DETECTION AND MOLECULAR CHARACTERIZATION OF GASTROINTESTINAL CANCERS AND PREMALIGNANT LESIONS

Thesis

Kateryna Nesteruk

ISBN: 978-94-6332-783-1
Cover and layout: K. Nesteruk
Printing: GVO drukkers & vormgevers B.V.

The printing of this thesis has been financially supported by the Department of Gastroenterology and Hepatology, Erasmus University Medical Center, Rotterdam, Erasmus University Rotterdam

DETECTION AND MOLECULAR CHARACTERIZATION OF GASTROINTESTINAL CANCERS AND PREMALIGNANT LESIONS

DETECTIE EN MOLECULAIRE KARAKTERISERING VAN GASTRO-INTESTINALE KANKERS EN PREMALIGNE LAESIES

Thesis

to obtain the degree of Doctor from the
Erasmus University Rotterdam
by command of the
rector magnificus

Prof. dr. A.L. Bredenoord

and in accordance with the decision of the Doctorate Board.
The public defence shall be held on

Wednesday 10 November 2021 at 13.00 hrs
by

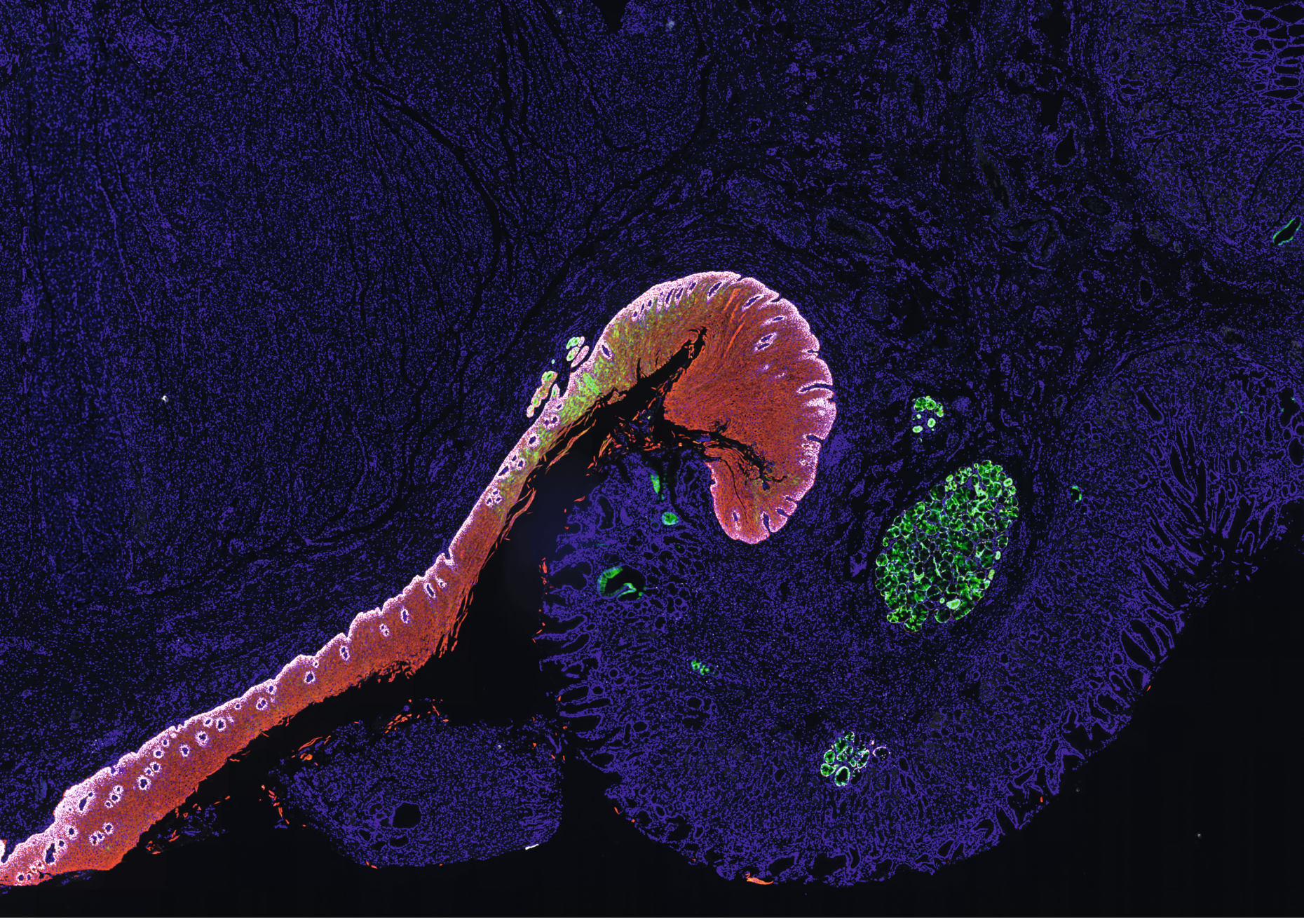
Kateryna Nesteruk
born in Ivano-Frankivsk, Ukraine

Erasmus University Rotterdam

Doctoral Committee:

Promotor:	Prof. dr. M.P. Peppelenbosch
Other members:	Prof. dr. M.J. Bruno
	Prof. dr. S. Krishnadath
	Prof. dr. J. Gribnau
Co-promotor:	Dr. G.M. Fuhler

CONTENT



Chapter 1

Introduction and outline of the thesis

Based on: Achalasia and associated esophageal cancer risk: what lessons can we learn from the molecular analysis of Barrett's-associated adenocarcinoma? K. Nesteruk, M. C. W. Spaander, I. Leeuwenburgh, M. P. Peppelenbosch, G. M. Fuhler.

Biochim Biophys Acta Rev Cancer. 2019

1. DEVELOPMENT OF GASTROINTESTINAL CANCERS

Gastrointestinal (GI) cancers remain a burden on society. They contribute to 26% of the global cancer incidence and 35% of all cancer-related deaths [1]. Moreover, rates of GI cancer will increase globally due to the aging and growth of the world population: by 2040, number of new cases will rise by a predicted 58% and deaths by 73% respectively [1]. The most common malignant conditions of the GI tract are cancers of the colorectum (1.8 million cases), stomach (approximately 1.0 million new cases in 2018), liver (840,000 cases), esophagus (570,000 cases), and pancreas (460,000 cases) [1].

In the GI tract and other locations, cells are considered neoplastic when they develop the well-defined characteristics which Hanahan and Weinberg described as “the hallmark of cancer”: maintaining proliferative signaling (even in the absence of extracellular signals), evading suppressors of cell growth, resisting cell death, enabling replicative immortality, inducing angiogenesis, initiation of invasion and metastasis, reprogramming of energy metabolism and avoiding immune destruction [2, 3]. Cells gain these features, necessary for carcinogenesis, in a stepwise evolutional process which is comparable to Darwinian natural selection but at micro-level [4, 5]. Similarto whole organisms, neoplastic cells compete for restricted space and resources of the microenvironment, limiting the growth of a solid tumor at every stage of its progression. Most neoplastic cells might either die before being able to proliferate or their micro-environment restricts their proliferation, as solid neoplasms (even aggressive ones) double in size much slower than individual neoplastic cells (months vs days) [6]. When genetic or epigenetic alterations in individual neoplastic cells provide them with a competitive advantage in a specific micro-environment, those cells undergo a positive selection, or in case of a competitive disadvantage – negative selection [7]. As a result, clones with a combination of favorable alterations (bringing them a competitive advantage) proliferate and expand [5, 8]. This process is termed a clonal evolution of cancer [5, 8].

Although most cancers, including those of the GI tract, arise from individual single-cells by clonal evolution, tumors form complex adaptive systems with a high level of genetic and phenotypic heterogeneity and diversity [7]. Phenotypic heterogeneity can occur not only due to genetic alterations but also as result of stochastic events in gene expression and protein stability, epigenetic divergence, and micro-environmental variations [7, 9, 10]. As prominent illustration, genetically homogeneous sub-clones of colorectal cancer may respond differently to chemotherapy [9, 11]. Genetic heterogeneity in cancer is a result of high frequencies of DNA mutations (amplifications, deletions, translocations, and other structural changes) induced by genomic instability [12]. Thus, although mutation itself is a random and undirected event, the accelerated mutagenesis is typical for cancer. Cancers generally have from 10,000 to 100,000 somatic mutations and epigenetic changes as identified with sequencing technologies [13]. Most of them are neutral non-transforming mutations (passengers), but some are driver mutations leading to cancer initiation or progression [14, 15]. Driver mutations belong to one of two types: (a) gain-of-function mutations activating pathways of proliferation and growth; (b) loss-of-function mutations inactivating tumor suppressors and apoptotic pathways [16]. For instance, the activating mutation of the proto-oncogene *KRAS* and the inactivation of tumor suppressor genes such as *CDKN2A*, *TP53*, *SMAD4*, and *BRCA2* are the most frequent genetic mutations in pancreatic cancer

[17]. Driving epigenetic alterations can lead to both aberrant gene expression and silencing and, thus, epigenetic alterations are of great importance for carcinogenesis, although they are less studied [16]. Identifying driver mutations and epigenetic alterations is a key step to understand tumor biology and develop targeted therapies [15].

GI neoplasms can be sporadic, occurring due to exposure to carcinogens or chronic inflammatory conditions, or be inherited [18-21]. Most GI neoplasms are sporadic in nature and their frequencies increase with age [22]. Thus, it is tempting to speculate that such neoplasms emerge because (future cancer-initiating) cells during decades accumulate random DNA mutations that eventually result in neoplastic transformation [8, 23]. Nevertheless, multiple observations are reporting that certain exposures, habits, diets, and backgrounds increase or decrease the risk for certain sporadic GI cancers, furthermore, genetic polymorphisms may increase susceptibility to dietary and environmental components [22]. Thus, some practices might be preventative. Chemical carcinogens also play an important role, as during ingestion GI epithelium is exposed continuously to potentially noxious chemicals [24]. For instance, esophageal squamous cell carcinoma (ESCC) can develop due to alcohol consumption and tobacco use, a diet without fruits and vegetables, high nitrate consumption, or even ingestion of very hot food and beverages, etc. [25, 26]. Different processes in GI tract can induce chronic inflammation including infections such as *Helicobacter pylori* (chronic gastritis) or Hepatitis B and C viruses (chronic hepatitis), recurrent chemical (gastroesophageal reflux disease – GERD) or enzymatic injury (recurrent pancreatitis), as well as autoimmune processes (Crohn's disease, ulcerative colitis). Each of these chronic inflammatory conditions increases the risk for cancer in individuals with the specific disorder [27, 28]. The contribution of chronic inflammation to carcinogenesis is discussed below for esophageal cancer in more detail.

Compared to sporadic, inherited GI tract neoplasms are generally uncommon and develop at a younger age [29]. Gene variants increasing the risk of neoplasm are cancer predisposition genes or cancer susceptibility genes [29]. Specific germline mutations in cancer predisposition genes have been identified for the most common familial cancer syndromes in humans [29]. In total, mutations in more than 100 human genes increase the risks of cancer development at least 2 fold, with at least 5% penetrance, from which about 52 genes play roles in GI carcinogenesis [29]. Most mutations in cancer predisposition genes inactivate the function of tumor suppressors, but gain-of-function mutations in for instance the oncogenes *RET, MET, KIT* and *ALK* activate kinases which promote cancer predisposition and malignancy [30]. The discovery of cancer predisposition gene mutations substantially impacted clinical practice, as surveillance and/or risk-reducing measures can be implemented to mitigate or prevent cancer in high-risk individuals [29]. Testing for mutations in cancer predisposition genes can identify people who will benefit from surveillance programs and those who do not have a known familial mutation and, thus, save cost for screening for familiar cancer, prevent interventions and reduce anxiety in these individuals [29]. Interventions usually include surgical removal of the at-risk tissue were possible in individuals at very high risk, such as the stomach in *CDH1* mutation carriers and the colon in *APC* mutation carriers [31-33]. Although rare, chemoprevention can take place: daily aspirin significantly reduces the risk of colorectal cancer in carriers of mismatch repair gene mutations [34].

Mortality of GI cancers is often close to the incidence due to late-stage of most diagnoses and therapy-resistance of cancers [22, 35]. The exception in this trend is colorectal cancer, for which both incidence and mortality are declining as a consequence of advances in both treatment and early detection. Specifically, because national screening programs in many countries detect and surgically remove precancerous polyps with fecal occult blood testing, sigmoidoscopy, and colonoscopy [36, 37]. Colonoscopy is included in guidelines and becoming the gold standard for colorectal cancer screening [38]. This positive result in colorectal cancer emphasizes the importance of early detection and accurate diagnostic tools in oncology. In contrast, for pancreatic cancer there was no much improvement over the past years [39]. There is only notable progress with surgery and adjuvant chemotherapy for respectable tumors, the progress made in perioperative and critical care, and in the standardization of surgical techniques, thus, complications after pancreatic resections decreased and surveillance increased [39-42]. However, most diagnosed PDAC are metastatic or locally advanced, which precludes curative surgery and results in an overall 5-year survival rate of less than 10% [43]·[43-46]. Other treatments are of limited efficacy. One of the reasons for the therapy to fail, is that heterogeneity in cancer cells gives the chance to escape for variants resistant to treatment. Those dormant resistant variants may start to expand when treatment massively kills competing cells [6]. Furthermore, cells surviving genotoxic treatment may have mutated further, and have improved malignant potential. Another reason of the PDAC resistance to therapeutics is an enrichment of PDAC with stromal and inflammatory factors influencing drug pharmacokinetics (e.g. drug accessibility, half-life). Thus, we cannot yet effectively control or eradicate advanced or metastatic malignancies [6]. Detection in a timely manner of GI cancer is a very important strategy to improve the outcomes along with the prevention and development of novel remedies. An essential step for this is to enable prediction of cancer behavior through understanding the key molecular processes underlying carcinogenesis. Nevertheless, the underlying molecular process of GI carcinogenesis are likely to be highly complex and vary greatly between cancers and individuals. The presence of mutations in cancer susceptibly genes by itself might be not transforming and the second normal allele must be inactivated or silenced. Besides cancer susceptibility genes, many other genetic components might be important and undiscovered (e.g common variants with small effects or mosaic mutations, particularly in individuals with multiple cancers) [47-49]. Genetic and environmental factors or stochastic events interact during cancer development. One type of GI cancer can be caused by multiple factors (e.g. genetic, environmental) and *vice versa* a mutation in a single susceptibly gene may lead to cancers in different organs. Thus, while of great importance for prevention and treatment, discerning the mechanisms underlying GI cancers has been difficult.

2. PANCREATIC LESIONS

2.1. Early detection of pancreatic cancer

In line with other cancers, GI carcinogenesis is a multistep process with cancers typically growing from benign dysplastic lesions. They are commonly apparent histologically and classified based on specific

pathologic criteria [50-52]. For instance, most pancreatic ductal adenocarcinoma (PDAC) arises from pancreatic intraepithelial neoplasia (PanIN), a minority of PDAC arises from intraductal papillary mucinous neoplasms (IPMNs), mucinous cystic neoplasms (MCNs), or intraductal tubular papillary neoplasms (ITPNs) [51, 52]. Normal pancreatic duct tissue develops into PanIN-1A lesion, then to PanIN-12 lesion, PanIN-2 lesion, PanIN-3 lesion, and eventually to invasive pancreatic carcinoma. Similarly, IPMN and MCN gradually transform to low-grade dysplasia, intermediate dysplasia, high-grade dysplasia and invasive pancreatic carcinoma [53]. As dysplastic lesions are prone to malignant transformation, for many of these, surveillance programs were implemented.

While a minority of pancreatic cancers represent neuroendocrine tumors, most are PDAC (85%) and this will be the tumor type further discussed in this chapter. The guidelines for surveillance for pancreatic cancer are created by multidisciplinary teams of experts and discuss the next questions regarding the surveillance: what is the main goal of the surveillance, who should be screened, at what age and with what tests [54]. Experts may re-evaluate the guidelines when we accumulate unbiased data about the long-term outcomes for high-risk individuals participating in pancreatic surveillance programs, the potential harms of over-diagnosis, cost-effectiveness, or develop more accurate diagnostic tools. The International cancer of the Pancreas Screening Consortium agreed in 2018 that the primary goal of pancreas surveillance is to detect and treat high-grade dysplasia before it evolves and gains the ability to metastasize or to detect and treat T1N0M0 pancreatic cancer. This is in an effort to prevent pancreatic cancer and death from it [54]. Multidisciplinary teams should perform surveillance in a research setting in medical centers with appropriate expertise [54].

Screening for pancreatic cancer should be performed in selected high-risk individuals rather than in a general population of a specific age as is currently performed for colorectal cancer (usually above 55) [55]. Otherwise, we may expect a high false-positive rate because pancreatic cancer is less prevalent compared to colonic and diagnosis is less accurate. In individuals with false-positive screening results, further invasive tests and unnecessary surgeries will lead to high cost and morbidity [55]. Thus, it is important to use risk stratification and accurately identify individuals that will benefit from the surveillance programs. This also highlights how critical is to understand the risk factors of pancreatic cancer.

The major criteria to determine eligibility for pancreas surveillance are family history and germline mutation status, detected potentially malignant cysts, and age [54]. Other risk features, such as diabetes, chronic pancreatitis, obesity, smoking status, other cancer family history could potentially can be helpful in the future along with gene variants identified through genome-wide meta-analysis and circulating biomarkers, particularly if risk models could be developed and validated [55]. Clinical features of pancreatic cancer are relatively broad and non-specific [17]. Those that occurs most frequently at the time of diagnosis include abdominal pain (40–60%), abnormal liver function tests (~50%), jaundice (~30%), dyspepsia (~20%), nausea or vomiting (~16%), back pain (~12%) and weight loss (~10%) [56]. Moreover, most of those symptoms commonly present themselves at the later stage, and, thus, are useless for the risk stratification for early detection of pancreatic cancer.

A *first* cohort which is eligible for the present pancreas surveillance are the carriers of germline deleterious variants in cancer susceptibility genes [57-60]: *BRCA2, ATM, BRCA1, PALB2, CDKN2A, STK1, MLH1* and *MSH2.* Recommendations for age and family history vary by gene and start no earlier than age 50 or 10 years earlier than the youngest relative with pancreatic cancer [61]. Surveillance for *CDKN2A* and *STK11* (Peutz-Jegher syndrome) mutation should be irrespective of patient family history of pancreatic cancer, because of their high lifetime risk. *Second,* pancreas surveillance is recommended for individuals with strong family histories of pancreatic cancer with no relevant germline mutation known (80% of familiar pancreatic cancer) [60]. Having 1 first-degree relative with PDAC increases the risk of developing the disease up to 2- to 5-fold and having 2 first-degree relatives with the disease increases the risk to 6.4-fold [62].

A *third* cohort eligible for pancreases surveillance includes individuals with high-risk pancreatic cystic lesions. The broad use of high-quality imaging techniques increased substantially the incident detection rate of pancreatic cysts to up to 45%. Individuals with familiar history and germline mutations commonly have pancreatic abnormalities: up to 50% will have pancreatic cysts depending on age and other risk factors; many also have subtle non-specific EUS parenchymal abnormalities, only a minority will develop concerning lesions [63, 64]. Most cystic lesion are non-neoplastic, thus should not be treated or watched if asymptomatic: lymphoepithelial cyst, mucinous non-neoplastic cyst, enterogeneous cyst, retention cyst/dysontogenetic cyst, peri-ampullary duodenal wall cyst, endometrial cyst, congenital cyst (in malformation syndromes); non-epithelial non-neoplastic pancreatitis-associated pseudocyst and parasitic cyst [65]. But some cysts harbor malignant potential and can evolve to cancer asymptomatically such as: IPMN, MCN, serous cystic neoplasm, serous cystadenocarcinoma, cystic neuroendocrine tumor G1−2, acinar cell cystadenoma, cystic acinar cell carcinoma, solid pseudopapillary neoplasm, accessory-splenic epidermoid cyst, cystic hamartoma, cystic teratoma (dermoid cyst), cystic ductal adenocarcinoma, cystic pancreatoblastoma, cystic metastatic epithelial neoplasm; some benign non-epithelial neoplasm (e.g. lymphangioma), malignant non-epithelial neoplasms (e.g., sarcomas) [65-67]. IPMNs comprise 24–82% of these cysts. Mucin produced by proliferative papillary epithelial cells of IPMN dilate the duct [67]. Depending on localization and extent of dilation, there are three types of IPMNs: main-duct, branch-duct and mixed-type IPMN. Every subtype exhibits a certain risk of malignancy and requires a specific therapeutic approach based on imaging characteristics of IPMN. For instance, main duct diameter ≥ 10 mm is considered an absolute indication for surgical removal of the IPMN [68].

At the moment, the possible modalities of imaging for the detection of pancreatic cancer are abdominal ultrasound, endoscopic ultrasound (EUS), endoscopic retrograde cholangiopancreatography (ERCP), CT, magnetic resonance imaging (MRI)/magneticretrograde cholangiopancreatography (MRCP), and positron emission tomography (PET). Of these, most pancreas surveillance protocols for high-risk individuals recommend pancreatic imaging with MRI/ MRCP and EUS due to the highest sensitivity [54, 55, 69-71]. The reported accuracy for identifying the specific type of pancreatic cystic lesions is between 40% and 95% for MRI/MRCP comparatively to 40% and 81% for CT [72-75]. MRI/MRCP is very sensitive for identifying whether a patient has one or more PCN, with the latter indicative of a diagnosis

of multifocal side-branch IPMN [74, 76, 77]. EUS can identify subtle non-specific parenchymal abnormalities [78-80]. However, EUS diagnostic yield is highly operator dependent [81]. CT is less preferable as high-risk individuals may require lifelong imaging follow-up and repeated exposure to ionizing radiation following CT increases the risk of malignancy, but it is an alternative option for individuals unable to have MRI or EUS. Abdominal ultrasound, although minimally invasive and easily available, is not accurate in identifying the pancreatic lesions, with sensitivity usually below 70% due to the location of the pancreas in the retroperitineum [82].

