.05$). Otherwise, no significant associations between a specific VNTR length and pancreatic cancer were observed (Tables 2 and 3).

Next, we classified the *CEL* VNTR lengths into groups of short (S, 4–15 repeats), normal (N, 16) and long (L, 17–23) alleles. Also in this pooled analysis, the VNTR distribution did not differ significantly between pancreatic cancer patients and controls (Tables 2 and 3). Finally, we analyzed the *CEL* VNTR length polymorphisms according to genotypes. Each case and control subject was classified into one of the following seven groups: SS, SN, NN, NL, LL, SL and samples harboring three VNTR lengths. The genotypes SN and NN were by far the two most common genotypes in both populations (Tables 4 and 5). None of the differences in genotype frequencies between cases and controls reached statistical significance. When the different genotypes of samples with three VNTR lengths were grouped (e.g. as SSN) and tested, no significant differences in distribution between cases and controls were observed (Supplementary Tables 2 and 3, and data not shown).

4. Discussion

Variants of *CEL* are known to cause a rare monogenic disease of the pancreas and to influence the risk of chronic pancreatitis [11,19,32]. In contrast, little is known about the role of this gene as

Table 1
CNVs of the *CEL* locus in pancreatic cancer.

Cohort	CNV type	Cases	Controls	*P*-value	Odds ratio	95% CI
German	*CEL-HYB*	0.023 (6/263)	0.008 (4/495)	0.10	2.87	0.80–10.25
	CEL-DUP	0.042 (11/265)	0.052 (26/502)	0.60	0.79	0.39–1.36
Norwegian	*CEL-HYB*	0.000 (0/197)	0.003 (1/380)	–	–	–
	CEL-DUP	0.056 (11/197)	0.026 (10/380)	0.10	2.19	0.91–5.25

Table 2
Distribution of *CEL* VNTR lengths in the German pancreatic cancer cohort.

		Allele frequency		*P*-value	OR	95% CI
		Cases ($n = 265$)	Controls ($n = 497$)			
Number of VNTR repeats	11	0.007 (4)	0.015 (15)	0.24	0.49	0.16–1.49
	12	0.013 (7)	0.018 (18)	0.53	0.72	0.30–1.74
	13	0.052 (28)	0.053 (53)	1.00	0.98	0.61–1.57
	14	0.118 (64)	0.115 (116)	0.85	1.03	0.75–1.43
	15	0.155 (84)	0.138 (139)	0.35	1.15	0.86–1.54
	16	0.619 (335)	0.624 (629)	0.85	0.98	0.79–1.22
	17	0.031 (17)	0.035 (35)	0.77	0.90	0.50–1.63
	20	–	0.001 (1)	–	–	–
	21	0.002 (1)	0.001 (1)	1.00	1.87	0.12–29.87
	23	0.002 (1)	0.001 (1)	1.00	1.87	0.12–29.87
	Sum	1.000 (541)	1.000 (1008)	–	–	–
Pooled VNTR lengths	Short (<16)	0.346 (187)	0.338 (341)	0.77	1.03	0.83–1.29
	Normal (16)	0.619 (335)	0.624 (629)	0.85	0.98	0.79–1.22
	Long (>16)	0.035 (19)	0.038 (38)	0.89	0.93	0.53–1.63
	Sum	1.000 (541)	1.000 (1008)	–	–	–

Table 3
Distribution of *CEL* VNTR lengths in the Norwegian pancreatic cancer cohort.

		Allele frequency		*P*-value	OR	95% CI
		Cases ($n = 197$)	Controls ($n = 378$)			
Number of VNTR repeats	4	–	0.001 (1)	–	–	–
	6	–	0.001 (1)	–	–	–
	9	–	0.001 (1)	–	–	–
	11	0.007 (3)	0.005 (4)	0.70	1.44	0.32–6.46
	12	0.012 (5)	0.008 (6)	0.53	1.60	0.49–5.29
	13	0.055 (22)	0.057 (44)	0.89	0.96	0.56–1.62
	14	0.100 (40)	0.091 (70)	0.63	1.10	0.73–1.66
	15	0.142 (57)	0.113 (87)	0.15	1.30	0.91–1.86
	16	0.621 (249)	0.661 (508)	0.17	0.84	0.65–1.08
	17	0.045 (18)	0.044 (34)	1.00	1.02	0.57–1.82
	18	0.002 (1)	0.009 (7)	0.28	0.27	0.03–2.22
	21	0.002 (1)	0.004 (3)	1.00	0.64	0.07–6.15
	23	0.012 (5)	0.003 (2)	0.05	4.84	0.93–25.04
	Sum	1.000 (401)	1.000 (768)	–	–	–
Pooled VNTR lengths	Short (<16)	0.317 (127)	0.279 (214)	0.17	1.20	0.92–1.56
	Normal (16)	0.621 (249)	0.661 (508)	0.17	0.84	0.65–1.08
	Long (>16)	0.062 (25)	0.060 (46)	0.87	1.04	0.63–1.73
	Sum	1.000 (401)	1.000 (768)	–	–	–

risk factor in pancreatic cancer, which is the 4-5th most common cause of cancer-related death in Western societies [33]. Martinez et al. recently analyzed a French material for the SNP rs488087 located within the *CEL* VNTR, and reported an association between this variant and pancreatic cancer [34]. Their cohort was, however, extremely small (n = 36) and also included pancreatic neuroendocrine neoplasms, so the observation needs independent verification before this SNP can be considered a pancreatic cancer risk factor.

GWAS of pancreatic cancer have never revealed associations between *CEL* and this disease [27,35,36]. It is to be noted, though, that the *CEL* gene is extremely polymorphic and that SNP sets used

Table 4
Genotype distribution of *CEL* VNTR lengths in the German pancreatic cancer cohort.

Genotypes of pooled alleles[a]	Frequency		*P*-value	OR	95% CI
	Cases ($n = 265$)	Controls ($n = 497$)			
SS	0.094 (25)	0.111 (55)	0.54	0.84	0.51–1.38
SN	0.442 (117)	0.400 (199)	0.27	1.18	0.88–1.60
SL	0.015 (4)	0.016 (8)	1.00	0.94	0.28–3.14
NN	0.377 (100)	0.394 (196)	0.65	0.93	0.69–1.27
NL	0.030 (8)	0.048 (24)	0.26	0.61	0.27–1.39
LL	–	0.002 (1)	–	–	–
3 VNTR[b]	0.042 (11)	0.028 (14)	0.39	1.49	0.67–3.34
Sum	1.000 (265)	1.000 (497)	–	–	–

[a] Short (S); 4–15 repeats, Normal (N); 16 repeats, Long (L); 17–23 repeats.
[b] Samples harboring three copies of the VNTR region. Their genotypes are listed in Supplementary Table 2.

Table 5
Genotype distribution of *CEL* VNTR lengths in the Norwegian pancreatic cancer cohort.

Genotypes of pooled alleles[a]	Frequency		*P*-value	OR	95% CI
	Cases ($n = 197$)	Controls ($n = 378$)			
SS	0.107 (21)	0.071 (27)	0.15	1.55	0.85–2.82
SN	0.360 (71)	0.341 (129)	0.65	1.09	0.76–1.56
SL	0.015 (3)	0.026 (10)	0.56	0.57	0.16–2.09
NN	0.411 (81)	0.452 (171)	0.34	0.85	0.60–1.20
NL	0.061 (12)	0.071 (27)	0.73	0.84	0.42–1.70
LL	0.010 (2)	0.005 (2)	0.61	1.93	0.27–13.79
3 VNTR[b]	0.036 (7)	0.032 (12)	0.81	1.12	0.44–2.90
Sum	1.000 (197)	1.000 (378)	–	–	–

[a] Short (S); 4–15 repeats, Normal (N); 16 repeats, Long (L); 17–23 repeats.
[b] Samples harboring three copies of the VNTR region. Their genotypes are listed in Supplementary Table 3.

in genotyping arrays are inadequate to cover the extensive genetic variation of the *CEL* locus [19]. Possible disease associations of *CEL* must therefore be explored also by a targeted approach. Accordingly, we set out to investigate whether *CEL* CNVs or VNTR length influence pancreatic cancer risk. To improve the power of our study, we included two independently collected cohorts of pancreatic cancer patients from Germany and Norway.

In recent years, accumulating evidence has defined pre-existing longstanding CP as a strong risk factor for pancreatic neoplastic disease. During a 20 year period around 5% of patients with CP will present with cancer of the pancreas. The risk of developing this cancer form appears to be highest in rare types of pancreatitis with an early onset, such as hereditary pancreatitis [37], whereas the most common type, alcoholic CP, does not necessarily involve higher pancreatic cancer susceptibility [38]. Interestingly, cystic fibrosis, a recessive early-onset genetic disorder affecting many organ systems including the pancreas, has been associated with a ~5-fold increased risk of pancreatic cancer, often occurring at a relatively young age [39]. Therefore, *CEL* and any other new gene causing or being associated with pancreatic inflammatory disease should be investigated also in the context of pancreas cancer.

We were, however, unable to reveal an association between the *CEL* CNVs and pancreatic cancer, even though the *CEL-HYB* allele is a CP risk factor that we also had detected in a pedigree with a history of CP and pancreatic cancer. For this family, other pancreatic cancer risk factors (genetic and/or environmental) might be present. There are also previous examples where CP genetic risk factors have failed to be linked to increased pancreatic cancer risk [40,41]. The population carrier frequency of the *CEL-HYB* allele is low and its associated CP risk is intermediate [19]. It may therefore not be surprising that *CEL-HYB* is a rare event in cohorts of pancreatic cancer. Nevertheless, this allele could confine an increased risk in those carriers that develop CP, especially if pancreatitis onset is relatively early in life. Additional epidemiologic studies in large materials may be able to clarify this question. These studies should also take into account how smoking might interact with *CEL-HYB* and other *CEL* variants. Unfortunately, data on smoking status were available only for about half of our cases and unknown for the controls, which prohibited us from drawing any conclusions about how smoking may have influenced our results.

To further investigate *CEL* in the context of pancreatic cancer, we studied *CEL* VNTR length polymorphisms. VNTR lengths pooled into groups of short, normal and long alleles did not show any association. When analyzing the VNTR alleles one by one, the most notable observation was an enrichment of the 23-repeat allele among Norwegian cases (Table 3). This association was borderline ($P = 0.05$), it was not found in the German material, and will not be significant if correction for multiple testing is performed. We therefore conclude that it is a spurious finding. Nevertheless, it should be noted that extreme-length VNTR alleles (<13 and >17 repeats) are rare, and analysis in very large materials would be needed to gain sufficient power to firmly conclude about their effects on human disease.

When analyzing *CEL* as genetic risk factor one also needs to consider that some alleles may be population-specific. Zou et al. recently attempted to replicate the association between *CEL-HYB* and CP in Indian, Chinese and Japanese cohorts, but discovered that this allele was completely absent in these Asian populations [42]. In this respect it is noteworthy that the distribution of VNTR allele sizes between our Norwegian and German control materials differed. Short alleles were significantly more frequent among German than in Norwegian controls (0.338 vs. 0.279; $P = 0.007$), whereas long alleles were less common among Germans (0.038 vs. 0.060; $P = 0.029$; Pearson's Chi-square test, data taken from Tables 2 and 3). Moreover, the frequencies of the *CEL-DUP* allele were close to being significantly different in the two control materials (0.052 vs. 0.026; $P = 0.06$).

Another caveat when genotyping *CEL* VNTR lengths is the presence of CNV alleles where the complete VNTR itself has been duplicated [12]. During our analyses we identified a total of 44 individuals who carried three copies of the *CEL* VNTR region (Supplementary Tables 2 and 3). Interestingly, only 18 of these subjects tested positive for the *CEL-DUP* allele. This strongly suggests that there exist duplication CNVs of the *CEL* locus that remain to be characterized.

Moreover, in the current VNTR assay, any sample harboring three complete VNTRs of which two are of equal lengths (e.g. 16, 16, 18) is not distinguishable from a standard heterozygous sample (e.g. 16, 18). In total, we identified 58 subjects who tested positive in the *CEL-DUP*-specific assay (Table 1), but when performing VNTR analysis only 18 of these samples resulted in a genotype pattern consisting of three different peaks (Supplementary Fig. 2B, panel iv). Thus, the true number of individuals with an extra *CEL* VNTR copy could be substantially higher in our populations than what recorded in the present study. This observation shows that there is a need for better methods to genotype CNV alleles of the *CEL* locus.

Taken together, our results do not indicate any association between the two analyzed *CEL* CNVs or *CEL* VNTR length polymorphisms and pancreatic cancer risk. However, *CEL* is a highly polymorphic gene where yet uncharacterized CNVs are very likely to exist and where also some disease-associated alleles may be confined to specific populations. It is therefore too early to rule out that genetic variants of *CEL* could play a role in pancreatic cancer.

Acknowledgements

This study was supported by a PhD fellowship to M.D from the University of Bergen, and by grants from the Meltzer Fund, KG

Jebsen Foundation and Western Norway Regional Health Authority (grant numbers: 911831, 912057).

The authors declare no competing financial interests.

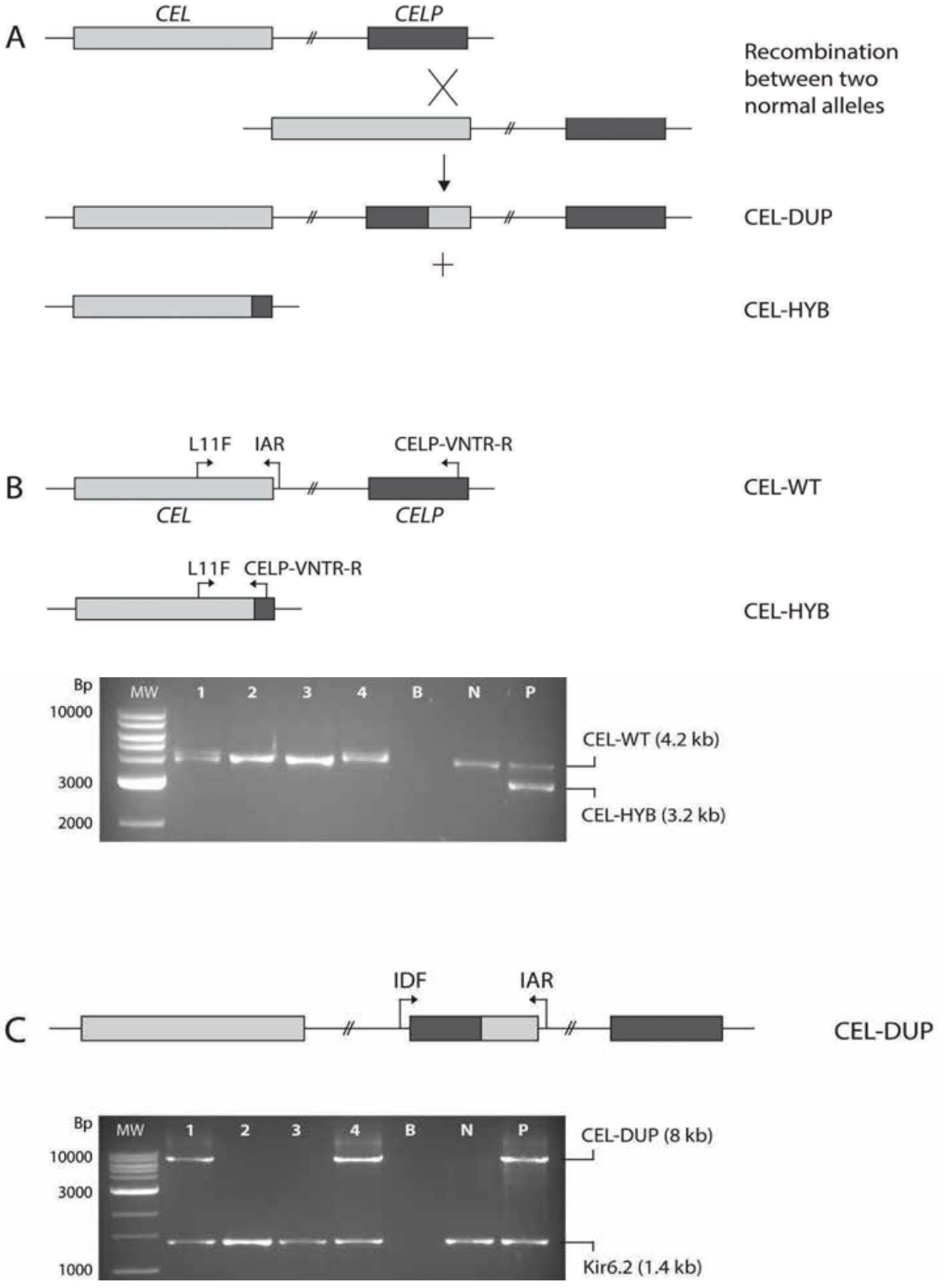
Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.pan.2016.10.006.

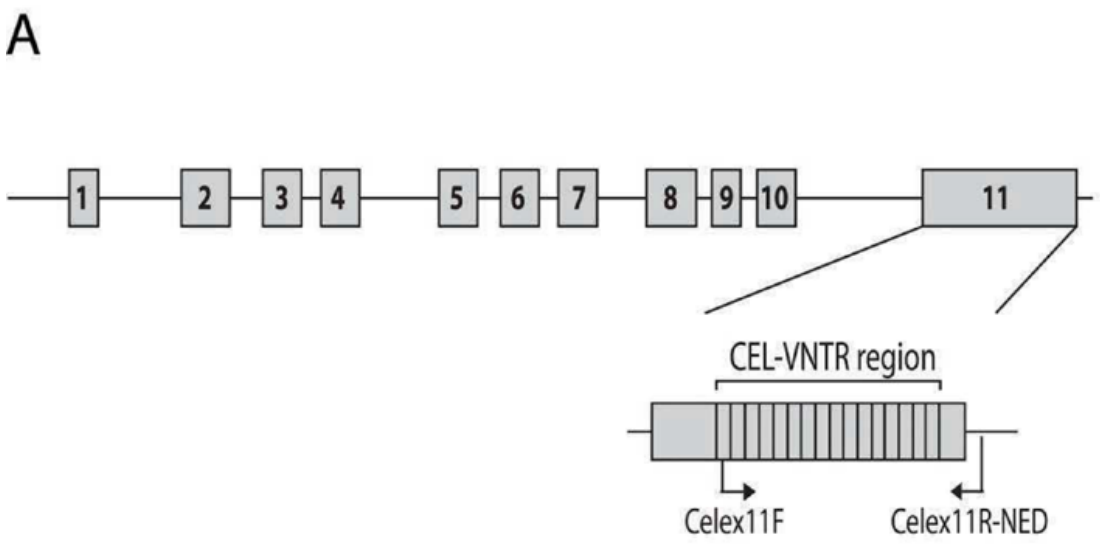
References

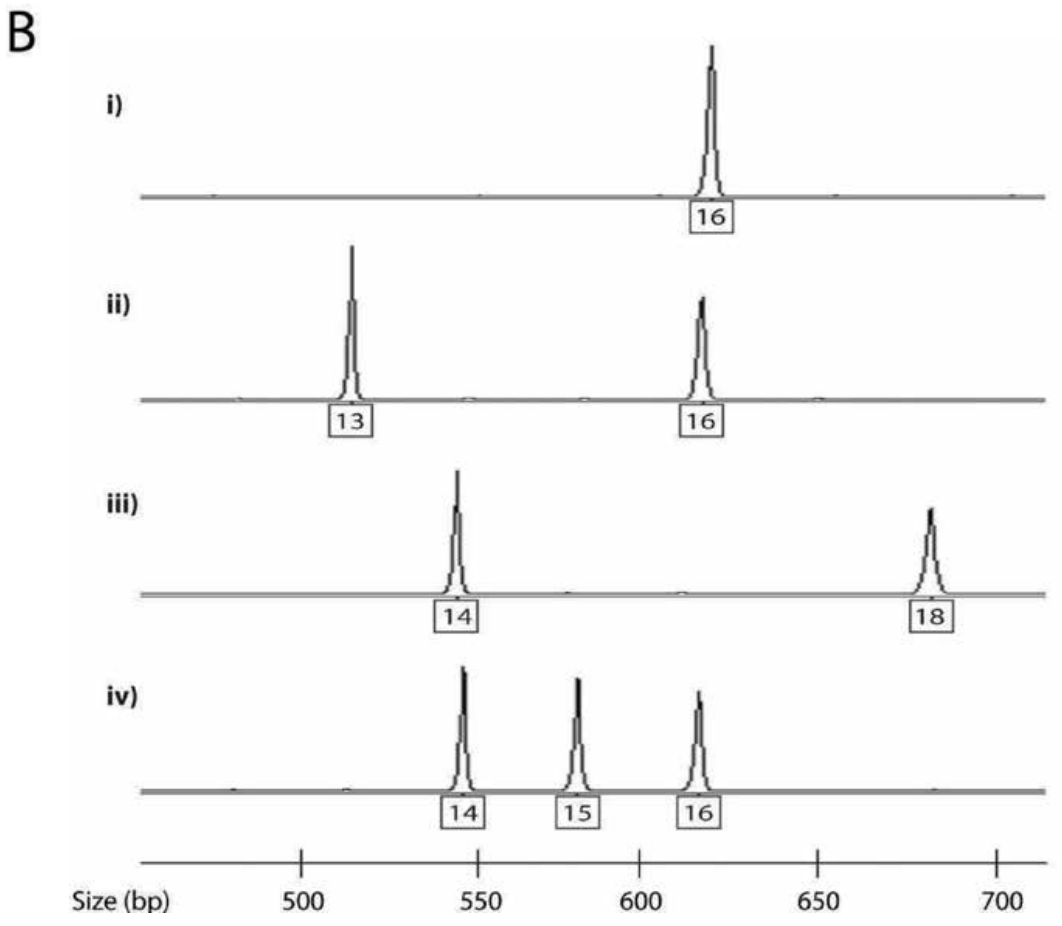
[1] Roudani S, Miralles F, Margotat A, Escribano MJ, Lombardo D. Bile salt-dependent lipase transcripts in human fetal tissues. Biochim Biophys Acta 1995;1264:141–50.
[2] Blackberg L, Angquist KA, Hernell O. Bile-salt-stimulated lipase in human milk: evidence for its synthesis in the lactating mammary gland. FEBS Lett 1987;217:37–41.
[3] Lombardo D, Guy O, Figarella C. Purification and characterization of a carboxyl ester hydrolase from human pancreatic juice. Biochim Biophys Acta 1978;527:142–9.
[4] Lombardo D, Guy O. Studies on the substrate specificity of a carboxyl ester hydrolase from human pancreatic juice. II. Action on cholesterol esters and lipid-soluble vitamin esters. Biochim Biophys Acta 1980;611:147–55.
[5] Lombardo D. Bile salt-dependent lipase: its pathophysiological implications. Biochim Biophys Acta 2001;1533:1–28.
[6] 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;10:425–31.
[7] 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;13:630–40.
[8] Madeyski K, Lidberg U, Bjursell G, Nilsson J. Structure and organization of the human carboxyl ester lipase locus. Mamm Genome 1998;9:334–8.
[9] Higuchi S, Nakamura Y, Saito S. Characterization of a VNTR polymorphism in the coding region of the CEL gene. J Hum Genet 2002;47:213–5.
[10] Lindquist S, Blackberg L, Hernell O. Human bile salt-stimulated lipase has a high frequency of size variation due to a hypervariable region in exon 11. Eur J Biochem 2002;269:759–67.
[11] 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. Nat Genet 2006;38:54–62.
[12] 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. Hum Genet 2010;127:55–64.
[13] Ragvin A, Fjeld K, Weiss FU, Torsvik J, Aghdassi A, Mayerle J, et al. The number of tandem repeats in the carboxyl-ester lipase (CEL) gene as a risk factor in alcoholic and idiopathic chronic pancreatitis. Pancreatology 2013;13:29–32.
[14] 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;286:34593–605.
[15] 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;289:29097–111.
[16] 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. Eur J Hum Genet 2004;12:627–32.
[17] Stax MJ, Kootstra NA, van 't Wout AB, Tanck MW, Bakker M, Pollakis G, et al. HIV-1 disease progression is associated with bile-salt stimulated lipase (BSSL) gene polymorphism. PLoS One 2012;7:e32534.
[18] Miyasaka K, Ohta M, Takano S, Hayashi H, Higuchi S, Maruyama K, et al. Carboxylester lipase gene polymorphism as a risk of alcohol-induced pancreatitis. Pancreas 2005;30:e87–91.
[19] 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;47:518–22.
[20] Lowenfels AB, Maisonneuve P, DiMagno EP, Elitsur Y, Gates Jr LK, Perrault J, et al. Hereditary pancreatitis and the risk of pancreatic cancer. International hereditary pancreatitis study group. J Natl Cancer Inst 1997;89:442–6.
[21] Nitsche C, Simon P, Weiss FU, Fluhr G, Weber E, Gartner S, et al. Environmental risk factors for chronic pancreatitis and pancreatic cancer. Dig Dis 2011;29:235–42.
[22] Rebours V, Boutron-Ruault MC, Schnee M, Ferec C, Maire F, Hammel P, et al. Risk of pancreatic adenocarcinoma in patients with hereditary pancreatitis: a national exhaustive series. Am J Gastroenterol 2008;103:111–9.
[23] Weiss FU, Schurmann C, Guenther A, Ernst F, Teumer A, Mayerle J, et al. Fucosyltransferase 2 (FUT2) non-secretor status and blood group B are associated with elevated serum lipase activity in asymptomatic subjects, and an increased risk for chronic pancreatitis: a genetic association study. Gut 2015;64:646–56.
[24] Whitcomb DC, LaRusch J, Krasinskas AM, Klei L, Smith JP, Brand RE, et al. Common genetic variants in the CLDN2 and PRSS1-PRSS2 loci alter risk for alcohol-related and sporadic pancreatitis. Nat Genet 2012;44:1349–54.
[25] Wolpin BM, Rizzato C, Kraft P, Kooperberg C, Petersen GM, Wang Z, et al. Genome-wide association study identifies multiple susceptibility loci for pancreatic cancer. Nat Genet 2014;46:994–1000.
[26] Amundadottir L, Kraft P, Stolzenberg-Solomon RZ, Fuchs CS, Petersen GM, Arslan AA, et al. Genome-wide association study identifies variants in the ABO locus associated with susceptibility to pancreatic cancer. Nat Genet 2009;41:986–90.
[27] Petersen GM, Amundadottir L, Fuchs CS, Kraft P, Stolzenberg-Solomon RZ, Jacobs KB, et al. A genome-wide association study identifies pancreatic cancer susceptibility loci on chromosomes 13q22.1, 1q32.1 and 5p15.33. Nat Genet 2010;42:224–8.
[28] Xu HL, Cheng JR, Zhang W, Wang J, Yu H, Ni QX, et al. Re-evaluation of ABO gene polymorphisms detected in a genome-wide association study and risk of pancreatic ductal adenocarcinoma in a Chinese population. Chin J Cancer 2014;33:68–73.
[29] Gao F, Liu QC, Zhang S, Zhuang ZH, Lin CZ, Lin XH. PRSS1 intron mutations in patients with pancreatic cancer and chronic pancreatitis. Mol Med Rep 2012;5:449–51.
[30] McWilliams RR, Petersen GM, Rabe KG, Holtegaard LM, Lynch PJ, Bishop MD, et al. Cystic fibrosis transmembrane conductance regulator (CFTR) gene mutations and risk for pancreatic adenocarcinoma. Cancer 2010;116:203–9.
[31] Hamoir C, Pepermans X, Piessevaux H, Jouret-Mourin A, Weynand B, Habyalimana JB, et al. Clinical and morphological characteristics of sporadic genetically determined pancreatitis as compared to idiopathic pancreatitis: higher risk of pancreatic cancer in CFTR variants. Digestion 2013;87:229–39.
[32] Molven A, Fjeld K, Lowe ME. Lipase genetic variants in chronic pancreatitis: when the end is wrong, all's not well. Gastroenterology 2016;150:1515–8.
[33] Maisonneuve P, Lowenfels AB. Epidemiology of pancreatic cancer: an update. Dig Dis 2010;28:645–56.
[34] Martinez E, Silvy F, Fina F, Bartoli M, Krahn M, Barlesi F, et al. Rs488087 single nucleotide polymorphism as predictive risk factor for pancreatic cancers. Oncotarget 2015;24(6):39855–64.
[35] Childs EJ, Mocci E, Campa D, Bracci PM, Gallinger S, Goggins M, et al. Common variation at 2p13.3, 3q29, 7p13 and 17q25.1 associated with susceptibility to pancreatic cancer. Nat Genet 2015;47:911–6.
[36] Wu C, Miao X, Huang L, Che X, Jiang G, Yu D, et al. Genome-wide association study identifies five loci associated with susceptibility to pancreatic cancer in Chinese populations. Nat Genet 2012;44:62–6.
[37] Raimondi S, Lowenfels AB, Morselli-Labate AM, Maisonneuve P, Pezzilli R. Pancreatic cancer in chronic pancreatitis; aetiology, incidence, and early detection. Best Pract Res Clin Gastroenterol 2010;24:349–58.
[38] Midha S, Sreenivas V, Kabra M, Chattopadhyay TK, Joshi YK, Garg PK. Genetically determined chronic pancreatitis but not alcoholic pancreatitis is a strong risk factor for pancreatic cancer. Pancreas 2016;45:1478–84. http://dx.doi.org/10.1097/MPA.0000000000000684.
[39] Maisonneuve P, Marshall BC, Lowenfels AB. Risk of pancreatic cancer in patients with cystic fibrosis. Gut 2007;56:1327–8.
[40] Schubert S, Traub F, Brakensiek K, von Kopylow K, Marohn B, Maelzer M, et al. CFTR, SPINK1, PRSS1, and CTRC mutations are not associated with pancreatic cancer in German patients. Pancreas 2014;43:1078–82.
[41] Matsubayashi H, Fukushima N, Sato N, Brune K, Canto M, Yeo CJ, et al. Polymorphisms of SPINK1 N34S and CFTR in patients with sporadic and familial pancreatic cancer. Cancer Biol Ther 2003;2:652–5.
[42] 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;150:1558–60.

Supplementary Figure 1



Supplementary Figure 2





Supplementary Figure Legends

Supplementary Figure 1 – Long-PCR assays for detection of *CEL* CNVs. A) Schematic model of how the *CELHYB* and *CEL-DUP* alleles have arisen by non-allelic, homologous recombination. B) Principle of the *CEL-HYB* screening assay. Primer binding sites are marked in the upper panel. The L11F and *CELP* VNTR-R primers amplify a 3.2 kb product specific for the *CEL-HYB* allele, whereas primers L11F and IAR amplify a 4.2 kb PCR product from the normal *CEL* gene (internal control). A typical image of the PCR products analyzed by agarose gel electrophoresis is shown in the lower panel. Note that there is a double band for sample 1, probably reflecting the difference of its *CEL* VNTR lengths (13 and 23 repeats). C) Principle of the *CEL-DUP* screening assay. Primer binding sites are marked in the upper panel. The IDF and IAR primers amplify an 8.0 kb product specific for the *CEL-DUP* allele. Primers Kir6.2 1F and Kir6.2 6R amplify a 1.4 kb product from the unrelated *KCNJ11* gene (not shown) and serve as internal control for the assay. A typical image of the PCR products analyzed by agarose gel electrophoresis is shown in the lower panel.

Supplementary Figure 2 – Analysis of *CEL* VNTR length. A) Location of primer binding sites in *CEL* exon 11. The forward primer (Celex11F) binds partly into the first repeat of the *CEL*-VNTR whereas the NED-labelled revers primer (Celex11R-NED) binds downstream of the exon. B) Examples of capillary electrophoresis spectra after DNA fragment analysis. Peak sizes (base pairs) of the NED-labelled fragments reflect the repeat number of each *CEL* VNTR allele, as indicated under the peaks. The illustrated spectra are screenshots of four samples corresponding to i) homozygous 16-repeat alleles, ii) heterozygous 13- and 16-repeat alleles, iii) heterozygous 14- and 18-repeat alleles and iv) three copies of VNTRs with 14, 15 and 16 repeats.

Supplementary Table 1. Primers used for *CEL* genotyping

Name	Sequence (5'-3')
L11F	GTCCCTCACTCATTCTTCTATGGCAAC
CELP VNTR-R	CTGTGGAGGGCATGGAACT
IAR	TCCAAAGCCCTAGCAGTAACGA
IDF	CTGCCCACCAGTCCTTTATGCT
Kir6.2 1F	CCGAGAGGACTCTGCAGTGA
Kir6.2 6R	CAAGTGCCTTGTAACACCCT
Celex11F	ACCGACCAGGAGGCCACCC
Celex11R-NED	CCTGGGGTCCCACTCTTGT

Supplementary Table 2. Genotypes of observed *CEL* duplication alleles in the German pancreatic cancer cohort

Materials	Number of VNTR repeats			S/N/L[a]	*CEL-DUP*
Cases	12	12	16	SSN	-
(n=11)	13	14	16	SSN	+
	14	15	16	SSN	-
	14	15	16	SSN	-
	14	15	16	SSN	-
	14	15	17	SSL	+
	14	16	17	SNL	+
	14	16	17	SNL	-
	14	16	17	SNL	+
	15	16	17	SNL	+
	16	17	23	NLL	+
Controls	11	14	16	SSN	+
(n=14)	12	14	16	SSN	+
	13	14	16	SSN	-
	13	15	16	SSN	-
	14	15	16	SSN	-
	14	15	16	SSN	-
	14	15	16	SSN	-
	14	15	16	SSN	-
	14	15	16	SSN	+
	14	15	16	SSN	-
	14	15	16	SSN	-
	15	16	17	SNL	-
	15	16	17	SNL	+
	16	17	23	NLL	+

[a] Short (S), 4-15 repeats; Normal (N), 16 repeats; Long (L), 17-23 repeats

Supplementary Table 3. Genotypes of observed *CEL* duplication alleles in the Norwegian pancreatic cancer cohort

Materials	Number of VNTR repeats			S/N/L[a]	*CEL-DUP*
Cases	13	14	15	SSS	-
(n=7)	11	14	16	SSN	-
	14	15	16	SSN	-
	14	15	16	SSN	+
	14	17	23	SLL	+
	15	17	23	SLL	+
	16	17	23	NLL	+
Controls	13	14	15	SSS	-
(n=12)	13	14	16	SSN	-
	13	14	16	SSN	-
	13	14	16	SSN	-
	13	14	16	SSN	-
	13	15	16	SSN	-
	14	15	16	SSN	-
	14	15	16	SSN	-
	14	15	16	SSN	+
	14	16	17	SNL	-
	15	17	23	SLL	+
	16	17	23	NLL	+

[a] Short (S), 4-15 repeats; Normal (N), 16 repeats; Long (L), 17-23 repeats

II

Cancer Medicine

Open Access

ORIGINAL RESEARCH

Associations between ABO blood groups and pancreatic ductal adenocarcinoma: influence on resection status and survival

Khadija El Jellas[1,2,3], Dag Hoem[4], Kristin G Hagen[5], May Britt Kalvenes[1], Sura Aziz[1,2], Solrun J Steine[1], Heike Immervoll[1], Stefan Johansson[3,6] & Anders Molven[1,2,6]

[1]Gade Laboratory for Pathology, Department of Clinical Medicine, University of Bergen, Bergen, Norway
[2]Department of Pathology, Haukeland University Hospital, Bergen, Norway
[3]Center for Medical Genetics and Molecular Medicine, Haukeland University Hospital, Bergen, Norway
[4]Department of Gastrointestinal Surgery, Haukeland University Hospital, Bergen, Norway
[5]Department of Immunology and Transfusion Medicine, Haukeland University Hospital, Bergen, Norway
[6]KG Jebsen Center for Diabetes Research, Department of Clinical Science, University of Bergen, Bergen, Norway

Keywords
ABO, blood group, FUT2, glycosyltransferase, pancreatic ductal adenocarcinoma, risk factor

Correspondence
Anders Molven, Gade Laboratory for Pathology, Department of Clinical Medicine, University of Bergen, Bergen, Norway.
Tel: 0047 55 973169; Fax: 0047 55 973158;
E-mail: anders.molven@uib.no

Received: 22 January 2017; Revised: 12 April 2017; Accepted: 20 April 2017

***Cancer Medicine* 2017; 6(7):1531–1540**

doi: 10.1002/cam4.1097

Abstract

Both serology-based and genetic studies have reported an association between pancreatic cancer risk and ABO blood groups. We have investigated this relationship in a cohort of pancreatic cancer patients from Western Norway ($n = 237$) and two control materials (healthy blood donors, $n = 379$; unselected hospitalized patients, $n = 6149$). When comparing patient and blood donor *ABO* allele frequencies, we found only the A_1 allele to be associated with significantly higher risk for pancreatic ductal adenocarcinoma (PDAC) (23.8% vs. 17.9%; OR = 1.43, $P = 0.018$). Analyzing phenotypes, blood group A was more frequent among PDAC cases than blood donors (50.8% vs. 40.6%; OR = 1.51, $P = 0.021$), an enrichment fully explained by the A_1 subgroup. Blood group O frequency was lower in cases than in blood donors (33.8% vs. 42.7%; OR = 0.69, $P = 0.039$). This lower frequency was confirmed when cases were compared to hospitalized patients (33.8% vs. 42.9%; OR = 0.68, $P = 0.012$). Results for blood group B varied according to which control cohort was used for comparison. When patients were classified according to surgical treatment, the enrichment of blood group A was most prominent among unresected cases (54.0%), who also had the lowest prevalence of O (28.7%). There was a statistically significant better survival ($P = 0.04$) for blood group O cases than non-O cases among unresected but not among resected patients. Secretor status did not show an association with PDAC or survival. Our study demonstrates that pancreatic cancer risk is influenced by ABO status, in particular blood groups O and A_1, and that this association may reflect also in tumor resectability and survival.

Introduction

Pancreatic cancer is one of the most dreaded malignant diseases. It ranks the fourth most common cause of cancer-related death in many Western countries and has a remarkably dismal prognosis with a 5-year survival rate of <5% [1]. The manifestation of symptoms occurs relatively late in the disease process. Most pancreatic cancer cases are therefore diagnosed at an unresectable stage and only 15–20% of the patients undergo surgical resection, which is the only potentially curative treatment.

Histologically, the large majority of pancreatic cancer cases are exocrine and classified as pancreatic ductal adenocarcinoma (PDAC), in which the characteristic morphological pattern consists of abundant duct-like neoplastic structures embedded in a dense desmoplastic stroma [2]. Accepted risk factors for PDAC include advanced age, cigarette smoking [3], and long-standing chronic

pancreatitis [4], whereas some studies also have implicated *Helicobacter pylori* infection [5] and diabetes mellitus [6]. In addition, inherited susceptibility plays a role in the disease, both as high-risk gene variants in the context of familial cancer syndromes [7] and as the presence of variants with modest effect, usually discovered by genome-wide association studies (GWAS) [8].

In a landmark GWAS paper, the Pancreatic Cancer Cohort Consortium (PanScan) reported that the statistically most significant variants associating with pancreatic cancer risk belonged to the *ABO* locus on chromosome 9q34 [9]. *ABO* codes for a glycosyltransferase that gives rise to the histo-blood group antigens of the ABO system. Single-nucleotide polymorphisms (SNPs) of this gene determine the specificity of the enzyme [10, 11]. Hence, by adding either *N*-acetyl-D-galactosamine or D-galactose to the precursor H antigen, the A and B glycosyltransferases produce A or B antigens, respectively, on cellular surfaces and secretions. A frequent *ABO* variant is a one-base pair deletion that inactivates the encoded enzyme, leaving the H antigen unaltered and corresponding to the O phenotype [10]. This deletion is in strong linkage disequilibrium with the T allele of SNP rs505922, which was identified as being associated with decreased susceptibility for pancreatic cancer [9]. Accordingly, individuals with blood group O have a lower risk for this disease than those with other blood groups [12].

There are two major subtypes of the *ABO* allele determining the blood group A, namely A_1 and A_2. Data from PanScan demonstrated that, among all common *ABO* variants, the greatest risk of pancreatic cancer was conferred by the A_1 allele [13] which gives rise to the ABO protein with highest enzymatic activity [14]. This finding suggests that it is the glycosyltransferase activity itself that is linked to cancer risk rather than actions of other nearby genes on chromosome 9q34. Somewhat surprisingly though, the association between blood group and pancreatic cancer was not influenced by the secretor phenotype, that is, a person's ability to secrete A, B, and H antigens into body fluids [13]. This property is determined by *FUT2*, the gene encoding the glycosylating enzyme fucosyltransferase 2.

The initial reports [9, 12, 13] were followed up by genetic studies in various populations that confirmed the influence of ABO blood group alleles on pancreatic cancer risk [15–17]. The finding that blood group O confers protection was also consistent with older papers that had reported an association between ABO phenotype and gastrointestinal cancers including pancreatic cancer [18–20]. A meta-analysis including over 20 studies, both genetic and serology-based, concluded that all non-O blood groups have elevated risks for pancreatic cancer as compared with the O phenotype [16].

In this study, we aimed at evaluating the link between pancreatic cancer and ABO histo-blood groups in patients from Norway, a population in which this association has not yet been investigated. Our patients were carefully characterized to exclude non-PDAC cases and also classified according to tumor resectability and survival. Two different sets of controls were included for statistical comparisons.

Materials and Methods

Study population

The study was performed according to the Helsinki Declaration and all patients gave their written informed consent. The project was approved by the Regional Ethical Committee of Western Norway. The patient cohort consisted of 237 cases of pancreatic adenocarcinoma seen at Haukeland University Hospital, Bergen, Norway between the years 1998 and 2012 (Table 1). Medical records, pathology reports, and/or tissue sections were examined by two pathologists (HI and SA) for diagnosis confirmation. Final classification resulted in 195 PDAC cases and

Table 1. Overview of pancreatic cancer patients and controls included in the study.

		Females			Males		
Cohort	Total *n*	*n*	%	Median age	*n*	%	Median age
Cases							
All pancreatic adenocarcinoma cases	237	119	50.2	69	118	49.8	69
Pancreatic ductal adenocarcinoma (PDAC)	195	97	49.7	69	98	50.3	69
Resected	108	50	46.3	70	58	53.7	68
Not resected	87	47	54.0	70	40	46.0	67
Other adenocarcinomas[1]	42	22	52.4	67	20	47.6	68
Controls							
DNA-typed blood donors	379	189	49.9	39	190	50.1	44
Serotyped hospital patients	6149	2805	45.6	66	3344	54.4	64

[1]See Materials and Methods for description.

42 other adenocarcinomas located within the pancreas. The latter group consisted of intraductal papillary mucinous neoplasm (5 cases) or mucinous cystic neoplasm with malignant component (1 case), intrapancreatic adenocarcinoma of the ampulla/papilla of Vateri (19 cases) or of ductus choledochus (9 cases), and unspecified adenocarcinoma of the pancreas (8 cases).

For the statistical comparison, two different control groups were employed (Table 1). One consisted of 379 healthy blood donors from Haukeland University Hospital (49.9% females) that were genotyped in the same way as the cases. The other control group contained 6149 patients (45.6% females) born before 1.1.1970 and admitted to the same hospital during a randomly chosen period of six consecutive months in 2007. All patients had been blood-typed by serological means as part of their health care. No selection with regard to diagnosis was done.

DNA extraction and genotyping

EDTA-blood, frozen tissue and formalin-fixed paraffin-embedded (FFPE) tissue blocks were used for DNA extraction from the 237 pancreatic adenocarcinoma cases. DNA from frozen buffy coats from EDTA-blood (195 cases) were purified using MagAttract DNA Blood Midi M48 kit on the BioRobot M48 workstation (both from Qiagen, Hilden, Germany) or manually processed with the E.Z.N.A DNA extraction kit (Omega Bio-Tek, Norcross, GA, USA), according to the manufacturers' protocol. For 38 cases, only FFPE tissue samples were available; 10-micron sections were then sliced and subjected to manual deparaffinization with xylene, followed by ethanol washes before overnight incubation at 56°C with Proteinase K (Qiagen) and processing with the E.Z.N.A. DNA extraction kit. For the final four cases, only fresh-frozen tissue samples were available. These were incubated directly with Proteinase K and then processed as the FFPE samples. DNA from the blood donor controls was isolated from EDTA-blood buffy coats using the same purification system as for the patient blood samples.

Genotyping was performed using TaqMan predesigned genotyping assays (Cat. No. 4351379; Applied Biosystems, Foster City, CA). Each sample was tested for three common SNPs at the *ABO* locus (Table S1): rs8176704 (intron 3) for the A_2 allele (Assay ID: C_30336657_10), rs8176746 (exon 7) for the B allele (Assay ID: C_25610772_20), and rs505922 (intron 1) for the O allele (Assay ID: C_2253769_10). The samples were also screened for the *FUT2* variant rs601338 (Assay ID: C_2405292_10), which determines secretor status of ABH antigens.

The genotyping assays were performed on the 7900 Fast Real-time PCR System with the corresponding 7900 Fast System SDS 2.4 Software (Applied Biosystems). Positive and negative controls were included to ensure appropriate clustering. Each assay was performed using 10 ng template DNA, TaqMan Universal Master Mix buffer (Applied Biosystems), and 20x primer and probe mix as recommended by the manufacturer. Thermal cycling was performed by first activating the DNA polymerase at 95°C for 10 min and then running 40 amplification cycles, each consisting of denaturation at 92°C for 15 sec and combined annealing/extension at 60°C for 1 min.

Quality control

Misclassification of ABO genotypes was expected to be minimal, as genotyping results from the three different ABO SNPs matched the haplotype phasing of the *ABO* gene (Table 1 in [13]) for all analyzed samples. In addition, we compared our genotyping results for 40 cases (20 extracted from blood and 20 from FFPE tissues) with serologically determined ABO status as stated in the patients' medical records. The concordance rate was 100%. Genotyping quality for the blood donor control group was assessed by testing for Hardy–Weinberg equilibrium using the Haploview Software [21]. The distribution of genotypes was as expected from the SNP frequencies.

Statistical analysis

Each SNP was tested under various genetic models (dominant, codominant, recessive, additive) using the software PLINK (http://pngu.mgh.harvard.edu/~purcell/plink). Assessment of differences in genotype or phenotype distributions between cases and controls was carried out by the two-tailed Fisher`s exact test or Pearson`s chi-square test. Odds ratios (OR) and 95% confidence intervals (CI) were calculated using two-by-two contingency table analysis on the SISA webpage (http://www.quantitativeskills.com/sisa). Survival analysis was performed using the software package STATISTICA version 12 (StatSoft, Tulsa, OK). The Product–Limit (Kaplan–Meier) Analysis Module was employed for comparing survival between groups by log-rank test of significance. Survival times versus cumulative proportion surviving, according to breakdown by blood group, were plotted. In all tests, $P \leq 0.05$ was chosen for statistical significance.

Results

Patient characteristics

From our biobank of patients with pancreatic tumors, we initially selected the patients diagnosed with adenocarcinoma of the exocrine gland. This cohort consisted of 237 cases (50.2% females), with a median age at

diagnosis of 69 years in both sexes (Table 1). We reviewed all cases to identify those that were consistent with a diagnosis of PDAC (n = 195, 82.3%). The cases were also classified according to whether or not the tumor had been judged resectable at the time of diagnosis (Table 1).

Association between ABO blood group and pancreatic cancer risk

The genotype frequencies for all adenocarcinoma cases and blood donor controls are given in Table S2. In both groups, the most common ABO genotypes were A_1O and OO, and the least frequent were BB and A_2A_2. We first compared allele frequencies of A_1, A_2, B, and O in the whole adenocarcinoma cohort with frequencies observed in the blood donors (Table 2, 'All cases'). The A_1 allele frequency was higher among the patients (22.4% vs. 17.9%) but the difference did not quite reach statistical significance (P = 0.057). When the analysis was limited to PDAC cases only, there was a significant difference in A_1 frequency (23.8% vs. 17.9%; OR = 1.43, CI = 1.06–1.93; P = 0.018). Interestingly, the A_2 frequency appeared almost identical between the groups compared (7.6–7.8%). The B and O allele frequencies varied, but were not statistically different.

For the further analyses, we restricted our analysis to the PDAC cases only. We deduced ABO phenotypes from the genotype data of the cases and the blood donors. The blood group distributions are shown in Table 3. The blood group A prevalence was clearly different (50.8% vs. 40.6%; OR = 1.51, CI = 1.06–2.13; P = 0.021,) and, in keeping with the data of Table 2, the subgroup A_1 frequencies fully explained the observed difference (42.6% vs. 29.3%; OR = 1.79, CI = 1.25–2.56; P = 0.001). Moreover, the prevalence of blood group O was lower in cases than in controls (33.8% vs. 42.7%; OR = 0.69, CI = 0.48–0.98; P = 0.039). Blood group B did not show a statistically significant difference in distribution. Neither did blood group AB, although in this case, the number of subjects was too small for meaningful comparisons to be made.