Screening for early-stage pancreatic cancer remains extremely challenging because diagnostic differentiation of the pancreatic lesions is imperfect. At the same time, it is of great importance for decision-making as possible steps after detection of pancreatic lesions are: discharge patients with a cyst without malignant potential, follow up of patients with cyst with low-grade dysplasia, or resect cyst high-grade dysplasia with following consequences in cost of surveillance and outcome for the patient [83]. Currently obtained results are suboptimal as about 25% of resected pancreatic lesions would never have progressed to cancer according to pathology [84]. Moreover, it is difficult to find evidence of high-grade dysplasia in microscopic PanIN lesions, which are multifocal, flat, or papillary lesion arising in the small intralobular pancreatic ducts [17, 85]. Cells of PanIN may be columnar to cuboidal, they produce varying amounts of mucin and show in cytology varying extend of atypia [86]. PanIN usually are <5 mm and not detectable by imaging, which is a major issue as most PDAC originate from PaniN, in case of both sporadic and inherited/familiar disease [17]. Furthermore, pancreatic cancer may metastasize at sizes <1 cm [87], which makes detecting smaller lesions of paramount importance. Another important contributor to diagnostic failure is a rapid progression of PDAC. Stage I pancreatic cancers can progress to stage IV disease within 1 year [88], which also may explain why pancreatic cancers keep being diagnosed despite the annual surveillance, even when concerning lesions (worrisome features or solid lesions) were absent on prior scans [65]. Thus, we urgently need to develop non-invasive, accurate diagnostic tools (e.g., biomarkers) for diagnosis of high-grade dysplastic pancreatic lesions.

The main current molecular markers for pancreatic cancer are serum levels of carbohydrate antigen 19-9 (CA19-9) and carcinoembryonic antigen (CEA) which have sensitivities for pancreatic cancer of 70–80%, 30–60% respectively [89, 90]. However, these markers might not be positive until pancreatic cancer reaches an advanced stage. In addition, the false-positivity rate for CA19-9 is relatively high at 20–30% [91, 92]. CA19-9 could have diagnostic value in individuals in whom the pre-test probability of pancreatic cancer is significant, although this question requires further investigation. Current consensus is that level of CA19-9 should be determined in individuals with suspected pancreatic cancer, for example when worrisome features are found on pancreatic imaging [54]. Additionally, experts recommend testing possible new-onset diabetes in high-risk individuals (measure fasting glucose or HbA1C) [54] because epidemiological studies show that 0.4% to 0.8% of patients with new-onset diabetes aged ≥50 will be diagnosed with pancreatic cancer within 3 years [93-95]. However, there is no evidence yet that patients would benefit from the testing [54] as when glucose levels reach diabetic levels, pancreatic tumor diameter may already be of 1.6–2.5 cm [96]. Moreover, heterogeneity of pancreatic cancer within the tumor and between individuals is very high [97]. Consequently, single

biomarkers will probably not have high sensitivity for detection of pancreatic cancer, and robust panels of biomarkers will be required.

Potential candidate sources of biomarkers for PDAC detection are serum, urine, saliva, and pancreatic juice (PJ), fine-needle biopsy (FNB) and fine-needle aspiration (FNA). While PJ are relatively less studied, we consider PJ is a promising candidate biomarker source as it is in direct contact with the pancreatic ductal epithelial lining from which PDAC arises. PJ potentially contains information from all tumour clones present. Thus, compared to blood, PJ markers may be more pancreas-specific. PJ secretion can be stimulated by intravenous secretin administration and collected from the duodenal lumen during EUS. In contrast with cytology and histology based on tissue biopsies, collection of PJ from the duodenal lumen is less invasive. In addition, tissue sampling relies on a visible mass, while PJ likely enables detection of invisible lesions.

2.2. Treatment of pancreatic cancer

For patients with advanced disease, systemic chemotherapy remains the mainstay of treatment and includes FOLFIRINOX (5-fluorouracil, folinic acid [leucovorin], irinotecan, and oxaliplatin) and gemcitabine plus nab-paclitaxel [17]. Despite an increasing understanding of the genetic, epigenetic and metabolic complexity in cancer, and the importance of the interaction of cancer with surrounding cells (e.g., stromal cells, immune cells and endothelial cells), we have not yet improved dramatically the overall outcome for patients with PDAC [98]. There is urgent need in better treatment options for resistant pancreatic cancer [98]. Radiation, chemotherapy, immunotherapy, and the use of targeted drugs just mildly increase survival rate and reduce cancer-related symptoms [99]. Chemotherapy agent gemcitabine increases median survival to 6 months of patients with advanced stage tumors while median survival for untreated patients of about 3 months [100, 101]. The improvement in survival is rarely beyond 6 months in numerous randomized controlled trials for novel chemotherapy agents and their combinations. Thus, new approaches and findings in diagnosis and cure are very valuable.

Recently, new molecule metavert was designed and synthesized for PDAC management [102]. It is a dual inhibitor targeting both glycogen synthase kinase 3 beta (GSK3B) and histone deacetylases (HDACs), enzymes relevant for PDAC progression [103, 104]. Inhibition of HDAC prevents epithelial-to-mesenchymal transition of cancer cells which may still occurs when only GSK3B is inhibited [105]. Metavert increases survival in mouse model, slows tumor growth, prevent tumor metastasis, decrease tumor infiltration by tumor-associated macrophages, and decreases blood levels of pro-inflammatory cytokines [102]. Moreover, metavert has synergistic effects with gemcitabine. Interestingly, an upregulation of β-catenin was observed after treatment with metavert, suggesting a role for Wnt3a signaling in the cytotoxic effect of metavert in line with study where excessive activation of Wnt/β-catenin signaling was toxic for cells of KRAS-dependent tumors [106]. However, Wnt/β-catenin signaling is a known driver of many malignancies [107-110]. Thus, what role plays activation of Wnt/ β -catenin signaling by metavert – pro-cancerous or anti-cancerous – is not clear.

2.3. Cellular fate in pancreatic cancer

One specific cell acquiring the cancer-promoting mutations can be considered as the cancer-initiating cell or cell of origin [111]. Distinct cells of origin may initiate different histological subtypes of cancer within the same organ [111]. It should be noted however, that the cell of origin, the normal cell, is not necessarily related to the cancer stem cell or cancer-propagating cells, the subset within the tumor that uniquely maintains malignant growth [111]. Lineage-tracing studies in mouse genetic models have identified probable cells of origin of intestinal, basal cell carcinomas, as well as PDAC [111]. For colonic cancers, potential cells of origin are two types of $LGR5^+$ stem cells having different position in the crypt [111]. These gave rise to tumorigenesis when scientists manipulated their WNT signaling cells [112, 113] (constitutive WNT signaling pathway activation causes the vast majority of colorectal cancer [114]). Similarly, $LGR5^+$ stem cells seeded small adenomas in stomach and are likely to be the target population for WNT-driven tumorigenesis there [115]. In pancreas, however, PDAC and PanIN have a ductal morphology, suggesting that they develop from pancreatic duct cells [116], while transdifferentiated acinar cells could grow into PanIN [116, 117], and insulin-positive endocrine cells and PDX1-expressing progenitor cells into PDAC under activation of an oncogenic Kras [116]. Thus, pancreatic carcinogenesis seems to be context-dependent and complicated. While the cell of origin of GI cancers remains often unknown, the identification of them may allow earlier detection of malignancies and better prediction of tumor behavior because activation of the same oncogenic pathway in different cellular compartments or contexts may greatly affect malignant potential [111, 118].

3. ESOPHAGEAL CANCER AND PREMALIGNANT LESIONS

The normal esophageal mucosa consists of a nonkeratinizing, stratified squamous epithelium, lamina propria, and muscularis mucosae [119]. Gastroesophageal reflux is a normal physiological process in humans, occurring after a meal. In addition to gastric acid, the refluxate contains pepsin, bile, pancreatic enzymes, ingested foods and their metabolites [120]. Anti-reflux and tissue resistance mechanisms are in place to protect the esophageal mucosa against these abrasive fluids [121]. However, these physiological defense mechanisms may no longer be sufficient when poor closure of the LES occurs and a subsequently increased frequency of gastroesophageal reflux causes gastroesophageal reflux disease (GERD) [121]. Esophageal epithelial barrier function is disrupted in GERD, with decreased expression of tight junction proteins resulting in increased barrier permeability compared with healthy subjects [122]. GERD, itself characterized by squamous hyperplasia, elevated presence of intraepithelial inflammatory infiltrate, epithelial cell necrosis and lack of surface maturation, is a precursor to Barrett's esophagus (BE). BE is characterized by the replacement of normal squamous epithelium of the lower esophagus with metaplastic columnar epithelium. This transformation is called intestinal metaplasia, and poses a risk factor for the development of esophageal adenocarcinoma (EAC). The annual EAC incidence rate in BE cohorts varies from 0.12 to 3.55% in different studies [123] while the global EAC incidence rate is 0.7 per 100,000 but varies greatly across countries [124].

The metaplastic columnar epithelium of BE appears to be more resistant to reflux-induced injury than the native squamous cells [125] making it tempting to speculate that selection pressure contributes to the development of BE. It is composed of mucinous columnar epithelial cells arranged in surface and crypt epithelia, and contains a variable number of scattered goblet cells, enterocytes, Paneth cells, endocrine cells, and cells with combined gastric/intestinal or intestinal/squamous-cell features [126].

Another precursor lesion for esophageal cancer is achalasia. Achalasia is an uncommon motility disorder of the esophagus characterized by impaired esophageal peristalsis and reduced lower esophageal sphincter (LES) relaxation [127] leading to the impeded flow of ingested food and secretions from the esophagus into the stomach [128]. Current evidence suggests that an initial inflammation in the myenteric plexus causes an autoimmune response in genetically susceptible individuals. This results in a degeneration of the myenteric ganglion neurons that control esophageal motility [127]. As achalasia is a disorder with poorly studied etiology, available treatments aim to alleviate symptoms by diminishing the LES pressure [127]. Effective treatment of achalasia may prompt significant sphincter insufficiency, resulting in GERD and its complications such as chronic inflammation of the esophagus (esophagitis) and BE [129]. Treatment of achalasia is associated with an increased EAC risk, with an incidence of 21.23 (StDev31.6) cases per 100,000 patient-years at risk compared to 3.2 cases/100,000 patient-years in the general population in this study [130].

On the other hand, in suboptimal treated or non-treated achalasia patients, bacterial overgrowth and chemical irritation from the ongoing decomposition of food and saliva can also lead to chronic hyperplastic esophagitis and malignant transformation of esophageal epithelial cells to esophageal squamous cell carcinoma (ESCC) [131]. A recent review and meta-analysis determined the risk of ESCC in achalasia patients to be 312.4 (StDev 429.16) cases per 100,000 patient-years at risk, compared to 4.3 cases/100,000 patient-years in the general population in this study [130]. Worldwide, ESCC accounts for around 90% of the 456,000 cases of esophageal cancers seen each year, with a global incidence rate of 5.2 per 100 000 [124]. Also ESCC in BE patients can occur, although very rare [132]. Thus, achalasia is a rare esophageal disease that increases the risk of development of two types of cancer. Achalasia can progress to ESCC or BE and then to EAC.

The majority of patients with esophageal carcinoma have a poor prognosis as they are often diagnosed at advanced stages, no longer eligible for curative surgery. Approximately 80% of patients are inoperable at initial diagnosis [133]. Radiotherapy, chemotherapy or chemoradiotherapy in both EAC and ESCC are less efficient resulting in a 5-year survival rate of 19% for esophageal cancer and only 0.9% for advanced esophageal cancer [134]. A better insight into the molecular pathways governing esophageal cancer development in achalasia and BE may be of use to identify patients at risk, better inform patients on associated neoplastic progression risk after dilatation treatment aiming to improve surveillance and treatment strategies. Both precursor lesions for esophageal cancer, achalasia and BE, are associated with chronic inflammation that can contribute to neoplastic transformation, however, BE is relatively well characterized in terms of genetic alterations, molecular pathways and microbiota changes.

Specific risk factors for GI cancers vary across world regions and this variation may determine the prevalent type or particular histological subtype of cancer in the particular region. For instance, ESCC is prevalent in Asia while esophageal adenocarcinoma (EAC) is more common in the Western countries, partly because polymorphism of *Glu504Lys*, which decreases aldehyde dehydrogenase (ALDH2) activity, is more common in Eastern countries [135-138]. ALDH2 is of importance as it detoxifies common mutagenic and carcinogenic acetaldehyde in the liver, which is an alcohol metabolite but can also be produced by oral microbiota or ingested with food [139-143]. Furthermore, ESCC incidence rates are declining, but rates of EAC are rising likely by the reason that smoking and alcohol abuse relevant for ESCC are decreasing while obesity rates predisposing to EAC is increasing [144]. In the future, EAC may surpass ESCC in many mainly high-income countries [144].

For the esophagus, Barrett's intestinal-type metaplasia and squamous dysplasia of the esophagus and achalasia are recognizable tissue abnormalities that occur before of EAC and ESCC respectively [145-147]. Thus, these lesions may warrant surveillance. Current guidelines recommend screening for BE in men with 5-year history of GERD and minimum two other risk factors (e.g. age over 50, obesity, smoking history, Caucasian ethnicity, or family history of BE and in women only when multiply risk factors are present. When BE is diagnosed, surveillance for EAC in non-dysplastic BE should be performed every 3 to 5 years, while in BE with low-grade dysplasia every 12 months [148-150]. BE with high-grade dysplasia should be resected. Guidelines for screening of squamous dysplasia are available for regions with high prevalence of ESCC as in Northern China with 1 screening endoscopy at age 50 years or 3 screening endoscopy at age 40 depending on region [150]. Surveillance for low low-grade dysplasia should be performed every 5-year follow-up, for high-grade dysplasia every 3-years.

3.1. Cellular fate in BE, EAC and ESSC

There is a lot of debate regarding the cell of origin of BE. So far, six hypotheses have been suggested: 1) transdifferentiation of esophageal squamous epithelial cells, 2) expansion of submucosal glandular epithelium, 3) expansion of gastric cardia cells, 4) differentiation of circulating bone marrow cells, 5) expansion of residual embryonic cells located at the squamous-columnar junction, 6) p63+KRT5+KRT7+ basal cells in a transitional zone between the epithelium of the esophagus and cardia [151] (**Table 1, Figure 1**).

The suggestion is made that whatever the cell of origin, this is also the origin of the subsequent progression to EAC. However, while it is clear that EAC originates from glandular cells near the stomach, BE consists of many cell types, and thus it remains uncertain which cells are the main drivers of EAC. In contrast, ESCC is derived from squamous epithelial cells and appears to be driven by carcinogenic environmental influences, but may also go through dysplastic precursor lesions (Figure 1) [152]. Squamous dysplasia is characterized by the presence of nuclear atypia (enlargement, pleomorphism, and hyperchromasia), loss of normal cell polarity, and abnormal tissue maturation without invasion of epithelial cells through the basement membrane. Over the full follow-up period, ESCC developed in 8% of participants with normal histology but 24% with mild dysplasia, 50% with

moderate dysplasia, 74% with severe dysplasia, 58% with dysplasia NOS, and 75% with carcinoma in situ [153]. Cell types present in the esophageal mucosa in described esophageal disorders are summarized in **Table 1**.

3.2. The possible role of chronic inflammation in the progression of BE toward esophageal cancer

Chronic inflammation is associated with an increased risk of malignant disease. Around 20% of human cancers are related to chronic inflammation caused by infections, exposure to irritants or autoimmune disease [156]. Inflammation may contribute to cancer development through numerous mechanisms, including DNA damage, angiogenesis, promotion of cellular proliferation, and inhibition of apoptosis [157]. Indeed, inflammatory conditions of the esophagus, specifically reflux esophagitis and BE, have been implicated in the development of esophageal adenocarcinoma [158]. Metaplasia can be accompanied by acute and chronic inflammation of the lower esophagus resulting in increased release of proinflammatory mediators [159]. Key mediators connecting inflammation and BE carcinogenesis include ROS, NFκB pathway activation, inflammatory cytokines, prostaglandins, and immune modulatory microRNAs [159]. For instance, IL-1β, a pleiotropic pro-inflammatory cytokine upstream of inflammatory IL-6 and TNF-α signaling cascades, is overexpressed in BE. Clinical studies have suggested that polymorphisms in the *IL-1β* gene cluster are associated with BE, suggesting that genetic factors predisposing for altered immune regulation contribute to BE susceptibility [160]. In addition, inflammation markers, particularly C-reactive protein and IL-6, were proposed as potential markers for patients with a higher risk of progression to EAC [161]. Furthermore, expression of TNF-α as well as its receptor TNF-R1 are progressively increased from normal squamous mucosa to BE and EAC [162]. The inflammatory link with esophageal adenocarcinoma is further strengthened by the observation that regular use of nonsteroidal anti-inflammatory drugs and aspirin is correlated with decreased risk of cancer development [163].

To what extent inflammation plays a role in achalasia and its progression to ESCC or BE-EAC is less clear. Histological analysis of the full-layer mucosa in early and advanced achalasia showed that inflammation was present in early achalasia, but histological esophagitis with findings of increased inflammatory cell infiltration and dilated intercellular spaces were also observed in patients with late achalasia [155]. Furthermore, in patients with end-stage achalasia, the squamous mucosa is consistently altered compared with control specimens and closely resembles that seen in GERD with different grades of esophagitis [164]. Thus, chronic esophagitis is present in achalasia patients, and as this is the main risk factor for ESCC development, it may also contribute to the increased risk of ESCC in patients with achalasia [165].

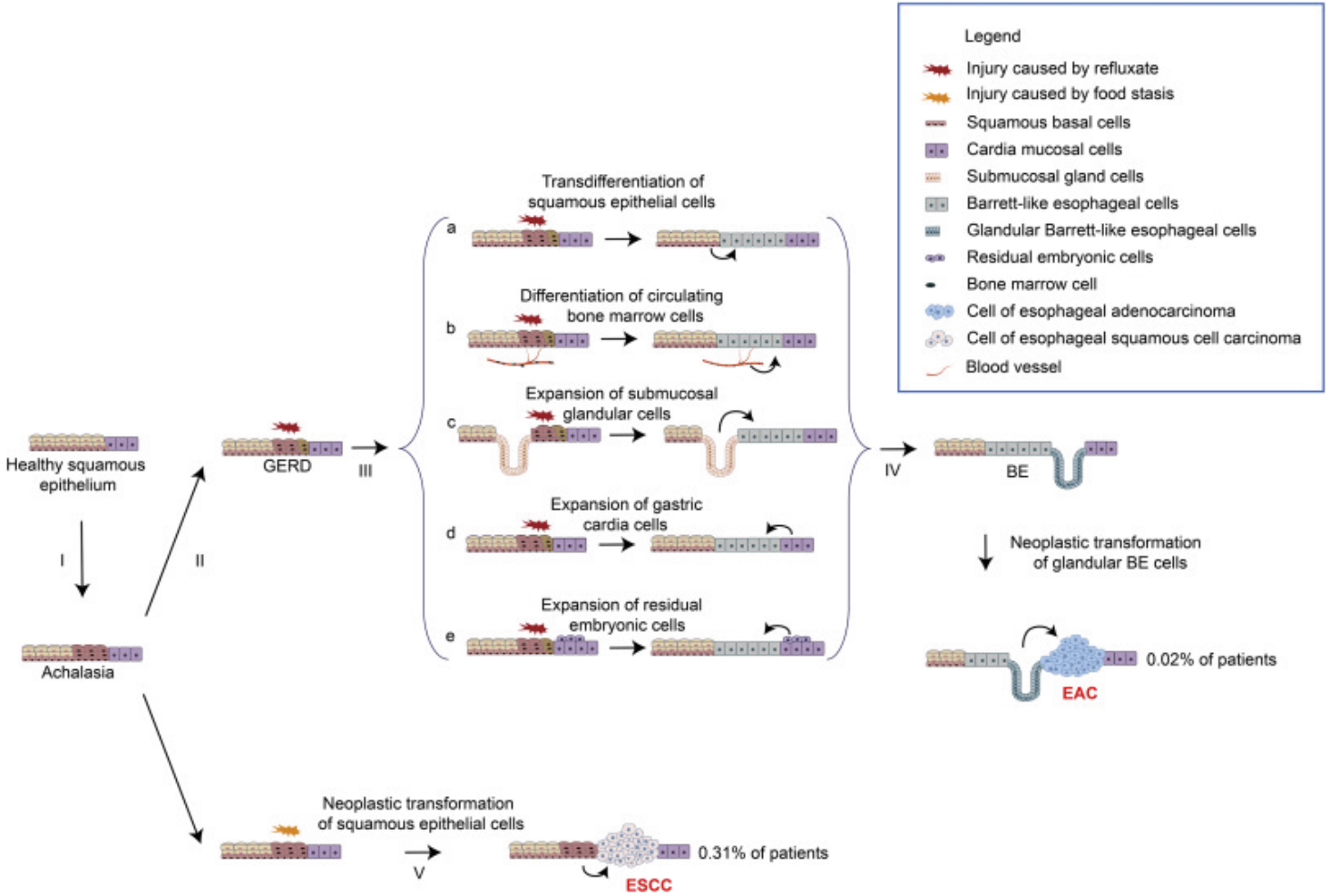

Legend
Injury caused by refluxate
Injury caused by food stasis
Squamous basal cells
Cardia mucosal cells
Submucosal gland cells
Barrett-like esophageal cells
Glandular Barrett-like esophageal cells
Residual embryonic cells
Bone marrow cell
Cell of esophageal adenocarcinoma
Cell of esophageal squamous cell carcinoma
Blood vessel
Transdifferentiation of squamous epithelial cells
a
Differentiation of circulating bone marrow cells
b
Expansion of submucosal glandular cells
c
Expansion of gastric cardia cells
d
Expansion of residual embryonic cells
e
Healthy squamous epithelium
I
Achalasia
II
GERD
III
IV
BE
Neoplastic transformation of glandular BE cells
0.02% of patients
EAC
Neoplastic transformation of squamous epithelial cells
V
0.31% of patients
ESCC

Figure 1. Cells of origin of Barrett's esophagus (BE), esophageal adenocarcinoma (EAC) and esophageal squamous cell carcinoma (ESCC) in achalasia patients (extended from Jiang et al. 2017 [154]. I. The esophageal epithelium of achalasia patients is different from normal squamous epithelium: it is inflamed, has dilated intercellular spaces and increased infiltration of inflammatory cell [155]. II. Treatment of achalasia can lead to gastro-esophageal reflux disease (GERD) and BE (III). Several hypotheses are suggested to explain cells of origin of BE: a) BE epithelium arises through transdifferentiation of stratified squamous esophageal epithelium; b) circulating bone marrow cells transdifferentiate to BE epithelium; c) BE arises from expanded esophageal submucosal gland cells; d) BE originates from stem and progenitor cells (Lgr5+) in the cardia mucosa; e) BE originates from quiescent residual embryonic cells (REC) at the squamous-columnar junction. IV. BE can lead to EAC wich originates from glandular cells near the stomach (0.02% of achalasia patients/year). Sporadically, ESCC can develop from BE (not shown). V. ESCC is derived from squamous epithelial cells (0.31% of achalasia patients /year).