Healthy blood donors may not always serve as an optimal control group in case–control studies [22]. We therefore collected information on ABO blood group distribution from a large control cohort of unselected hospitalized patients (see Materials) from the same geographical region as our pancreatic cancer patients. The subgroup A_1/A_2 distribution was not known for the hospitalized patient cohort. Notably, the hospitalized patients had a significantly higher blood group A prevalence than the blood donors (45.8% vs. 40.6%; OR = 1.24, CI = 1.00–1.53; P = 0.049). Similarly, blood group B was significantly less frequent (7.7% vs. 12.1%; OR = 0.60, CI = 0.44–0.83; P = 0.002). Blood group O had very similar prevalence in the two control groups (42.7% and 42.9%).

When the PDAC blood group distribution was compared with that of the hospitalized patients, the enrichment of blood group A among the PDAC cases no longer reached statistical significance (50.8% vs. 45.8%; OR = 1.22, CI = 0.92–1.62; P = 0.173) (Table 4). On the other hand, the blood group B difference was now significant (12.8% vs. 7.7%; OR = 1.76, CI = 1.15–2.71; P = 0.009). The lower frequency of blood group O in the PDAC cohort remained significant with almost identical odds ratio (33.8% vs. 42.9%; OR = 0.68, CI = 0.50–0.92; P = 0.012).

Tumor resectability

To further explore the association between blood group frequencies and PDAC, the patients were stratified into two subgroups: those who had their pancreatic tumor resected and those who were considered surgically unresectable at the time of cancer diagnosis (Table 1). The latter group consisted of patients with locally advanced tumors with encasement of adjacent large blood vessels (Clinical stage III) or with metastatic disease at the time of diagnosis (Clinical stage IV) [23]. We observed that

Table 2. *ABO* and *FUT2* allele frequencies of blood donor controls compared with pancreatic cancer cases.

	Controls (n = 758)	All cases (n = 474)			PDAC cases only (n = 390)		
Allele	%	%	P	OR (95% CI)	%	P	OR (95% CI)
ABO							
A_1	17.9	22.4	0.057	1.32 (0.99–1.75)	23.8	**0.018**	1.43 (1.06–1.93)
A_2	7.8	7.6	1.000	0.97 (0.63–1.50)	7.7	1.000	0.99 (0.63–1.56)
B	8.7	7.6	0.491	0.86 (0.56–1.32)	7.7	0.556	0.87 (0.56–1.37)
O	65.6	62.4	0.266	0.87 (0.69–1.11)	60.8	0.109	0.81 (0.63–1.05)
FUT2							
Se	51.5	51.5	1.000	1.00 (0.79–1.26)	52.1	0.847	1.02 (0.80–1.31)
Se^0	48.5	48.5	1.000	1.00 (0.79–1.26)	47.9	0.847	0.98 (0.77–1.25)

n, number of genotyped alleles; PDAC, pancreatic ductal adenocarcinoma; P, P -value from chi-square test (df = 1); OR (95% CI), odds ratio (95% confidence interval). Significant P-value is shown in bold face.

Table 3. ABO blood group and secretor phenotype frequencies of blood donor controls compared with PDAC cases.

Phenotype	Controls (n = 379) %	PDAC cases (n = 195)			PDAC cases according to resection status					
					Resected (n = 108)			Unresected (n = 87)		
		%	P	OR (95% CI)	%	P	OR (95% CI)	%	P	OR (95% CI)
ABO										
A	40.6	50.8	**0.021**	1.51 (1.06–2.13)	48.1	0.163	1.36 (0.88–2.09)	54.0	**0.023**	1.72 (1.07–2.74)
A_1	29.3	42.6	**0.001**	1.79 (1.25–2.56)	40.7	**0.024**	1.66 (1.07–2.59)	44.8	**0.005**	1.96 (1.22–3.16)
A_2	11.3	8.2	0.241	0.70 (0.38–1.27)	7.4	0.238	0.63 (0.29–1.37)	9.2	0.562	0.79 (0.36–1.75)
B	12.1	12.8	0.814	1.06 (0.63–1.79)	11.1	0.771	0.91 (0.46–1.78)	14.9	0.478	1.27 (0.65–2.47)
AB[1]	4.5	2.6	0.359	0.56 (0.20–1.54)	2.8	0.586	0.61 (0.18–2.12)	2.3	0.548	0.50 (0.11–2.21)
O	42.7	33.8	**0.039**	0.69 (0.48–0.98)	38.0	0.374	0.82 (0.53–1.27)	28.7	**0.016**	0.54 (0.33–0.90)
FUT2										
Secretor	77.6	76.4	0.753	0.94 (0.62–1.41)	72.2	0.248	0.74 (0.46–1.22)	81.6	0.410	1.28 (0.71–2.32)
Non-secretor	22.4	23.6	0.753	1.07 (0.71–1.61)	27.8	0.248	1.33 (0.82–2.16)	18.4	0.410	0.78 (0.43–1.41)

PDAC, pancreatic ductal adenocarcinoma; P, P -value from chi-square test (df=1); OR (95% CI), odds ratio (95% confidence interval). Significant P -values are shown in bold face.

[1]P -values from two-tailed Fisher`s exact test.

Table 4. ABO phenotypic frequencies of hospital patient controls compared with PDAC cases.

Blood types	Controls (n = 6149) %	PDAC cases (n = 195)			PDAC cases according to resection status					
					Resected (n = 108)			Unresected (n = 87)		
		%	P	OR (95% CI)	%	P	OR (95% CI)	%	P	OR (95% CI)
A	45.8	50.8	0.173	1.22 (0.92–1.62)	48.1	0.632	1.10 (0.75–1.61)	54.0	0.128	1.39 (0.91–2.12)
B	7.7	12.8	**0.009**	1.76 (1.15–2.71)	11.1	0.188	1.50 (0.82–2.75)	14.9	**0.012**	2.11 (1.16–3.83)
AB[1]	3.5	2.6	0.690	0.72 (0.29–1.76)	2.8	1.000	0.78 (0.25–2.47)	2.3	0.771	0.64 (0.16–2.62)
O	42.9	33.8	**0.012**	0.68 (0.50–0.92)	38.0	0.301	0.81 (0.55–1.20)	28.7	**0.008**	0.54 (0.34–0.86)

PDAC, pancreatic ductal adenocarcinoma; P, P-value from chi-square test (df=1); OR (95% CI), odds ratio (95% confidence interval). Significant P -values are shown in bold face.

[1]P-values from two-tailed Fisher`s exact test

the blood group A and subgroup A_1 prevalences were highest among the unresected cases (54.0% and 44.8%, respectively) and both were significantly different from the frequencies found in blood donors (A: OR = 1.72, CI = 1.07–2.74, P = 0.023; A_1: OR = 1.96, CI = 1.22–3.16, P = 0.005) (Table 3). Moreover, the unresected cases had the lowest frequency of blood group O (28.7% vs. 42.7%; OR = 0.54, CI = 0.33–0.90; P = 0.016). A comparison with the cohort of hospitalized patients (Table 4) revealed significant differences for the unresected patients, both with regard to blood group B (14.9% vs. 7.7%; OR = 2.11, CI = 1.16–3.83; P = 0.012) and blood group O (28.7% vs. 42.9%; OR = 0.54, CI = 0.34–0.86; P = 0.008).

Survival

Data on survival after time of diagnosis was available for all PDAC cases. As expected, survival was significantly better among resected than among unresected cases (median survival 18.1 months vs. 5.6 months, respectively; $P < 0.001$) (Fig. 1A). Given that blood group is a risk factor for PDAC and that it also may influence resection status, we also analyzed survival according to ABO phenotype. When all 195 patients were classified as O or non-O cases, survival of the two groups was not significantly different (median 10.4 vs. 9.3 months, respectively; P = 0.23) (Fig. 1B). We then looked at resected and unresected cases separately. Among the resected cases, survival did not differ between patients with O and non-O blood group (P = 0.93) (Fig. 1C). However, in the group of unresected cases, patients with blood group O survived longer than non-O patients (median 6.7 vs. 5.5 months, respectively; P = 0.04) (Fig. 1D). When the non-O cases were split into blood group A or B and compared to the O cases, the difference in survival reached significance for blood group B (median 2.7 months; P = 0.03), but not for A (median 5.6 months; P = 0.14) (Fig. 1E–F). Although it should be noted that the number of A_2 cases

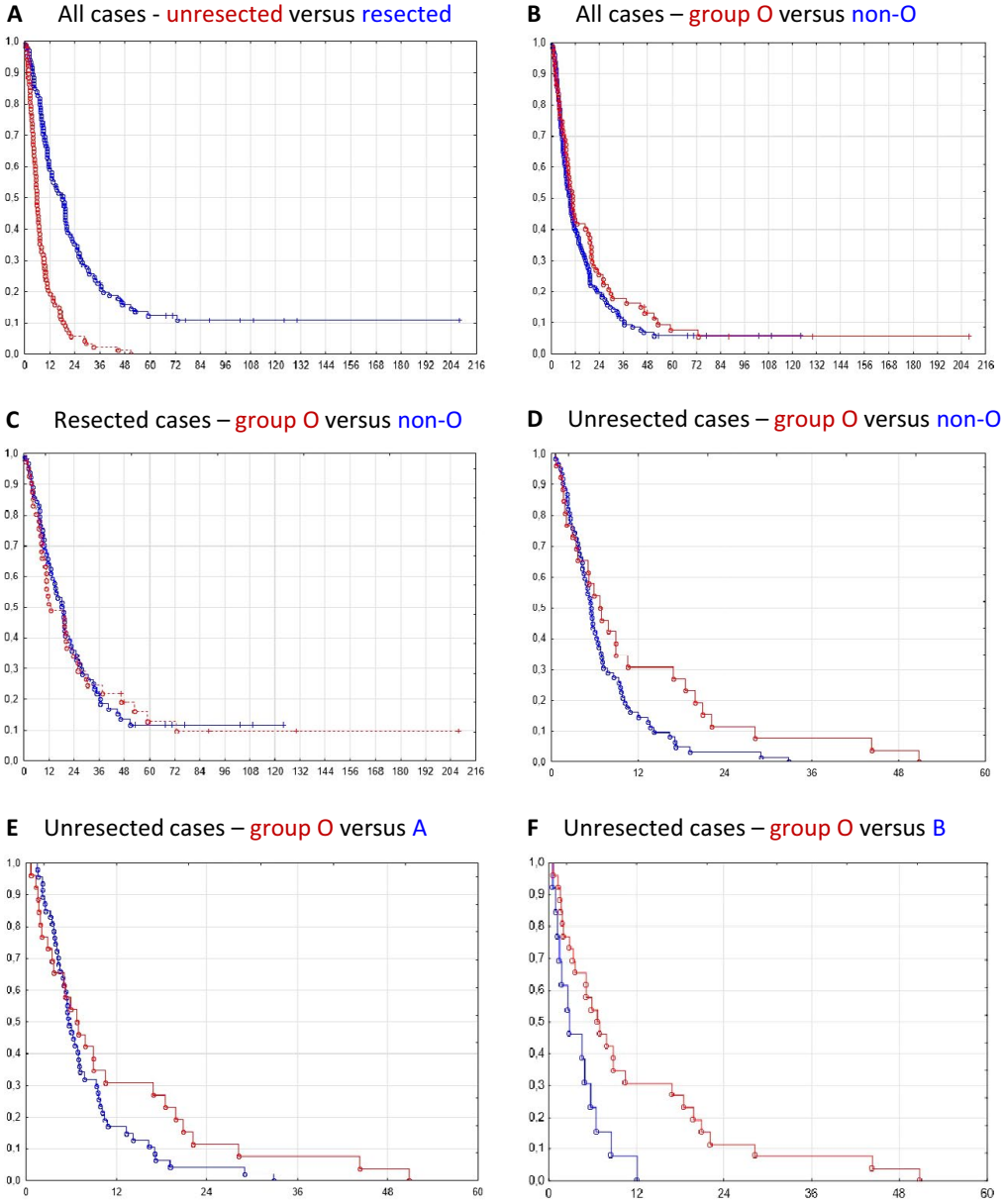


Figure 1. Cumulative proportion survival (Kaplan–Meier) plot for the 195 pancreatic ductal adenocarcinomas according to breakdown by resection status and blood group phenotype. (A)-(F) Pairwise comparison of subgroups as specified in the heading of each panel. The observed survival times (months along the *X*-axis) are indicated by circles (complete) or crosses (censored observations). In panel A, survival curves for resected and unresected cases are shown by a blue and a red line, respectively. In all other panels, survival curves for blood group O and the comparison group are shown by a red and blue line, respectively.

is relatively small (16/195 cases), we also examined whether there was a survival difference between patients of blood group A_1 and A_2. Neither among all cases nor when cases were classified according to resection status was any statistically significant difference observed (data not shown).

Secretor status

Because the secretion of soluble H antigen is associated with susceptibility to multiple pathogens through adherence to the gastrointestinal mucosa we looked at the secretor phenotype determined by the rs601338 *FUT2* polymorphism (Tables S1, S2). Consistent with another study [13], there were no significant differences in prevalence of alleles or phenotypes when compared with the blood donor controls (Tables 2, 3). We further checked whether *FUT2* could have an effect on resection status by comparing the groups of resected and unresected PDAC cases in Table 3 directly against each other. The difference in secretor status was not significant (72.2% vs. 81.6%; OR=0.59, CI=0.30–1.16; $P = 0.124$). Finally, we examined if secretor status might associate with survival in our patient cohort. No such association was seen (Fig. S1).

Discussion

A variety of biological traits have been linked to ABO histo-blood group phenotypes. Reported associations include, among others, plasma lipid levels [24], susceptibility to *H. pylori*–associated disorders [25], cardiovascular disease [26], and cancer [27]. Here, we have investigated *ABO* genetic variants and phenotypes in the context of pancreatic cancer. A strength of our study is that a re-evaluation of the recorded pancreatic adenocarcinomas was performed to select only those cases consistent with a diagnosis of PDAC. Thus, from our initial cohort of adenocarcinomas with an anatomical location in the pancreas, 42 (17.7%) cases were excluded. Materials of pancreatic head cancers may contain a significant number of adenocarcinomas originating from the distal bile duct, ampulla, or duodenum [28], and proper classification of cases collectively regarded as pancreatic cancer should always be attempted.

One limitation of our study is the relatively low number of cases and blood donor controls. Another relates to the selection of control groups, which may pose a challenge for ABO phenotype and other traits with a distribution that varies by geography and ethnicity. The optimal control group for our study would have been randomly drawn individuals from the population of Western Norway, matched to cases by age, sex and county of residence. However, such a cohort could not be obtained within the restrictions of the present study.

We confirmed the association between the blood group O phenotype and reduced susceptibility to pancreatic exocrine cancer (frequency of 33.8% in cases versus 42.7% and 42.9% in the two control groups). When comparing with blood donors, blood group A was significantly over-represented among cases. We made the interesting observation that the association with blood group A was explained by the A_1 allele/blood group only, a finding previously reported by a few papers [13, 29]. Serology-based analyses and many genetic studies of ABO blood groups in pancreatic cancer do not distinguish between A_1 and A_2 subgroups. However, this distinction might be of biological importance as A_1 and A_2 glycosyltransferases display different catalytic activities, the A_2 isoform having a higher K_m and an estimated enzyme activity 30–50 times lower than A_1 [14]. The finding that A_1, but not A_2, associates with pancreatic cancer could therefore suggest a role of blood group A glycosyltransferase activity in carcinogenesis.

Studies of associations between ABO blood groups and human disease are challenging for at least two reasons. Firstly, across the globe there are important variations in *ABO* frequencies, reflecting human migration movements and possibly also selection pressures from the environment [30, 31]. Within Norway, the *ABO* distribution differs according to geographical localization with the highest blood group O frequency observed in the Western part of the country and the highest B frequency in the northernmost region [32]. Secondly, healthy blood donors might not always serve as appropriate controls for epidemiological studies in general and, studies of ABO in particular. There could, for example, be a bias in ABO distribution among donors because of the association of blood groups with a variety of diseases. Another potential selective effect could arise in that certain blood groups might be preferred as donors.

In the literature, it is well established that blood group O individuals have a lower risk of pancreatic cancer [33]. For the other blood types, the data are not always consistent. Some authors report an association with blood group A, but not with B [29, 34] whereas others describe an association with B, but not with A [20] or even find AB to be protective [35]. One reason for this discrepancy might therefore be the way that control groups have been selected. We employed two different sets of controls: a limited number of blood donors and a large number of random patients from the same geographical area as our patients. The control groups exhibited very similar blood group O frequencies (close to 43%). Thus, we found a very consistent protective effect of blood group O on PDAC risk. On the other hand, the frequencies of blood groups A and B varied considerably among the two control groups, resulting in a significant association

of PDAC risk with blood group A when blood donors were used and with blood group B when employing hospital patients. Interestingly, in a recent large study where 1.6 million blood donors from Denmark and Sweden were followed over an average of 17 years, there was increased risk for pancreatic cancer in subjects with blood group A when compared with O subjects [36]. Blood group B did not associate with increased risk, that is, a result similar to what we observed when using blood donors as controls.

We also analyzed ABO blood group distribution in the context of resectability at the time of diagnosis (Tables 3 and 4). In this respect, it should be noted that our cases were recruited at a surgical department to which patients usually are referred when they are to be evaluated for operation. Thus, those PDAC patients that at the time of diagnosis obviously are unresectable will be clearly underrepresented. This explains the relatively high fraction of operated patients in our material (55.4%) compared with the general low resection rate for pancreatic adenocarcinomas, which is around 20%.

We divided our PDAC patients into those who underwent surgery to remove the primary tumor (Clinical stages I and II) and those who presented with an unresectable tumor (Stages III and IV). In the latter group, the O phenotype frequency dropped to 28.7% and the A_1 and B frequencies increased to 44.8% and 14.9%, respectively (Table 3). Thus, the ABO frequency differences that we observed for the whole PDAC group versus the control groups appeared more prominent when analyzing unresected cases only. This suggests an association between ABO phenotype and clinical stage which again could indicate that blood group O subjects diagnosed with pancreatic cancer have the best prognosis. Previous analyses in this regard are discrepant. In a study involving Han Chinese patients, blood group O subjects had better outcome that those with non-O blood group [37]. However, in a German material, although the incidence of blood group O was significantly lower among the investigated cohort of pancreatic cancer patients, no association was seen between ABO status and survival [38]. Intriguingly, when we analyzed unresected cases only, there was a significant association between having blood group O and better survival (Fig. 1D–F). This effect was not seen when the group of resected cases were analyzed according to blood group O status (Fig. 1C). Overall, the data from our material indicate that individuals with blood group O may have somewhat less aggressive PDACs than patients having other blood groups.

The implication of ABO glycosyltransferases in pancreatic cancer is still enigmatic. There is, however, strong evidence that the expression of ABH antigens in the gastrointestinal tract affects the anchoring efficiency of certain pathogenic strains to the mucosa and influences bacterial and/or viral colonization and infectivity [39]. In this context, *H. pylori* infection as risk factor in pancreatic pathologies has been extendedly researched [40], and it was shown that the increased risk of non-O individuals for PDAC became even greater if these subjects were positive for the CagA-negative *H. pylori* strain [41]. *FUT2* secretor status seems to affect the composition of the intestinal microbiota [42] and the susceptibility to certain gastrointestinal infections [43]. Nevertheless, we did not observe a difference in *FUT2* allele distribution or phenotype among pancreatic cancer cases and controls, similar to what a previous report showed [13]. Similarly, there was no association between secretor phenotype and survival (Fig. S1).

Naturally occurring alloantibodies (i.e., isoagglutinins) could also play a role in the association between ABO blood groups and pancreatic cancer. These antibodies provide immunity against pathogens expressing blood group-like antigens on their surface. Anti-A and anti-B isoagglutinins, present in the plasma of blood group O individuals, have been shown to react toward the Tn and T pancarcinoma antigens [44]. Tn and T may be structurally related to A and B blood group antigens, respectively [45] and could possibly render cancer cells immunologically less recognized in blood group A and B individuals.

In conclusion, although the choice of control material involves some particular challenges when studying ABO blood groups as a disease risk factor, our results confirm both a genetic and phenotypic association with pancreatic ductal adenocarcinoma. We found that the increased susceptibility connected with blood group A is likely to be caused by the subtype A_1 only, and we suggest that also clinical stage at the time of diagnosis and survival could be influenced by blood group status. However, the biology behind the fascinating ABO-pancreatic cancer link is poorly understood, and studies are warranted on how ABO glycosyltransferase activity may influence biological aggressiveness of pancreatic neoplastic cells.

Acknowledgements

We are very grateful to all patients who took part in this study, to Jorunn Vadheim for helping with extraction of blood type data from the hospitalized patients and to Paal Henning Borge for assistance with the blood donor samples. This project was funded by a PhD fellowship (grant no. 911831) to K.E.J. from the Western Norway Regional Health Authority (Helse Vest).

Conflict of Interest

None declared.

References

1. Available at: http://www.cancerresearchuk.org/about-cancer/type/pancreatic-cancer/ (accessed 12 April 2017).
2. Esposito, I., B. Konukiewitz, A. M. Schlitter, and G. Kloppel. 2014. Pathology of pancreatic ductal adenocarcinoma: facts, challenges and future developments. World J. Gastroenterol. 20:13833–13841.
3. Bosetti, C., E. Lucenteforte, D. T. Silverman, G. Petersen, P. M. Bracci, B. T. Ji, et al. 2012. Cigarette smoking and pancreatic cancer: an analysis from the International Pancreatic Cancer Case-Control Consortium (Panc4). Annals Oncol. 23:1880–1888.
4. Raimondi, S., A. B. Lowenfels, A. M. Morselli-Labate, P. Maisonneuve, and R. Pezzilli. 2010. Pancreatic cancer in chronic pancreatitis; aetiology, incidence, and early detection. Best Practice Res. Clin. Gastroenterol. 24:349–358.
5. Schulte, A., N. Pandeya, J. Fawcett, L. Fritschi, H. A. Risch, P. M. Webb, et al. 2015. Association between Helicobacter pylori and pancreatic cancer risk: a meta-analysis. Cancer Causes Control 26:1027–1035.
6. Xu, C. X., H. H. Zhu, and Y. M. Zhu. 2014. Diabetes and cancer: associations, mechanisms, and implications for medical practice. World J. Diabetes. 5:372–380.
7. Roberts, N. J., A. L. Norris, G. M. Petersen, M. L. Bondy, R. Brand, S. Gallinger, et al. 2016. Whole genome sequencing defines the genetic heterogeneity of familial pancreatic cancer. Cancer Discov. 6:166–175.
8. Childs, E. J., E. Mocci, D. Campa, P. M. Bracci, S. Gallinger, M. Goggins, et al. 2015. Common variation at 2p13.3, 3q29, 7p13 and 17q25.1 associated with susceptibility to pancreatic cancer. Nat. Genet. 47:911–916.
9. Amundadottir, L., P. Kraft, R. Z. Stolzenberg-Solomon, C. S. Fuchs, G. M. Petersen, A. A. Arslan, et al. 2009. Genome-wide association study identifies variants in the ABO locus associated with susceptibility to pancreatic cancer. Nat. Genet. 41:986–990.
10. Yamamoto, F., H. Clausen, T. White, J. Marken, and S. Hakomori. 1990. Molecular genetic basis of the histo-blood group ABO system. Nature 345(6272):229–233.
11. Clausen, H., E. P. Bennett, and N. Grunnet. 1994. Molecular genetics of ABO histo-blood groups. Transfus. Clin. Biol. 1:79–89.
12. Wolpin, B. M., P. Kraft, M. Gross, K. Helzlsouer, H. B. Bueno-de-Mesquita, E. Steplowski, et al. 2010. Pancreatic cancer risk and ABO blood group alleles: results from the pancreatic cancer cohort consortium. Cancer Res. 70:1015–1023.
13. Wolpin, B. M., P. Kraft, M. Xu, E. Steplowski, M. L. Olsson, A. A. Arslan, et al. 2010. Variant ABO blood group alleles, secretor status, and risk of pancreatic cancer: results from the pancreatic cancer cohort consortium. Cancer Epidemiol. Biomarkers Preven. 19:3140–3149.
14. Yamamoto, F., P. D. McNeill, and S. Hakomori. 1992. Human histo-blood group A2 transferase coded by A2 allele, one of the A subtypes, is characterized by a single base deletion in the coding sequence, which results in an additional domain at the carboxyl terminal. Biochem. Biophys. Res. Commun. 187(1):366–374.
15. Nakao, M., K. Matsuo, S. Hosono, S. Ogata, H. Ito, M. Watanabe, et al. 2011. ABO blood group alleles and the risk of pancreatic cancer in a Japanese population. Cancer Sci. 102:1076–1080.
16. Risch, H. A., L. Lu, J. Wang, W. Zhang, Q. Ni, Y. T. Gao, et al. 2013. ABO blood group and risk of pancreatic cancer: a study in Shanghai and meta-analysis. Am. J. Epidemiol. 177:1326–1337.
17. Xu, H. L., J. R. Cheng, W. Zhang, J. Wang, H. Yu, Q. X. Ni, et al. 2014. Re-evaluation of ABO gene polymorphisms detected in a genome-wide association study and risk of pancreatic ductal adenocarcinoma in a Chinese population. Chin. J. Cancer. 33:68–73.
18. Aird, I., D. R. Lee, and J. A. Roberts. 1960. ABO blood groups and cancer of oesophagus, cancer of pancreas, and pituitary adenoma. Br. Med. J. 1:1163–1166.
19. Newell, G. R., J. E. Gordon, A. P. Monlezun, and J. S. Horwitz. 1974. ABO blood groups and cancer. J. Natl Cancer Inst. 52:1425–1430.
20. Annese, V., M. Minervini, A. Gabbrielli, G. Gambassi, and R. Manna. 1990. ABO blood groups and cancer of the pancreas. Int. J. Pancreatol. 6:81–88.
21. Barrett, J. C., B. Fry, J. Maller, and M. J. Daly. 2005. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21:263–265.
22. Golding, J., K. Northstone, L. L. Miller, G. Davey Smith, M. Pembrey. 2013. Differences between blood donors and a population sample: implications for case-control studies. Int. J. Epidemiol. 42:1145–1156.
23. Bosman, F. T., F. Carneiro, R. H. Hruban, and N. D. Theise, eds. 2010. P. 280 WHO Classification of Tumors of the Digestive system, 4th ed.. International Agency for Research on Cancer, Lyon.
24. Chen, Y., C. Chen, X. Ke, L. Xiong, Y. Shi, J. Li, et al. 2014. Analysis of circulating cholesterol levels as a mediator of an association between ABO blood group and coronary heart disease. Circ. Cardiovasc. Genet. 7:43–48.
25. Jaff, M. S. 2011. Relation between ABO blood groups and Helicobacter pylori infection in symptomatic patients. Clin. Exper. Gastroenterol. 4:221–226.
26. Zhang, H., C. J. Mooney, and M. P. Reilly. 2012. ABO blood groups and cardiovascular diseases. Int. J. Vasc. Med. 2012:641917.

27. Rummel, S. K., and R. E. Ellsworth. 2016. The role of the histoblood ABO group in cancer. Future Sci. OA. 2:FSO107.
28. Pomianowska, E., K. Grzyb, A. Westgaard, O. P. Clausen, and I. P. Gladhaug. 2012. Reclassification of tumour origin in resected periampullary adenocarcinomas reveals underestimation of distal bile duct cancer. Eur. J. Surg. Oncol. 38:1043–1050.
29. Rizzato, C., D. Campa, R. Pezzilli, P. Soucek, W. Greenhalf, G. Capurso, et al. 2013. ABO blood groups and pancreatic cancer risk and survival: results from the PANcreatic Disease ReseArch (PANDoRA) consortium. Oncol. Rep. 29:1637–1644.
30. Seymour, R. M., M. J. Allan, A. Pomiankowski, and K. Gustafsson. 2004. Evolution of the human ABO polymorphism by two complementary selective pressures. Proceed. Royal Soc. B: Biol. Sci. 271:1065–1072.
31. Cserti, C. M., and W. H. Dzik. 2007. The ABO blood group system and Plasmodium falciparum malaria. Blood 110:2250–2258.
32. Kornstad, L. 1997. Frequency of the blood group antigen K and the A1A2BO groups in the Norwegian counties. Gene Geog. 11:37–46.
33. Liumbruno, G. M., and M. Franchini. 2014. Hemostasis, cancer, and ABO blood group: the most recent evidence of association. J. Thromb. Thrombolysis 38:160–166.
34. Pelzer, U., F. Klein, M. Bahra, M. Sinn, B. Dorken, P. Neuhaus, et al. 2013. Blood group determinates incidence for pancreatic cancer in Germany. Front. Physiol. 4:118.
35. Engin, H., C. Bilir, H. Ustun, and A. Gokmen. 2012. ABO blood group and risk of pancreatic cancer in a Turkish population in Western Black Sea region. Asian Pac. J. Cancer Prev. 13:131–133.
36. Vasan, S. K., J. Hwang, K. Rostgaard, O. Nyren, H. Ullum, O. B. Pedersen, et al. 2016. ABO blood group and risk of cancer: a register-based cohort study of 1.6 million blood donors. Cancer Epidemiol. 44:40–43.
37. Ben, Q., K. Wang, Y. Yuan, and Z. Li. 2011. Pancreatic cancer incidence and outcome in relation to ABO blood groups among Han Chinese patients: a case-control study. Int. J. Cancer 128:1179–1186.
38. Rahbari, N. N., U. Bork, U. Hinz, A. Leo, J. Kirchberg, M. Koch, et al. 2012. AB0 blood group and prognosis in patients with pancreatic cancer. BMC Cancer 12:319.
39. Liumbruno, G. M., and M. Franchini. 2013. Beyond immunohaematology: the role of the ABO blood group in human diseases. Blood Transfus. 11(4):491–499.
40. Xiao, M., Y. Wang, and Y. Gao. 2013. Association between Helicobacter pylori infection and pancreatic cancer development: a meta-analysis. PLoS ONE 8:e75559.
41. Risch, H. A., H. Yu, L. Lu, and M. S. Kidd. 2010. ABO blood group, Helicobacter pylori seropositivity, and risk of pancreatic cancer: a case-control study. J. Natl Cancer Inst. 102:502–505.
42. Wacklin, P., H. Makivuokko, N. Alakulppi, J. Nikkila, H. Tenkanen, J. Rabina, et al. 2011. Secretor genotype (FUT2 gene) is strongly associated with the composition of Bifidobacteria in the human intestine. PLoS ONE 6:e20113.
43. Carlsson, B., E. Kindberg, J. Buesa, G. E. Rydell, M. F. Lidon, R. Montava, et al. 2009. The G428A nonsense mutation in FUT2 provides strong but not absolute protection against symptomatic GII.4 Norovirus infection. PLoS ONE 4:e5593.
44. Hofmann, B. T., A. Stehr, T. Dohrmann, C. Gungor, L. Herich, J. Hiller, et al. 2014. ABO blood group IgM isoagglutinins interact with tumor-associated O-glycan structures in pancreatic cancer. Clin. Cancer Res. 20:6117–6126.
45. Ju, T., V. I. Otto, and R. D. Cummings. 2011. The Tn antigen-structural simplicity and biological complexity. Angew. Chem. Int. Ed. Engl. 50:1770–1791.

Supporting Information

Additional supporting information may be found in the online version of this article:

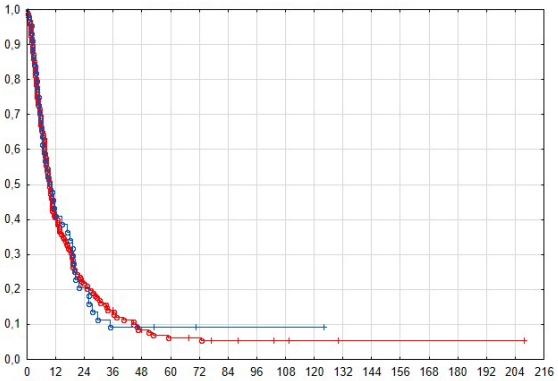
Figure S1. Cumulative proportion survival (Kaplan–Meier) plot for the 195 pancreatic ductal adenocarcinomas according to breakdown by *FUT2* secretor phenotype and resection status.

Table S1. SNPs used for *ABO* and *FUT2* allele genotyping.

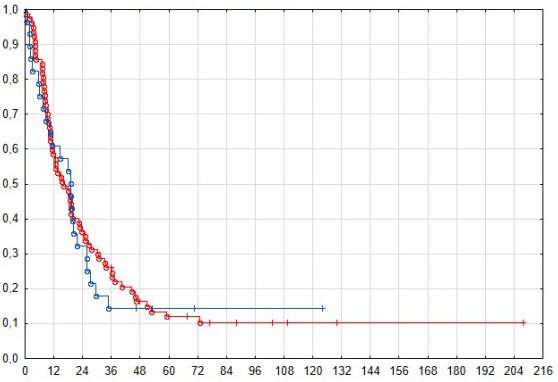
Table S2. *ABO* and *FUT2* genotype frequencies of blood donor controls and pancreatic cancer cases.

Supplementary Figure 1

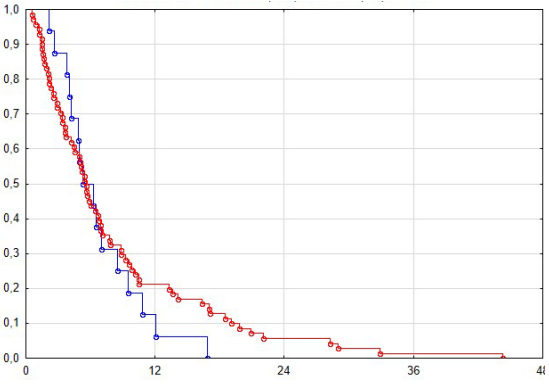
A: All cases - Secretor vs. Non-secretor



B: Resected cases - Secretor vs. Non-secretor



C: Unresected cases - Secretor vs. Non-secretor



Supplementary Table 1. SNPs used for *ABO* and *FUT2* allele genotyping

	ABO SNPs			***FUT2 SNP***
Allele	rs8176704	rs8176746	rs505922	rs601338
A_1	G	C	C	-
A_2	A	C	C	-
B	G	A	C	-
O	G	C	T	-
Se	-	-	-	G
Se^0	-	-	-	A

Se and Se^0 refer to secretor and non-secretor allele, respectively

Supplementary Table 2. *ABO* and *FUT2* genotype frequencies of blood donor controls and pancreatic cancer cases

	Controls (*n* =379)		All cases (*n* =237)		PDAC cases only (*n* =195)	
Genotype	***n***	**%**	***n***	**%**	***n***	**%**
ABO						
A_1A_1	15	4.0	6	2.5	6	3.1
A_1A_2	5	1.3	14	5.9	13	6.7
A_2A_2	4	1.1	0	0.0	0	0.0
A_1O	91	24.0	75	31.6	64	32.8
A_2O	39	10.3	20	8.4	16	8.2
BB	3	0.8	0	0.0	0	0.0
BO	43	11.3	29	12.2	25	12.8
A_1B	10	2.6	5	2.1	4	2.1
A_2B	7	1.8	2	0.8	1	0.5
OO	162	42.7	86	36.3	66	33.8
FUT2						
SeSe	96	25.3	64	27.0	64	27.0
$SeSe^0$	198	52.2	116	48.9	116	48.9
Se^0Se^0	85	22.4	57	24.1	57	24.1

III

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

Khadija El Jellas[1,2,3], Bente B. Johansson[3,4], Karianne Fjeld[3,5], Heike Immervoll[1,6], Man H. Choi[1,2,3], Dag Hoem[7], Mark E. Lowe[8], Dominique Lombardo[9], Pål R. Njølstad[3,4], Anne Dell[10], Eric Mas[9], Stuart M. Haslam[10], Anders Molven[1,2,3] *

Running title: *Carboxyl-ester lipase contains ABO blood group determinants*

From the: [1]Gade Laboratory for Pathology, Department of Clinical Medicine, University of Bergen, Norway; [2]Department of Pathology, Haukeland University Hospital, Bergen, Norway; [3]KG Jebsen Center for Diabetes Research, Department of Clinical Science, University of Bergen, Norway; [4]Department of Pediatrics and Adolescent Medicine, Haukeland University Hospital, Bergen, Norway; [5]Center for Medical Genetics and Molecular Medicine, Haukeland University Hospital, Bergen, Norway; [6]Department of Pathology, Ålesund Hospital, Ålesund, Norway; [7]Department of Gastrointestinal Surgery, Haukeland University Hospital, Bergen, Norway; [8]Department of Pediatrics, Washington University School of Medicine, St. Louis, MO 63110; [9]Aix-Marseille University, INSERM, CRO2, Center for Research in Biological Oncology and Oncopharmacology, Marseille, France; [10]Department of Life Sciences, Imperial College London, South Kensington Campus, London, SW7 2AZ, United Kingdom

*To whom correspondence should be addressed: Gade Laboratory for Pathology, Department of Clinical Medicine, University of Bergen, Jonas Lies vei 87, N-5021 Bergen, Norway; E-mail: anders.molven@uib.no; Tel.: +47-55973169.

Key words: CEL, BSDL, ABO, blood group, pancreatic cancer, Tn, O-glycans

ABSTRACT

Carboxyl-ester lipase (CEL) is a pancreatic fat-digesting enzyme associated with human disease. Rare mutations in the *CEL* gene cause a syndrome of pancreatic exocrine and endocrine dysfunction denoted MODY8, whereas a recombined *CEL* allele increases the risk for chronic pancreatitis. Moreover, CEL has been linked to pancreatic ductal adenocarcinoma (PDAC) through a postulated oncofetal CEL variant termed feto-acinar pancreatic protein (FAPP). The monoclonal antibody mAb16D10 was previously reported to detect a glycotope in the highly O-glycosylated, mucin-like C-terminus of CEL/FAPP. We here assessed the expression of human CEL in malignant pancreatic lesions and cell lines. CEL was not detectably expressed in neoplastic cells, implying that FAPP is unlikely to be a glycoisoform of CEL in pancreatic cancer. Testing of the mAb16D10 antibody in glycan microarrays then demonstrated that it recognized structures containing terminal GalNAc-α1,3(Fuc-α1,2)Gal (blood group A antigen) and also repeated protein sequences containing GalNAc residues linked to Ser/Thr (Tn antigen), findings that were supported by immunostainings of human pancreatic tissue. To examine whether the CEL glycoprotein might be modified by blood group antigens, we used high-sensitivity MALDI-TOF mass spectrometry to characterize the released O-glycan pool of CEL immunoprecipitated from human pancreatic juice. We found that the O-glycome of CEL consisted mainly of core 1/2 structures with a composition depending on *FUT2* and *ABO* gene polymorphisms. Thus, among digestive enzymes secreted by the pancreas, CEL is a glycoprotein

with some unique characteristics, supporting the view that it could serve additional biological functions to its cholesteryl esterase activity in the duodenum.

INTRODUCTION

Carboxyl-ester lipase (CEL; EC 3.1.1.13), also designated bile salt-dependent lipase (BSDL), is one of four major duodenal lipases secreted by pancreatic acinar cells (1). In addition, mammalian breast glands secrete CEL, alternatively called bile salt-stimulated lipase (BSSL), into mother's milk (2). CEL is involved in the hydrolysis and absorption of dietary fat, cholesteryl esters and fat-soluble vitamins. It constitutes a significant fraction of pancreatic juice, around 4% of the total protein content (3).

The eleven exons of the *CEL* gene code for a short signal peptide and a globular N-terminal domain of 538 amino acid residues in which the bile salt-binding site and the catalytic activity of the enzyme reside (4). The last exon contains a variable number of tandem repeats (VNTR) that gives rise to a flexible C-terminus protruding from the globular core (5). The VNTR encodes 11-amino acid segments that are repeated from 3 to 23 times in humans (6), with16 repeats being the most common number (7-9). The variable length of the *CEL* VNTR makes this gene and its protein product highly polymorphic in human populations. In particular, rare mutational events affecting the VNTR region cause an inherited syndrome of diabetes and pancreatic exocrine dysfunction (MODY8) (10), most likely due to protein aggregation and endoplasmatic reticulum (ER) stress resulting from altered repeat sequences (11-13). Moreover, a recombined allele between the VNTR regions of CEL and the neighboring pseudogene *CELP* is associated with a significantly increased risk for chronic pancreatitis (14).

The mature form of CEL is heavily glycosylated. Its globular domain contains a N-linked glycan at the conserved residue Asn210 (15-17), whereas the VNTR domain has a high occupation of *O*-linked glycans on its threonine, and probably serine, residues (15,18), similar to that found in mucinous glycoproteins (19). As a consequence of the VNTR length variability and possible differences in the glycan pool, the molecular weight of the CEL glycoprotein may span from 60 to 140 kDa (7,11).

Glycosylation changes are a hallmark of disease, especially prominent in inflammation and cancer (20). Some studies have indirectly implicated glycovariants of CEL in malignant disease by using antisera against embryonic pancreas extracts to search for acinar cell factors that are absent in the adult pancreas and re-expressed during oncogenesis (21,22). Among many antibodies produced against fetal pancreatic tissue, the monoclonal antibody mAbJ28 detected a concanavalin A-reactive pancreatic glycoprotein with estimated molecular weight of 110 kDa (23). This acinar cell component was given the name feto-acinar pancreatic protein (FAPP) (23,24). FAPP was later reported to have the same amino acid composition as the N-terminal globular domain of CEL (25) and postulated to be a CEL variant expressed in embryogenesis, inflammation and cancer. However, whether oncofetal forms of CEL exist has not been confirmed.

Another antibody that seemingly targets a CEL-related glycotope on FAPP is mAb16D10. This monoclonal antibody was raised against CEL purified from pancreatic juice of a PDAC patient (26). In SOJ-6 pancreatic cancer cells, mAb16D10 has been found to specifically recognize an antigen present on the plasma membrane (27), to inhibit the growth of tumor xenografts (27), and to induce cell death by the p53/caspase-dependent apoptotic pathway (26). In human pancreatic tissue sections, mAb16D10 was reported to discriminate pancreatic cancer from non-neoplastic tissue and other cancers (28).

We sought to confirm and further characterize the postulated onco-glycoforms of CEL by 1) evaluating human CEL expression in PDAC surgical

specimens; 2) determining the identity of the 16D10 glycotope by means of glycan arrays; and 3) identifying the O-glycan structures at the CEL mucinous tail by means of high-sensitivity MALDI-TOF mass spectrometry. However, we could not detect CEL expression in neoplastic cells, and found that mAb16D10 did not target an epitope specific for CEL or for pancreatic cancer tissue. Instead, the antibody showed a high affinity for blood group A antigens, which led to the finding that ABO blood group determinants are present on the C-terminus of CEL.

Results

CEL expression in normal pancreas, PDAC and pancreatic cancer cell lines

Initially, we performed fluorescent *in situ* hybridization and immunohistochemistry for simultaneous detection of CEL mRNA and protein, respectively, in morphologically normal pancreas (Fig. 1). Our staining method did not detect CEL transcripts or protein in ductal cells or in islets of Langerhans (Fig. 1A-B). The acinar cells were strongly positive, as expected, with CEL transcripts located basally and the protein accumulating apically in the zymogen granules (Fig. 1C). Having established the specificity of our method, we evaluated CEL mRNA and protein expression in adjacent sections from PDAC tumors by chromogenic staining (Fig. 2). CEL was detectable in preserved and atrophic parenchyma (Fig. 2A, B, E, F). In contrast, the epithelial lining of acinar-to-ductal metaplasia (ADM), pancreatic intraepithelial neoplasia (PanIN) and malignant ducts was consistently negative (Fig. 2B-D, F-H).

Next, RT-qPCR was performed on selected areas of unstained pancreatic FFPE sections, which were scraped off the glass slide after comparison with H&E-stained parallel sections. Neoplastic regions from nine different PDAC patients were compared with morphologically normal regions from four patients having non-neoplastic disease. Very low *CEL* mRNA levels were detected in the neoplastic areas compared with levels in normal pancreatic parenchyma (Fig. 2I). In addition, the commonly used pancreatic cancer cell lines MIAPaCa-2, PANC-1 and BxPC-3 were tested for CEL expression (Fig. 3). MIAPaCa-2 and PANC-1 exhibited very low mRNA levels (comparable to the basal levels in HeLa cells). BxPC-3 cells had a CEL mRNA level comparable to that of native HEK293 cells. Immunoblots stained with an anti-CEL antibody detected a strong protein band in the control HEK293_CEL line and a weak band in BxPC-3 cells, but no signal in the other cell lines (Fig. 3B).

Characterization of the 16D10 glycotope

The mAb16D10 antibody had been produced against purified CEL (26) and was shown to react well with pancreatic cancer tissue sections (28). Since we did not detect CEL mRNA or protein in neoplastic cells of PDAC tumors, we next sought an explanation for the reactivity of the mAb16D10 antibody. To this end, an aliquot of mAb16D10 was analyzed on a glycan microarray (see Experimental Procedures). By comparing bound and unbound structures, we concluded that mAb16D10 had a strong reactivity towards the structural motif GalNAc-α1,3(Fuc-α1,2)Gal, which corresponds to the blood group A antigen (Fig. 4). To verify this result, the antibody was used to stain normal pancreatic parenchyma from subjects of blood group A and O. There was strong positivity towards normal acinar cells from individuals with blood group A, but not in blood group O pancreatic tissue (Fig. 5A, C). Similarly, mAb16D10 specifically stained both erythrocytes and endothelium of blood group A specimens (Fig. 5B, D).

We then examined the reactivity of mAb16D10 on pancreatic tissue sections from PDAC cases of different ABO phenotypes. Blood group A cases exhibited a virtually identical staining pattern when parallel sections were incubated with either anti-A antibody or mAb16D10 (Fig. 6A-H). There was strong reactivity with acinar cells, ducts, blood

vessels and red blood cells in areas of preserved morphology (Fig. 6A, E), as well as with epithelial cells of ADM, PanIN and malignant ducts (Fig. 6B-D, F-H). However, mAb16D10 also stained sections from PDAC patients with blood group O. Weak positivity, in a dot-like apical pattern, was observed in normal-looking acinar cells close to the tumor area (Fig. 6I). When ADM was present nearby, expression of the mAb16D10 glycotope seemed to increase in the small intralobular ducts connecting the acini (Fig. 6J). Relatively strong positivity was observed in PanINs and malignant ducts (Fig. 6K, L).

Based on the observed staining pattern in blood group O neoplastic tissue (Fig. 6L), it was clear that mAb16D10 also recognized some non-A epitope(s). Thus, to further characterize its specificity, mAb16D10 was tested on a second microarray, consisting of peptides with a varying number of GalNAc residues conjugated to Ser or Thr residues (Fig. 7). Among GalNAc-containing peptides the intensity of the binding increased as the number of GalNAc-α-Ser residues increased, indicating that mAb16D10 binds strongly to repeating poly-Ser-α-GalNAc residues (Fig. 7A). Blood group A antigen bears similarity to the cancer-associated Tn antigen as they both carry a terminal GalNAc-α residue (Fig. 7B). The glycan array experiment was therefore repeated using an anti-Tn antibody. A certain overlap between the reactivity of this antibody and that of mAb16D10 was indeed observed (Fig. 7A). We therefore assumed that mAb16D10 could recognize the Tn antigen in cancerous tissue, and parallel staining of PDAC sections for these two antigens showed very similar patterns in blood group O specimens (Fig. 7C). Moreover, we also found that mAb16D10 reacted with gastric and breast cancer tissue sections from blood group O individuals (Suppl. Fig. 1).

mAb16D10 reactivity in pancreatic juice

We were now left with the conundrum that mAb16D10 recognized the blood group A antigen very well, despite being produced against purified CEL from pancreatic juice. The most straightforward explanation would be that CEL, being a heavily O-glycosylated protein (29), can be decorated with blood group A antigens, which then had stimulated mAb16D10 production. To investigate this hypothesis, we performed immunoprecipitation of CEL directly from pancreatic juice collected from pancreatic cancer patients of either blood type A or O. All four patients had functioning α1,2-fucosyltransferase (as determined by their *FUT2* genotype) and were therefore able to synthesize the H antigen of the ABO blood group system on secreted glycoproteins. Such subjects are referred to as secretors and comprise 80% of the Caucasian population (30,31).