Table 1. Different cell types present in esophageal lining of the esophagus in health and disease. **GERD:** gastro-esophageal reflux disease, **BE:** Barret's esophagus, **EAC:** esophageal adenocarcinoma, **ESCC:** esophageal squamous cell carcinoma

Heath/disease	Cell types	Disease manifestation
Healthy mucosa	Squamous epithelial cells	-
GERD	Squamous epithelial cells	Immune infiltrate; barrier defect
BE	Mucinous columnar epithelial cells; enterocytes; Paneth cells; endocrine cells; cells with combined gastric/intestinal or intestinal/squamous cell features	Immune infiltrate; spatial mislocalisation of intestinal barrier cells
Achalasia	Squamous epithelial cells	Immune infiltrate
EAC	Derives from glandular cells near the stomach	Gland-forming tumor with variable grade of differentiation (as defined by gland formation or mucinous differentiation)
ESCC	Derives from squamous epithelial cells	Squamous cell hyperproliferation with variable degree of differentiation (as defined by keratinization)

3.3. Genetic alterations and molecular pathways involved in cancer development of EAC

In order to identify possible therapeutic targets for prevention and treatment of esophageal carcinoma, the molecular pathways involved in the malignant progression from BE to EAC have received vast attention [123]. Improvements in high-throughput genomic technologies have led to a better understanding of the molecular basis underlying the development of EAC and ESCC [166] .

Analysis of gene mutations revealed that in EAC, 26 genes are frequently and significantly mutated. Among these genes are tumor suppressors such as *TP53* (72% of cases) and *p16/CDKN2A* (12% of cases) as well as bacterial recognition receptor *TLR4* mutations (6% of cases) [167]. Interestingly, BE tissue appears to be highly mutated even prior to the occurrence of dysplasia, with a mutation rate superior to many other tumors at an advanced stage of development (6.76 mutations/Mb) [168]. Thus, an accumulation of mutations appears to underlie the BE-to-EAC sequence, which is already initiated at early BE stages. This is seen for instance for *TP53* mutations, which are scarce in BE, but accumulate in EAC [169]. However, while a shared mutational context suggests that the same mutational trigger underlies both BE and EAC, it has also been shown that the mutations in BE are clonal, and the specific mutations observed in different clones do not overlap greatly with those found in EAC (for example,

mutation of *EYS*, *ARID1A,* and *ABCB1* genes was only shared in 28% of paired Barrett's and EAC samples) [168]. Furthermore, while *TP53* and *P16* mutations are homogeneously present within EACs, and appear to represent early events during carcinogenesis, clonality within EAC also exists, with loss of heterozygosity of *SMAD4* and *APC* not evenly distributed within the tumor [170]. Longitudinal genetic analysis of BE patients suggests that the number and diversity of clones within BE segments changes little over time [171], however, patients who progress to EAC during their lifetimes (<5%) develop signs of chromosome instability with gene losses and gains, genomic heterogeneity, selection of somatic chromosome abnormalities and catastrophic genome doublings [172]. Esophageal cancer development is also associated with a clear increase in copy number alterations (CNAs), which are much less frequent in BE. Some of these molecular abnormalities can be used to predict the neoplastic progression risk of BE [173]. For instance, high clonal diversity was associated with increased progression risk of BE [171].

In addition to gene mutations, altered gene transcription patterns are observed in BE and EAC. Based on this pattern, prediction models for progression have been developed with a 90-gene signature showing promise as a biomarker for low grade dysplasia in BE [174]. Within this signature, one third of genes was regulated by the proto-oncogene c-MYC, with other candidates HNF1-α, SP-1, NF-Y, E2F1, TP53, ESR1 and HIF1A following suit [174]. A recent review and meta-analysis confirmed the use of p53 immunohistochemical staining to improve risk stratification in BE surveillance [175].

3.3.1. Spatial regulation in BE

In BE, there is impairment of location-specific tissue phenotype as distal gut phenotype (intestine-like) is present in proximal gut (esophagus), which is reminiscent of homeotic transformations – when during embryogenesis one body part transforms into something that resembles another body part in transgenic animal models. Homeotic transformations might be a consequence of mutations in the molecules governing formation of proximal-distal body axis and determine location-specific identity of cells. The impairment of location-specific phenotype in BE indicates that these molecules might have a role in the pathogenesis of BE. The main regulators of a region-specific gene expression are two highly conserved families of transcriptional factors: Caudal-related Homeobox (*CDX*) genes and Homeobox (*HOX*) cluster. Although relatively well studied in BE, there is no evidence that CDXs on their own are sufficient to induce an intestinal phenotype in the esophagus [176-181].

HOX transcriptional factors control organogenesis, maintain tissue homeostasis, and are key drivers of developmental processes [182, 183]. *HOX* genes have a strong link with homeotic transformations, but we know little about them in context of BE [184]. Clustering of 4 clusters of *HOX* genes (A, B, C and D) determine their functions as the 3' to 5' sequence of the *HOX* genes in a single *HOX* cluster, i.e. paralogues, corresponds to the sequence in which the paralogues act along body axes. Thus, location of Hox genes along the chromosome corresponds with their expression patterns along the proximal-distal body axis. This process is termed collinearity. *Hox* gene expression pattern exist along the murine embryonic gut [185], and their ectopic expression in mice can interfere with intestinal organogenesis

[186, 187]. *HOX* expression gradient also exists along the adult human gut [188]. Furthermore, in BE, the mid cluster *HOXB* gene expression resembles *HOXB* expression in the colon [189], however, GERD *in vitro* model did not induce *HOXB* genes [189]. Another cluster with the potential for being implicated in BE seems to be cluster A [190-192]. Moreover, expression of HOX proteins is associated with worse prognosis for patients with cancers of upper GI tract [191, 192]. Taking together, these observations prompt further investigation into the role of *HOX* genes in the BE and EAC.

Other genes involved in regional differentiation during embryogenesis are the homeobox1 gene *BARX1* and the forkhead box genes *FOXP1* and *FOXF1.* Interestingly, a large epidemiological genome-wide association studies (GWAS) identified variants in these three genes that had association with BE and EAC [193], all three of which were confirmed in a the separate GWAS study [194].

The minor G allele of rs11789015 (9q22) is protective for BE and EAC and lies in the intron of *BARX1* and decreases its expression [195]. The *BARX1* transcription factor plays a role in differentiation of esophagus, trachea, and stomach in murine embryos [196]. In the mouse stomach, *BARX1* stimulates expression of Wnt3a antagonists such as secreted frizzled proteins thereby inhibiting Wnt/β-catenin signaling, an important pathway in embryogenesis. Barx1 loss prevents stomach epithelial differentiation and squamous cell differentiation in the esophagus [196, 197], although its expression is not at such high levels in esophagus as observed in the stomach [197]. In the human GI tract, the same locus confers risk to ESCC, which has increased expression of *BARX1* [195]. BARX1 knock out inhibited ESCC cell proliferation, migration and invasion [195]. However, *BARX1* expression and function in human GI tract, BE and EAC have not been investigated yet.

FOXP1 and *FOXF1* are transcription factors with DNA-protein and protein-protein binding domains. *FOXP1* is the nearest gene to the peak SNP on chromosome 3 (rs2687201 in 16q24) associated with BE and EAC [193]. Near *FOXF1* there are six SNPs reported in different studies to be associated with BE, EAC or BE and EAC: rs9936833, rs3111601, rs1728400, rs3950627, rs2178146, rs13332095 [195]. FOXP1 regulates lung development (together with FOXP2) and esophagus development in mouse embryos [198, 199]. Foxp2−/−Foxp1−/+ mice have defects in the muscle surrounding the esophagus with loss of both skeletal and smooth muscle development [199]. In adult mice, loss of 1 Foxp1 allele leads to a pronounced atrophy of the tunica muscularis in the esophagus and colon, accompanied by a motility dysfunction [200]. In humans, *FOXP1* is detected in various tissues and linked with cancer development [198, 201]. There was a significantly higher expression level of *FOXP1* in EAC than in the adjacent normal tissues [202], but expression of *FOXP1* in BE is not known yet.

FOXF1 is downstream of the Hedgehog signaling pathway, which is an essential determinant of foregut separation. Depletion of *FOXF1* causes a similar phenotype as a disruption of Hedgehog signaling. Mice heterozygous for a Foxf1 null allele have major structural abnormalities, including a narrow esophageal lumen, aberrant connection to the trachea (tracheoesophageal fistula) and failure of the esophagus to join to the stomach (esophageal atresia) [203]. Deletion of Foxf1 from mesenchyme caused embryonic lethality due to numerous defects in the heart, lung, liver and esophagus [204].

Deletion of Foxf1 from smooth muscle cell lineage caused hyper-extension of esophagus and trachea, loss of tracheal and esophageal muscle, mispatterning of esophageal epithelium and decreased proliferation of smooth muscle cells [204].

Thus, these genes are relatively better studied in mouse embryogenesis while little is known about expression and function in adult human GI tract. Interestingly, all three genes are critical for the formation and development of esophagus. As carcinogenesis can be seen as an aberrant form of organogenesis, it is not unexpected that these transcription factors may also contribute carcinogenic pathways. Together with GWAS data, these indicate plausible functional role of *BARX1, FOXF1, FOXP1* in the etiology of BE and EAC, but this warrants further confirmation.

3.4. Markers for development of ESCC

Interestingly, esophagectomy specimens from achalasia patients also display a heightened frequency of p53 immunoreactivity, indicative of early changes related to ESCC risk [205]. Aberrant expression of the p53 protein correlated with grade of inflammation in idiopathic achalasia [206] and increased with progressive grades of dysplasia. A recent study further showed that patients with achalasia and retention esophagitis have higher positive rates of p53 and p16 expression (a key regulator at the G1-S checkpoint in the cell cycle often deregulated in cancers) than those from achalasia patients without retention esophagitis and control groups [207]. These data suggest that achalasia-associated chronic inflammation may mediate clonal evolution by generating a mutagenic pressure or providing a selective advantage to those clones able to survive an inflammatory insult [208]. However, aside from the above mentioned *TP53* mutations, the mutation burden in the mucosa of patients with achalasia is relatively uncharacterized and the exact genetic evolution from achalasia to esophageal squamous dysplasia and ESCC remains unknown. ESCC itself is characterized by aneuploidy of chromosomes 7, 11, and 17 as well as *TP53* gene deletion [209]. Aneuploidy was also reported to be present in achalasia and chagasic megaesophagus patients, with chromosome 7 monosomy or trisomy and chromosome 17 monosomy or trisomy being the most frequently occurring aneuploidies [209], suggestive of an achalasia-to-ESCC carcinogenic sequence. Mutation of the *PIK3CA* gene was reported to be associated with Chagas disease and ESCC, but there is no evidence regarding its association with idiopathic achalasia [210].

As for EAC and BE, clonal expansion of ESCC and its premalignant lesion esophageal squamous dysplasia are implicated by their highly heterogeneous and polyclonal nature [211]. Dysplasia is heavily mutated and harbors most of the driver events reported in ESCC, with *TP53* mutations a prerequisite for progression to ESCC. However, unlike BE to EAC progression, copy number alterations are already common in dysplastic stages and persist during the ESCC progression. Whether these copy number alterations are already present at achalasia stages remains unanswered.

Despite the presence of common denominators, including the aforementioned *TP53* point mutations, studies have indicated that ESCC is genetically more similar to other squamous cancers, such as head and neck, than to EAC [167, 212]. Risk factors for ESCC include tobacco and alcohol consumption and this tumor is more common in the upper and mid-esophagus, whereas the EAC predominates in the

lower esophagus and is associated with obesity and GERD [26]. A comparison of copy number alterations as well as DNA methylation, mRNA and microRNA expression patterns between 90 ESCCs and 72 EACs revealed a clear separation between these types of esophageal cancer [213]. Although both diseases share similarly high frequencies of overall and clinically relevant genomic alterations, different genetic mutations associated with specific cellular pathways, such as cell cycle, apoptosis, DNA repair mechanisms, growth factor receptors, have been identified in esophageal squamous cell cancers (see Table 2) [212]. When achalasia progresses to EAC, the oncogenic events are likely to be different compared to the progression to ESCC. Again, the mutational sequence from achalasia to BE remains unknown, and it would be of interest to compare mutational burden in achalasia-associated EAC to EAC that is not associated with achalasia.

4. WNT/B-CATENIN SIGNALING AND GI CANCER

The Wnt/ β-catenin pathway has already been mentioned above, and deserves further consideration due to its important role in GI development, homeostasis and tumorigenesis. Disruptions in Wnt/β-catenin-dependent signaling can be a key transforming event in many GI cancers [220]. Activation of this pathway can occur 1) as result of signal transduction from Wnt receptors (FZD-Frizzled, LRP5 - low-density lipoprotein receptor-related protein 5, or LRP6) or 2) deregulation of proteolytic cleavage of the transcriptional co-activator β-catenin (e.g. mutations in negative regulators of the pathway). In the absence of Wnt stimuli, a multisubunit destruction complex captures and phosphorylates cytosolic β-catenin which drives its ubiquitination and degradation in the proteasome; thus, this prevents β-catenin signaling under basal conditions [221-223]. The destruction complex consist of the tumor suppressors AXIN1 and APC (adenomatosis polyposis coli) and the kinases CK1 (casein kinase 1) and GSK3β (glycogen synthase kinase 3β). When Wnt activates its receptors, these receptors changes their conformation, recruit the effector protein DVL (Dishevelled) and then proteins of destruction complex get recruited to the cell surface [220, 222, 224]. This in turn prevents phosphorylation and, hence, degradation of β-catenin [220, 222, 224]. As a result, β-catenin accumulates, translocates to the nucleus where it bind the TCF/LEF family (T cell factor/lymphoid enhancer-binding factor) of transcription factors and induces the transcription of Wnt target genes [225]. Furthermore, RNF43 and ZNRF3 (membrane-bound E3 ligases ring finger protein 43 and zinc and ring finger protein 3 respectively) regulate this pathway. They mediate ubiquitylation of Wnt receptors, which drives their internalization and lysosomal degradation, thereby attenuating the sensitivity of cells to incoming Wnt [226, 227].

Wnt/β-catenin pathway has a prominent role in embryogenic development, carcinogenesis, adult stem cell self-renewal and cell-fate specification in GI tract [220, 228-232]. Thus, it is of great importance, that numerous cancer suppressors tightly control Wnt/β-catenin signaling activity during tissue homeostasis [220, 221, 226, 227]. Inactivation of those cancer suppressor genes frequently occurs in various cancers [233]. In line with this, constitutive activation Wnt/β-catenin pathway provides self-renewing growth properties to cancer cells and, furthermore, might cause therapy resistance [234, 235]. Mutations in negative regulators of this signaling pathway (often in *APC*, *AXIN1* and *AXIN2*) and β-

catenin (encoded by *CTNNB1)* itself are frequent events in cancer, driving Wnt-independent tumor growth [220]. As seen from the Cancer Genome Atlas, in colorectal cancer (CRC; $n = 594$), the largest fraction of sporadic tumors accumulates *APC* mutations (67%), followed by lower fractions of *RNF43* (8%), *CTNNB1* (6%) and *AXIN2* (5%) mutations. By contrast, liver cancer cases ($n = 372$) preferentially acquire *CTNNB1* (25%) and *AXIN1* (8%) mutations, whereas pancreatic cancers ($n = 184$) favor mutations in *RNF43* (6%), and adrenocortical cancer ($n = 92$) links to mutations in *ZNRF3* (20%) or *CTNNB1* (15%) [220]. In the esophagus, there was no robust nuclear accumulation of β-catenin shown in nondysplastic BE [236] and the common mutations of the pathway's components found in other cancers, such as β-catenin and APC, are not frequently detected [237], although some transcription activity present according to the recent study [238]. Wnt/β-catenin signaling progressively increases in the metaplasia–dysplasia–carcinoma sequence of BE [236, 239, 240] provably due to mutations in Wnt inhibitory factor 1 and secreted frizzled receptor proteins, as well as induction of Wnt-2 expression, which are expected to increase signaling along the Wnt axis [239, 241, 242].

We also know that some oncogenes show expression level-dependent effects, with superabundant expression levels stimulating intrinsic tumor suppressive programs. For example, an excessive accumulation of β-catenin leads to apoptosis in normal and carcinoma cell lines [243]. Similar results are also reported for *Myc*, one of the main target genes activated by the Wnt/β-catenin signalling pathway [244]. The "*Just-right signaling*' model describing the relevance of β-catenin signaling dosage in tumor growth has become widely accepted. According to this hypothesis, each tumor type selects for an optimal level of β-catenin signaling that is ideal for tumor initiation and progression. Cells characterized by β-catenin signaling levels above a given threshold undergo apoptosis and hence will not contribute to tumor growth [245-247]. In support, it has been shown that β-catenin upregulation at least partly mediates killing effects of GSK3 inhibitors in KRAS-dependent tumors [106]. Taken together, these data suggest that increased levels of β-catenin can play either pro-oncogenic or anti-oncogenic roles.

Table 2. Dysregulated genes observed in Barrett's esophagus (BE), esophageal adenocarcinoma (EAC) and esophageal squamous cell carcinoma (ESCC). When known, frequencies are reported (adapted from: [214-219])

Gene	BE (%)	EAC (%)	ESCC (%)
Receptor tyrosine kinases			
ERBB2	01-13	32	3
EGFR	0-4	15	19
KRAS	present, frequency not reported	14	7
PIK3CA	0-4	3	13
Cell cycle regulators			
CDKN2A	30-42	76	76
CCND1	present, frequency not reported	15	57
CCNE1	present, frequency not reported	14	4
RB	0-8	0	9
Proliferation and differentiation			
MYC	present, frequency not reported	32	23
SMAD4	0	24	8
GATA4	present, frequency not reported	19	1
GATA6	present, frequency not reported	21	3
TP63 OR SOX2		11	48
Chromatin remodeling			
KMT2D	4-13	1	14
Cell death			
TP53	2.5-72	75	69
Cell adhesion, migration, cytoskeleton organization			
TTN		55	34
MUC16		31	14
SYNE1	3-4	30	11
Other mutated genes			
TLR4	13	5	1
LRP1B	0-4	25	11

OUTLINE OF THE THESIS

Better characterization, novel remedies and improved diagnostic tools for GI cancer is needed. This thesis is divided into two parts. The aim of first part is development of biomarkers for early detection of pancreatic cancer (**Chapters 2-5**). The aim of the second part is to better understand the role of embryogenesis genes in the etiology of esophageal cancer (**Chapters 5-8**).

In **Chapter 2,** we test with what technique to collect secretin-stimulated pancreatic juice from the duodenal lumen during endoscopic ultrasound to receive the highest yield of organoids and possible biomarkers for detection of pancreatic cancer. We compare protein, RNA, DNA and extracellular vesicle and organoid yield of two suction techniques and three time frames of collection. After determining the optimal method to obtain high quality PJ, in **Chapter 3** and **4** we employ this method of collection to investigate potential biomarkers for pancreatic cancer detection. In **Chapter 3** we analyze the vesicular composition of pancreatic juice of 54 individuals with pancreatic cancer and 117 controls with nanoparticle tracking analysis with the aim of determining whether size and concentration of extracellular vesicles could provide a potential biomarker tool. In **Chapter 4**, we isolate microRNA from extracellular vesicles (EV) in PJ and serum and analyze the expression of EV-miR-21, EV-miR-25, EV-miR-210 and EV-miR-16 by qPCR to investigate whether these markers in pancreatic juice can be a better biomarker source for pancreatic cancer than serum markers alone. In **Chapter 5** we turn our attention to a potential novel treatment of pancreatic cancer and investigate the activity of the Wnt/β-catenin pathway upon inhibition of GSK3 and/or HDAC and determine its role in PDAC cell cytotoxicity.

In **Chapter 6** we study *HOX* genes in esophageal carcinogenesis. To this end, we investigate HOX gene expression patterns across the GI tract and its heterotopias. In particular *HOXA13* overexpression appears to be a candidate gene to explain both the phenotype and the oncogenic potential of Barrett's esophagus. We next investigated whether this gene could potentially also be involved in driving ESCC. In **Chapter 7** we therefor overexpressed *HOXA13* in the non-transformed esophageal cell line EPC2-hTERT and found that *HOXA13* indeed may drive early stages of esophageal cancer. In **Chapter 8** we study the expression of *BARX1, FOP1* and *FOXF1* in gastrointestinal tract in physiology and in Barrett's esophagus to establish whether other positional identity genes may also play a role in driving metaplasia and progression to more malignant phenotypes.