The immunoprecipitated complexes were subjected to SDS-PAGE and immunoblotted with anti-CEL, mAb16D10, anti-A and anti-Tn antibodies. We compared two experiments, one where the immunoprecipitates were washed with PBS only (Fig. 8, upper panel) and another where more stringent washings with RIPA buffer were performed, aiming at dissociating possible complexes that CEL might form with other proteins (Fig. 8, lower panel). The anti-CEL antibody detected the protein in all samples as a band of around 110 kDa. After staining with mAb16D10 or anti-A antibody a band of similar size were detected in blood group A samples but not in samples from blood group O individuals (Fig. 8). This observation strongly supported that CEL could contain terminal blood group A antigens. Anti-Tn staining was negative with regard to detecting a band that corresponded to CEL in molecular weight. A high-molecular weight band at ~170 kDa was seen for all antibodies except anti-CEL, but only when mild washing was applied (Fig. 8, upper panel). This might indicate that a CEL-binding protein is present in pancreatic juice and that this protein contains structures with terminal α-GalNAc residues that can be recognized by all three antibodies.

Characterization of the O-glycome of CEL

Finally, we attempted to detect blood group A epitopes directly on CEL after immunoprecipitation from pancreatic juice as described above. The protein band at 110 kDa was excised. As quality control, a small fraction of this band was used for protein identification by MS after trypsin-cleaved peptides/glycopeptides had been extracted. CEL was the major protein species identified (not shown). Then, mass-spectrometric profiling of O-glycans in the excised band was performed using MALDI-TOF-TOF after beta-reductive elimination of glycans.

The CEL O-glycome was analyzed in three samples (Fig. 9): two *FUT2*-positive cases (secretors) of blood group O and A, respectively, and one *FUT2*-negative case (non-secretor) of blood group O. The latter subject contained a mutation in the *FUT2* gene that leads to a non-functional α1,2-fucosyltransferase meaning that there are no H antigens on secreted glycoproteins to be further converted into A and B antigens. As described below, the identified O-glycan pool was in all three cases consistent with *ABO* and *FUT2* genotyping of genomic DNA extracted from the patients' blood samples (31).

The blood group O/secretor sample (Fig. 9A) was dominated by the molecular ion m/z 708 which is consistent with a blood group O antigen (or H antigen) core 1 structure. Core 2 structures containing the H antigen were also seen with a single (m/z 1159, 1332 and 1519) or two or more LacNAc units (m/z 1607, 1781, 1968, 2056 and 2230) with or without sialic acid capping. In the blood group A/secretor sample (Fig. 9B), the two highest peaks corresponded to the molecular ion at m/z 953, which is consistent with a core 1 blood group A structure, and its precursor at m/z 708 (H antigen). A higher molecular mass structure containing one blood group A antigen on each antenna was detected at m/z 1822. LacNAc extensions of two units or more were not seen in this sample.

When CEL was isolated from the pancreatic juice of a non-secretor patient (Fig. 9C), the glycan pool lacked the molecular ion m/z 708. As a consequence, the Gal residue of the precursor Gal-β1,3-GalNAc (m/z 534, not detected) could only undergo sialylation and therefore m/z 895 (mono-sialylated) and 1257 (di-sialylated) were the two most abundant structures. Core 2 structures were also seen at m/z 983, 1344, 1433, 1519, 1607, 1794, 1882 and 1968. Fucoses were only part of core 2 structures (m/z 1519, 1607 and 1968) attached most likely to a GlcNAc residue.

Structural assignment was facilitated by all molecular ions from the MS spectra being examined by MS/MS analysis. In Fig. 10, fragmentation patterns of the selected ions m/z 1781, 1822 and 1794 (from spectra in Figs. 9A, 9B and 9C, respectively) are shown. Moreover, three peaks observed in the blood group O secretor and non-secretor samples had identical m/z values (1519, 1607 and 1968; Fig. 9A, C). The fragmentation pattern showed that they corresponded to glycans of the same composition but with different architecture due to the position of the fucose residue. Such differences are exemplified in Suppl. Fig. 2 where two diverging fragmentation patterns of m/z 1519 are showed.

Discussion

CEL has been linked to the inherited pancreatic disease MODY8 (10) and to idiopathic chronic pancreatitis (14). However, whether CEL has a role in pancreatic malignancies is still an open question. No genome-wide association studies of pancreatic cancer have so far identified SNPs in the *CEL* gene as risk factor (32,33). Neither the *CEL-HYB* variant (the allele predisposing for chronic pancreatitis) nor specific *CEL* VNTR lengths were enriched among pancreatic cancer cases compared to controls (8,34). Nevertheless, the identification and characterization of FAPP, a postulated oncofetal variant of CEL, has raised the possibility that glycoisoforms of CEL

could serve as diagnostic markers or even targets for treatment in pancreatic cancer (26-28,35).

In the current study, by using complementary techniques, we found that CEL was not detectably expressed in either neoplastic cells of PDAC tissue sections or in the pancreatic cell lines examined. The only exception was BxPC-3 cells, a cell line derived from a moderately well-differentiated adenocarcinoma with wild type *KRAS* gene (36,37), and therefore not fully representative for PDAC tumor cells. We observed that CEL expression disappeared when acinar-to-ductal metaplasia was present, suggesting that the protein is a marker of pancreatic acinar cells only when these remain polarized and in acini-like arrangement. Our results are consistent with those of Reuss et al. (38). When they examined 25 ductal or mucinous adenocarcinomas from the pancreas, all cases were negative for CEL mRNA by *in situ* hybridization. Notably, a single case of pancreatic acinar cell carcinoma showed slightly positive expression (38). Our results support the conclusion that CEL is unlikely to correspond to the oncofetal marker FAPP.

Since acinar pancreatic markers expressed only during embryogenesis and oncogenesis have been identified and we could not detect CEL in neoplastic tissue or cells, we wondered which antigen the anti-CEL antibody mAb16D10 did detect in pancreatic neoplastic cells (21,22). Knowing that mAb16D10 reacts with the highly O-glycosylated C-terminal of CEL (27), we tested the antibody against an array of more than six hundred different mammalian glycan structures. Glycans containing the blood group A antigen consistently showed the highest binding (Fig. 4). The presence of blood group antigens in the O-glycome of CEL was then verified directly on CEL protein isolated from pancreatic juice (Figs. 8, 9). The samples investigated, as well as the juice sample used for mAb16D10 production, all stemmed from patients operated for pancreatic cancer. Aberrant glycosylation is a common characteristic for malignant transformation. It will therefore be necessary to analyze pancreatic juice from healthy subjects to determine whether the same glycan structures are found and, in particular, whether A antigens attached to CEL are a common feature of blood group A subjects. Similarly, based on the finding of H and A antigens attached to CEL, we predict that the protein when isolated from blood group B individuals will be modified by blood group B antigens.

It is perhaps not surprising that CEL's mucin-like C-terminal domain contains blood group antigens. Mucinous O-glycans consist mostly of core 1-4 structures further elongated with lactosamine chains and terminated by fucose and sialic acid (39). Therefore, the presence of ABH or Lewis-type blood group antigens are not unusual on such proteins (40-42). In studies of BSSL, the milk counterpart of CEL, the threonine residues flanking the consensus sequence PVPP of the mucinous domain were proposed to carry O-glycans (18), and terminal Lewis antigens Le^a, Le^b and Le^x were suggested to be present on this protein region (18,43). Nevertheless, as far as we know, no previous study has detected blood group antigens directly on pancreatic CEL using high-sensitivity mass spectrometry.

The antibody mAb16D10 also showed reactivity against Tn motifs. Truncation of the O-glycans in CEL and exposure of the first GalNAc residue (GalNAcα1-O-Ser/Thr) might have explained the reactivity of mAb16D10 towards both A and Tn antigens. However, in the MS analysis of purified CEL protein we did not detect any peak at m/z 330 corresponding to Tn. A peak at m/z 691 initially suggested a sialylated Tn structure (= STn antigen), but this was not consistent with the predicted glycan fragmentation pattern obtained when MS/MS was performed. Moreover, in blots of immunoprecipitates, no band matching CEL in size was seen with the anti-Tn antibody, although this antigen might be present in a higher molecular weight protein of ~170 kDa protein that co-

precipitated with CEL (Fig. 8, upper lane). Taken together, we find that mAb16D10 has reactivity towards Tn, but that CEL purified from pancreatic juice of PDAC patients does not seem to carry this antigen in detectable amounts. We speculate that the staining of mAb16D10 observed in gastric and breast cancer (Suppl. Fig. 1) could be due to Tn positivity and therefore consistent with the pancarcinoma nature of this antigen (44,45).

Our finding that the C-terminal mucin-like domain of CEL can be modified by blood group antigens should stimulate discussions about this heavily glycosylated region and its significance for CEL function. Neither chicken nor fish CEL contain the VNTR region (46), and truncated human CEL variants (devoid of the VNTR) are enzymatically active (14,47,48). Thus, the C-terminal domain is clearly dispensable for the catalytic activity of the protein although a role in substrate specificity cannot be ruled out. It is conceivable that O-glycosylation serves to increase the stability of CEL as the C-terminal region bears amino acid similarity to protein sequences that can be a signal for proteolytic degradation (49,50).

Another possibility is that CEL, through its C-terminal region, contributes to the mucosal barrier in the gastrointestinal tract. A main role of intestinal mucins is to provide lubrication and create a barrier that will prevent infection by the large number of microorganisms that populate the gut (51). This function is exerted through the ability of O-glycans to retain water and form mucus gels (52). Bearing in mind that CEL comprises up to 4% of the pancreatic acinar proteins (3), the molecule could to some extent add to the gastrointestinal mucosal barrier, with the potential of interacting with intestinal microorganisms (41,53). Even the presence of blood group antigens on CEL might be understood in this perspective. In the analyzed juice samples, the identified glycan pools of CEL were fully consistent with the patients' *ABO* and *FUT2* genotypes, with the H antigen core 1 structure being one of the most abundant motifs detected in the secretor samples. ABO blood groups are genetically determined factors that modulate the interaction between host and intestinal microbiota as well as the flora composition itself, by providing carbohydrate antigens which act as anchoring sites for commensal bacteria (53-55).

CEL is, in some respects, unique among digestive enzymes of the pancreas. The protein is expressed in several non-acinar cell types and has been linked to non-pancreatic pathogenic processes (56-58). The *CEL* gene is exceptionally polymorphic involving both copy number variants, VNTR length variations and single-base VNTR insertions (4). Moreover, the repeat region has expanded during mammalian evolution into an O-glycosylated mucin-like domain (46), which our current study shows is modified by blood group antigens. Although it might very well be a pure coincidence, we also note that the *ABO* and *CEL* genes are positioned in close vicinity of each other on the chromosome 9q terminal end (59). This opens possibilities of interactions at the gene structural level such as in the regulation of expression.

Intriguingly, ABO blood group status convincingly has been shown to associate with varying risk for developing pancreatic cancer (31,60). The mechanistic basis for this association remains to be explained. Whether variation in the glycosylation patterns of secreted gastrointestinal proteins such as CEL plays a role in the susceptibility to pancreatic cancer is therefore an intriguing matter that should be explored further.

Experimental Procedures

Antibodies — Three different anti-CEL antibodies were used: a polyclonal rabbit antiserum (against peptide sequence 299-427 of the CEL globular domain) for immunostainings (Sigma-Aldrich, Cat. HPA052701), a polyclonal rabbit antiserum for immunoblotting (generated against the truncated

CEL form pV562X; recently published in (13)), and the mouse monoclonal antibody As20.1 for immunoprecipitation (raised against the CEL globular domain; generous gift from Prof. Olle Hernell, Umeå University, Sweden). The monoclonal antibody mAb16D10 has been described previously (26-28). Anti-Tn antibody was a generous gift from Prof. Ulla Mandel (Copenhagen University, Denmark). Mouse monoclonal anti-A antigen (sc-69951) and goat polyclonal anti-GAPDH antibodies (sc-20357) were from Santa Cruz. Secondary antibodies were horseradish peroxidase (HRP)-donkey anti-mouse (Invitrogen, Cat. 626520), HRP-donkey anti-goat (Novusbio, Cat. NB7379), FITC-donkey anti-rabbit (Jackson ImmunoResearch, Cat. 711-095-152) and HRP-goat anti-rabbit (Invitrogen, Cat. 656120), Alexa 488 anti-mouse (Invitrogen, Cat. A-10680). MACH3 anti-mouse (Cat. M3M530H) or anti-rabbit HRP-conjugated polymer (Cat. M3R531H) were from Biocare Medical.

Human pancreatic cancer samples ─ Formalin-fixed, paraffin-embedded (FFPE) pancreatic tissue blocks from either blood group A or blood group O subjects were selected from a cohort of quality-controlled PDAC cases that had been genotyped to determine *ABO* and *FUT2* status, as described earlier (31). Pancreatic juice was obtained from patients undergoing Whipple's procedure for resection of pancreatic head tumors. After transection of the pancreas, the distal, dilated pancreatic duct was cannulated, and a sample of juice was suctioned out and immediately stored at -80 °C. COmplete protease inhibitor cocktail (Roche, Cat. 11697498001) was added when samples were thawed for analysis. Juice samples for the study were selected based on the patients' ABO and secretor phenotype. The patients had consented to the study, which was approved by the Regional Ethical Committee of Western Norway and performed according to the Helsinki Declaration.

Cell cultures ─ Human pancreatic cancer cell lines (BxPC-3, MIA PaCa-2, PANC-1) and HeLa cells (CCL-2) were obtained from American Type Culture Collection. HEK293 cells (Clontech Laboratories) were stably transfected with pcDNA3.1/V5-His plasmids containing either wild type CEL or no insert (11,12). PANC-1, Hela and HEK293 cells were cultured in Dulbecco's modified Eagle's medium (DMEM; Sigma-Aldrich) with 500 µg Geneticin (G-418; Invitrogen), 10% fetal bovine serum (FBS; Invitrogen) and 100 units/ml penicillin/streptomycin. BxPC-3 were maintained in Roswell Park Memorial Institute (RPMI; Sigma-Aldrich) 1640 medium supplemented with 10% FBS and 100 units/ml penicillin/streptomycin. MIA PaCa-2 cells were cultured in DMEM supplemented with 10% FBS, 2.5% horse serum (ATCC, Cat. 30-2040) and 100 units/ml penicillin/streptomycin. All cells were cultured in a standard humidified incubator at 37 °C in a 5% CO_2 atmosphere.

Immunohistochemistry ─ FFPE pancreatic tissue sections (3-5 µm) were placed onto SuperFrost Plus Adhesion slides (Menzel Gläser Laboratories), dried overnight at 56°C, deparaffinized in xylene, gradually rehydrated with decreasing concentrations of ethanol, and washed with distilled water and PBS containing Tween 0.05% (v/v). The slides were then incubated in Tris-EDTA buffer (pH 9) in a pressurized heating chamber at 120°C for 1 min. Next, the slides were cooled down under running tap water and incubated at 4°C overnight with the antibodies mAb16D10 (1:100 dilution), anti-A (1:2000) and anti-CEL (Sigma-Aldrich; 1:100), or for 1 h at room temperature with anti-Tn (undiluted supernatant). Blocking of endogenous peroxidase activity was done by incubating the sections in an aqueous solution of 3% hydrogen peroxide (v/v) for 5 min at room temperature. Primary antibody detection was performed with MACH3 anti-mouse or anti-rabbit HRP-conjugated polymers (Biocare Medical) using two incubations of 20 min with vigorous washing after each step on a rocking platform. Staining was visualized by developing

with 3,3′-diaminobenzidine (DAB) as substrate. The sections were counterstained with Mayer's haematoxylin (Dako) for 1 min and then dehydrated in alcohol, cleared in xylene, and eventually mounted with permanent mounting media.

In situ hybridization ─ Using a custom-made *CEL*-specific probe covering exons 2–7, the RNAscope 2.0HD assay was employed according to the supplier's instructions (Advanced Cell Diagnostics). Briefly, after deparaffinization and rehydration, tissue sections were sequentially treated with endogenous hydrogen peroxidase block solution for 10 minutes at room temperature followed by an immersion in a 100°C-preheated retrieval solution for 15 min and subjected to protease digestion at 40°C for 30 min, rinsing with distilled water after each step. The sections were then hybridized with the *CEL* probe at 40°C for 2 hours in a HybEZ oven (Advanced Cell Diagnostics). After wash buffer steps, signal detection and amplification were performed using the RNAscope 2.0HD detection reagents (Advanced Cell Diagnostics, Cat. 322310), starting with the application of the signal enhancer solutions, then washing vigorously between each step and eventually adding DAB as substrate for HRP. The sections were counterstained with Mayer's haematoxylin for 1 min and then dehydrated in alcohol, cleared in xylene, and eventually mounted with permanent mounting media. Images were acquired at various magnifications using a MC170HD camera attached to a DM2000LED microscope (Leica) and processed using LAS V4.8 software.

When *in situ* hybridization was combined with immunofluorescence, the RNAscope 2.0HD detection system with alkaline phosphatase (Advanced Cell Diagnostics, Cat. 310036) was first used in conjugation with the Vector Red fluorescent substrate (Vector Laboratories, Cat. SK-5100). The sections were then incubated with anti-CEL antibody (Sigma-Aldrich; 1:100) overnight, and thereafter with secondary antibody (donkey anti-rabbit IgG-FITC; 1:50) for 1 h. All washing steps were performed using TBS containing Tween 0.05% (w/v) when alkaline phosphatase was used instead of HRP. For mounting, Gold Antifade Solution with DAPI (Invitrogen) was used. Images were collected using a SP5 AOBS confocal microscope (Leica Microsystems) with 40x/1.4 NA and 63x/1.4 NA HCX Plan-Apochromat oil immersion objectives, ~1.2 airy unit pinhole aperture, and appropriate filter combinations. Images were acquired with 405 Diode and Argon ion/Argon Krypton lasers (Leica). The obtained images were merged and processed using Photoshop CC and Adobe illustrator CC (Adobe Systems).

RNA isolation and real-time quantitative PCR (RT-qPCR) ─ FFPE tissue sections (3-5 sections, 10 µm each) from the tumor, corresponding to either neoplastic areas or morphologically preserved pancreatic parenchyma, were scraped off from the glasses after comparison with parallel sections stained with haematoxylin and eosin (H&E). After deparaffinization, RNA was isolated using the RNeasy FFPE Kit (Qiagen, Cat. 73504) or the RNeasy Micro kit (Qiagen, Cat. 74004), according to the manufacturer's instructions. For cDNA synthesis, reverse transcription was performed on 100 ng RNA utilizing the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems/Thermo Fisher Scientific). Specific predesigned PCR primers/probe sets were purchased from Applied Biosystems (*CEL*: Hs 01068709_m1; *QARS*: Hs00192530_m1 and *GAPDH*: Hs02758991_g1 as normalizing control). RT-qPCR was performed using TaqMan Universal PCR Master Mix (Applied Biosystems, Cat. 4440040) following the manufacturer´s instructions. The PCR conditions were 40 cycles of denaturation at 95°C for 15 sec followed by combined annealing/extension at 60°C for 1 min. Fluorescent intensity was measured, and results were processed with SDS software, version 2.2 (Applied Biosystems).

Cell lysis and western blot analyses ─ Pellets from cultured cells were washed with ice-cold PBS, treated with ice-cold RIPA lysis buffer (Thermo Fisher Scientific, Cat. 89900) and centrifuged for 30 min at maximum speed at 4 degrees. Fifteen µg of total protein was mixed with sample buffer (Life Technologies, Cat. NP0007) and reducing agent (Invitrogen, Cat. NP009), boiled at 90°C for 5 min and resolved in NuPAGE Novex 7% Tris-acetate gels. Separated proteins were transferred to PVDF membrane following standard protocols, then immunoblotted with polyclonal rabbit anti-CEL antibody (13) (1:10000). Staining with anti-GAPDH antibody (1:500) was used as loading control. Blots were developed using Amersham ECL Prime Western Blotting Detection Reagent (GE Healthcare, Cat. RPN2232) and images were detected on the Syngene G:Box XR5 chemiluminescent imaging system (ISS).

Glycan arrays ─ An aliquot of the monoclonal antibody mAb16D10 was analyzed by glycan microarray technology at the Consortium for Functional Glycomics at Emory University, Atlanta, GA. The antibody was incubated at a concentration of 50 µg/ml on a Mammalian Printed Glycan Array Version 5.2 (609 structures, available on www.functionalglycomics.org/glycomics/publicdata/primaryscreen.jsp) or on a glycopeptide array (56 compounds, Suppl. Table 1), and probed for binding using fluorescent Alexa 488 anti-mouse antibody (Invitrogen, Cat. A-10680). On the latter array, also an anti-Tn antibody was probed. The arrays were processed using published methods (61). Binding experiments were done in replicates of six, removing the highest and lowest values from each replicate set and calculating the mean binding efficiency from the remaining four values.

Immunoprecipitation ─ CEL glycoprotein was purified from pancreatic juice with the Pierce co-immunoprecipitation kit (Thermo Fisher Scientific, Cat. 26149) as follows: 50-400 µl pancreatic juice (protein concentration between 0.8 and 5.4 µg per µl) were incubated in a spin-column containing immobilized antibody to the agarose-resin (monoclonal As20.1 antibody for CEL or unspecific mouse immunoglobulin for negative control). The sealed columns were incubated overnight at 4°C on a rotating wheel. Washing steps were performed using either PBS or RIPA lysis buffer, and elution was achieved using 50 mM Tris-HCl (pH 8) with 2.5% w/v SDS. Five µl eluate was separated by standard SDS polyacrylamide gel electrophoresis (SDS-PAGE) and immunoblotted with polyclonal rabbit anti-CEL antibody (13) (1:10000), mAb16D10 (1:100), anti-A (1:500) or anti-Tn (1:5) antibody. All incubations were done overnight at 4°C on a rocking platform. For mass spectrometry analysis, the rest of the eluted protein was subjected to SDS-PAGE electrophoresis and the resulting gel stained with SimplyBlue SafeStain solution (Thermo Fisher Scientific, Cat. LC6065) and washed thoroughly with ultrapure water prior to in-gel digestion.

In-gel digestion of purified glycoprotein ─ Excised bands corresponding to CEL were first destained in a 50% acetonitrile (MeCN) (v/v) solution in 50 mM ammonium bicarbonate for 10 min, followed by incubation in 10 mM dithiothreitol solution for 30 min at 56 °C. The gel pieces were desiccated, followed by incubation with a 55 mM solution of iodoacetic acid for 30 min in darkness at room temperature. The gel pieces were desiccated again prior to incubation with 1 µg of porcine trypsin (Sigma-Aldrich, Cat. S8045) in 50 mM ammonium bicarbonate, pH 8.5 (adjusted with ammonia), overnight at 37 °C. Following extraction of the tryptic glycopeptides from the gel in 0.1% trifluoroacetic acid/MeCN (1:2 v/v), the volume was reduced by a vacuum concentrator and the samples were freeze-dried by complete lyophilizing before proceeding to reductive elimination. All solvents used for glycan purification and analysis were HPLC-grade or higher.

Release of O-glycans by reductive elimination ─ O-glycans were released from the mixture of peptides/glycopeptides by reductive elimination in 400 µL of potassium borohydride (55 mg/mL in 0.1 M potassium hydroxide) at 45 °C for 16 h. After the reaction had been terminated by dropwise addition of glacial acetic acid, the mixture was passed through a Dowex-beads (Sigma-Aldrich, Cat. 217492) chromatography column and dried under a stream of nitrogen with repeated addition of 10% methanolic acetic acid to remove borates.

Permethylation of dried O-glycans ─ Two mL of a slurry composed of freshly crushed sodium hydroxide pellets and DMSO was added to the sample, followed by 600 µL of methyl iodide. The mixture was vigorously mixed on an automatic shaker for 45 min at room temperature. One mL of water was added to terminate the reaction, followed by another 1 mL of chloroform to recover the permethylated glycans. The aqueous phase was washed three times to remove impurities, and the chloroform layer was then dried under a nitrogen stream. Permethylated O-glycans were purified using a Sep-Pak Classic C18 cartridge (Waters). The cartridge was conditioned successively with methanol, water, MeCN and water. The samples were dissolved in 1:1 (v/v) methanol-water, loaded onto the cartridge, washed with water and 15% (v/v) aqueous MeCN solution, and then eluted using a 35% (v/v) aqueous MeCN solution. The organic solvent was removed on a vacuum concentrator and samples were lyophilized overnight prior to mass spectrometry (MS) analyses.

MALDI-TOF mass spectrometry ─ MALDI-TOF-MS data on permethylated samples were acquired in the reflectron positive-ion mode using a 4800 MALDI-TOF/TOF mass spectrometer (Applied Biosystems). The instrument was calibrated externally using the Calmix 4700 calibration standard, samples were dissolved in 10 µL of methanol, and 1 µL was mixed at a 1:1 ratio (v/v) with 20 mg/mL 3,4-diaminobenzophenone (DMBP) in 75% (v/v) MeCN in water as matrix. The samples were then spotted onto a 384-well sample plate and dried at room temperature. Data were acquired using 4000 Series Explorer instrument control software and were processed using the software Data Explorer (Version 4.9, Applied Biosystems). MS spectra were assigned and annotated with the help of the GlycoWorkbench software (62).

Statistical analyses ─ Analyses were performed using GraphPad prism 5.03 for Windows (GraphPad Software, San Diego California, USA; www.graphpad.com).

Acknowledgments:
We would like to acknowledge Dong Li Lu for MALDI-TOF-MS and MS/MS data acquisition. This project was funded by a PhD fellowship (911831) and grant (912057) from the Western Norway Regional Health Authority (Helse Vest) and by grants from the Gade Legacy and the Research Council of Norway (FRIMEDBIO program). We acknowledge the participation of the Protein-Glycan Interaction Resource of the Consortium for Functional Glycomics (grant R24 GM098791) for the glycan array experiments. A. D. and S. M. H. were supported by grant BB/K016164/1 from the Biotechnology and Biological Sciences Research Council.

Conflict of interest:
The authors declare that they have no conflicts of interest with the contents of this article

References

1. Whitcomb, D. C., and Lowe, M. E. (2007) Human Pancreatic Digestive Enzymes. *Digestive Diseases and Sciences* **52**, 1-17
2. Hernell, O., and Blackberg, L. (1994) Human milk bile salt-stimulated lipase: functional and molecular aspects. *The Journal of pediatrics* **125**, S56-61
3. Lombardo, D., Guy, O., and Figarella, C. (1978) Purification and characterization of a carboxyl ester hydrolase from human pancreatic juice. *Biochimica et biophysica acta* **527**, 142-149
4. Johansson, B. B., Fjeld, K., El Jellas, K., Gravdal, A., Dalva, M., Tjora, E., Raeder, H., Kulkarni, R. N., Johansson, S., Njolstad, P. R., and Molven, A. (2017) The role of the carboxyl ester lipase (CEL) gene in pancreatic disease. *Pancreatology* **18**(1): 12-19.
5. Chen, J. C., Miercke, L. J., Krucinski, J., Starr, J. R., Saenz, G., Wang, X., Spilburg, C. A., Lange, L. G., Ellsworth, J. L., and Stroud, R. M. (1998) Structure of bovine pancreatic cholesterol esterase at 1.6 A: novel structural features involved in lipase activation. *Biochemistry* **37**, 5107-5117
6. Torsvik, J., Johansson, S., Johansen, A., Ek, J., Minton, J., Raeder, H., Ellard, S., Hattersley, A., Pedersen, O., Hansen, T., Molven, A., and Njolstad, P. R. (2010) Mutations in the VNTR of the carboxyl-ester lipase gene (CEL) are a rare cause of monogenic diabetes. *Human genetics* **127**, 55-64
7. Lindquist, S., Blackberg, L., and Hernell, O. (2002) Human bile salt-stimulated lipase has a high frequency of size variation due to a hypervariable region in exon 11. *European journal of biochemistry* **269**, 759-767
8. Dalva, M., El Jellas, K., Steine, S. J., Johansson, B. B., Ringdal, M., Torsvik, J., Immervoll, H., Hoem, D., Laemmerhirt, F., Simon, P., Lerch, M. M., Johansson, S., Njolstad, P. R., Weiss, F. U., Fjeld, K., and Molven, A. (2017) Copy number variants and VNTR length polymorphisms of the carboxyl-ester lipase (CEL) gene as risk factors in pancreatic cancer. *Pancreatology* **17**, 83-88
9. Ragvin, A., Fjeld, K., Weiss, F. U., Torsvik, J., Aghdassi, A., Mayerle, J., Simon, P., Njolstad, P. R., Lerch, M. M., Johansson, S., and Molven, A. (2013) The number of tandem repeats in the carboxyl-ester lipase (CEL) gene as a risk factor in alcoholic and idiopathic chronic pancreatitis. *Pancreatology* **13**, 29-32
10. Raeder, H., Johansson, S., Holm, P. I., Haldorsen, I. S., Mas, E., Sbarra, V., Nermoen, I., Eide, S. A., Grevle, L., Bjorkhaug, L., Sagen, J. V., Aksnes, L., Sovik, O., Lombardo, D., Molven, A., and Njolstad, P. R. (2006) Mutations in the CEL VNTR cause a syndrome of diabetes and pancreatic exocrine dysfunction. *Nature genetics* **38**, 54-62
11. Johansson, B. B., Torsvik, J., Bjorkhaug, L., Vesterhus, M., Ragvin, A., Tjora, E., Fjeld, K., Hoem, D., Johansson, S., Raeder, H., Lindquist, S., Hernell, O., Cnop, M., Saraste, J., Flatmark, T., Molven, A., and Njolstad, P. R. (2011) 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. *The Journal of biological chemistry* **286**, 34593-34605
12. Torsvik, J., Johansson, B. B., Dalva, M., Marie, M., Fjeld, K., Johansson, S., Bjorkoy, G., Saraste, J., Njolstad, P. R., and Molven, A. (2014) Endocytosis of secreted carboxyl ester lipase in a syndrome of diabetes and pancreatic exocrine dysfunction. *The Journal of biological chemistry* **289**, 29097-29111
13. Xiao, X., Jones, G., Sevilla, W. A., Stolz, D. B., Magee, K. E., Haughney, M., Mukherjee, A., Wang, Y., and Lowe, M. E. (2016) A Carboxyl Ester Lipase (CEL) Mutant Causes Chronic Pancreatitis by Forming Intracellular Aggregates That Activate Apoptosis. *The Journal of biological chemistry* **291**, 23224-23236
14. Fjeld, K., Weiss, F. U., Lasher, D., Rosendahl, J., Chen, J. M., Johansson, B. B., Kirsten, H., Ruffert, C., Masson, E., Steine, S. J., Bugert, P., Cnop, M., Grutzmann, R., Mayerle, J., Mossner, J., Ringdal, M., Schulz, H. U., Sendler, M., Simon, P., Sztromwasser, P., Torsvik, J., Scholz, M., Tjora, E., Ferec, C., Witt, H., Lerch, M. M., Njolstad, P. R., Johansson, S., and Molven, A. (2015)

A recombined allele of the lipase gene CEL and its pseudogene CELP confers susceptibility to chronic pancreatitis. *Nature genetics* **47**, 518-522
15. Baba, T., Downs, D., Jackson, K. W., Tang, J., and Wang, C. S. (1991) Structure of human milk bile salt activated lipase. *Biochemistry* **30**, 500-510
16. Mechref, Y., Chen, P., and Novotny, M. V. (1999) Structural characterization of the N-linked oligosaccharides in bile salt-stimulated lipase originated from human breast milk. *Glycobiology* **9**, 227-234
17. Sugo, T., Mas, E., Abouakil, N., Endo, T., Escribano, M. J., Kobata, A., and Lombardo, D. (1993) The structure of N-linked oligosaccharides of human pancreatic bile-salt-dependent lipase. *European journal of biochemistry* **216**, 799-805
18. Wang, C. S., Dashti, A., Jackson, K. W., Yeh, J. C., Cummings, R. D., and Tang, J. (1995) Isolation and characterization of human milk bile salt-activated lipase C-tail fragment. *Biochemistry* **34**, 10639-10644
19. Hanisch, F. G. (2001) O-glycosylation of the mucin type. *Biological chemistry* **382**, 143-149
20. Dube, D. H., and Bertozzi, C. R. (2005) Glycans in cancer and inflammation--potential for therapeutics and diagnostics. *Nature reviews. Drug discovery* **4**, 477-488
21. Escribano, M. J., Carre-Llopis, A., and Loridon-Rosa, B. (1985) Expression of oncofoetal pancreatic antigens in hamster adult pancreas during experimental carcinogenesis. *British journal of cancer* **51**, 187-193
22. Escribano, M. J., Cordier, J., Nap, M., Ten Kate, F. J., and Burtin, P. (1986) Differentiation antigens in fetal human pancreas. Reexpression in cancer. *International journal of cancer* **38**, 155-160
23. Albers, G. H., Escribano, M. J., Gonzalez, M., Mulliez, N., and Nap, M. (1987) Fetoacinar pancreatic protein in the developing human pancreas. *Differentiation; research in biological diversity* **34**, 210-215
24. Albers, G. H., Escribano, M. J., Daher, N., and Nap, M. (1990) An immunohistologic study of the feto-acinar pancreatic protein (FAP) in the normal pancreas, chronic pancreatitis, pancreatic adenocarcinoma, and intraabdominal metastases of adenocarcinomas. *American journal of clinical pathology* **93**, 14-19
25. Mas, E., Abouakil, N., Roudani, S., Miralles, F., Guy-Crotte, O., Figarella, C., Escribano, M. J., and Lombardo, D. (1993) Human fetoacinar pancreatic protein: an oncofetal glycoform of the normally secreted pancreatic bile-salt-dependent lipase. *Biochemical Journal* **289**, 609-615
26. Crescence, L., Beraud, E., Sbarra, V., Bernard, J. P., Lombardo, D., and Mas, E. (2012) Targeting a novel onco-glycoprotein antigen at tumoral pancreatic cell surface by mAb16D10 induces cell death. *Journal of immunology (Baltimore, Md. : 1950)* **189**, 3386-3396
27. Panicot-Dubois, L., Aubert, M., Franceschi, C., Mas, E., Silvy, F., Crotte, C., Bernard, J. P., Lombardo, D., and Sadoulet, M. O. (2004) Monoclonal antibody 16D10 to the C-terminal domain of the feto-acinar pancreatic protein binds to membrane of human pancreatic tumoral SOJ-6 cells and inhibits the growth of tumor xenografts. *Neoplasia* **6**, 713-724
28. Benkoel, L., Bernard, J. P., Payan-Defais, M. J., Crescence, L., Franceschi, C., Delmas, M., Ouaissi, M., Sastre, B., Sahel, J., Benoliel, A. M., Bongrand, P., Silvy, F., Gauthier, L., Romagne, F., Lombardo, D., and Mas, E. (2009) Monoclonal antibody 16D10 to the COOH-terminal domain of the feto-acinar pancreatic protein targets pancreatic neoplastic tissues. *Molecular cancer therapeutics* **8**, 282-291
29. Mas, E., Abouakil, N., Roudani, S., Franc, J. L., Montreuil, J., and Lombardo, D. (1993) Variation of the glycosylation of human pancreatic bile-salt-dependent lipase. *European journal of biochemistry* **216**, 807-812
30. Wolpin, B. M., Kraft, P., Xu, M., Steplowski, E., Olsson, M. L., Arslan, A. A., Bueno-de-Mesquita, H. B., Gross, M., Helzlsouer, K., Jacobs, E. J., LaCroix, A. *et al.* (2010) Variant ABO blood group alleles, secretor status, and risk of pancreatic cancer: results from the pancreatic cancer cohort consortium. *Cancer Epidemiol Biomarkers Prev* **19**, 3140-3149

31. El Jellas, K., Hoem, D., Hagen, K. G., Kalvenes, M. B., Aziz, S., Steine, S. J., Immervoll, H., Johansson, S., and Molven, A. (2017) Associations between ABO blood groups and pancreatic ductal adenocarcinoma: influence on resection status and survival. *Cancer Medicine* **6**, 1531-1540
32. Wolpin, B. M., Rizzato, C., Kraft, P., Kooperberg, C., Petersen, G. M., Wang, Z., Arslan, A. A., Beane-Freeman, L., Bracci, P. M., Buring, J *et al.* (2014) Genome-wide association study identifies multiple susceptibility loci for pancreatic cancer. *Nature genetics* **46**, 994-1000
33. Childs, E. J., Mocci, E., Campa, D., Bracci, P. M., Gallinger, S., Goggins, M., Li, D., Neale, R. E., Olson, S. H., Scelo, G., Amundadottir, L. T. *et al.* (2015) Common variation at 2p13.3, 3q29, 7p13 and 17q25.1 associated with susceptibility to pancreatic cancer. *Nature genetics* **47**, 911-916
34. Shindo, K., Yu, J., Suenaga, M., Fesharakizadeh, S., Tamura, K., Almario, J. A. N., Brant, A., Borges, M., Siddiqui, A., Datta, L., Wolfgang, C. L., Hruban, R. H., Klein, A. P., and Goggins, M. (2017) Lack of association between the pancreatitis risk allele CEL-HYB and pancreatic cancer. *Oncotarget* **8**, 50824-50831
35. Martinez, E., Crenon, I., Silvy, F., Del Grande, J., Mougel, A., Barea, D., Fina, F., Bernard, J. P., Ouaissi, M., Lombardo, D., and Mas, E. (2017) Expression of truncated bile salt-dependent lipase variant in pancreatic pre-neoplastic lesions. *Oncotarget* **8**, 536-551
36. Tan, M. H., Nowak, N. J., Loor, R., Ochi, H., Sandberg, A. A., Lopez, C., Pickren, J. W., Berjian, R., Douglass, H. O., Jr., and Chu, T. M. (1986) Characterization of a new primary human pancreatic tumor line. *Cancer investigation* **4**, 15-23
37. Deer, E. L., Gonzalez-Hernandez, J., Coursen, J. D., Shea, J. E., Ngatia, J., Scaife, C. L., Firpo, M. A., and Mulvihill, S. J. (2010) Phenotype and Genotype of Pancreatic Cancer Cell Lines. *Pancreas* **39**, 425-435
38. Reuss, R., Aberle, S., Klingel, K., Sauter, M., Greschniok, A., Franke, F. E., Padberg, W., and Blin, N. (2006) The expression of the carboxyl ester lipase gene in pancreas and pancreatic adenocarcinomas. *International journal of oncology* **29**, 649-654
39. Stanley P., Cummings R.D., Esko J.D., *et al.* Essentials of Glycobiology. 3rd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2015-2017. Chapter 14. Available from: https://www.ncbi.nlm.nih.gov/books/NBK453042/
40. Rossez, Y., Maes, E., Lefebvre Darroman, T., Gosset, P., Ecobichon, C., Joncquel Chevalier Curt, M., Boneca, I. G., Michalski, J. C., and Robbe-Masselot, C. (2012) Almost all human gastric mucin O-glycans harbor blood group A, B or H antigens and are potential binding sites for Helicobacter pylori. *Glycobiology* **22**, 1193-1206
41. Marionneau, S., Ruvoen, N., Le Moullac-Vaidye, B., Clement, M., Cailleau-Thomas, A., Ruiz-Palacois, G., Huang, P., Jiang, X., and Le Pendu, J. (2002) Norwalk virus binds to histo-blood group antigens present on gastroduodenal epithelial cells of secretor individuals. *Gastroenterology* **122**, 1967-1977
42. Ruvoen-Clouet, N., Mas, E., Marionneau, S., Guillon, P., Lombardo, D., and Le Pendu, J. (2006) Bile-salt-stimulated lipase and mucins from milk of 'secretor' mothers inhibit the binding of Norwalk virus capsids to their carbohydrate ligands. *The Biochemical journal* **393**, 627-634
43. Landberg, E., Pahlsson, P., Krotkiewski, H., Stromqvist, M., Hansson, L., and Lundblad, A. (1997) Glycosylation of bile-salt-stimulated lipase from human milk: comparison of native and recombinant forms. *Archives of biochemistry and biophysics* **344**, 94-102
44. Ju, T., Wang, Y., Aryal, R. P., Lehoux, S. D., Ding, X., Kudelka, M. R., Cutler, C., Zeng, J., Wang, J., Sun, X., Heimburg-Molinaro, J., Smith, D. F., and Cummings, R. D. (2013) Tn and sialyl-Tn antigens, aberrant O-glycomics as human disease markers. *Proteomics. Clinical applications* **7**, 618-631
45. Springer, G. F., Desai, P. R., Ghazizadeh, M., and Tegtmeyer, H. (1995) T/Tn pancarcinoma autoantigens: fundamental, diagnostic, and prognostic aspects. *Cancer detection and prevention* **19**, 173-182

46. Holmes, R. S., and Cox, L. A. (2011) Comparative Structures and Evolution of Vertebrate Carboxyl Ester Lipase (CEL) Genes and Proteins with a Major Role in Reverse Cholesterol Transport. *Cholesterol* **2011**, 781643
47. Hansson, L., Blackberg, L., Edlund, M., Lundberg, L., Stromqvist, M., and Hernell, O. (1993) Recombinant human milk bile salt-stimulated lipase. Catalytic activity is retained in the absence of glycosylation and the unique proline-rich repeats. *The Journal of biological chemistry* **268**, 26692-26698
48. Downs, D., Xu, Y. Y., Tang, J., and Wang, C. S. (1994) Proline-rich domain and glycosylation are not essential for the enzymic activity of bile salt-activated lipase. Kinetic studies of T-BAL, a truncated form of the enzyme, expressed in Escherichia coli. *Biochemistry* **33**, 7979-7985
49. Rogers, S., Wells, R., and Rechsteiner, M. (1986) Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. *Science* **234**, 364-368
50. Loomes, K. M., Senior, H. E., West, P. M., and Roberton, A. M. (1999) Functional protective role for mucin glycosylated repetitive domains. *European journal of biochemistry* **266**, 105-111
51. McGuckin, M. A., Linden, S. K., Sutton, P., and Florin, T. H. (2011) Mucin dynamics and enteric pathogens. *Nature reviews. Microbiology* **9**, 265-278
52. Bergstrom, K. S. B., and Xia, L. (2013) Mucin-type O-glycans and their roles in intestinal homeostasis. *Glycobiology* **23**, 1026-1037
53. Tailford, L. E., Crost, E. H., Kavanaugh, D., and Juge, N. (2015) Mucin glycan foraging in the human gut microbiome. *Frontiers in genetics* **6**, 81
54. Mäkivuokko, H., Lahtinen, S. J., Wacklin, P., Tuovinen, E., Tenkanen, H., Nikkilä, J., Björklund, M., Aranko, K., Ouwehand, A. C., and Mättö, J. (2012) Association between the ABO blood group and the human intestinal microbiota composition. *BMC Microbiology* **12**, 94-94
55. Johansson, M. E., Larsson, J. M., and Hansson, G. C. (2011) The two mucus layers of colon are organized by the MUC2 mucin, whereas the outer layer is a legislator of host-microbial interactions. *Proceedings of the National Academy of Sciences of the United States of America* **108** Suppl 1, 4659-4665
56. La Rosa, S., Vigetti, D., Placidi, C., Finzi, G., Uccella, S., Clerici, M., Bartolini, B., Carnevali, I., Losa, M., and Capella, C. (2010) Localization of carboxyl ester lipase in human pituitary gland and pituitary adenomas. *The journal of histochemistry and cytochemistry : official journal of the Histochemistry Society* **58**, 881-889
57. Kodvawala, A., Ghering, A. B., Davidson, W. S., and Hui, D. Y. (2005) Carboxyl ester lipase expression in macrophages increases cholesteryl ester accumulation and promotes atherosclerosis. *The Journal of biological chemistry* **280**, 38592-38598
58. Li, F., and Hui, D. Y. (1998) Synthesis and secretion of the pancreatic-type carboxyl ester lipase by human endothelial cells. *The Biochemical journal* **329 (Pt 3)**, 675-679
59. Yamamoto, F. (2017) Evolutionary divergence of the ABO and GBGT1 genes specifying the ABO and FORS blood group systems through chromosomal rearrangements. *Scientific reports* **7**, 9375
60. Amundadottir, L., Kraft, P., Stolzenberg-Solomon, R. Z., Fuchs, C. S., Petersen, G. M., Arslan, A. A., Bueno-de-Mesquita, H. B., Gross, M., Helzlsouer, K., Jacobs, E. J., *et al.* (2009) Genome-wide association study identifies variants in the ABO locus associated with susceptibility to pancreatic cancer. *Nature genetics* **41**, 986-990
61. Blixt, O., Head, S., Mondala, T., Scanlan, C., Huflejt, M. E., Alvarez, R., Bryan, M. C., Fazio, F., Calarese, D., Stevens, J., Razi, N., Stevens, D. J., Skehel, J. J., van Die, I., Burton, D. R., Wilson, I. A., Cummings, R., Bovin, N., Wong, C.-H., and Paulson, J. C. (2004) Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 17033-17038
62. Ceroni, A., Maass, K., Geyer, H., Geyer, R., Dell, A., and Haslam, S. M. (2008) GlycoWorkbench: a tool for the computer-assisted annotation of mass spectra of glycans. *Journal of proteome research* **7**, 1650-1659

63. Varki, A., Cummings, R. D., Esko, J. D., Freeze, H. H., Stanley, P., Marth, J. D., Bertozzi, C. R., Hart, G. W., and Etzler, M. E. (2009) Symbol nomenclature for glycan representation. *Proteomics* **9**, 5398-5399

Figure legends

FIGURE 1. **CEL expression in the pancreas is restricted to acinar cells.** *A*, double fluorescent labeling of CEL mRNA (red) and protein (green) in normal human pancreatic parenchyma. Acinar cells are strongly positive, whereas ductal cells are negative. The dotted line circumscribes a representative intralobular duct. White arrowheads indicate the epithelial lining of the duct lumen. *B*, a representative islet of Langerhans (circumscribed by dotted line) devoid of CEL expression. *C*, high magnification of an acinus with one well-polarized acinar cell highlighted by the dotted line. Note the basal labeling of the mRNA and the apical position of the zymogen granules staining positive for CEL protein. Cell nuclei are stained blue with DAPI. A single z-plane is shown (0.19 µm). Scale bars represent 100 µm.

FIGURE 2. **Pancreatic premalignant and malignant lesions are negative for CEL expression.** *A-D,* chromogenic *in situ* labeling of CEL mRNA in sections from representative pancreatic cancer cases. *E-H*, chromogenic immunostaining of CEL protein in adjacent sections of *A-D*. *A* and *E*, preserved pancreatic parenchyma with normal morphology. Acinar cells contain high levels of CEL mRNA and protein located basally and apically, respectively. *B* and *F*, areas of atrophy with acinar-to-ductal metaplasia (ADM, marked by red stars). Both CEL mRNA and protein cease to be expressed when a lumen is visible. *C* and *G*, a pancreatic intraepithelial neoplasia lesion (PanIN, circumscribed by dotted line). *Insets*, magnification of one papillary protrusion showing negative epithelial cells. The arrow in *G* indicates secreted CEL protein present in the duct lumen. *D* and *H*, malignant pancreatic ducts embedded in desmoplastic stroma. All structures in these areas were devoid of detectable levels of both CEL mRNA and protein. Scale bars in *A-H* represent 100 µm. *I*, RT-qPCR comparing *CEL* transcript level in PDAC sections (areas corresponding to *D*/*H*; n=9 different patients) with level in normal pancreatic tissue from patients with non-pancreatic pathologies (n=4). *QARS* gene expression was used as normalizing control. Values are expressed as mean ± SEM.

FIGURE 3. **CEL expression is low or absent in pancreatic cancer cell lines**. *A*, *CEL* mRNA levels of the pancreatic cancer cell lines BxPC-3, MIA PaCa-2 and PANC-1 compared with levels in HeLa cells and in stably transfected HEK293 cells (positive control: HEK293_CEL, transfected with a CEL-expressing plasmid construct; negative control: HEK293_EV, transfected with empty plasmid vector). *GAPDH* gene expression was used as loading control. Y-axis scale is logarithmic. Error bars represents standard deviation from three experimental replicates. *B*, a representative western blot (n=2) of protein lysates from the above cell lines stained with anti-CEL antibody. GAPDH protein levels in the lower panel was used as loading control.

FIGURE 4. **The monoclonal antibody mAb16D10 binds strongly to blood group A-containing structures in a glycan array.** *A*, mAb16D10 screened against 609 different glycans on a printed microarray from the Consortium for Functional Glycomics (see Experimental Procedures). Binding was detected via Alexa Fluor 488-labeled anti-mouse secondary antibody. *B*, the 18 glycan structures with the highest affinity shown in decreasing order of binding. Numbers refer to position in the array in *A*. *C*, the structural motif common to all binding structures. This was the terminal glycan having the sequence GalNAc-α1,3(Fuc-α1,2)Gal, which corresponds to the blood group A antigen.