REFERENCES

1. Arnold, M., et al., *Global Burden of 5 Major Types of Gastrointestinal Cancer.* Gastroenterology, 2020. 159(1): p. 335-349 e15.
2. Hanahan, D. and R.A. Weinberg, *The hallmarks of cancer.* Cell, 2000. 100(1): p. 57-70.
3. Hanahan, D. and R.A. Weinberg, *Hallmarks of cancer: the next generation.* Cell, 2011. 144(5): p. 646-74.
4. Nowell, P.C., *The clonal evolution of tumor cell populations.* Science, 1976. 194(4260): p. 23-28.
5. Cahill, D.P., et al., *Genetic instability and darwinian selection in tumours.* Trends Cell Biol, 1999. 9(12): p. M57-60.
6. Greaves, M. and C.C. Maley, *Clonal evolution in cancer.* Nature, 2012. 481(7381): p. 306-313.
7. Marusyk, A. and K. Polyak, *Tumor heterogeneity: causes and consequences.* Biochim Biophys Acta, 2010. 1805(1): p. 105-17.
8. Nowell, P.C., *The clonal evolution of tumor cell populations.* Science, 1976. 194(4260): p. 23-8.
9. Kreso, A., et al., *Variable clonal repopulation dynamics influence chemotherapy response in colorectal cancer.* Science, 2013. 339(6119): p. 543-8.
10. Meads, M.B., R.A. Gatenby, and W.S. Dalton, *Environment-mediated drug resistance: a major contributor to minimal residual disease.* Nature Reviews Cancer, 2009. 9(9): p. 665-674.
11. Gascoigne, K.E. and S.S. Taylor, *Cancer cells display profound intra- and interline variation following prolonged exposure to antimitotic drugs.* Cancer Cell, 2008. 14(2): p. 111-22.
12. Burrell, R.A., et al., *The causes and consequences of genetic heterogeneity in cancer evolution.* Nature, 2013. 501(7467): p. 338-345.
13. Stratton, M.R., *Exploring the genomes of cancer cells: progress and promise.* Science, 2011. 331(6024): p. 1553-8.
14. Stratton, M.R., P.J. Campbell, and P.A. Futreal, *The cancer genome.* Nature, 2009. 458(7239): p. 719-724.
15. Pon, J.R. and M.A. Marra, *Driver and passenger mutations in cancer.* Annu Rev Pathol, 2015. 10: p. 25-50.
16. Ponder, B.A., *Cancer genetics.* Nature, 2001. 411(6835): p. 336-41.
17. Mizrahi, J.D., et al., *Pancreatic cancer.* The Lancet, 2020. 395(10242): p. 2008-2020.
18. Sadikovic, B., et al., *Cause and consequences of genetic and epigenetic alterations in human cancer.* Curr Genomics, 2008. 9(6): p. 394-408.
19. Stoffel, E.M., *Heritable Gastrointestinal Cancer Syndromes.* Gastroenterol Clin North Am, 2016. 45(3): p. 509-27.
20. Blackadar, C.B., *Historical review of the causes of cancer.* World J Clin Oncol, 2016. 7(1): p. 54-86.
21. Colotta, F., et al., *Cancer-related inflammation, the seventh hallmark of cancer: links to genetic instability.* Carcinogenesis, 2009. 30(7): p. 1073-81.
22. Katona, B.W. and J.P. Lynch, *Chapter 66 - Mechanisms of Gastrointestinal Malignancies*, in *Physiology of the Gastrointestinal Tract (Sixth Edition)*, H.M. Said, Editor. 2018, Academic Press. p. 1615-1642.
23. Vogelstein, B. and K.W. Kinzler, *Cancer genes and the pathways they control.* Nat Med, 2004. 10(8): p. 789-99.
24. Chandra, S.A., M.W. Nolan, and D.E. Malarkey, *Chemical carcinogenesis of the gastrointestinal tract in rodents: an overview with emphasis on NTP carcinogenesis bioassays.* Toxicologic pathology, 2010. 38(1): p. 188-197.
25. Ferlay, J., et al., *Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012.* Int J Cancer, 2015. 136(5): p. E359-86.
26. Domper Arnal, M.J., A. Ferrandez Arenas, and A. Lanas Arbeloa, *Esophageal cancer: Risk factors, screening and endoscopic treatment in Western and Eastern countries.* World J Gastroenterol, 2015. 21(26): p. 7933-43.
27. Fox, J.G. and T.C. Wang, *Inflammation, atrophy, and gastric cancer.* J Clin Invest, 2007. 117(1): p. 60-9.
28. Ullman, T.A. and S.H. Itzkowitz, *Intestinal inflammation and cancer.* Gastroenterology, 2011. 140(6): p. 1807-16.
29. Rahman, N., *Realizing the promise of cancer predisposition genes.* Nature, 2014. 505(7483): p. 302-308.
30. Dixit, A., et al., *Sequence and structure signatures of cancer mutation hotspots in protein kinases.* PLoS One, 2009. 4(10): p. e7485.
31. Wells, S.A., Jr., et al., *Multiple endocrine neoplasia type 2 and familial medullary thyroid carcinoma: an update.* J Clin Endocrinol Metab, 2013. 98(8): p. 3149-64.
32. Seevaratnam, R., et al., *A systematic review of the indications for genetic testing and prophylactic gastrectomy among patients with hereditary diffuse gastric cancer.* Gastric Cancer, 2012. 15(1): p. 153-163.
33. Seevaratnam, R., et al., *A systematic review of the indications for genetic testing and prophylactic gastrectomy among patients with hereditary diffuse gastric cancer.* Gastric Cancer, 2012. 15 Suppl 1: p. S153-63.
34. Soualy, A., et al., *Effect of chemoprevention by low-dose aspirin of new or recurrent colorectal adenomas in patients with Lynch syndrome (AAS-Lynch): study protocol for a multicenter, double-blind, placebo-controlled randomized controlled trial.* Trials, 2020. 21(1): p. 764.
35. Allemani, C., et al., *Global surveillance of trends in cancer survival 2000–14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries.* The Lancet, 2018. 391(10125): p. 1023-1075.
36. Hardcastle, J.D., et al., *Randomised controlled trial of faecal-occult-blood screening for colorectal cancer.* Lancet, 1996. 348(9040): p. 1472-7.
37. Walsh, J.M. and J.P. Terdiman, *Colorectal cancer screening: scientific review.* Jama, 2003. 289(10): p. 1288-96.
38. Smith, R.A., et al., *Cancer screening in the United States, 2011: A review of current American Cancer Society guidelines and issues in cancer screening.* CA Cancer J Clin, 2011. 61(1): p. 8-30.
39. Lambert, A., et al., *An update on treatment options for pancreatic adenocarcinoma.* Ther Adv Med Oncol, 2019. 11: p. 1758835919875568.
40. Roland, C.L., et al., *The Addition of Postoperative Chemotherapy is Associated with Improved Survival in Patients with Pancreatic Cancer Treated with Preoperative Therapy.* Ann Surg Oncol, 2015. 22 Suppl 3(Suppl 3): p. S1221-8.
41. Neoptolemos, J.P., et al., *Comparison of adjuvant gemcitabine and capecitabine with gemcitabine monotherapy in patients with resected pancreatic cancer (ESPAC-4): a multicentre, open-label, randomised, phase 3 trial.* Lancet, 2017. 389(10073): p. 1011-1024.
42. Conroy, T., et al., *FOLFIRINOX or Gemcitabine as Adjuvant Therapy for Pancreatic Cancer.* N Engl J Med, 2018. 379(25): p. 2395-2406.
43. Siegel, R.L., K.D. Miller, and A. Jemal, *Cancer statistics, 2019.* CA Cancer J Clin, 2019. 69(1): p. 7-34.
44. Konings, I., et al., *Surveillance for pancreatic cancer in high-risk individuals.* BJS Open, 2019. 3(5): p. 656-665.
45. Konings, I., et al., *Evolution of features of chronic pancreatitis during endoscopic ultrasound-based surveillance of individuals at high risk for pancreatic cancer.* Endosc Int Open, 2018. 6(5): p. E541-E548.

46. Konings, I.C., et al., *Prevalence and Progression of Pancreatic Cystic Precursor Lesions Differ Between Groups at High Risk of Developing Pancreatic Cancer.* Pancreas, 2017. 46(1): p. 28-34.
47. Antoniou, A.C., et al., *Common breast cancer-predisposition alleles are associated with breast cancer risk in BRCA1 and BRCA2 mutation carriers.* Am J Hum Genet, 2008. 82(4): p. 937-48.
48. Moorman, P.G., et al., *Evaluation of established breast cancer risk factors as modifiers of BRCA1 or BRCA2: a multi-center case-only analysis.* Breast Cancer Res Treat, 2010. 124(2): p. 441-51.
49. Turnbull, C., et al., *Gene-gene interactions in breast cancer susceptibility.* Hum Mol Genet, 2012. 21(4): p. 958-62.
50. Marqués-Lespier, J.M., M. González-Pons, and M. Cruz-Correa, *Current Perspectives on Gastric Cancer.* Gastroenterol Clin North Am, 2016. 45(3): p. 413-28.
51. Hruban, R.H., et al., *Pancreatic intraepithelial neoplasia: a new nomenclature and classification system for pancreatic duct lesions.* Am J Surg Pathol, 2001. 25(5): p. 579-86.
52. Cooper, C.L., S.A. O'Toole, and J.G. Kench, *Classification, morphology and molecular pathology of premalignant lesions of the pancreas.* Pathology, 2013. 45(3): p. 286-304.
53. Vullierme, M.P. and M. Lagadec, *Predisposing factors for pancreatic adenocarcinoma: What is the role of imaging?* Diagnostic and Interventional Imaging, 2016. 97(12): p. 1233-1240.
54. Goggins, M., et al., *Management of patients with increased risk for familial pancreatic cancer: updated recommendations from the International Cancer of the Pancreas Screening (CAPS) Consortium.* Gut, 2020. 69(1): p. 7.
55. Poruk, K.E., et al., *Screening for pancreatic cancer: why, how, and who?* Annals of surgery, 2013. 257(1): p. 17-26.
56. Schmidt-Hansen, M., S. Berendse, and W. Hamilton, *Symptoms of Pancreatic Cancer in Primary Care: A Systematic Review.* Pancreas, 2016. 45(6).
57. Goggins, M., et al., *Germline BRCA2 gene mutations in patients with apparently sporadic pancreatic carcinomas.* Cancer Res, 1996. 56(23): p. 5360-4.
58. Kastrinos, F., et al., *Risk of pancreatic cancer in families with Lynch syndrome.* Jama, 2009. 302(16): p. 1790-5.
59. Jones, S., et al., *Exomic sequencing identifies PALB2 as a pancreatic cancer susceptibility gene.* Science, 2009. 324(5924): p. 217.
60. Rustgi, A.K., *Familial pancreatic cancer: genetic advances.* Genes Dev, 2014. 28(1): p. 1-7.
61. Brune, K.A., et al., *Importance of age of onset in pancreatic cancer kindreds.* J Natl Cancer Inst, 2010. 102(2): p. 119-26.
62. Copur, M.S., et al., *Hereditary vs Familial Pancreatic Cancer: Associated Genetic Syndromes and Clinical Perspective.* Oncology (Williston Park), 2020. 34(6): p. 196-201.
63. Canto, M.I., et al., *Risk of Neoplastic Progression in Individuals at High Risk for Pancreatic Cancer Undergoing Long-term Surveillance.* Gastroenterology, 2018. 155(3): p. 740-751.e2.
64. Zerboni, G., et al., *Systematic review and meta-analysis: Prevalence of incidentally detected pancreatic cystic lesions in asymptomatic individuals.* Pancreatology, 2019. 19(1): p. 2-9.
65. European Study Group on Cystic Tumours of the, P., *European evidence-based guidelines on pancreatic cystic neoplasms.* Gut, 2018. 67(5): p. 789-804.
66. Brugge, W.R., et al., *Cystic neoplasms of the pancreas.* N Engl J Med, 2004. 351(12): p. 1218-26.
67. Dimastromatteo, J., T. Brentnall, and K.A. Kelly, *Imaging in pancreatic disease.* Nature Reviews Gastroenterology & Hepatology, 2017. 14(2): p. 97-109.
68. Yu, S., et al., *Validation of the 2012 Fukuoka Consensus Guideline for Intraductal Papillary Mucinous Neoplasm of the Pancreas From a Single Institution Experience.* Pancreas, 2017. 46(7): p. 936-942.
69. Canto, M.I., et al., *Frequent detection of pancreatic lesions in asymptomatic high-risk individuals.* Gastroenterology, 2012. 142(4): p. 796-804; quiz e14-5.
70. Harinck, F., et al., *A multicentre comparative prospective blinded analysis of EUS and MRI for screening of pancreatic cancer in high-risk individuals.* Gut, 2016. 65(9): p. 1505.
71. Khashab, M.A., et al., *EUS is still superior to multidetector computerized tomography for detection of pancreatic neuroendocrine tumors.* Gastrointest Endosc, 2011. 73(4): p. 691-6.
72. Song, S.J., et al., *Differentiation of intraductal papillary mucinous neoplasms from other pancreatic cystic masses: comparison of multirow-detector CT and MR imaging using ROC analysis.* J Magn Reson Imaging, 2007. 26(1): p. 86-93.
73. Visser, B.C., et al., *Diagnostic evaluation of cystic pancreatic lesions.* HPB (Oxford), 2008. 10(1): p. 63-9.
74. Sainani, N.I., et al., *Comparative performance of MDCT and MRI with MR cholangiopancreatography in characterizing small pancreatic cysts.* AJR Am J Roentgenol, 2009. 193(3): p. 722-31.
75. Lee, H.J., et al., *Relative accuracy of CT and MRI in the differentiation of benign from malignant pancreatic cystic lesions.* Clin Radiol, 2011. 66(4): p. 315-21.
76. Sahani, D.V., et al., *Diagnosis and management of cystic pancreatic lesions.* AJR Am J Roentgenol, 2013. 200(2): p. 343-54.
77. Waters, J.A., et al., *CT vs MRCP: optimal classification of IPMN type and extent.* J Gastrointest Surg, 2008. 12(1): p. 101-9.
78. Topazian, M., et al., *Interobserver agreement for EUS findings in familial pancreatic-cancer kindreds.* Gastrointest Endosc, 2007. 66(1): p. 62-7.
79. Thiruvengadam, S.S., et al., *Chronic pancreatitis changes in high-risk individuals for pancreatic ductal adenocarcinoma.* Gastrointestinal Endoscopy, 2019. 89(4): p. 842-851.e1.
80. Brune, K., et al., *Multifocal neoplastic precursor lesions associated with lobular atrophy of the pancreas in patients having a strong family history of pancreatic cancer.* Am J Surg Pathol, 2006. 30(9): p. 1067-76.
81. Wani, S., et al., *Setting minimum standards for training in EUS and ERCP: results from a prospective multicenter study evaluating learning curves and competence among advanced endoscopy trainees.* Gastrointestinal Endoscopy, 2019. 89(6): p. 1160-1168.e9.
82. Long, E.E., et al., *Computed tomography, endoscopic, laparoscopic, and intra-operative sonography for assessing resectability of pancreatic cancer.* Surg Oncol, 2005. 14(2): p. 105-13.
83. Pereira, S.P., et al., *Early detection of pancreatic cancer.* Lancet Gastroenterol Hepatol, 2020. 5(7): p. 698-710.
84. Valsangkar, N.P., et al., *851 resected cystic tumors of the pancreas: a 33-year experience at the Massachusetts General Hospital.* Surgery, 2012. 152(3 Suppl 1): p. S4-12.
85. Hruban, R.H., et al., *An Illustrated Consensus on the Classification of Pancreatic Intraepithelial Neoplasia and Intraductal Papillary Mucinous Neoplasms.* The American Journal of Surgical Pathology, 2004. 28(8).
86. Zamboni, G., et al., *Precancerous lesions of the pancreas.* Best Practice & Research Clinical Gastroenterology, 2013. 27(2): p. 299-322.
87. Tsunoda, T., et al., *Staging and treatment for patients with pancreatic cancer. How small is an early pancreatic cancer?* J Hepatobiliary Pancreat Surg, 1998. 5(2): p. 128-32.

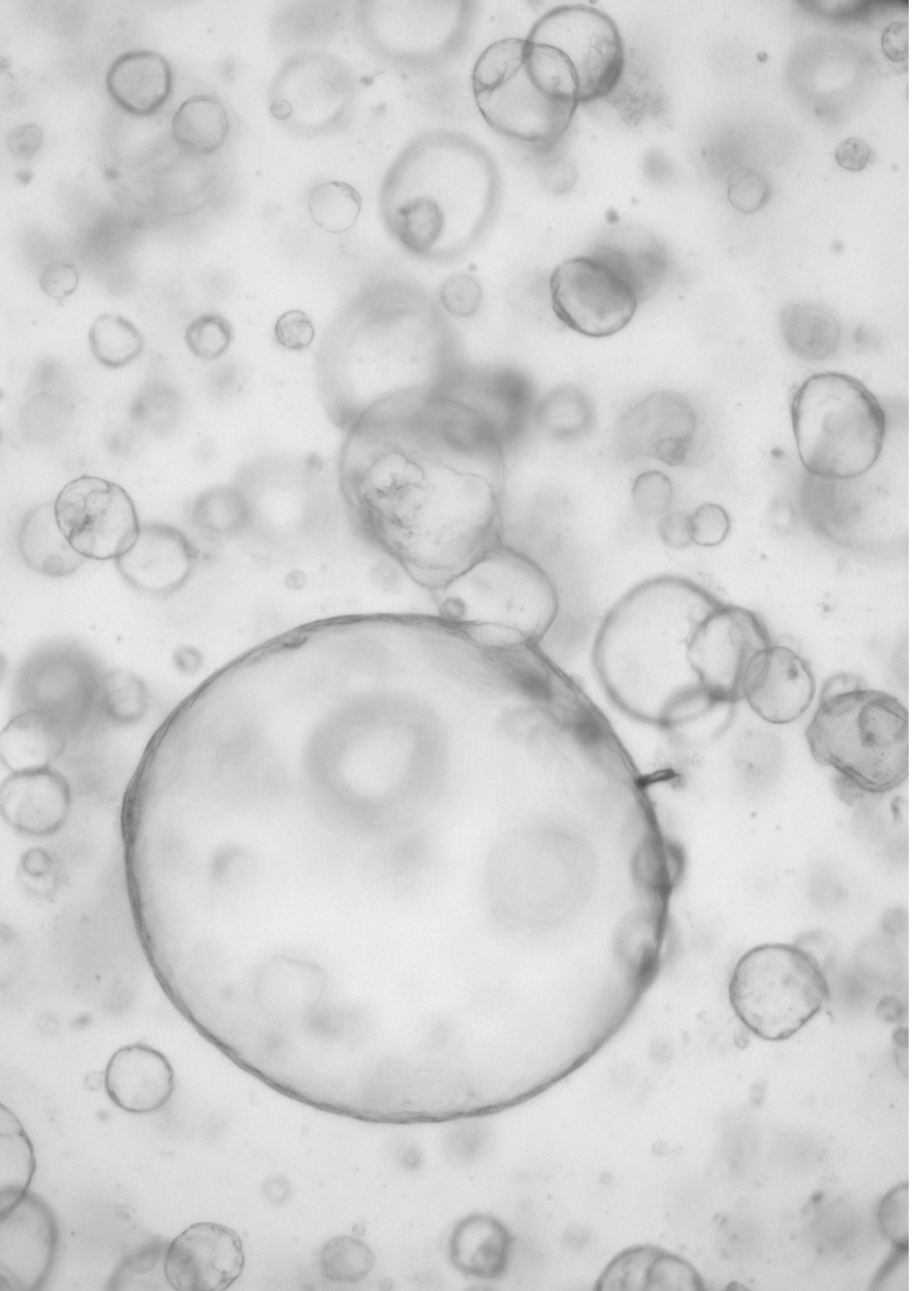
88. Yu, J., et al., *Time to progression of pancreatic ductal adenocarcinoma from low-to-high tumour stages.* Gut, 2015. 64(11): p. 1783.
89. Satake, K., et al., *A clinical evaluation of various tumor markers for the diagnosis of pancreatic cancer.* Int J Pancreatol, 1990. 7(1-3): p. 25-36.
90. Nazli, O., et al., *The diagnostic importance of CEA and CA 19-9 for the early diagnosis of pancreatic carcinoma.* Hepatogastroenterology, 2000. 47(36): p. 1750-2.
91. Sawabu, N., et al., *Serum Tumor Markers and Molecular Biological Diagnosis in Pancreatic Cancer.* Pancreas, 2004. 28(3).
92. Egawa, S., et al., *Clinicopathological Aspects of Small Pancreatic Cancer.* Pancreas, 2004. 28(3).
93. Setiawan, V.W., et al., *Pancreatic Cancer Following Incident Diabetes in African Americans and Latinos: The Multiethnic Cohort.* JNCI: Journal of the National Cancer Institute, 2019. 111(1): p. 27-33.
94. Chari, S.T., et al., *Probability of pancreatic cancer following diabetes: a population-based study.* Gastroenterology, 2005. 129(2): p. 504-11.
95. Chari, S.T., et al., *Pancreatic cancer-associated diabetes mellitus: prevalence and temporal association with diagnosis of cancer.* Gastroenterology, 2008. 134(1): p. 95-101.
96. Sharma, A., et al., *Model to Determine Risk of Pancreatic Cancer in Patients With New-Onset Diabetes.* Gastroenterology, 2018. 155(3): p. 730-739 e3.
97. Jenkinson, C., et al., *Biomarkers for early diagnosis of pancreatic cancer.* Expert Rev Gastroenterol Hepatol, 2015. 9(3): p. 305-15.
98. Neoptolemos, J.P., et al., *Therapeutic developments in pancreatic cancer: current and future perspectives.* Nature Reviews Gastroenterology & Hepatology, 2018. 15(6): p. 333-348.
99. Adamska, A., A. Domenichini, and M. Falasca, *Pancreatic Ductal Adenocarcinoma: Current and Evolving Therapies.* International journal of molecular sciences, 2017. 18(7): p. 1338.
100. Burris, H.A., 3rd, et al., *Improvements in survival and clinical benefit with gemcitabine as first-line therapy for patients with advanced pancreas cancer: a randomized trial.* J Clin Oncol, 1997. 15(6): p. 2403-13.
101. Park, J.K., et al., *Survival and prognostic factors of unresectable pancreatic cancer.* J Clin Gastroenterol, 2008. 42(1): p. 86-91.
102. Edderkaoui, M., et al., *An Inhibitor of GSK3B and HDACs Kills Pancreatic Cancer Cells and Slows Pancreatic Tumor Growth and Metastasis in Mice.* Gastroenterology, 2018. 155(6): p. 1985-1998.e5.
103. Schneider, G., et al., *Targeting histone deacetylases in pancreatic ductal adenocarcinoma.* J Cell Mol Med, 2010. 14(6A): p. 1255-63.
104. Ben-Josef, E., et al., *Glycogen Synthase Kinase 3 Beta Predicts Survival in Resected Adenocarcinoma of the Pancreas.* Clin Cancer Res, 2015. 21(24): p. 5612-8.
105. Zheng, H., et al., *Glycogen synthase kinase-3 beta regulates Snail and β-catenin expression during Fas-induced epithelial–mesenchymal transition in gastrointestinal cancer.* European Journal of Cancer, 2013. 49(12): p. 2734-2746.
106. Kazi, A., et al., *GSK3 suppression upregulates β-catenin and c-Myc to abrogate KRas-dependent tumors.* Nature Communications, 2018. 9(1): p. 5154.
107. Sano, M., et al., *Activation of WNT/beta-Catenin Signaling Enhances Pancreatic Cancer Development and the Malignant Potential Via Up-regulation of Cyr61.* Neoplasia, 2016. 18(12): p. 785-794.
108. Wang, L., et al., *Oncogenic function of ATDC in pancreatic cancer through Wnt pathway activation and beta-catenin stabilization.* Cancer Cell, 2009. 15(3): p. 207-19.
109. Wang, L., et al., *ATDC induces an invasive switch in KRAS-induced pancreatic tumorigenesis.* Genes Dev, 2015. 29(2): p. 171-83.
110. Yu, M., et al., *RNA sequencing of pancreatic circulating tumour cells implicates WNT signalling in metastasis.* Nature, 2012. 487(7408): p. 510-3.
111. Visvader, J.E., *Cells of origin in cancer.* Nature, 2011. 469(7330): p. 314-322.
112. Barker, N., et al., *Crypt stem cells as the cells-of-origin of intestinal cancer.* Nature, 2009. 457(7229): p. 608-11.
113. Sangiorgi, E. and M.R. Capecchi, *Bmi1 is expressed in vivo in intestinal stem cells.* Nature Genetics, 2008. 40(7): p. 915-920.
114. Clevers, H., *Wnt/β-Catenin Signaling in Development and Disease.* Cell, 2006. 127(3): p. 469-480.
115. Barker, N., et al., *Lgr5(+ve) stem cells drive self-renewal in the stomach and build long-lived gastric units in vitro.* Cell Stem Cell, 2010. 6(1): p. 25-36.
116. Gidekel Friedlander, S.Y., et al., *Context-dependent transformation of adult pancreatic cells by oncogenic K-Ras.* Cancer Cell, 2009. 16(5): p. 379-89.
117. Habbe, N., et al., *Spontaneous induction of murine pancreatic intraepithelial neoplasia (mPanIN) by acinar cell targeting of oncogenic Kras in adult mice.* Proc Natl Acad Sci U S A, 2008. 105(48): p. 18913-8.
118. Perez-Losada, J. and A. Balmain, *Stem-cell hierarchy in skin cancer.* Nature Reviews Cancer, 2003. 3(6): p. 434-443.
119. DeNardi, F.G. and R.H. Riddell, *The normal esophagus.* Am J Surg Pathol, 1991. 15(3): p. 296-309.
120. Jenkins, G.J., et al., *The bile acid deoxycholic acid has a non-linear dose response for DNA damage and possibly NF-kappaB activation in oesophageal cells, with a mechanism of action involving ROS.* Mutagenesis, 2008. 23(5): p. 399-405.
121. Wang, R.-H., *From reflux esophagitis to Barrett's esophagus and esophageal adenocarcinoma.* World Journal of Gastroenterology : WJG, 2015. 21(17): p. 5210-5219.
122. Bjorkman, E.V., et al., *Esophageal barrier function and tight junction expression in healthy subjects and patients with gastroesophageal reflux disease: functionality of esophageal mucosa exposed to bile salt and trypsin in vitro.* Scand J Gastroenterol, 2013. 48(10): p. 1118-26.
123. Schoofs, N., R. Bisschops, and H. Prenen, *Progression of Barrett's esophagus toward esophageal adenocarcinoma: an overview.* Ann Gastroenterol, 2017. 30(1): p. 1-6.
124. Arnold, M., et al., *Global incidence of oesophageal cancer by histological subtype in 2012.* Gut, 2015. 64(3): p. 381.
125. Shaheen, N.J., *Barrett Esophagus: Prevalence, Treatment, and the Link to Esophageal Adenocarcinoma.* Gastroenterology & Hepatology, 2015. 11(1): p. 61-63.
126. Odze, R.D., *Update on the diagnosis and treatment of Barrett esophagus and related neoplastic precursor lesions.* Arch Pathol Lab Med, 2008. 132(10): p. 1577-85.
127. Boeckxstaens, G.E., *The lower oesophageal sphincter.* Neurogastroenterol Motil, 2005. 17 Suppl 1: p. 13-21.
128. Vantrappen, G., et al., *Treatment of achalasia with pneumatic dilatations.* Gut, 1971. 12(4): p. 268-275.
129. Kauer, W.K. and H.J. Stein, *Role of acid and bile in the genesis of Barrett's esophagus.* Chest Surg Clin N Am, 2002. 12(1): p. 39-45.