FIGURE 5. **mAb16D10 cross-reacts with blood group A antigens in normal pancreas tissue.** *A* and *C*, normal pancreatic parenchyma of subjects with blood group A and O, respectively, immunostained with mAb16D10. Only specimens from blood group A subjects reacted with the antibody. *B* and *D*, blood vessels of the same specimens. The black arrows point at red blood cells, whereas the arrowheads point at the endothelial lining of the vessels. Pancreatic islets (unstained region in *A*; circumscribed by dotted line in *C*) were negative for both blood groups. Scale bars represent 100 µm.

FIGURE 6. **mAb16D10 cross-reacts with blood group A antigens in pancreatic cancer tissue.** *A-H*, adjacent tissue sections from a patient with blood group A immunostained either with mAb16D10 or an anti-A antibody. *I-L*, tissue sections from a patient with blood group O immunostained with mAb16D10. *A*, *E* and *I*, preserved pancreatic parenchyma with normal morphology. The arrowhead points to pancreatic secretions within a small duct. *B, F* and *J*, areas of atrophy with acinar-to-ductal metaplasia (ADM, marked by red stars). *C, G* and *K*, pancreatic intraepithelial neoplasia lesions (PanIN). *Inset* in *K*, magnification of one papillary protrusion. *D, H* and *L*, malignant pancreatic ducts embedded in desmoplastic stroma. Scale bars represent 100 µm.

FIGURE 7. **The mAb16D10 antibody can recognize the cancer-associated Tn antigen.** *A*, Comparative binding profile of mAb16D10 (upper part) and an anti-Tn antibody (lower part) on an array of synthetic Tn glycopeptides from the Consortium for Functional Glycomics (see Experimental Procedures). The array consisted of 56 compounds including the blood group A glycan structure (red box). Structures that bound both antibodies are highlighted by yellow boxes. Only mAb16D10 bound to blood group A antigen structures with high affinity. Note that peptides containing the sequence DVPVEG[S_n]TSTVAPANK showed increasing binding intensities as the number of GalNAc-conjugated serine residues was increased (array positions 45, 47, 49), whereas the same peptide sequences devoid of the sugar bound very weakly (positions 46, 48, 50). Compounds on the array and their binding affinities are listed in Suppl. Table 2. *B*, the terminal GalNAc-α residue shared by the Tn (GalNAc-α-Ser/Thr) and blood group A (GalNAc-α1,3(Fuc-α1,2)Gal) antigens. *C*, immunostaining with mAb16D10 or the anti-Tn antibody in adjacent sections of malignant pancreatic ducts embedded in desmoplastic stroma (blood group O individual). Scale bars represent 100 µm.

FIGURE 8. **Immunoprecipitated CEL protein from pancreatic juice reacts with mAb16D10 and an anti-A antibody but not with anti-Tn.** Co-immunoprecipitation (*upper panel*) and immunoprecipitation (*lower panel*) of CEL were performed on four juice samples, followed by parallel immunoblotting with anti-CEL, mAb16D10, anti-A, or anti-Tn antibodies. CEL was detected at around 110 kDa in all samples. The mAb16D10 and anti-A antibodies gave the same staining pattern in samples from blood group A cases (juices 3 and 4). Tn did not react with CEL but showed reactivity towards a protein of higher molecular weight regardless of blood group (*upper panel*). Two different anti-CEL antibodies were used: the mouse monoclonal antibody As20.1 for immunoprecipitation and a rabbit polyclonal antibody (13) for immunoblotting. The dashed line indicates that the membrane part to the left was cut and exposed for a longer time in order to achieve better signals.

FIGURE 9. **The mucinous domain of CEL contains glycans reflecting the patients' *ABO* and *FUT2* status.** The O-glycan pool from the pancreatic juice of three PDAC cases was investigated by MALDI-TOF-MS after alkaline beta-elimination and permethylation. The O-glycan population was composed of

core 1 and core 2 structures. *A,* spectrum showing O-linked glycans from a blood group O *FUT2*-positive (secretor) case. *Inset*, 34X magnification in the m/z range 1500-2400 showing core 2 extended structures of less abundance. *B*, corresponding spectrum from a blood group A *FUT2*-positive (secretor) case, with 15X magnification in the m/z range 1190-2400. *C*, corresponding spectrum from a blood group O *FUT2*-negative (non-secretor) case, with a 52X magnification in the m/z range 1500-2400. All molecular ions are present in singly charged sodiated form ([M+Na]+). Major peaks are annotated with their proposed carbohydrate structure, according to the symbolic nomenclature adopted by the CFG (63). Putative structures are based on monosaccharide composition, data from the MS/MS analysis of all molecular ions and knowledge of the biosynthetic pathways for O-linked glycans. MS and MS/MS spectra were annotated manually with the aid of the semi-automated tool Glycoworkbench (62). When structures have not been unequivocally defined from MS/MS information, the monosaccharide is shown after a bracket.

FIGURE 10. **MALDI-TOF/TOF-MS/MS fragmentation spectras confirm the annotation of selected precursor ions from the CEL O-glycome.** Molecular ions m/z 1781, 1822 and 1794 from Fig. 9A, B and C, respectively underwent collision-induced dissociation and the resulting spectra are shown. Arrows pointing from the dashed line indicate the loss of a specific fragment from the precursor ion. All fragment ions are $[M+Na]^+$. When structures have not been unequivocally defined from MS/MS information, the monosaccharide is shown after a bracket.

SUPPLEMENTARY FIGURE 1. **mAb16D10 stains gastric and breast cancer tissue.** *A, B,* sections from gastric and breast cancer, respectively, both from non-A individuals. Scale bars represent 100 µm.

SUPPLEMENTARY FIGURE 2. **MALDI-TOF/TOF-MS/MS fragmentation spectra of the selected molecular ion m/z 1519 from two samples with different secretor status.** *A*, fragmentation spectra of the molecular ion 1519 from the blood group O/*FUT2*-positive (secretor) case; *B*, same from the blood group O/*FUT2*-negative (non-secretor) case. All fragment ions are [M+Na]+. When structures have not been unequivocally defined from MS/MS information, the monosaccharide is shown after a bracket.

SUPPLEMENTARY TABLE 1. **List of compounds in the Tn microarray and their observed binding affinities to mAb16D10 and the anti-Tn antibody.** Chart ID refers to the structures' position in the array; STDV is the SD value calculated from 4 technical replicates. An asterisk within the sequence indicates that the residue before it contains a GalNAc residue, and the number between parentheses indicates how many times the preceding residue is repeated.

FIGURE 1

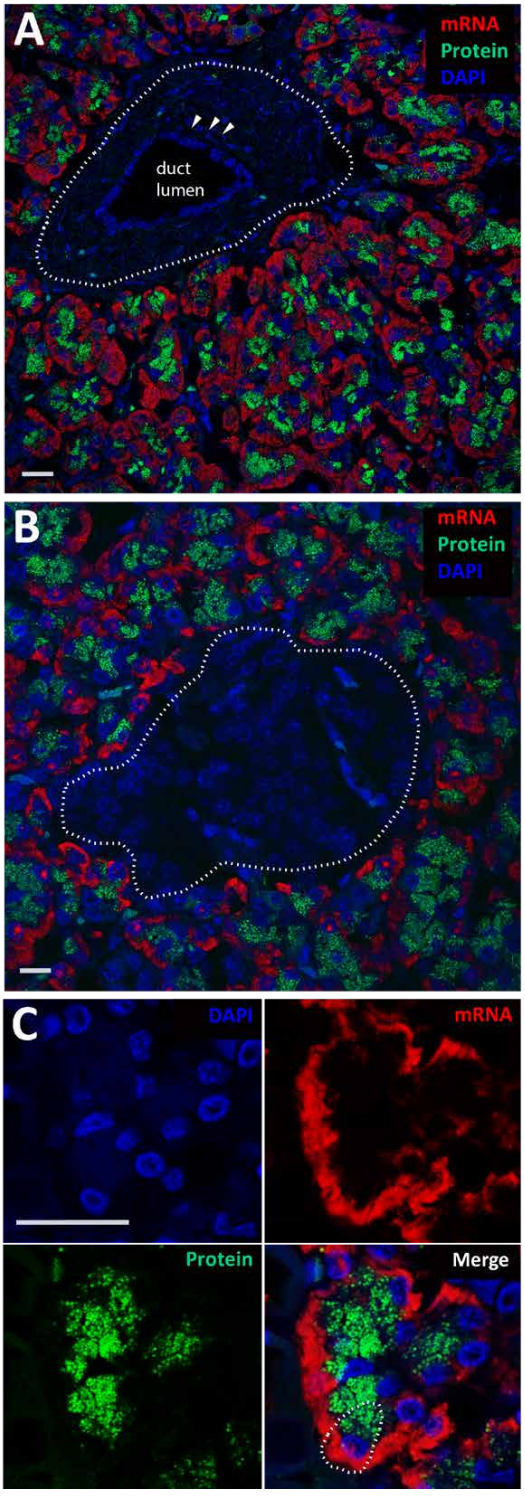

A
mRNA
Protein
DAPI
duct
lumen
B
mRNA
Protein
DAPI
C
DAPI
mRNA
Protein
Merge

FIGURE 2

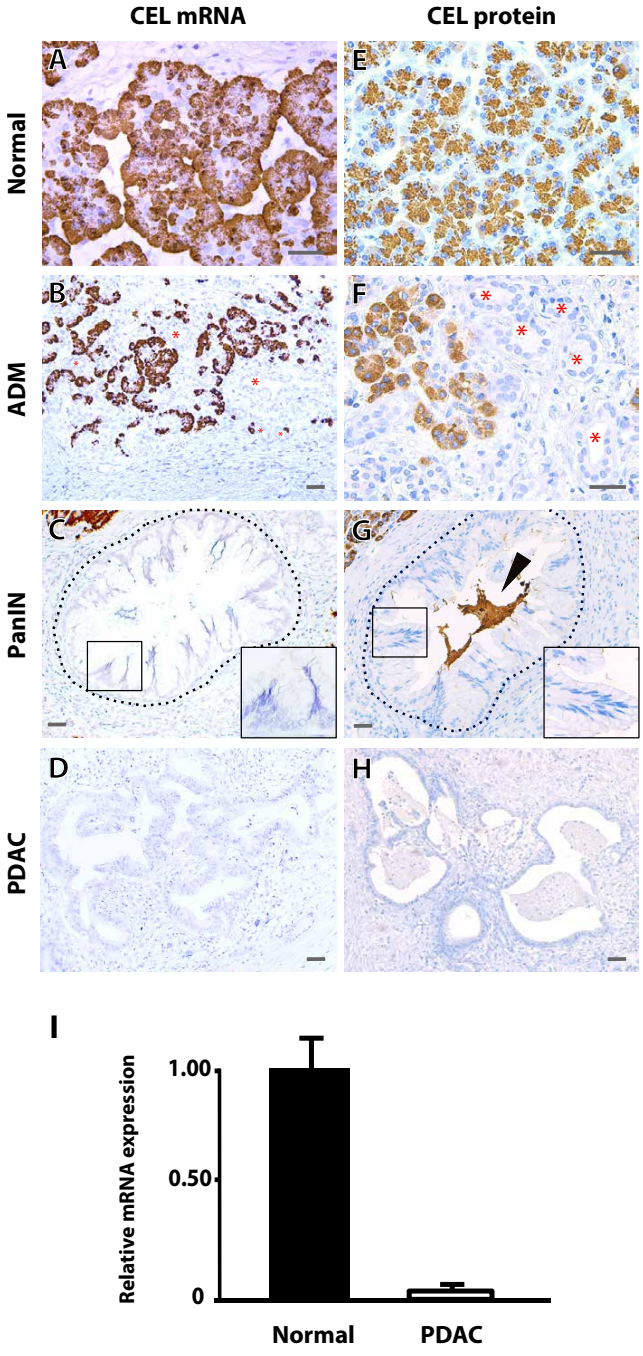

CEL mRNA
CEL protein
Normal
ADM
PanIN
PDAC
A
B
C
D
E
F
G
H
I
Relative mRNA expression
1.00
0.50
0
Normal
PDAC

FIGURE 3

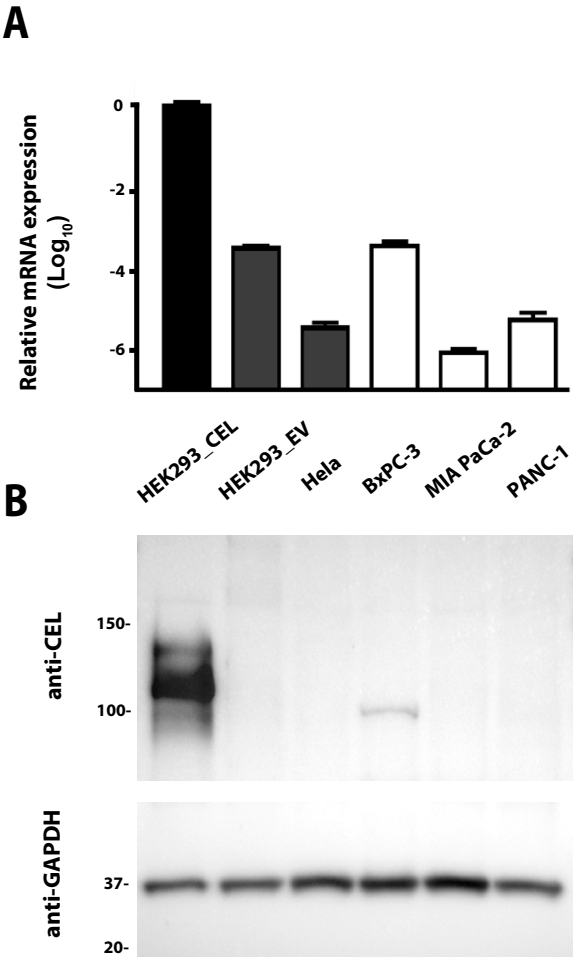
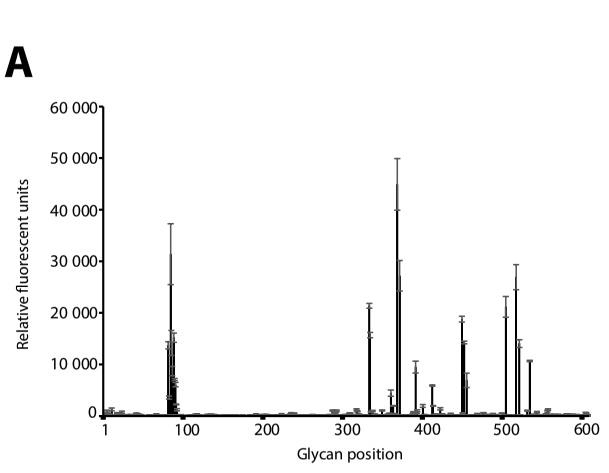
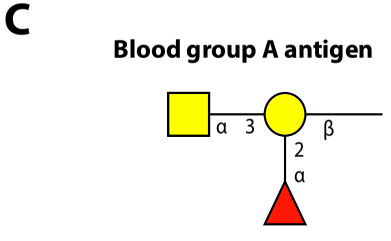

A
Relative mRNA expression
(Log10)
0
-2
-4
-6
HEK293_CEL
HEK293_EV
Hela
BxPC-3
MIA PaCa-2
PANC-1
B
anti-CEL
150-
100-
anti-GAPDH
37-
20-

FIGURE 4


A
Relative fluorescent units
60 000
50 000
40 000
30 000
20 000
10 000
0
1
100
200
300
400
500
600
Glycan position


C
Blood group A antigen
α
3
β
2
α

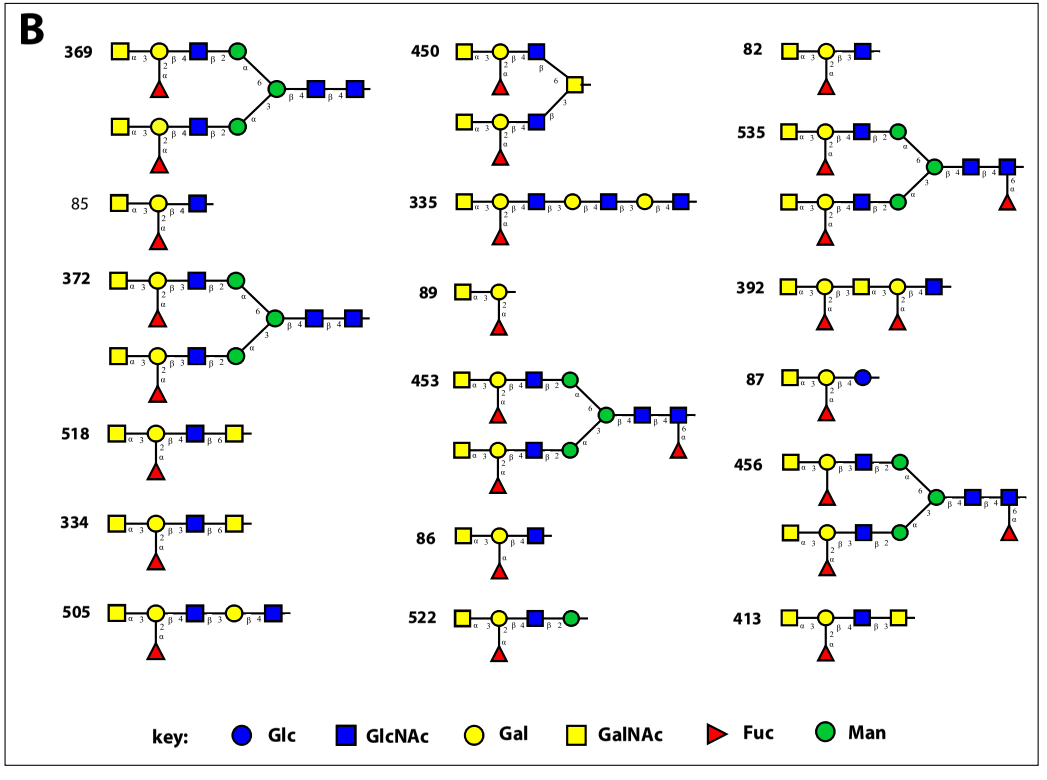

B
369
85
372
518
334
505
450
335
89
453
86
522
82
535
392
87
456
413
key:
Glc
GlcNAc
Gal
GalNAc
Fuc
Man

FIGURE 5

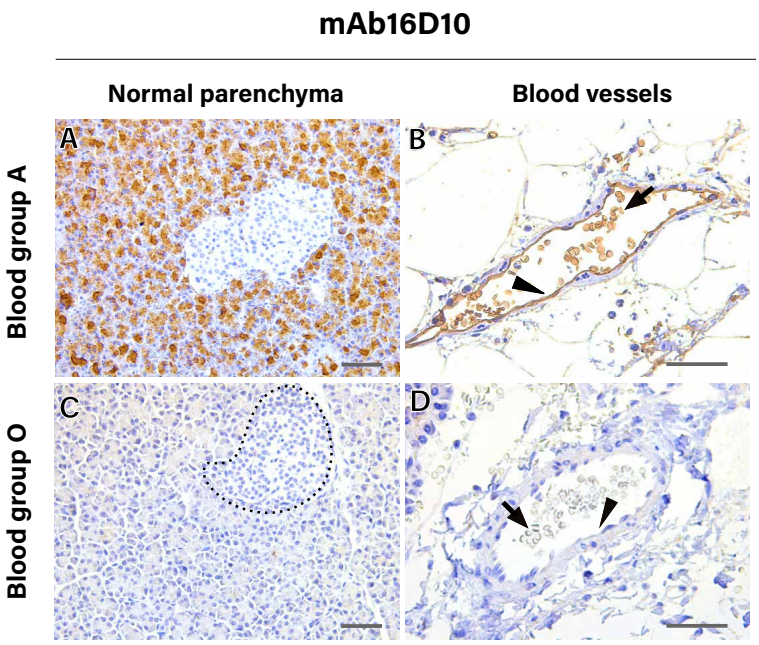

mAb16D10
Normal parenchyma
Blood vessels
Blood group A
Blood group O
A
B
C
D

FIGURE 6

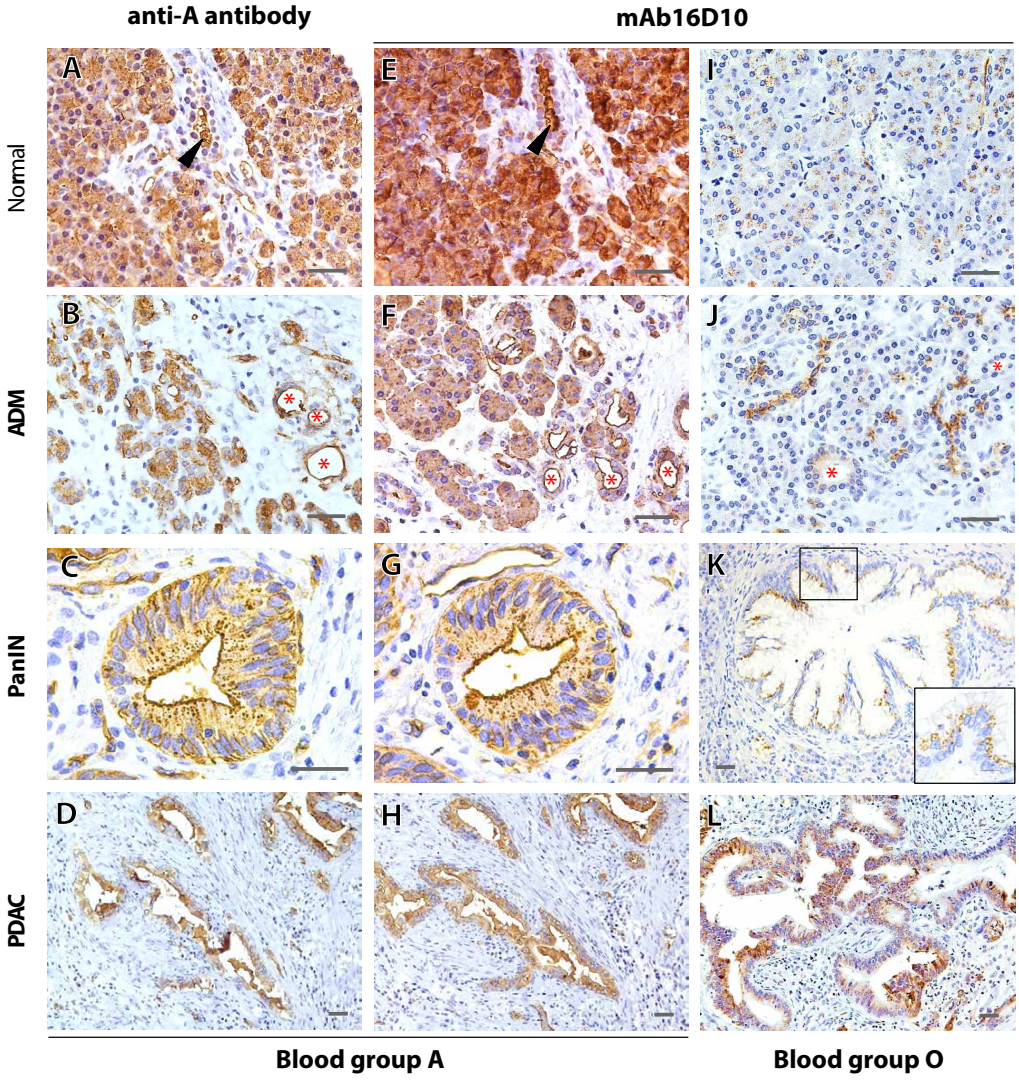
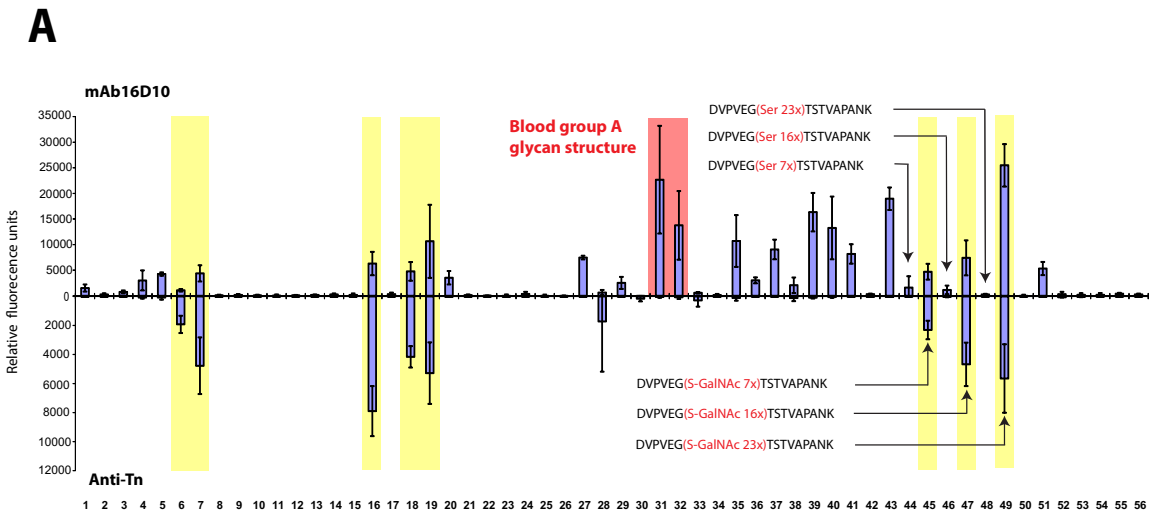
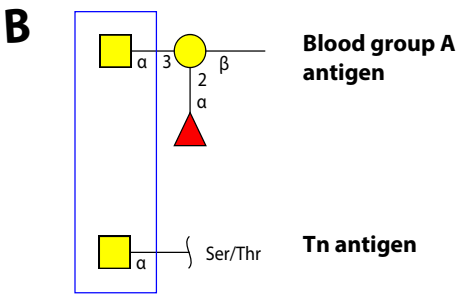

anti-A antibody
mAb16D10
Normal
ADM
PanIN
PDAC
A
E
I
B
F
J
C
G
K
D
H
L
Blood group A
Blood group O