130. Tustumi, F., et al., *Esophageal achalasia: a risk factor for carcinoma. A systematic review and meta-analysis.* Dis Esophagus, 2017. 30(10): p. 1-8.
131. Leeuwenburgh, I., et al., *Long-term risk of oesophagitis, Barrett's oesophagus and oesophageal cancer in achalasia patients.* Scandinavian Journal of Gastroenterology, 2006. 41(sup243): p. 7-10.
132. Streppel, M.M., et al., *Squamous cell carcinoma in Barrett's esophagus: field effect versus metastasis.* Diseases of the Esophagus, 2012. 25(7): p. 630-637.
133. Park, J.C., et al., *Achalasia Combined with Esophageal Cancer Treated by Concurrent Chemoradiation Therapy.* Gut and Liver, 2009. 3(4): p. 329-333.
134. Smyth, E.C., et al., *Oesophageal cancer.* Nat Rev Dis Primers, 2017. 3: p. 17048.
135. Arnold, M., et al., *Global incidence of oesophageal cancer by histological subtype in 2012.* Gut, 2015. 64(3): p. 381-387.
136. Yoshida, A., I.Y. Huang, and M. Ikawa, *Molecular abnormality of an inactive aldehyde dehydrogenase variant commonly found in Orientals.* Proc Natl Acad Sci U S A, 1984. 81(1): p. 258-61.
137. Goedde, H.W., et al., *Distribution of ADH2 and ALDH2 genotypes in different populations.* Hum Genet, 1992. 88(3): p. 344-6.
138. Muto, M., et al., *Association between aldehyde dehydrogenase gene polymorphisms and the phenomenon of field cancerization in patients with head and neck cancer.* Carcinogenesis, 2002. 23(10): p. 1759-65.
139. Brooks, P.J., et al., *The alcohol flushing response: an unrecognized risk factor for esophageal cancer from alcohol consumption.* PLoS Med, 2009. 6(3): p. e50.
140. Uebelacker, M. and D.W. Lachenmeier, *Quantitative determination of acetaldehyde in foods using automated digestion with simulated gastric fluid followed by headspace gas chromatography.* J Autom Methods Manag Chem, 2011. 2011: p. 907317.
141. Homann, N., et al., *High acetaldehyde levels in saliva after ethanol consumption: methodological aspects and pathogenetic implications.* Carcinogenesis, 1997. 18(9): p. 1739-43.
142. Salaspuro, M.P., *Acetaldehyde, microbes, and cancer of the digestive tract.* Crit Rev Clin Lab Sci, 2003. 40(2): p. 183-208.
143. Muto, M., et al., *Acetaldehyde production by non-pathogenic Neisseria in human oral microflora: implications for carcinogenesis in upper aerodigestive tract.* Int J Cancer, 2000. 88(3): p. 342-50.
144. Arnold, M., et al., *Predicting the Future Burden of Esophageal Cancer by Histological Subtype: International Trends in Incidence up to 2030.* American Journal of Gastroenterology, 2017. 112(8): p. 1247-1255.
145. Spechler, S.J. and R.F. Souza, *Barrett's esophagus.* N Engl J Med, 2014. 371(9): p. 836-45.
146. Lewin, K.J., *Malignant and Premalignant Lesions of the Esophagus.* The Keio Journal of Medicine, 1992. 41(3): p. 177-183.
147. Taylor, P.R., C.C. Abnet, and S.M. Dawsey, *Squamous Dysplasia—The Precursor Lesion for Esophageal Squamous Cell Carcinoma.* Cancer Epidemiology Biomarkers & Prevention, 2013. 22(4): p. 540.
148. Michopoulos, S., *Critical appraisal of guidelines for screening and surveillance of Barrett's esophagus.* Ann Transl Med, 2018. 6(13): p. 259.
149. Falk, G.W., *Updated Guidelines for Diagnosing and Managing Barrett Esophagus.* Gastroenterology & hepatology, 2016. 12(7): p. 449-451.
150. Yang, J., et al., *Cost-benefit analysis of esophageal cancer endoscopic screening in high-risk areas of China.* World J Gastroenterol, 2012. 18(20): p. 2493-501.
151. Jiang, M., et al., *Transitional basal cells at the squamous-columnar junction generate Barrett's oesophagus.* Nature, 2017. 550(7677): p. 529-533.
152. Jain, S. and S. Dhingra, *Pathology of esophageal cancer and Barrett's esophagus.* Ann Cardiothorac Surg, 2017. 6(2): p. 99-109.
153. Wang, G.Q., et al., *Histological precursors of oesophageal squamous cell carcinoma: results from a 13 year prospective follow up study in a high risk population.* Gut, 2005. 54(2): p. 187.
154. Jiang, M., et al., *Transitional basal cells at the squamous–columnar junction generate Barrett's oesophagus.* Nature, 2017. 550: p. 529.
155. Sato, H., et al., *Full-layer mucosal histology in achalasia: Histological epithelial wave is characteristic in "pinstripe pattern"-positive achalasia.* Neurogastroenterol Motil, 2018. 30(1).
156. Mantovani, A., et al., *Cancer-related inflammation.* Nature, 2008. 454: p. 436.
157. Elinav, E., et al., *Inflammation-induced cancer: crosstalk between tumours, immune cells and microorganisms.* Nat Rev Cancer, 2013. 13(11): p. 759-71.
158. Reid, B.J., et al., *Barrett's oesophagus and oesophageal adenocarcinoma: time for a new synthesis.* Nature Reviews Cancer, 2010. 10: p. 87.
159. Poehlmann, A., et al., *Inflammation and Barrett's carcinogenesis.* Pathol Res Pract, 2012. 208(5): p. 269-80.
160. O'Riordan, J.M., et al., *Proinflammatory cytokine and nuclear factor kappa-B expression along the inflammation-metaplasia-dysplasia-adenocarcinoma sequence in the esophagus.* Am J Gastroenterol, 2005. 100(6): p. 1257-64.
161. Hardikar, S., et al., *Inflammation and oxidative stress markers and esophageal adenocarcinoma incidence in a Barrett's esophagus cohort.* Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology, 2014. 23(11): p. 2393-2403.
162. Tselepis, C., et al., *Tumour necrosis factor-α in Barrett's oesophagus: a potential novel mechanism of action.* Oncogene, 2002. 21: p. 6071.
163. Liao, L.M., et al., *Nonsteroidal anti-inflammatory drug use reduces risk of adenocarcinomas of the esophagus and esophagogastric junction in a pooled analysis.* Gastroenterology, 2012. 142(3): p. 442-452 e5; quiz e22-3.
164. Leeuwenburgh, I., et al., *Oesophagitis is common in patients with achalasia after pneumatic dilatation.* Aliment Pharmacol Ther, 2006. 23(8): p. 1197-203.
165. Loviscek, L.F., et al., *Early cancer in achalasia.* Diseases of the Esophagus, 1998. 11(4): p. 239-247.
166. Cancer Genome Atlas Research, N., et al., *Integrated genomic characterization of oesophageal carcinoma.* Nature, 2017. 541(7636): p. 169-175.
167. Lagergren, J., et al., *Oesophageal cancer.* Lancet, 2017. 390(10110): p. 2383-2396.
168. Ross-Innes, C.S., et al., *Whole-genome sequencing provides new insights into the clonal architecture of Barrett's esophagus and esophageal adenocarcinoma.* Nat Genet, 2015. 47(9): p. 1038-1046.
169. Weaver, J.M.J., et al., *Ordering of mutations in preinvasive disease stages of esophageal carcinogenesis.* Nature genetics, 2014. 46(8): p. 837-843.
170. van Nistelrooij, A.M., et al., *Molecular clonality analysis of esophageal adenocarcinoma by multiregion sequencing of tumor samples.* BMC Res Notes, 2017. 10(1): p. 144.

171. Martinez, P., et al., *Dynamic clonal equilibrium and predetermined cancer risk in Barrett's oesophagus.* Nature Communications, 2016. 7: p. 12158.
172. Nones, K., et al., *Genomic catastrophes frequently arise in esophageal adenocarcinoma and drive tumorigenesis.* Nat Commun, 2014. 5: p. 5224.
173. Illig, R., et al., *GERD—Barrett—Adenocarcinoma: Do We Have Suitable Prognostic and Predictive Molecular Markers?* Gastroenterology Research and Practice, 2013. 2013: p. 643084.
174. Varghese, S., et al., *Analysis of dysplasia in patients with Barrett's esophagus based on expression pattern of 90 genes.* Gastroenterology, 2015. 149(6): p. 1511-1518 e5.
175. Janmaat, V.T., et al., *Use of immunohistochemical biomarkers as independent predictor of neoplastic progression in Barrett's oesophagus surveillance: A systematic review and meta-analysis.* PLoS One, 2017. 12(10): p. e0186305.
176. Debruyne, P.R., et al., *Bile acids induce ectopic expression of intestinal guanylyl cyclase C Through nuclear factor-kappaB and Cdx2 in human esophageal cells.* Gastroenterology, 2006. 130(4): p. 1191-1206.
177. Eda, A., et al., *Aberrant expression of CDX2 in Barrett's epithelium and inflammatory esophageal mucosa.* Journal of Gastroenterology, 2003. 38(1): p. 14-22.
178. Gao, N., P. White, and K.H. Kaestner, *Establishment of intestinal identity and epithelial-mesenchymal signaling by Cdx2.* Developmental cell, 2009. 16(4): p. 588-599.
179. Liu, T., et al., *Regulation of Cdx2 expression by promoter methylation, and effects of Cdx2 transfection on morphology and gene expression of human esophageal epithelial cells.* Carcinogenesis, 2007. 28(2): p. 488-496.
180. Mari, L., et al., *A pSMAD/CDX2 complex is essential for the intestinalization of epithelial metaplasia.* Cell Rep, 2014. 7(4): p. 1197-210.
181. Silberg, D.G., et al., *Cdx2 ectopic expression induces gastric intestinal metaplasia in transgenic mice.* Gastroenterology, 2002. 122(3): p. 689-696.
182. Garcia-Fernàndez, J. and J. Garcia-Fernandez, *The genesis and evolution of homeobox gene clusters.* Nat Rev Genet, 2005. 6(12): p. 881-892.
183. Stern, C.D., et al., *Head-tail patterning of the vertebrate embryo: one, two or many unresolved problems?* Int J Dev Biol, 2006. 50(1): p. 3-15.
184. Pearson, J.C., D. Lemons, and W. McGinnis, *Modulating Hox gene functions during animal body patterning.* Nat Rev Genet, 2005. 6(12): p. 893-904.
185. Kawazoe, Y., et al., *Region-specific gastrointestinal Hox code during murine embryonal gut development.* Dev Growth Differ, 2002. 44(1): p. 77-84.
186. Zacchetti, G., D. Duboule, and J. Zakany, *Hox gene function in vertebrate gut morphogenesis: the case of the caecum.* Development, 2007. 134(22): p. 3967-73.
187. Gao, N., P. White, and K.H. Kaestner, *Establishment of intestinal identity and epithelial-mesenchymal signaling by Cdx2.* Dev Cell, 2009. 16(4): p. 588-99.
188. Yahagi, N., et al., *Position-specific expression of Hox genes along the gastrointestinal tract.* Congenit Anom (Kyoto), 2004. 44(1): p. 18-26.
189. di Pietro, M., et al., *Evidence for a functional role of epigenetically regulated midcluster HOXB genes in the development of Barrett esophagus.* Proc Natl Acad Sci U S A, 2012. 109(23): p. 9077-82.
190. Mallo, M., D.M. Wellik, and J. Deschamps, *Hox genes and regional patterning of the vertebrate body plan.* Dev Biol, 2010. 344(1): p. 7-15.
191. Gu, Z.D., et al., *HOXA13 promotes cancer cell growth and predicts poor survival of patients with esophageal squamous cell carcinoma.* Cancer Res, 2009. 69(12): p. 4969-73.
192. Han, Y., et al., *Identification and validation that up-expression of HOXA13 is a novel independent prognostic marker of a worse outcome in gastric cancer based on immunohistochemistry.* Med Oncol, 2013. 30(2): p. 564.
193. Levine, D.M., et al., *A genome-wide association study identifies new susceptibility loci for esophageal adenocarcinoma and Barrett's esophagus.* Nature Genetics, 2013. 45(12): p. 1487-1493.
194. Becker, J., et al., *Supportive evidence for FOXP1, BARX1, and FOXF1 as genetic risk loci for the development of esophageal adenocarcinoma.* Cancer Med, 2015. 4(11): p. 1700-4.
195. Yan, C., et al., *An esophageal adenocarcinoma susceptibility locus at 9q22 also confers risk to esophageal squamous cell carcinoma by regulating the function of BARX1.* Cancer Lett, 2018. 421: p. 103-111.
196. Woo, J., et al., *Barx1-mediated inhibition of Wnt signaling in the mouse thoracic foregut controls tracheo-esophageal septation and epithelial differentiation.* PLoS One, 2011. 6(7): p. e22493.
197. Kim, B.-M., et al., *The Stomach Mesenchymal Transcription Factor Barx1 Specifies Gastric Epithelial Identity through Inhibition of Transient Wnt Signaling.* Developmental Cell, 2005. 8(4): p. 611-622.
198. Koon, H.B., et al., *FOXP1: a potential therapeutic target in cancer.* Expert Opinion on Therapeutic Targets, 2007. 11(7): p. 955-965.
199. Shu, W., et al., *Foxp2 and Foxp1 cooperatively regulate lung and esophagus development.* Development, 2007. 134(10): p. 1991.
200. Fröhlich, H., et al., *Gastrointestinal dysfunction in autism displayed by altered motility and achalasia in Foxp1 (+/-) mice.* Proc Natl Acad Sci U S A, 2019. 116(44): p. 22237-22245.
201. Banham, A.H., et al., *The FOXP1 winged helix transcription factor is a novel candidate tumor suppressor gene on chromosome 3p.* Cancer Res, 2001. **61**(24): p. 8820-9.
202. Zhang, J., et al., *Genetic variants of FOXP1 and FOXF1 are associated with the susceptibility of oesophageal adenocarcinoma in Chinese population.* J Genet, 2018. **97**(1): p. 213-218.
203. Mahlapuu, M., S. Enerbäck, and P. Carlsson, *Haploinsufficiency of the forkhead gene Foxf1, a target for sonic hedgehog signaling, causes lung and foregut malformations.* Development, 2001. **128**(12): p. 2397-406.
204. Ustiyan, V., et al., *FOXF1 transcription factor promotes lung morphogenesis by inducing cellular proliferation in fetal lung mesenchyme.* Dev Biol, 2018. **443**(1): p. 50-63.
205. Lehman, M.B., et al., *Squamous mucosal alterations in esophagectomy specimens from patients with end-stage achalasia.* Am J Surg Pathol, 2001. **25**(11): p. 1413-8.
206. Safatle-Ribeiro, A.V., et al., *Integrated p53 histopathologic/genetic analysis of premalignant lesions of the esophagus.* Cancer Detect Prev, 2000. **24**(1): p. 13-23.
207. Kim, H., et al., *Retention Esophagitis as a Significant Clinical Predictor of Progression to Esophageal Cancer in Achalasia.* Clin Endosc, 2018.
208. Choi, C.R., et al., *Clonal evolution of colorectal cancer in IBD.* Nat Rev Gastroenterol Hepatol, 2017. **14**(4): p. 218-229.

209. Manoel-Caetano Fda, S., et al., *Cytogenetic alterations in chagasic achalasia compared to esophageal carcinoma.* Cancer Genet Cytogenet, 2004. **149**(1): p. 17-22.
210. Munari, F.F., et al., *PIK3CA mutations are frequent in esophageal squamous cell carcinoma associated with chagasic megaesophagus and are associated with a worse patient outcome.* Infectious agents and cancer, 2018. **13**: p. 43-43.
211. Chen, X.X., et al., *Genomic comparison of esophageal squamous cell carcinoma and its precursor lesions by multi-region whole-exome sequencing.* Nat Commun, 2017. **8**(1): p. 524.
212. Wang, K., et al., *Comprehensive Genomic Profiling of Advanced Esophageal Squamous Cell Carcinomas and Esophageal Adenocarcinomas Reveals Similarities and Differences.* The Oncologist, 2015. **20**(10): p. 1132-1139.
213. The Cancer Genome Atlas Research, N., *Integrated genomic characterization of oesophageal carcinoma.* Nature, 2017. **541**(7636): p. 169-175.
214. Lagergren, J., et al., *Oesophageal cancer.* The Lancet, 2017. **390**(10110): p. 2383-2396.
215. Song, Y., et al., *Identification of genomic alterations in oesophageal squamous cell cancer.* Nature, 2014. **509**(7498): p. 91-5.
216. Lin, D.C., et al., *Genomic and molecular characterization of esophageal squamous cell carcinoma.* Nat Genet, 2014. **46**(5): p. 467-73.
217. Cerami, E., et al., *The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data.* Cancer Discov, 2012. **2**(5): p. 401-4.
218. Gao, J., et al., *Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal.* Sci Signal, 2013. **6**(269): p. pl1.
219. Dulak, A.M., et al., *Exome and whole-genome sequencing of esophageal adenocarcinoma identifies recurrent driver events and mutational complexity.* Nat Genet, 2013. **45**(5): p. 478-86.
220. Bugter, J.M., N. Fenderico, and M.M. Maurice, *Mutations and mechanisms of WNT pathway tumour suppressors in cancer.* Nature Reviews Cancer, 2021. **21**(1): p. 5-21.
221. Behrens, J., et al., *Functional interaction of an axin homolog, conductin, with beta-catenin, APC, and GSK3beta.* Science, 1998. **280**(5363): p. 596-9.
222. van Kappel, E.C. and M.M. Maurice, *Molecular regulation and pharmacological targeting of the β-catenin destruction complex.* Br J Pharmacol, 2017. **174**(24): p. 4575-4588.
223. Rubinfeld, B., et al., *Association of the APC gene product with beta-catenin.* Science, 1993. **262**(5140): p. 1731-4.
224. Janda, C.Y., et al., *Structural basis of Wnt recognition by Frizzled.* Science, 2012. **337**(6090): p. 59-64.
225. Behrens, J., et al., *Functional interaction of beta-catenin with the transcription factor LEF-1.* Nature, 1996. **382**(6592): p. 638-42.
226. Koo, B.K., et al., *Tumour suppressor RNF43 is a stem-cell E3 ligase that induces endocytosis of Wnt receptors.* Nature, 2012. **488**(7413): p. 665-9.
227. Hao, H.X., et al., *ZNRF3 promotes Wnt receptor turnover in an R-spondin-sensitive manner.* Nature, 2012. **485**(7397): p. 195-200.
228. Nusse, R. and H. Varmus, *Three decades of Wnts: a personal perspective on how a scientific field developed.* Embo J, 2012. **31**(12): p. 2670-84.
229. Korinek, V., et al., *Depletion of epithelial stem-cell compartments in the small intestine of mice lacking Tcf-4.* Nat Genet, 1998. **19**(4): p. 379-83.
230. Pinto, D., et al., *Canonical Wnt signals are essential for homeostasis of the intestinal epithelium.* Genes Dev, 2003. **17**(14): p. 1709-13.
231. Tan, X., et al., *Conditional deletion of beta-catenin reveals its role in liver growth and regeneration.* Gastroenterology, 2006. **131**(5): p. 1561-72.
232. Nusse, R. and H. Clevers, *Wnt/β-Catenin Signaling, Disease, and Emerging Therapeutic Modalities.* Cell, 2017. **169**(6): p. 985-999.
233. Polakis, P., *Wnt signaling in cancer.* Cold Spring Harb Perspect Biol, 2012. **4**(5).
234. Reya, T. and H. Clevers, *Wnt signalling in stem cells and cancer.* Nature, 2005. **434**(7035): p. 843-50.
235. Zhong, Z. and D.M. Virshup, *Wnt Signaling and Drug Resistance in Cancer.* Mol Pharmacol, 2020. **97**(2): p. 72-89.
236. Bian, Y.-S., et al., *Nuclear Accumulation of beta-Catenin Is a Common and Early Event During Neoplastic Progression of Barrett Esophagus.* American Journal of Clinical Pathology, 2000. **114**(4): p. 583-590.
237. Choi, Y.W., et al., *Mutations in beta-catenin and APC genes are uncommon in esophageal and esophagogastric junction adenocarcinomas.* Mod Pathol, 2000. **13**(10): p. 1055-9.
238. Lyros, O., et al., *Wnt/β-Catenin Signaling Activation beyond Robust Nuclear β-Catenin Accumulation in Nondysplastic Barrett's Esophagus: Regulation via Dickkopf-1.* Neoplasia, 2015. **17**(7): p. 598-611.
239. Clément, G., et al., *Alterations of the Wnt signaling pathway during the neoplastic progression of Barrett's esophagus.* Oncogene, 2006. **25**(21): p. 3084-3092.
240. Moyes, L.H., et al., *Activation of Wnt signalling promotes development of dysplasia in Barrett's oesophagus.* The Journal of Pathology, 2012. **228**(1): p. 99-112.
241. Clément, G., et al., *Epigenetic alteration of the Wnt inhibitory factor-1 promoter occurs early in the carcinogenesis of Barrett's esophagus.* Cancer Science, 2008. **99**(1): p. 46-53.
242. Zou, H., et al., *Aberrant methylation of secreted frizzled-related protein genes in esophageal adenocarcinoma and Barrett's esophagus.* Int J Cancer, 2005. **116**(4): p. 584-91.
243. Kim, K., et al., *Overexpression of beta-catenin induces apoptosis independent of its transactivation function with LEF-1 or the involvement of major G1 cell cycle regulators.* Mol Biol Cell, 2000. **11**(10): p. 3509-23.
244. Murphy, D.J., et al., *Distinct thresholds govern Myc's biological output in vivo.* Cancer cell, 2008. **14**(6): p. 447-457.
245. Albuquerque, C., et al., *The 'just-right' signaling model: APC somatic mutations are selected based on a specific level of activation of the β-catenin signaling cascade.* Human Molecular Genetics, 2002. **11**(13): p. 1549-1560.
246. Albuquerque, C., et al., *Colorectal cancers choosing sides.* Biochimica et Biophysica Acta (BBA) - Reviews on Cancer, 2011. **1816**(2): p. 219-231.
247. Gaspar, C. and R. Fodde, *APC dosage effects in tumorigenesis and stem cell differentiation.* Int J Dev Biol, 2004. **48**(5-6): p. 377-86.