FIGURE 7


A
mAb16D10
Relative fluorescence units
Blood group A glycan structure
DVPVEG(Ser 23x)TSTVAPANK
DVPVEG(Ser 16x)TSTVAPANK
DVPVEG(Ser 7x)TSTVAPANK
DVPVEG(S-GalNAc 7x)TSTVAPANK
DVPVEG(S-GalNAc 16x)TSTVAPANK
DVPVEG(S-GalNAc 23x)TSTVAPANK
Anti-Tn


B
α
3
β
2
α
α
Ser/Thr
Blood group A antigen
Tn antigen

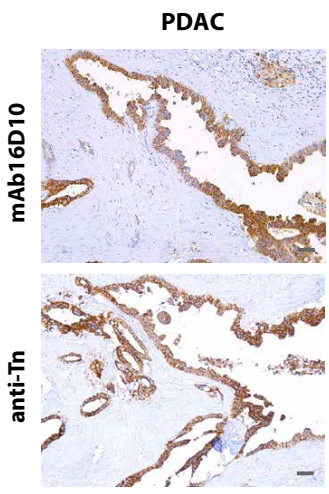
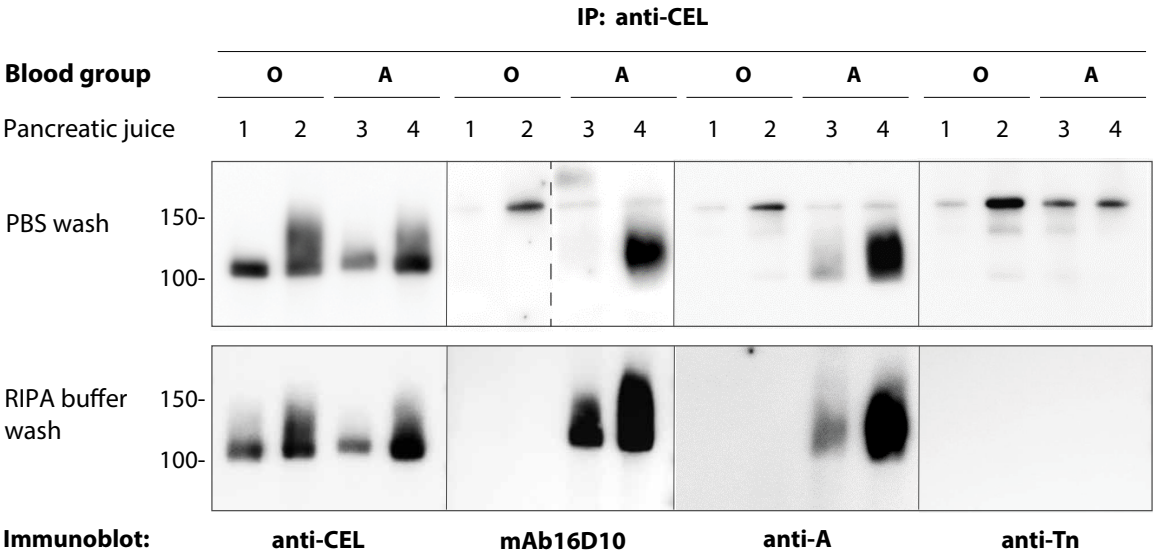

C
PDAC
mAb16D10
anti-Tn

FIGURE 8


IP: anti-CEL
Blood group
O
A
O
A
O
A
O
A
Pancreatic juice
1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4
PBS wash
150-
100-
RIPA buffer wash
150-
100-
Immunoblot:
anti-CEL
mAb16D10
anti-A
anti-Tn

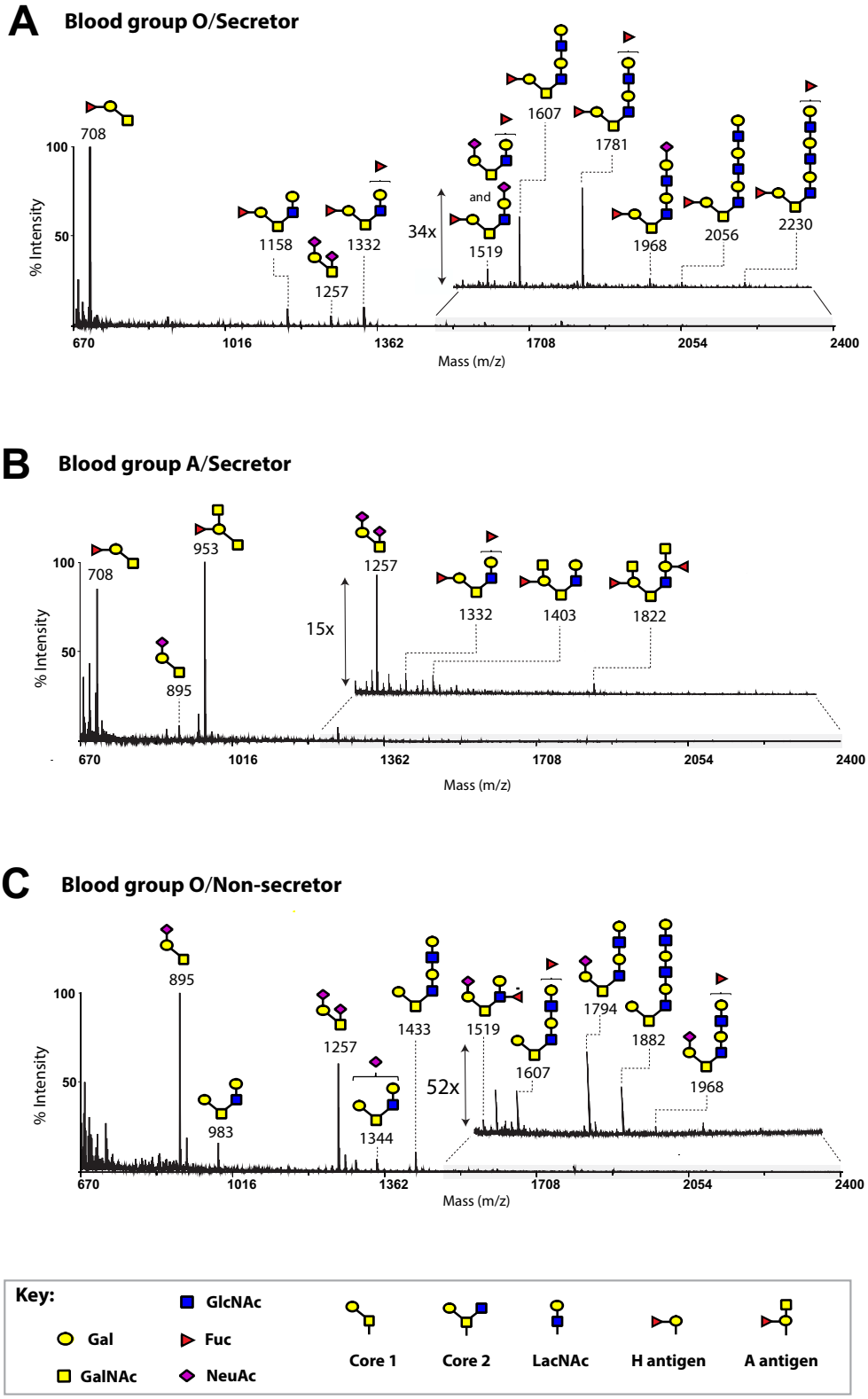
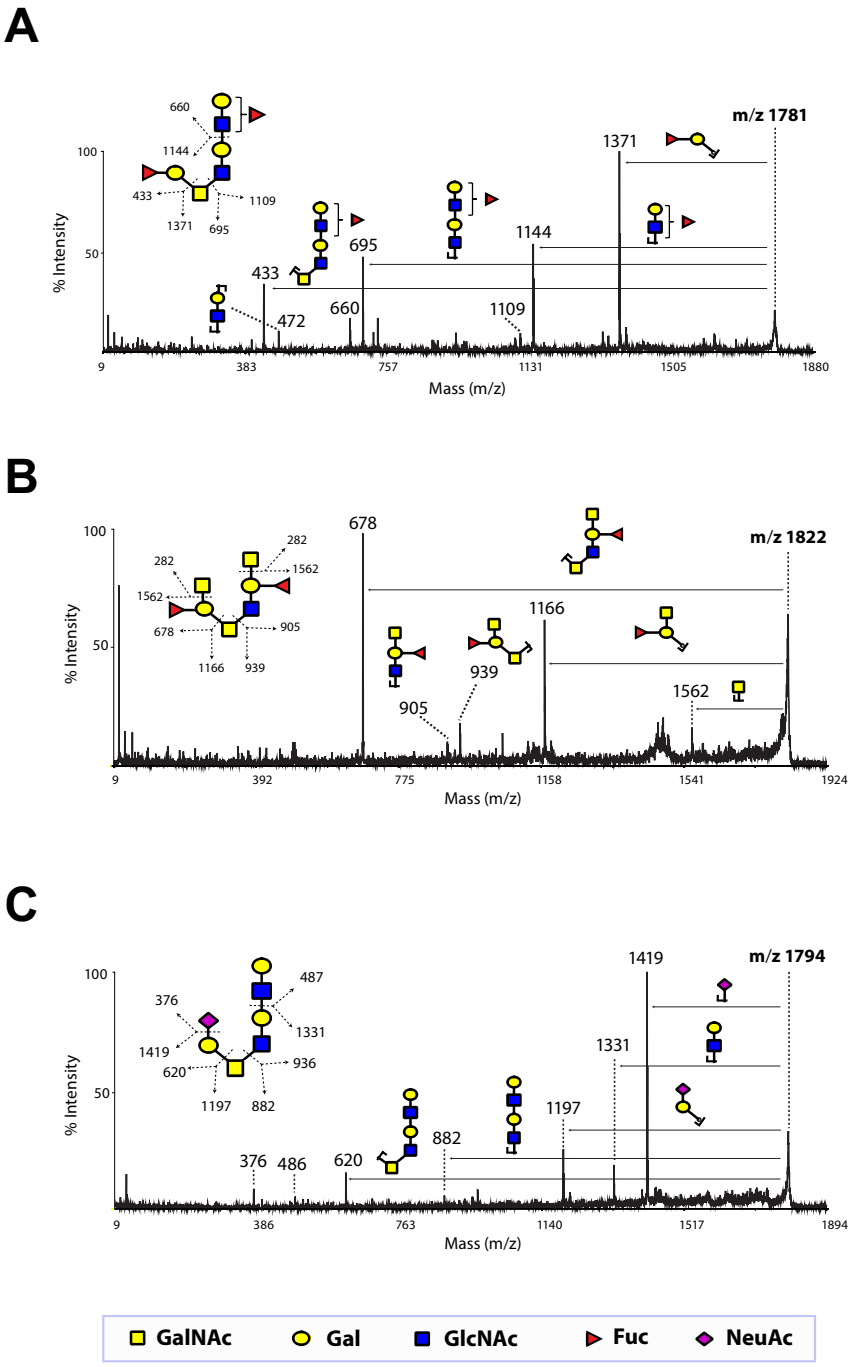
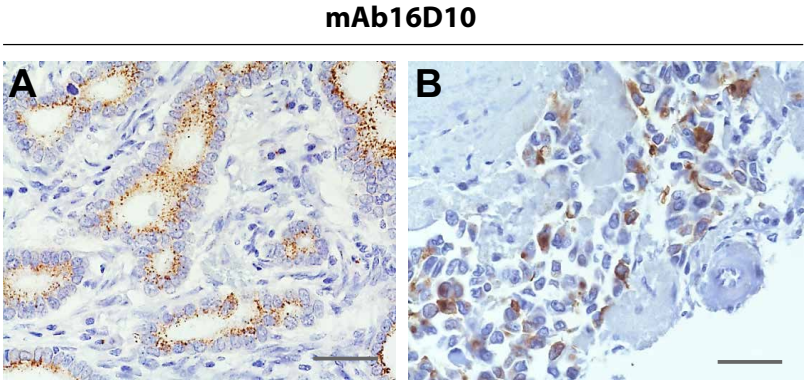
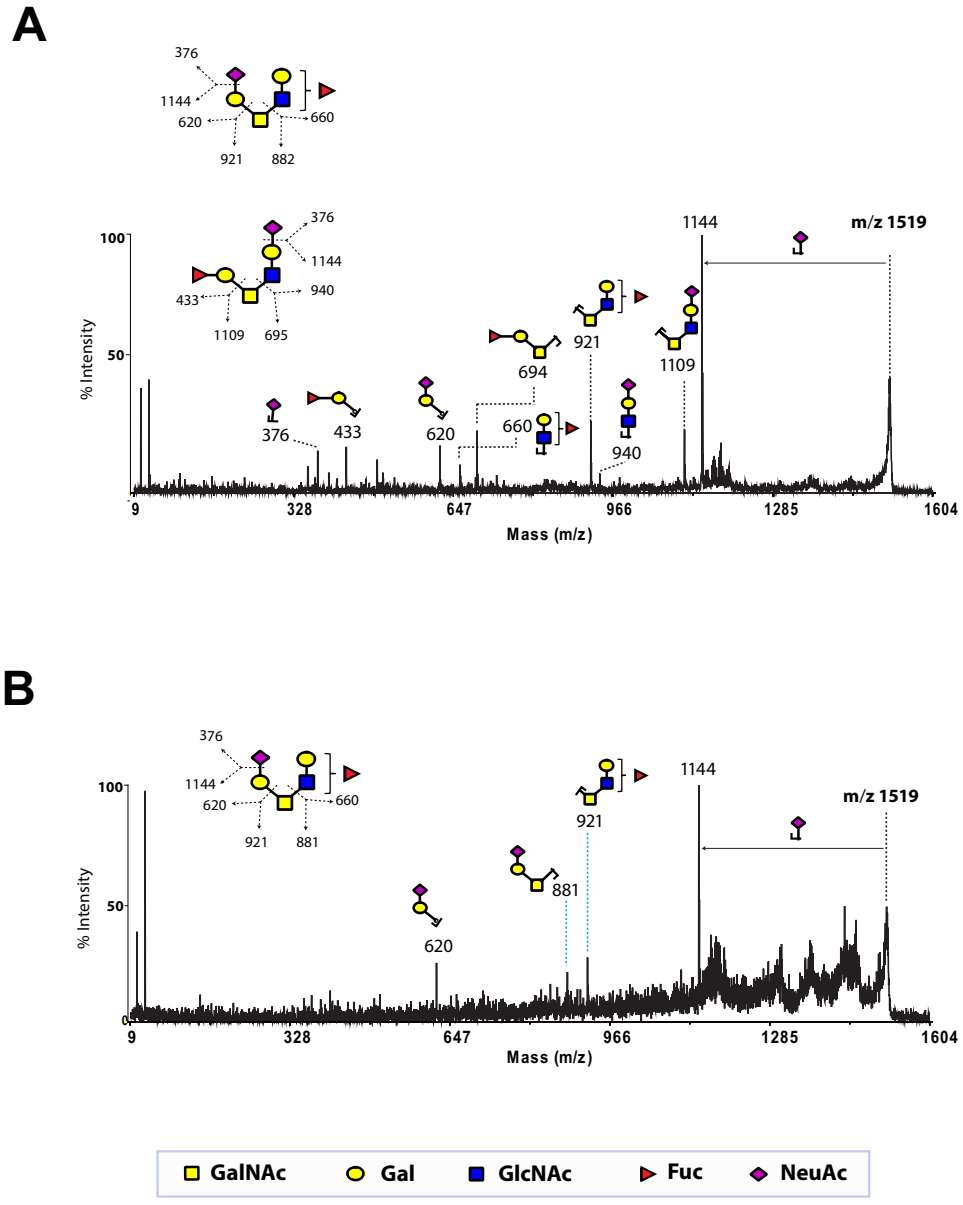

A
Blood group O/Secretor
708
1158
1257
1332
1519
and
34x
1607
1781
1968
2056
2230
100
50
% Intensity
670
1016
1362
1708
2054
2400
Mass (m/z)
B
Blood group A/Secretor
708
953
895
1257
15x
1332
1403
1822
100
50
% Intensity
670
1016
1362
1708
2054
2400
Mass (m/z)
C
Blood group O/Non-secretor
895
983
1257
1344
1433
1519
52x
1607
1794
1882
1968
100
50
% Intensity
670
1016
1362
1708
2054
2400
Mass (m/z)
Key:
Gal
GalNAc
GlcNAc
Fuc
NeuAc
Core 1
Core 2
LacNAc
H antigen
A antigen

FIGURE 10


A
m/z 1781
% Intensity
Mass (m/z)
B
m/z 1822
% Intensity
Mass (m/z)
C
m/z 1794
% Intensity
Mass (m/z)
GalNAc
Gal
GlcNAc
Fuc
NeuAc

SUPPLEMENTARY FIGURE 1




A
376
1144
620
660
921
882
100
50
% Intensity
376
1144
433
940
1109
695
1144
m/z 1519
921
694
1109
376
433
620
660
940
9
328
647
966
1285
1604
Mass (m/z)
B
376
1144
620
660
921
881
100
50
% Intensity
921
1144
m/z 1519
881
620
9
328
647
966
1285
1604
Mass (m/z)
GalNAc
Gal
GlcNAc
Fuc
NeuAc

SUPPLEMENTARY TABLE 1

			mAb16D10		anti-Tn	
Chart ID	**Detail**	**Sequence**	**Average**	**STDEV**	**Average**	**STDEV**
1	a-100uM	AcPT*TTPLKNH2	1509	688	1	1
2	b-100uM	AcPTT*TPLKNH2	272	215	2	2
3	c-100uM	AcPTTT*PLKNH2	743	288	6	11
4	d-100uM	AcPT*T*TPLKNH2	2996	1951	76	75
5	e-100uM	AcPT*TT*PLKNH2	4240	321	76	150
6	f-100uM	AcPTT*T*PLKNH2	1071	241	1925	594
7	g-100uM	AcPT*T*T*PLKNH2	4383	1581	4774	1952
8	R-100uM	AcPTTTPLKNH2	73	142	1	2
9	I-100uM	AcPPTTTTKKPNH2	166	172	2	1
10	II-100uM	H-GTTPSPVPT*TSTTSAP-OH	72	148	20	14
11	III-100uM	AcPTTDSTT*PAPTTKNH2	25	230	11	6
12	EA2 -1-100uM	Ac-PTTDSTTPAPTTK-HH2	60	113	3	5
13	EA2 -2-100uM	Ac-PPT*T*T*T*KKP-HN2	165	98	2	2
14	EA2 -3-100uM	NH2-TSAPDT*RDAP-NH2	252	204	1	1
15	EA2 -4-100uM	NH2-TSAPDTRPAP-NH2	152	293	5	6
16	G-8-Pep-100uM	H-APGS*T*APP-NH2	6287	2287	7898	1724
17	P-8-mer-100uM	H-APGSTAPP-NH2	321	308	2	3
18	PADRE-Tn3-100uM	C107H178N26O38 - Sequence not given	4757	1818	4158	738
19	Tn3-linker-100uM	C41H72N8O22 - Sequence not given	10640	7146	5284	2121
20	Tn-linker-100uM	C17H32N4O8 - Sequence not given	3521	1316	27	10
21	Peptide-4-100uM	AcHN-KTTT-CONH2	100	149	7	12
22	Peptide-5-100uM	AcHN-KTTTG-CONH2	8	55	6	3
23	Ser1-100uM	H-Ser(a-D-GalNAc)-NH2	-104	365	15	2
24	S2-100uM	H-Ser(a-D-GalNAc)-OH	357	418	6	8
25	Thr1-100uM	H-Thr(a-D-GalNAc)-NH2	-133	358	3	2
26	T1-100uM	H-Thr(a-D-GalNAc)-OH	-90	164	2	3
27	IgA-Pep01-100uM	KPVPST*PPT*PS*C	7466	369	11	7
28	IgA-Pep02-100uM	KPVPSTPPTPSC	630	507	1728	3449
29	S-GalNAc-100uM	S-GalNAc	2521	1181	0	1
30	T-GalNAc-100uM	T-GalNAc	-111	88	153	187
31	Bloodgroup A Tetra-AEAB-100uM		22673	10527	37	60
32	Bloodgroup A penta-AEAB-100uM		13755	6708	60	116
33	LNnT-100uM		573	144	274	429
34	Man5-100uM		117	230	6	9
35	Crypto peptide 01-50uM	H-ETS*EAAAT*VDLFAFT*LDGGK-NH2	10697	5063	102	192
36	Crypto peptide 02-50uM	H-ETSEAAAT*VDLFAFT*LDGGK-NH2	3014	576	12	19
37	Crypto peptide 03-50uM	H-ETS*EAAATVDLFAFT*LDGGK-NH2	9042	1926	13	6
38	Crypto peptide 04-50uM	H-ETSEAAATVDLFAFT*LDGGK-NH2	2056	1519	110	218
39	Crypto peptide 05-50uM	H-ETS*EAAAT*VDLFAFTLDGGK-NH2	16332	3763	89	58
40	Crypto peptide 06-50uM	H-ETSEAAAT*VDLFAFTLDGGK-NH2	13248	6132	45	47
41	Crypto peptide 07-50uM	H-ETS*EAAATVDLFAFTLDGGK-NH2	8165	1893	13	23
42	Crypto peptide 08-50uM	H-ETSEAAATVDLFAFTLDGGK-NH2	316	115	17	14
43	Crypto peptide 09-50uM	H-ETT*EAAAS*VDLFAFS*LDGGK-NH2	18958	2203	30	3
44	Crypto peptide 10-50uM	H-ETTEAAASVDLFAFSLDGGK-NH2	1608	2215	2	2
45	Crypto peptide 11-50uM	H-DVPVEGSS*(7)TSTVAPANK-NH2	4695	1544	2311	638
46	Crypto peptide 12-50uM	H-DVPVEGSS(7)TSTVAPANK-NH2	1143	841	27	36
47	Crypto peptide 13-50uM	H-DVPVEGSS*(16)TSTVAPANK-NH2	7407	3419	4676	1496
48	Crypto peptide 14-50uM	H-DVPVEGSS(16)TSTVAPANK-NH2	249	131	1	2
49	Crypto peptide 15-50uM	H-DVPVEGSS*(23)TSTVAPANK-NH2	25484	4147	5651	2361
50	Crypto peptide 16-50uM	H-DVPVEGSS(23)TSTVAPANK-NH2	-68	286	2	3
51	Cp17 protein-200ug/ml		5324	1297	15	11
52	Cp23 protein-200ug/ml		308	473	62	8
53	PBS		182	317	25	29
54	PBS		191	332	3	2
55	PBS		400	185	6	7
56	Biotin		267	166	3	3

Graphic design: Communication Division, UiB / Print: Skipnes Kommunikasjon AS

uib.no

ISBN: 978-82-308-3673-6