Pancreatic lesions



Chapter 2

Optimization of Pancreatic Juice Collection for Biomarker Discovery and Early Detection of Pancreatic Cancer

I.J.M. Levink*, K. Nesteruk*, D.I. Visser, A.M. Sieuwerts, C.J.C. Fernandes, M.P.M.H. Jansen, L.M.J.W. van Driel, J.W. Poley, M.P. Peppelenbosch, D.L. Cahen, G.M. Fuhler, M.J. Bruno.

American Journal of Gastroenterology. 2020

INTRODUCTION

Comprehensive surveillance programs in individuals with hereditary predisposition for pancreatic ductal adenocarcinoma (PDAC) or pancreatic cysts have proven that early detection by imaging is challenging; by the time neoplasia becomes detectable, many patients already developed advanced disease [1-5]. Recent research has shown that pancreatic juice (PJ) is a promising source of biomarkers for the detection of pancreatic dysplasia and cancer [6-9]. One of the reasons that PJ collection has not yet been implemented in routine surveillance programs is the lack of an optimized and standardized collection protocol.

A wash-out of PJ from the pancreatic ductal system can be provoked by intravenous secretin infusion during endoscopic ultrasound (EUS) and enables non-invasive collection from the duodenal lumen. PJ is, by virtue of its origin, in close contact with ductal cells – the location where PDAC develops – and therefore potentially rich in diagnostic biomarkers. In contrast to PJ collection by direct pancreatic duct cannulation, adverse events (e.g. pancreatitis) have not yet been described for secretin-stimulated PJ collection [10, 11]. However, there is no consensus on the best method to collect PJ. (See **Supplemental Figure S1**).

In this study, we investigated the feasibility of analyzing potential biomarkers in PJ, including cell-free DNA (cfDNA), exosomal microRNA (ex-miR) and cytokines, by comparing two collection techniques and three time periods. Furthermore, we explored the potential of culturing organoids from the cellular content of PJ, as organoids may serve as an unlimited diagnostic source for targeted personalized medicine.

MATERIAL AND METHODS

This prospective study involves patients who underwent EUS for suspected (sporadic) PDAC and high-risk individuals under surveillance for a hereditary predisposition for PDAC (FPC). For full details of all methods see **Supplemental Material and Methods.** We compared two collection techniques (utilization of a through-the-scope catheter [CATH] *vs* performing suction through the endoscope suction channel [END]) and three collection time periods (0-4 minutes *vs* 4-8 minutes [phase 1] *vs* 8-15 minutes [phase 2]; **Supplemental Figure S1 & S2**). Collected PJ samples were weighed as a proxy for yield of PJ volume. Juice origin was assessed based on color as well as Phospholipase A2 Group IB (PLA2G1B) and IgG concentration (as a proxy for pancreas-specific and blood-derived content, respectively). cfDNA was isolated using Maxwell kit. Mutational load (%mu*KRAS*) and DNA length (i.e. 75, 300bp) were determined with (digital) PCR. Extracellular vesicles (EVs) were isolated and characterized by Nanoparticle Tracking Analysis (NTA). miRNA expression in EVs (ex-miR) was calculated using the $2^{-\Delta\Delta Ct}$ method using ex-miR-16 as internal control. The total protein concentration in PJ was assessed by Lowry protein assay and PLA2G1B, IgG, Interleukin-8 (IL-8), IL-10, and interferon-γ (IFN-γ) were measured by ELISA. Furthermore, organoid culture was based on Broutier et al [12]. If possible, paired analyses were performed using the Wilcoxon Matched-Pairs (2 groups) or Friedman test (>2 groups). Alternatively, the Mann-Whitney U (2 groups) test or Wilcoxon Signed-Rank

test (>2 groups) was performed for unpaired comparisons. P-values of subgroup analyses were only displayed when significant (p≤0.05). Correlations were made using Spearman's correlation.

RESULTS

For phase 1, PJ was collected from 41 patients (23 FPC; 18 PDAC; **Supplementary Table S1**). Overall, the mean quantity of PJ per 8-min collection was 10.4 g (95% CI: 7.2-13.5) and did not depend on the endoscopist (**Supplemental Figure S3**). Sample colors ranged from bright green to dark red. IgG, as a measure of blood contamination, was highest in dark red samples (p=0.01) and was correlated to cfDNA (r=0.53; P=0.01) and long-segment DNA (300bp; r=0.47; P=0.03), but not to any of the other biomarkers investigated (**Supplemental Figure S4**). PLA2G1B, which is pancreas specific, was detected in 123/130 samples and correlated to IL-10 (r=0.31; p=0.008) and IFN-γ (r=0.24; p=0.04) levels (**Supplemental Figure S4**).

When comparing collection techniques (**Figure 1**), collection with endoscope suction channel resulted in higher volumes of juice (p=0.0005) that was more likely of pancreatic origin due to higher concentration of PLA2G1B (p=0.09). Blood contamination did not relate to collection technique (p=0.14). The total yield of cfDNA, in particular of short fragment length (75bp), was highest in PJ juice collected with endoscope suction channel (total cfDNA p=0.008, 75bp p=0.04). No differences were found for %mu*KRAS*, EV yield, ex-miR levels or cytokines.

When comparing time periods of collection (**Figure 2)**, higher volumes of juice were collected in the 4-8-minute timeframe (p=0.06). We noticed that despite an increasing contamination with blood during the 4-8-minute timeframe (IgG p=0.04; long-segment DNA; p=0.008), this did not affect overall PLA2B concentration (p=0.41), suggesting that pancreas-derived biomarkers should theoretically be measured equally well in both time periods. Indeed, no differences were found for any of the biomarkers (the number of EVs, ex-miR expression, cytokine concentration) besides protein concentration (p=0.006).

Addition of superase or protease inhibitors to the collected PJ did not result in improved detection of relevant biomarkers (**Supplemental Figures S5).** Details regarding alternative DNA isolation kits and ex-miR normalization methods can be found in **Supplemental Figures S6 & S7.**

For an additional ten patients (Phase 2, **Supplemental Figure S8 & Table S2**), juice collection was extended to 15 minutes. This third timeframe (t=8-15min) resulted in a lower yield of juice (volume/minute; p=0.003), while blood contamination increased (IgG p=0.006; long-segment DNA p=0.009). Additionally, levels of biomarkers (ex-miR-205, ex-miR-155, IL-10 and IFN-γ) decreased after 8 minutes.

To analyze the feasibility of organoid growth from PJ (**Figure 3 & Supplemental Figure S9**), the cell pellet of PJ was seeded into 3D cultures (Figure 3A shows culture optimization). Organoid growth was feasible for both collection techniques and time periods (0-4 *vs* 4-8 minutes). However, it was more efficient for PJ collected with CATH. Pancreatic organoid markers CK19, CK7, Axin2, Sox9 were expressed in all PJ-derived organoid lines.

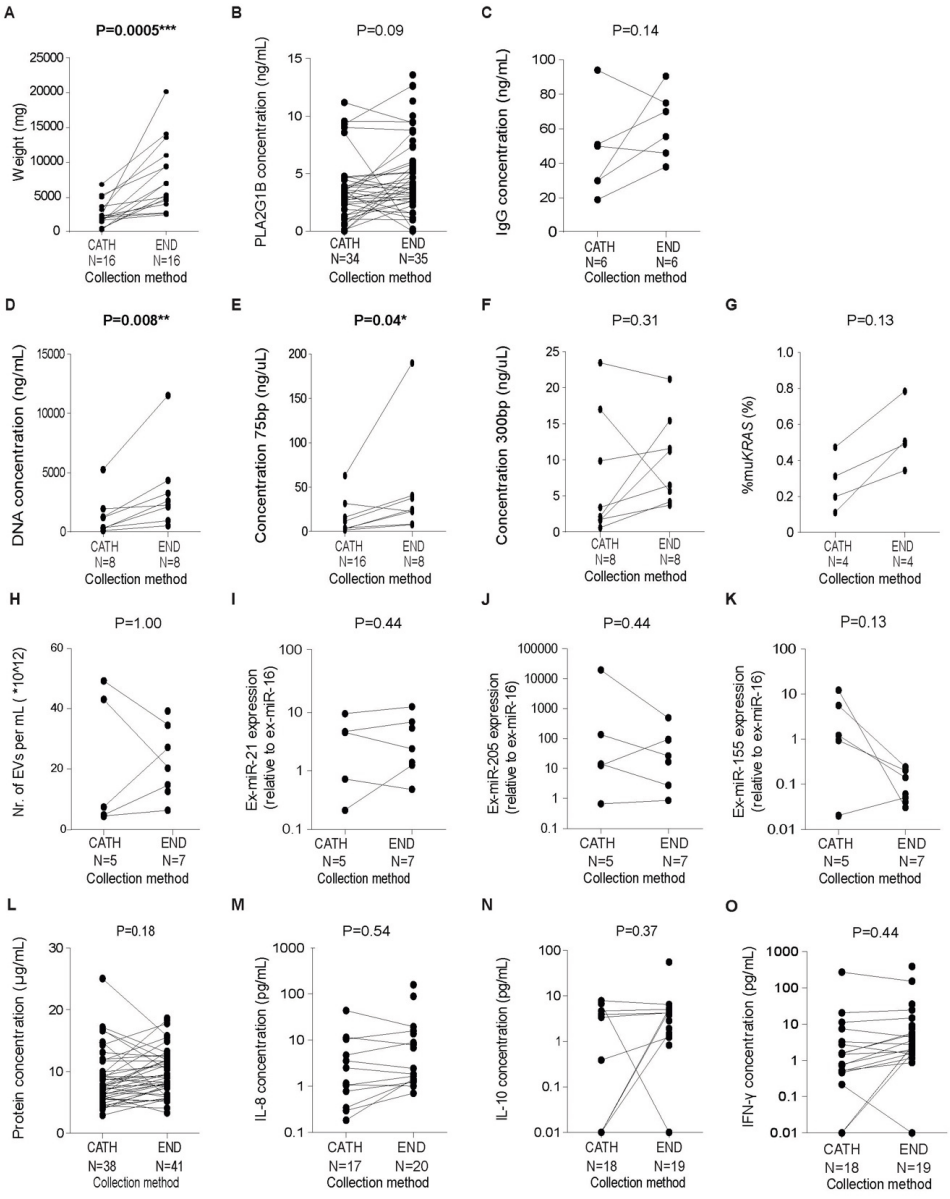


Figure 1: Comparison between collection techniques. (A) after collection, all PJ samples were weighed. The yield of juice was higher when performing suction with endoscope suction channel (END), than when using the catheter (CATH). (B) The concentration of PLA2G1B, as a measure of pancreatic origin, was similar for the two collection techniques, yet showed a trend in favor for the END. (C) The concentration of IgG, as a measure of blood contamination, did not differ between collection techniques; (D-F) The DNA concentration (total and short-segment) was higher when using END. (G) KRAS mutational load (%muKRAS) showed a trend in favor of END. (H-K) The yield of extracellular vesicles (EVs) isolated from PJ, determined by nanoparticle tracking analysis (NTA), and the expression of ex-miRs (relative to ex-miR-16) did not differ. (L-O) No differences in total protein and cytokine concentrations was found for the collection techniques.

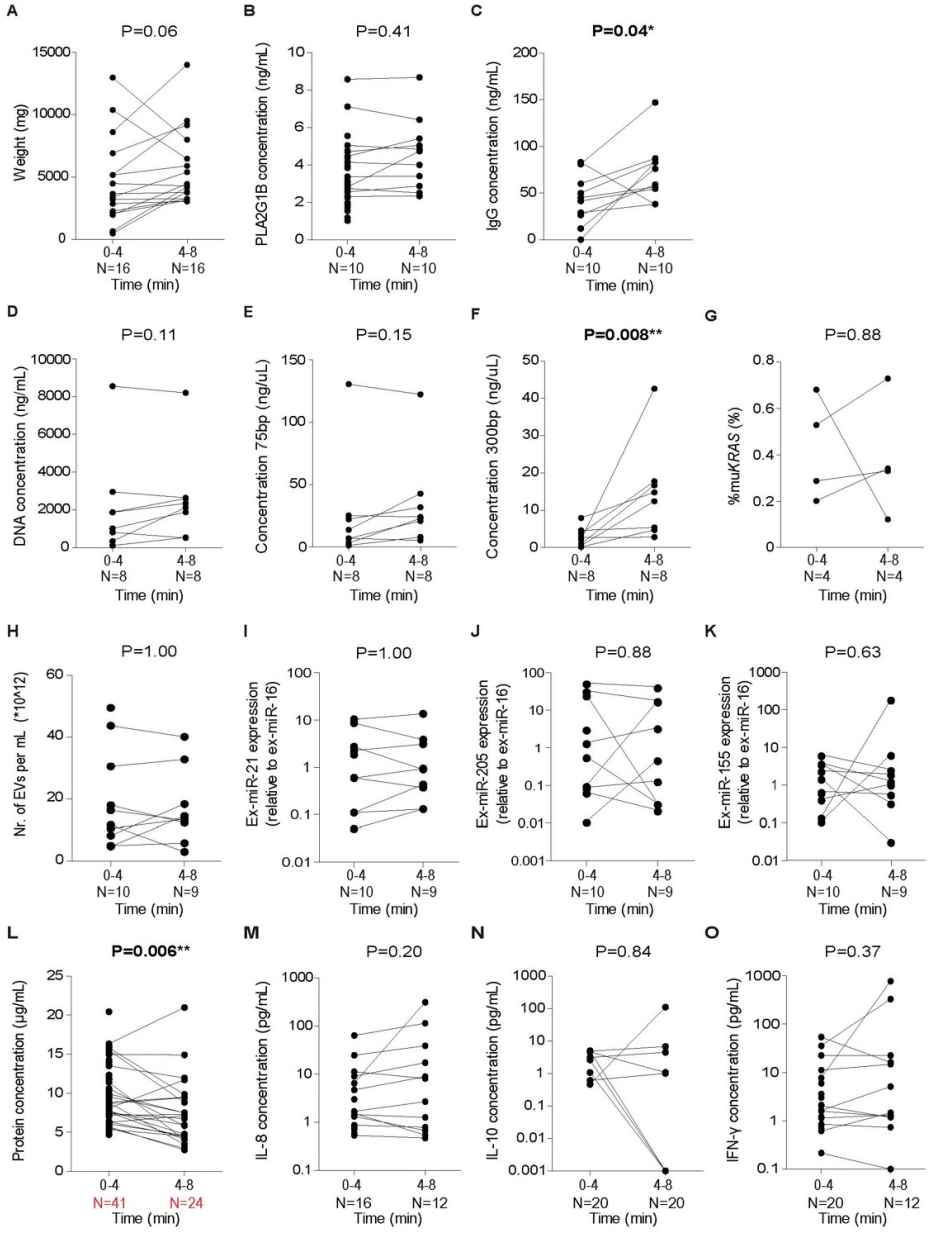


Figure 2: Comparison between two collection time periods. (A) Collection during the second timeframe resulted in a higher volume as compared to the first. (B) The concentration of PLA2G1B was similar for the two time periods. (C) The concentration of IgG was higher for the second timeframe. (D-F) The DNA concentration (isolated performed using the Maxwell kit (was not affected by the timeframe of PJ collection (D), yet concentration of longer DNA fragments (300bp) was higher for the second timeframe (F). (G) KRAS mutational load (%muKRAS) was not affected by the timeframe of PJ collection. (H-K) The yield of extracellular vesicles (EVs) isolated from PJ and the expression of ex-miRs (relative to ex-miR-16) did not differ. (L) The total protein level was higher for the first timeframe than the second timeframe. (M-O) No differences in cytokine concentrations were found for the three collection time periods.

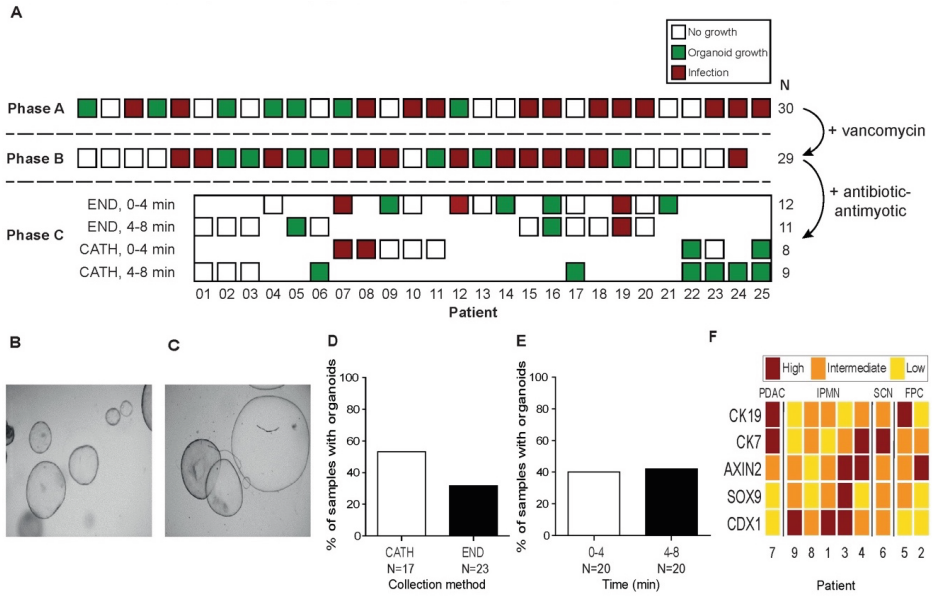


Figure 3: Workflow and yield of organoids during the sequential steps of organoid culture development. (A) During phase A, we established the use of Matrigel as a matrix (BME also worked), pre-treatment of the cell pellet with collagenase, and application of a cell strainer prior to seeding, as standard protocol. Due to a high level of infection (43% of cases), we added vancomycin to the culturing medium (Phase B). Culture-based analysis of infected samples revealed contamination with fungus and yeast, after addition of Antiobiotic-Antimyotic the infection rate decreased (phase C); (B-C) Representative picture of organoid culture from pancreatic tissue (B)and PJ (C); (D) PJ collection with catheter (CATH) results in a higher yield of organoids as compared to collection with endoscope suction channel (END); (E) Timeframe of PJ collection does not affect organoid growth; (F) Heatmap showing clinical diagnosis of patients and expression of genes reported in pancreatic (CK19, CK7, SOX9, Axin2) and intestinal (CDX1) organoids (PDAC=pancreatic ductal adenocarcinoma, IPMN = intraductal papillary mucinous neoplasm, SCN = serous cystic neoplasm, FPC=individual with hereditary risk of developing PDAC without morphologic abnormalities).

DISCUSSION

For secretin-stimulated PJ collection with an EUS-scope, we established that the most effective method is collection through the suction channel of the endoscope for no longer than 8 minutes. This resulted in the most optimal detection of a variety of potential biomarkers based on their yield, concentration or expression, and the ability to culture oganoids. This study provides a first step towards an optimized PJ collection protocol for implementation into a pancreas surveillance program. Future studies are needed to examine the potential of biomarkers to establish a reliable diagnostic test that accurately detects sub-centimeter PDAC and its precursor lesions.

ACKNOWLEDGMENTS

We would like to acknowledge L.G.B. Wermelink, D.M. De Jong, A.S. Audhoe, K.F. Meijsen, N.D.A Nieuwenhuijze, C.M. Beaufort.

REFERENCES

1. Konings, I.C., et al., *Prevalence and Progression of Pancreatic Cystic Precursor Lesions Differ Between Groups at High Risk of Developing Pancreatic Cancer.* Pancreas, 2017. **46**(1): p. 28-34.
2. Konings, I., et al., *Surveillance for pancreatic cancer in high-risk individuals.* BJS Open, 2019. **3**(5): p. 656-665.
3. Goggins, M., et al., *Management of patients with increased risk for familial pancreatic cancer: updated recommendations from the International Cancer of the Pancreas Screening (CAPS) Consortium.* Gut, 2020. **69**(1): p. 7-17.
4. Canto, M.I., et al., *Risk of Neoplastic Progression in Individuals at High Risk for Pancreatic Cancer Undergoing Long-term Surveillance.* Gastroenterology, 2018. **155**(3): p. 740-751 e2.
5. Vasen, H., et al., *Benefit of Surveillance for Pancreatic Cancer in High-Risk Individuals: Outcome of Long-Term Prospective Follow-Up Studies From Three European Expert Centers.* J Clin Oncol, 2016. **34**(17): p. 2010-9.
6. Suenaga, M., et al., *Pancreatic Juice Mutation Concentrations Can Help Predict the Grade of Dysplasia in Patients Undergoing Pancreatic Surveillance.* Clin Cancer Res, 2018. **24**(12): p. 2963-2974.
7. Wang, J., et al., *Circulating microRNAs in Pancreatic Juice as Candidate Biomarkers of Pancreatic Cancer.* J Cancer, 2014. **5**(8): p. 696-705.
8. Noh, K.W., et al., *Do cytokine concentrations in pancreatic juice predict the presence of pancreatic diseases?* Clin Gastroenterol Hepatol, 2006. **4**(6): p. 782-9.
9. Tanaka, M., et al., *Cytologic Analysis of Pancreatic Juice Increases Specificity of Detection of Malignant IPMN - A Systematic Review.* Clin Gastroenterol Hepatol, 2019.
10. Yamakawa, K., et al., *Evaluation of efficacy of pancreatic juice cytology for risk classification according to international consensus guidelines in patients with intraductal papillary mucinous neoplasm; a retrospective study.* Pancreatology, 2019. **19**(3): p. 424-428.
11. Hart, P.A., et al., *Endoscopic Pancreas Fluid Collection: Methods and Relevance for Clinical Care and Translational Science.* Am J Gastroenterol, 2016. **111**(9): p. 1258-66.
12. Broutier, L., et al., *Culture and establishment of self-renewing human and mouse adult liver and pancreas 3D organoids and their genetic manipulation.* Nat Protoc, 2016. **11**(9): p. 1724-43.

SUPPLEMENTARY DATA

Supplementary Material and Methods

Study design

The study comprised of two phases (**Figure 1**). During the first phase, we compared two collection techniques (utilization of a through-the-scope catheter [CATH] *vs* performing suction through the endoscopic channel [END]) and two collection timeframes (0-4 minutes *vs* 4-8 minutes), by assessing concentrations (or expression) of selected biomarkers. The collection timeframes were based on agreement by three experienced endoscopists, who judged an 8-minute collection as logistically feasible when performing multiple juice collections per half-day EUS program. During this first phase, we performed additional assessments; 1. Phospholipase A2 group1B (PLA2G1B; representing true pancreas-derived material), total IgG and albumin concentrations (representing blood and bile contamination); 2. Reproducibility between endoscopists; 3. The effect of adding a protease or superase inhibitor on the biomarker concentration; 4. The optimal way to isolate DNA from PJ (Nucleospin Kit vs. Maxwell Kit); and 5. The ability to grow organoids from PJ and development of an organoid culture protocol.

After determination of the optimal collection and analysis techniques in the first phase, in the second phase, we investigated whether biomarker detection improved with even longer collection. Taking into account the results from Suenaga et al.[1], we extended the collection duration to 15 minutes, and performed a paired-wise comparison – based on the same biomarkers as in phase 1 – between juice collected during the three timeframes (0-4, 4-8 and 8-15 minutes).

Pancreatic Juice Collection

PJ collection was performed by three experienced endoscopists. To reduce duodenal contamination, duodenal fluid was aspirated prior to juice collection. Next, a wash-out of PJ was stimulated by intravenous administration of human synthetic secretin (ChriRhoStim, Burtonsville, MD, 16µg/patient). Collection started immediately after injection and took eight (phase 1) or 15 minutes (phase 2). To enable comparison of the different timeframes, after each collection timeframe (at t=4 for phase 1, at t=4 and t=8 for phase 2), both the mucus extractor (END) and syringe (CATH) were replaced. The collection techniques were alternated every 30 seconds (**Figure 1**).

Collection of PJ in the first technique (END) was performed by applying suction with the endoscope (Pentax Medical, Tokyo, Japan). For this, a mucus extractor (Pennine Healthcare, Derby, United Kingdom, 15 mL) was attached to the proximal end of the endoscopic channel.

For the second technique (CATH), a catheter (Huijbregtse, 7 Fr, Cook Medical, Bloomington, IN) was passed through the scope and positioned close to the ampullary orifice (without cannulation) (**Supplementary Figure S1**). A three-way stopcock syringe was attached to the proximal end, to prevent efflux of juice back to the duodenum, through which suction was applied by the assisting nurse.

After collection, PJ for organoid culture was kept on ice and processed within 2h. For the other tests, PJ was aliquoted into 1.5 mL tubes and protease and superase inhibitors were added in 50% of the subjects (randomly). PJ was snap frozen within 10 minutes after collection. Samples were weighed as a proxy for the yield of collected PJ and stored at -80°C until further use.

cfDNA quantification and qualification

To investigate the optimal technique for cfDNA isolation, two extraction kits were used according to manufacturer's instructions: (1) the silica membrane-based NucleoSpin DNA kit (Bioké, Leiden, The Netherlands, #6181527); and (2) the automatic bead-based Maxwell RSC cfDNA Plasma kit (Promega, Fitchburg, WI, AX1115). To quantify the concentration of total (double stranded) DNA, the Quant-iT dsDNA High-Sensitivity (HS) Assay Kit was employed, according to manufacturer's instructions. To quantify 75 base-pair (bp), 150bp and 300bp DNA fragments, the ProNex® DNA QC Assay (Promega, Fitchburg, WI) a human-specific, multiplexed probe-based quantitative polymerase chain reaction (qPCR) was used, according to the manufacturer's protocol. This assay also includes an internal positive control to test for false-negative results that may occur in the presence of PCR inhibitors.

KRAS mutational load determination

To generate sufficient copies of DNA, cfDNA was pre-amplified with the Taqman PreAmp master mix (Thermo Fisher Scientific, Waltham, MA, #4488593). For this, the 20x primer-probe KRAS Screening assay (Bio-Rad, Hercules, CA, #186-3506) was diluted 100 times in LoTe (3 mM Tris-HCL (pH 8.0)/0.2 mM EDTA (pH 8.0)). A PreAmp reaction mix consisting of Taqman, Pre-Amp master mix (4 µL), 100-fold diluted KRAS Screening assay (2 µL) and DNA (0.1-4.0 ng in 2 µL) was prepared and PCR was performed under the following conditions: cycle at 95°C for 10 minutes, 15 cycles at 95°C for 15 seconds and 60°C for 4 minutes, followed by a cool-down to 4°C. Finally, 8 µL of pre-amplified product was diluted 10-fold in LoTe. KRAS copies present in the pre-amplified product were quantified in a regular quantitative PCR (primer sequences and PCR amplification program). For this, 2 µL of the 10-fold diluted pre-amplified sample was added to 2.5 µL PCR mastermix (GCBiotech, Waddinxveen, The Netherlands, #BIO-84020), 4.5 µL H_2O and 0.5 µL 20x KRAS-Screening assay. PCR conditions were as follows: 1 cycle at 95°C for 3 minutes, followed by 45 cycles at 92°C for 10 seconds and 60°C for 1 minute. After this quantification, 10 to 30 ng amplification product from the first round of PCR was used for digital polymerase chain reaction (dPCR, Thermo Fisher Scientific, Waltham, MA, quant3D studio). In each sample, KRAS hotspot mutations were assessed with the KRAS Screening Multiplex Kit (Bio-Rad, Hercules, CA,#186-3506). Primers (final concentration, 900 nM) and probes (final concentration 250 nM) present in this kit were designed to detect mutated G12S, G12D, G12R, G12V, G13D and wild-type (WT) KRAS. The fluorescent label FAM was used to quantify the number of mutated copies and HEX was used to determine the number of WT copies. Finally, the resulting cycle quantification values of HEX were used to calculate the optimal volume of sample to load into the dPCR chip, as described before[2]. dPCRs were performed with the QuantStudio 3D Digital PCR System (Thermo

Fisher Scientific, Waltham, MA), according to the manufacturer's protocol. For this, each pre-amplified diluted DNA sample was portioned into 20.000 wells of a QuantStudio 3D Digital PCR v2 Chip and run on a ProFlex 2x Flat PCR System (Thermo Fisher Scientific, Waltham, MA). The target-specific optimized PCR program was as follows: 10 minutes at 96°C, followed by 40 cycles of 30 seconds incubation at 98°C, and 2 minutes at 52°C, and a final pause at 10°C. Chips were read in a QuantStudio 3D dPCR instrument, and analyzed with web-based QuantStudio 3D dPCR Analysis Software version 3.4.1 (Thermo Fisher Scientific, Waltham, MA).

Exosome isolation and analysis

400 µL of PJ was centrifuged for 10 min at 4000 RPM 4°C to remove debris. Then, 100 µL of Total Exosome Isolation Reagent (Thermo Fisher Scientific, Waltham, MA, #4478359) was added to 200 µL of supernatant and kept on a rollerbank at 4°C overnight. After this, samples were centrifuged for 1h at 14000 RPM and the pellet was resuspended in 400 µL of PBS (filtered with 0.2µM filter). For Nanoparticle Tracking Analysis (NTA), samples were diluted 1:1000 in PBS. The size and concentration of the extracellular vesicles (EV) were detected by NanoSight NS300 (NTA 3.4 Build 3.4.003 software). Concentrated EVs were stored at -80°C until further analysis.

Exosomal miRNA analysis

miRNA (miR) was isolated from 200 uL EV preparation with QIAzol Lysis Reagent (Qiagen, Hilden, Germany, #79306) and miRNeasy Mini kit (Qiagen, Hilden, Germany, #217004) according to manufacturer's recommendations. miRNA-specific cDNA was prepared using the Taqman microRNA Reverse Transcription Kit (Thermo Fisher Scientific, Waltham, MA, #217004; miR-16, miR-21, miR-205, miR-155) as described before[3, 4]. In a modified protocol, every cDNA reaction consisted of 0.4 µL dNTP mix, 1.35 µL Multiscribe RT enzyme (500U/µL), 2.0 µL 10x RT Buffer, 0.25 µL RNase inhibitor, 1.0 µL of each RT primer, and 5 µL of diluted template RNA. The total reaction volume was adjusted to 20 µL with nuclease-free water. All cDNA and qPCR reactions were performed according to the manufacturer's instructions and carried out in duplicate. Each qPCR reaction consisted of 6 µL TaqMan Universal PCR Master Mix (Thermo Fisher Scientific, Waltham, MA, #4324018), 0.5 µL microRNA-specific PCR primer and 5.0 µL of the previously 1:5 diluted cDNA. The final volume of every PCR reaction was adjusted to 12 µL with nuclease-free water. MiRNA expression changes were calculated relative to miR-16 as a reference gene using the $2^{-\Delta\Delta Ct}$ method[5], as reported before[6-9]. Additionally, we explored the possibilities of normalization to exosome concentration (ΔCT) in this analysis.

ELISA

The total protein concentration in PJ was assessed by Lowry protein assay (Bio-Rad, Hercules, CA)[10]. Interleukin-8 (IL-8), interleukin-10 (IL-10), and interferon-γ (IFN-γ) were measured by Enzyme-Linked

Immuno Sorbent Assay (ELISA) according to manufacturer's protocol of the used kits (Thermo Fisher Scientific, Waltham, MA, #88-8086, #88-7106, #88-7316). Briefly, immunosorb plates (Nunc, Hardenberg, The Netherlands) were coated with cytokine-specific capture antibody overnight at 4 °C and plates were blocked with ELISA diluent for 1 hour at room temperature. PJ (75 µL per well) was incubated at 4 °C overnight, after which biotin conjugated detection antibody was added at room temperature for 1 hour. Following incubation of avidin-HRP for 30 minutes at room temperature TMB substrate was added. Reactions were stopped by addition of sulfiric and absorbance was read at 450 nm (Tecan Infinite200 pro plate reader). Assessment of the concentrations of PLA2G1B (pancreatic marker, MyBiosource, San Diego, #MBS703283) and IgG (Thermo Fisher Scientific, Waltham, MA, #BMS2091) were performed similarly, using pre-coated, pre-blocked plates according to the manufacturer's protocol. Albumin levels were detected using an in house designed protocol.

Organoid growth

Organoid culture was based on a protocol for culture of tissue-derived pancreatic organoids described by Broutier et al[11]. 2-4 mL of PJ was collected in a 15 mL tube containing 5 mL of basal medium supplemented with 100 µg/mL of vancomycin and 1x of Antibiotic-Antimycotic (100X) (Thermo Fisher Scientific, Waltham, MA, #15240062) and kept on ice until processing. PJ was incubated with collagenase II (0.1 mg/mL; Sigma-Aldrich, St. Louis, MO, #C9891-100MG) for 20 min on a shaker at 37°C to dissociate tissue clumps. To remove collagenase, samples were centrifuged at 1350 RPM for 5 min and pellet was washed with 5 mL of wash medium, which is DMEM (Thermo Fisher Scientific, Waltham, MA, #41965039) supplemented with 1% UltraGlutamine I (Alanyl-L-Glutamine, Westburg BV, Leusden, NL, #BE17-605E/U1), 1% sodium pyruvate (Life Technologies, Carlsbad, CA, #11360070), 1% FBS (Sigma-Aldrich, St. Louis, MO, #F7524-500ML), 100 µgmL of vancomycin, and 1x of Antibiotic-Antimycotic (100X). Next, cells were passed through a 70 µM cell strainer to remove clumps or debris and washed once with 10 mL of wash medium and once with 5 mL of basal medium (Advanced DMEM/F-12 [Invitrogen, Carlsbad, CA, #12634-028], with 1% GlutaMAX [Invitrogen, Carlsbad, CA, #35050-079] and 10mM HEPES [Invitrogen, Carlsbad, CA, #15630-056]). The pellet was seeded in a pre-warmed 24-well plate for cell suspension (Corning, Corning, NY) in a 50 µL drop of matrigel (BD bioscience, Franklin Lanes, New Jersey, USA #356231) or BME (Bio-Techne, Minneapolis, MN, #3533-010-02). Droplets were incubated at 37°C until solidified. 500 µL of pancreatic organoid isolation and expansion medium was added to each well: Advanced DMEM/F-12, 1% penicillin/streptomycin, 1% GlutaMAX, HEPES 10 mM, 1:50 B27 supplement (without vitamin A) (Invitrogen, Carlsbad, CA, #17504-044), 1:100 N2 supplement (Invitrogen, Carlsbad, CA, #17502-048), 1 mM N-acetylcysteine (Sigma-Aldrich, St. Louis, MO, #A9165-5G), 30% (v/v) Wnt3a-conditioned medium, 5% (v/v) Rspo1-conditioned medium, 10 mM nicotinamide (Sigma-Aldrich, St. Louis, MO, #N0636), 10 nM recombinant human [Leu15]-gastrin I (Sigma-Aldrich, St. Louis, MO, #G9145), 50 ng/mL recombinant human EGF (Peprotech EL Ltd Princeton, NJ, #AF10015), 100 ng/mL recombinant human FGF10 (PeproTech EC Ltd, Princeton, NJ, #100-26), 5% (v/v) Noggin-conditioned medium, 5 µM A83-01 (Tocris, Abingdon, UK, #2939), and 3 µM PGE-2 (Tocris, Abingdon, UK, #2296). 10 µg/mL vancomycin, 10 µM RhoK

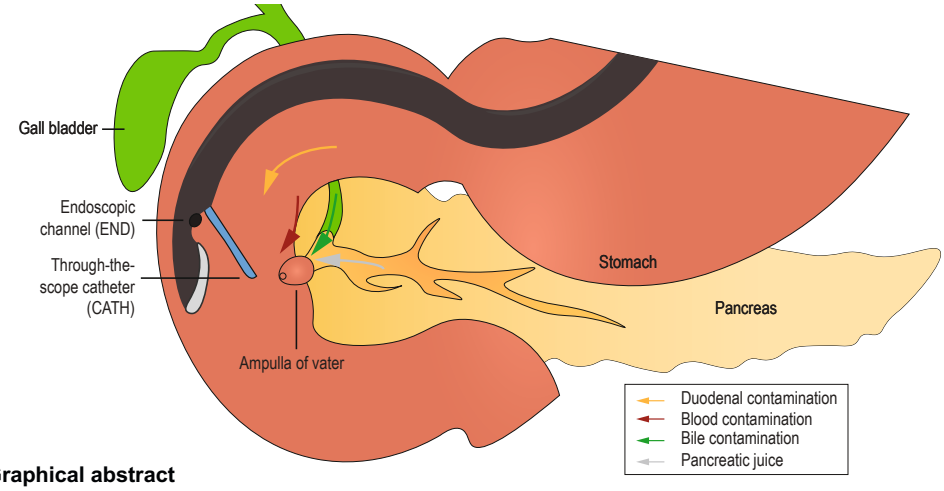
inhibitors (Y-27632; R&D Systems Europe, #1254/10) and 1X Antibiotic-Antimycotic (100X) and 10 µG/mL vancomycin were added only during the first 2-3 days post-seeding. Plates were kept under standard tissue culture conditions (37 °C, 5% CO2) for at least 2 weeks and checked for organoid growth. Medium was replaced 3 times per week. Organoids were passaged by mechanical disruption with a pipette or in combination with TrypLE™ Express Enzyme (1X) (Thermo Fisher Scientific, Waltham, MA, #12604013). All variations of the protocol are indicated in the result section.

The *qPCR* for CK19, CK7, SOX9, Axin2 and CDX1 was performed as described[12]. In short, total RNA from organoids was isolated for cDNA preparation. Primers used (first forward, then reverse:
CK19 (CTACAGCCACTACTACACGAC, CAGAGCCTGTTCCGTCTCAAA),
CK7 (GGGGACGACCTCCGGAATAC, CTTGGCACGCTGGTTCTTGA),
SOX9 (GGAAGTCGGTGAAGAACGGG, TGTTGGAGATGACGTCGCTG),
Axin2 (TATCCAGTGATGCGCTGAC, TTACTGCCCACACGATAAGG),
CDX1 (GTGGCAGCGGTAAGACTC, GTTCACTTTGCGCTCCTTTGC). Data are calculated based on $2^{-\Delta\Delta Ct}$ method [5] and presented as relative expression to RP2 (forward AAGCTGAGGATGCTCAAAG, reverse CCCATTAAACTCCAAGGCAA).

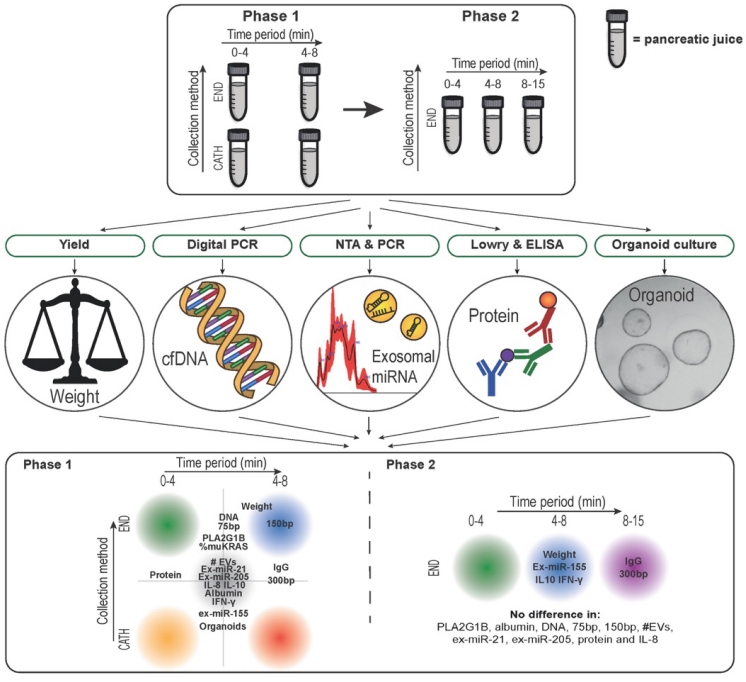
References for supplementary material and methods

1. Suenaga, M., et al., *The Effect of Pancreatic Juice Collection Time on the Detection of KRAS Mutations.* Pancreas, 2018. 47(1): p. 35-39.
2. Vitale, S.R., et al., *An Optimized Workflow to Evaluate Estrogen Receptor Gene Mutations in Small Amounts of Cell-Free DNA.* J Mol Diagn, 2019. 21(1): p. 123-137.
3. Verhoeven, C.J., et al., *MicroRNA profiles in graft preservation solution are predictive of ischemic-type biliary lesions after liver transplantation.* Journal of Hepatology, 2013. 59(6): p. 1231-1238.
4. Farid, W.R.R., et al., *Hepatocyte-derived microRNAs as serum biomarkers of hepatic injury and rejection after liver transplantation.* Liver Transplantation, 2012. 18(3): p. 290-297.
5. Livak, K.J. and T.D. Schmittgen, *Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method.* Methods, 2001. 25(4): p. 402-8.
6. Nakamura, S., et al., *Pancreatic Juice Exosomal MicroRNAs as Biomarkers for Detection of Pancreatic Ductal Adenocarcinoma.* Annals of Surgical Oncology, 2019.
7. Xu, J., et al., *Plasma miRNAs Effectively Distinguish Patients With Pancreatic Cancer From Controls: A Multicenter Study.* Ann Surg, 2016. 263(6): p. 1173-9.
8. Abue, M., et al., *Circulating miR-483-3p and miR-21 is highly expressed in plasma of pancreatic cancer.* Int J Oncol, 2015. 46(2): p. 539-47.
9. Wang, J., et al., *MicroRNAs in plasma of pancreatic ductal adenocarcinoma patients as novel blood-based biomarkers of disease.* Cancer Prev Res (Phila), 2009. 2(9): p. 807-13.
10. Lowry, O.H., et al., *Protein measurement with the Folin phenol reagent.* J Biol Chem, 1951. 193(1): p. 265-75.
11. Broutier, L., et al., *Culture and establishment of self-renewing human and mouse adult liver and pancreas 3D organoids and their genetic manipulation.* Nat Protoc, 2016. 11(9): p. 1724-43.
12. Janmaat, V.T., et al., *HOXA9 mediates and marks premalignant compartment size expansion in colonic adenomas.* Carcinogenesis, 2019. **40**(12): p. 1514-1524

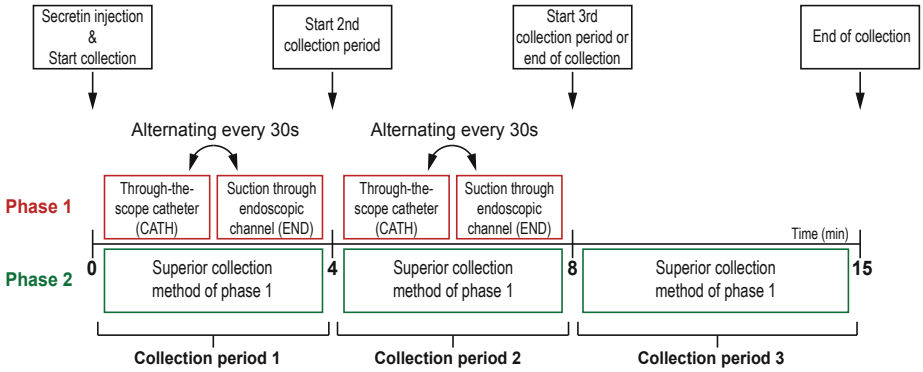
Supplementary figures



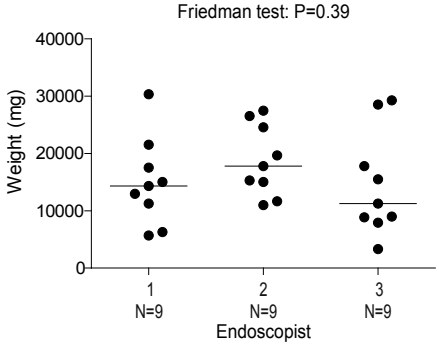
Graphical abstract



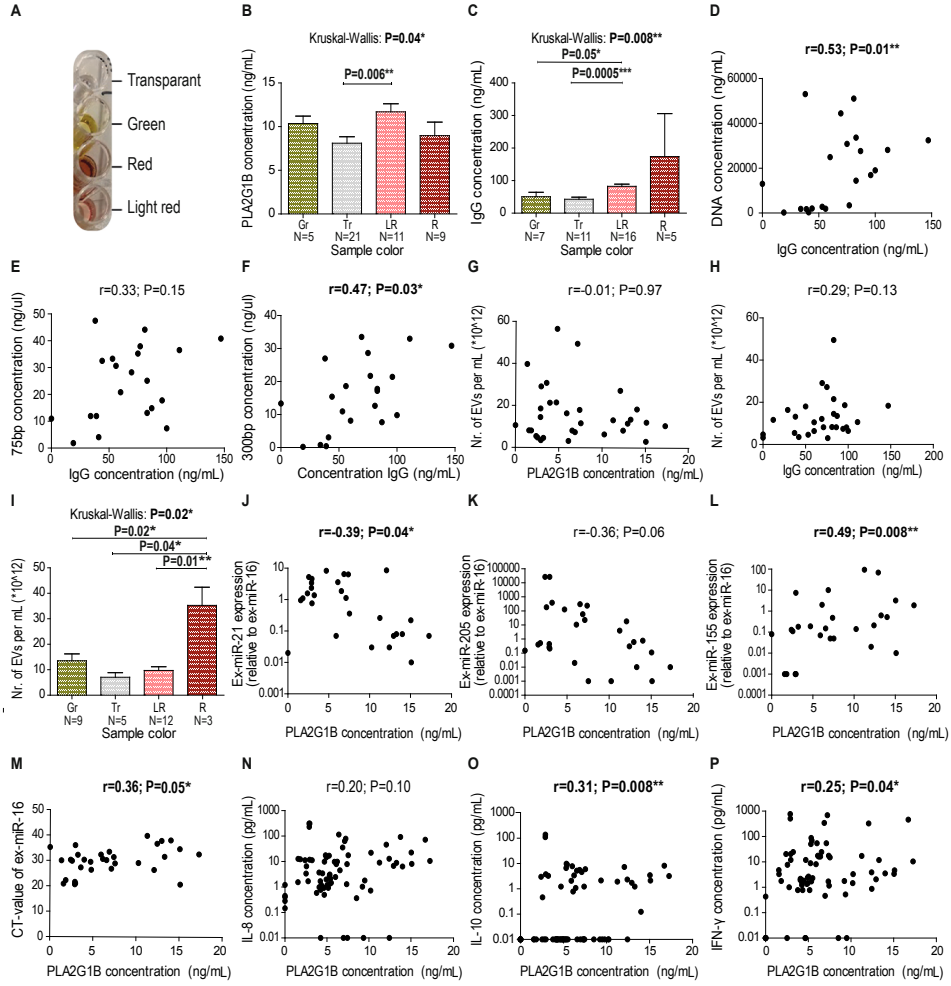
Supplemental Figure S1: Graphical representation of the positions of the two collection methods: performing suction using a trough-the-scope-catheter (CATH) or with the endoscope suction channel (END). A through-the-scope catheter may decrease (duodenal) contamination by precise positioning of the catheter close to the ampullary orifice, while performing suction with the endoscope suction channel may yield higher volumes. Additionally, the first flush after secretin injection may be stagnant remnants of earlier ejections, and unsuitable for biomarker detection. Alternatively, the first wash-out may be more concentrated, while prolonged collection may dilute the biomarkers of interest.



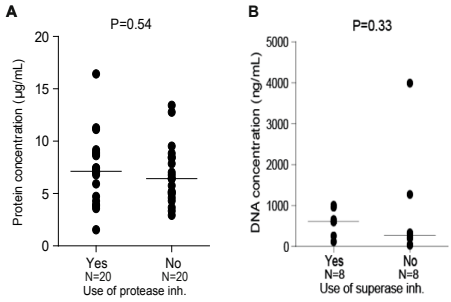
Supplemental Figure S2: Graphical overview of pancreatic juice collection during endoscopic ultrasound. Collection starts immediately after secretin injection. For phase 1, collection takes two times four minutes with alternating collection techniques. For phase 2, collection is prolonged to 15 minutes. The superior collection technique is used in this phase.



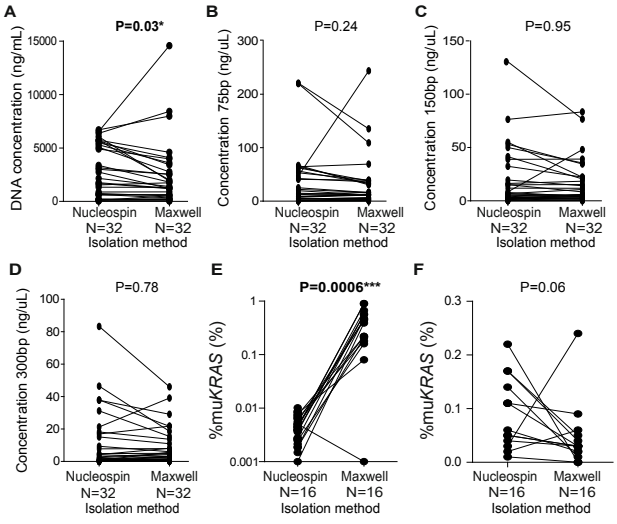
Supplemental Figure S3: The collected volume of pancreatic juice did not differ between the endoscopists. Groups were matched based on study cohort (FPC N=6; PDAC N=3).



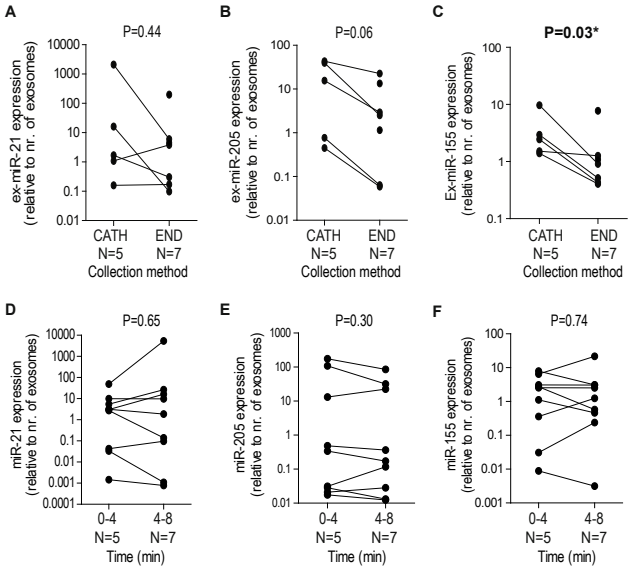
Supplemental Figure S4: Color and biomarker concentration in relation to PLA2G1B and IgG concentration. (A) Examples of PJ colors ranging from green (Gr) to transparent (Tr), light red (LR) and dark red (R). (B, C) The PLA2G1B and IgG concentrations in PJ stratified according to color. Red color and green color, as a potential measure of, respectively, blood and bile contamination did not result in lower PLA2G1B concentrations (B). The redder the color, the higher the IgG concentration (C). (D-F) A correlation was found for both total and long-segment (300bp) DNA concentration (D, F). DNA concentration did not correlate with PLA2G1B (not shown). (H, I) The total number of extracellular vesicles (EVs) was not significantly associated IgG concentration correlation (although a trend was seen), yet was associated with a dark red color. The number of EVs was not associated with PLA2G1B concentration (not shown). (J-L) The tested ex-miRs were correlated with PLA2G1B. Ex-miR-21 and Ex-miR-205 were, surprisingly, negatively correlated with PLA2G1B (J,K), while Ex-miR-155 was positively correlated (L). (M) The CT-value of ex-miR-16, used in this study as internal control was correlated with PLA2G1B. (N-P) IL-8 was not, yet IL-10 and IFN-γ were correlated with PLA2G1B concentration. Although cytokines are known to be present in blood, cytokine concentrations were not associated with IgG or sample color (not shown).



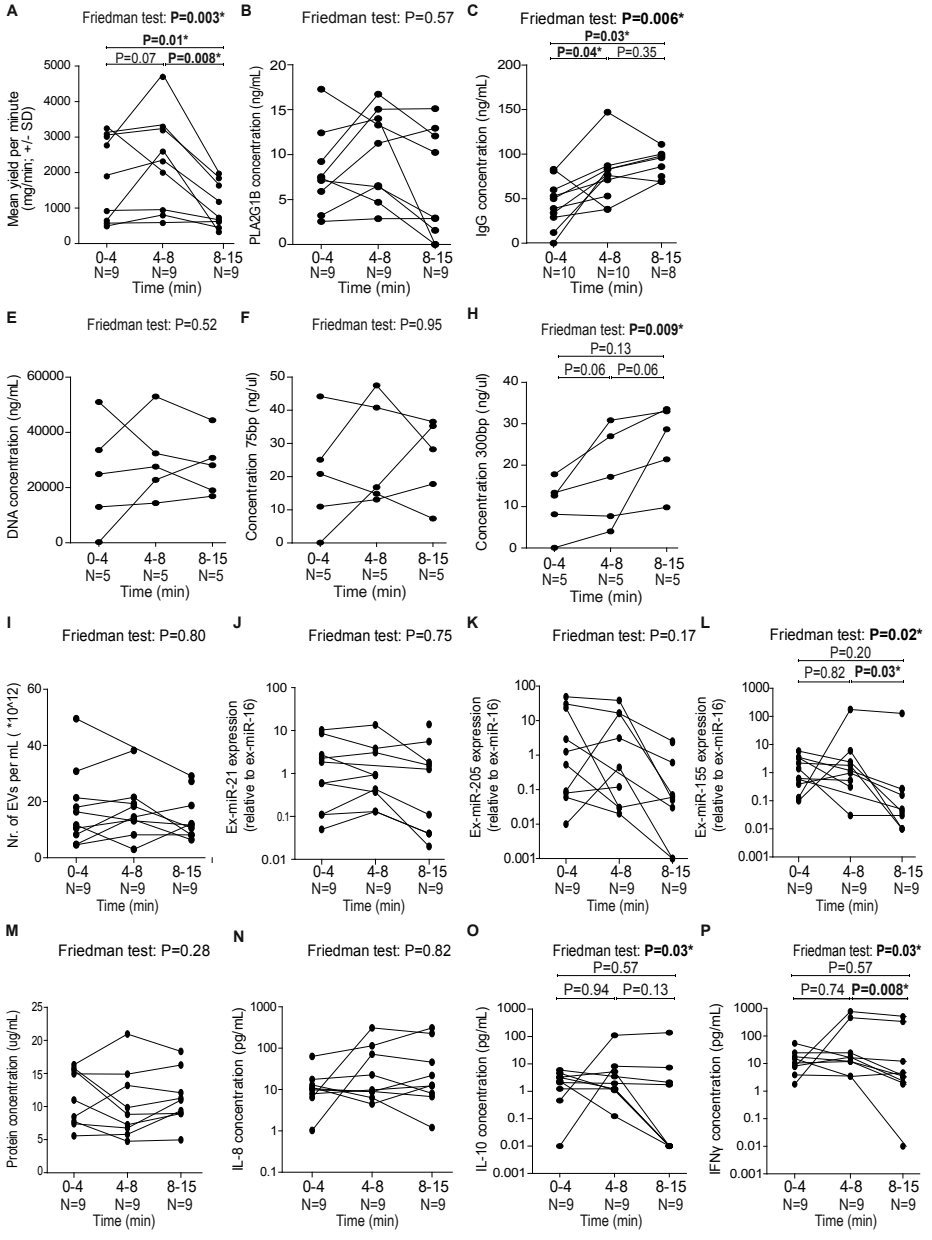
Supplemental Figure S5: The effect of a buffer on pancreatic juice. (A) The addition of protease inhibitor (inh.) does not affect protein concentration; (B) Superase inhibitor does not affect DNA concentration.



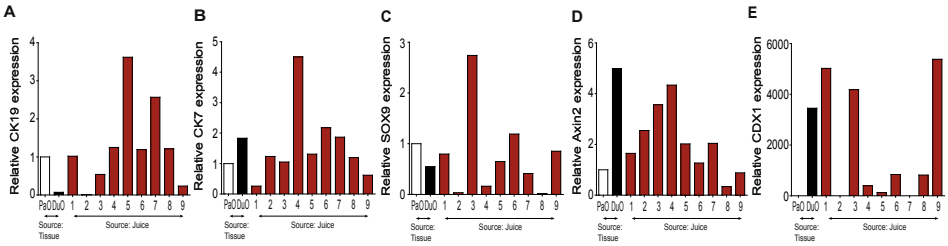
Supplemental Figure S6: Comparison of total DNA concentration and mutational component for two DNA isolation kits. The cfDNA was isolated Maxwell and Nucleospin cfDNA isolation kits, and the concentration of DNA and the percentage mutated KRAS (%muKRAS) were determined. (A) The total concentration of DNA was higher when using the Nucleospin kit. (B-C) The concentration of isolated 75bp, 150bp and 300bp fragments did not differ between the two cfDNA isolation methods. (E) A significantly higher %muKRAS was detected in cfDNA isolated by Maxwell kit. (F) KRAS mutation rate was determined by digital PCR at least two times for individual samples, and error was indicated by a standard deviation. The standard deviation was lower (indicating higher reproducibility) when measurements were performed after isolation with the Maxwell kit.



Supplemental Figure S7: Exosomal miRNA-155 (ex-miR-155) levels (relative to the number of exosomes) are higher in PJ collected through suction with the catheter. The expression of the selected ex-miRs was compared for the collection methods (A-C) and time periods (D-F). However, as both blood and bile contain exosomes, contamination of PJ with these fluids may bias results when normalizing towards the number of exosomes present in these juices.



Supplemental Figure S8: The yield of pancreatic juice and biomarkers for longer collection time. The yield of pancreatic juice, IL10 and IFN-γ and the expression of ex-miR-155 decreases after 8 minutes of collection. In contrast, the concentration of IgG and long-segment DNA (300bp) increases after 8 minutes. The change of yield of PJ (per minute; A), contamination (B-D), DNA (E-H), extracellular vesicles (I) proteins (K-N), and ex-miR (J-L) over time is shown.



Supplemental Figure 9. Barcharts showing expression of CK9, CK7, SOX9 and Axin2 (as reported for organoid culutured from pancreatic organoids; A-D) and CDX1 (intestinal marker; E)

Supplementary tables

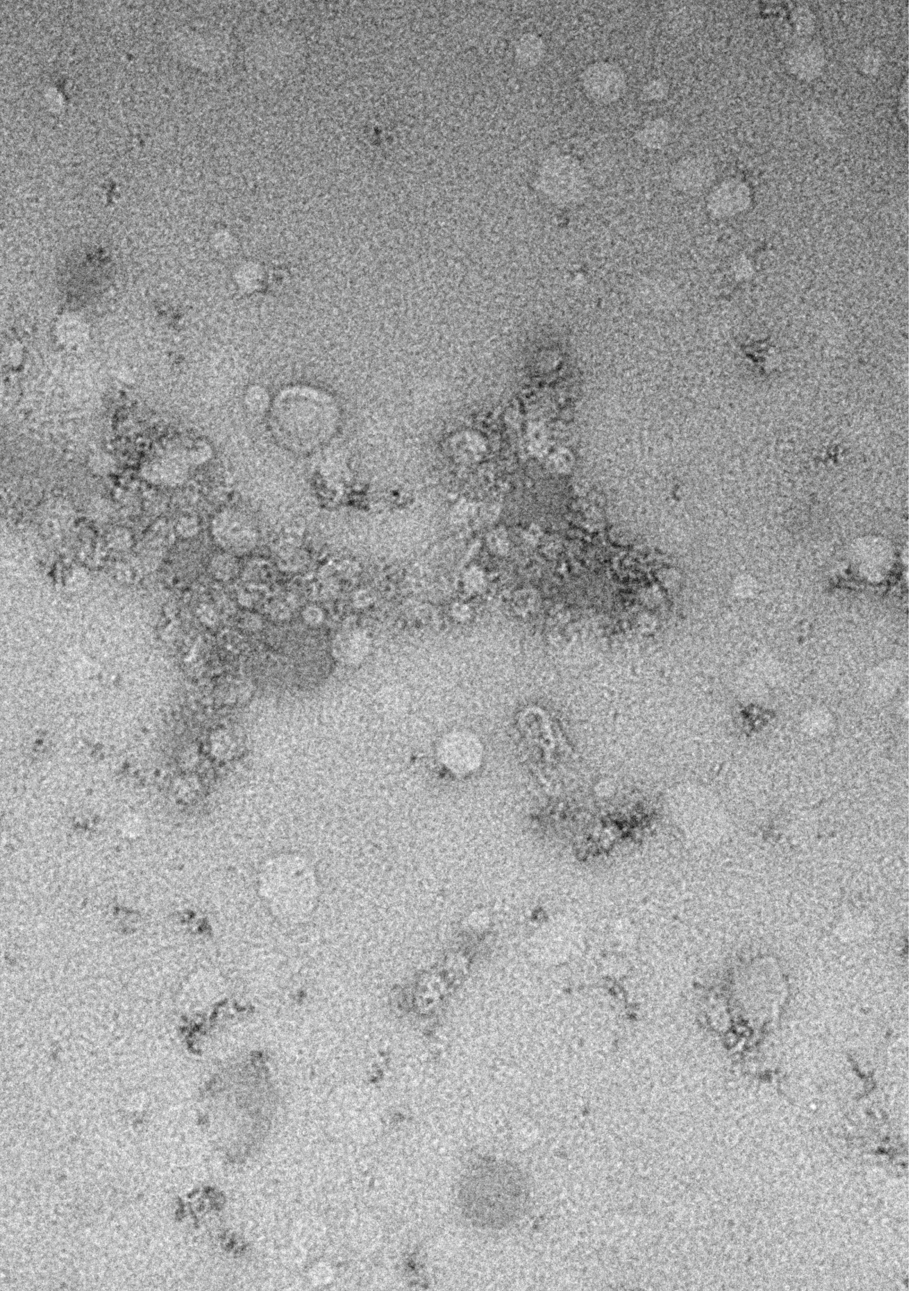
Supplemental Table S1: Characteristics of patients included in phase 1 of this study. (PDAC = Pancreatic ductal adenocarcinoma; FPC = Familial pancreatic cancer; SB-IPMN = Side-branch intraductal papillary mucinous neoplasm; IQR = interquartile range).

	FPC (N=23)	PDAC (N=18)
Age (years), median (IQR)	60 (48-67)	68 (63-72)
Gender, n female (%)	17 (74)	9 (50)
Indication		
Surveillance, n (%)	23 (100)	-
Suspected PDAC, n (%)	-	13 (72)
Fiducials placement, n (%)	-	5 (28)
Pathology proven PDAC, n (%)	-	18 (100)
Mutation carrier		
BRCA1, n (%)	1 (4)	-
BRCA2, n (%)	3 (13)	-
CDKN2A, n (%)	6 (26)	-
Peutz-Jeghers, n (%)	1 (4)	-
Pancreatic cyst		
Aspecific cyst, n (%)	7 (30)	-
SB-IPMN, n (%)	4 (17)	-
If cyst: Size (mm), median (IQR)	5 (2-9)	-
If cyst: absolute or relative indications for surgery, n (%) [1]	0 (0)	-

[1]. European evidence-based guidelines on pancreatic cystic neoplasms. Gut, 2018. 67(5): p. 789.

Supplemental Table S2: Characteristics of patients included in phase 2 of the study. (NA = not applicable; FPC = familial pancreatic cancer; PDAC = Pancreatic ductal adenocarcinoma; LAPC = locally advanced pancreatic cancer; SB-IPMN = Side-branch intraductal papillary mucinous neoplasm; MT = mixed-type intraductal papillary mucinous neoplasm; EUS = endoscopic ultrasound).

	Age (years)	**Gender**	**Mutation carrier**	**EUS indication**	**Morphology pancreas**	**Pathology proven (yes/no)**
FPC1	53	F	No	Surveillance	No abnormalities	No
FPC2	71	M	BRCA2	Surveillance	Multifocal SB-IPMN, largest cyst 14 mm, no worrisome features.	No
FPC3	63	F	No	Surveillance	Multifocal SB-IPMN, largest cyst 9 mm, no worrisome features	No
FPC4	66	F	No	Surveillance	No abnormalities	No
FPC5	40	F	No	Surveillance	No abnormalities	No
PDAC1	73	F	NA	Fiducial placement	Pathology-proven LAPC	Yes
PDAC2	72	F	NA	Diagnosis	Pathology-proven LAPC	Yes
Cyst1	53	F	NA	Surveillance	Multifocal SB-IPMN, largest cyst 12 mm, no worrisome features	No
Cyst2	48	F	NA	Surveillance	SB-IPMN, 24 mm, non-enhanced (thickened) septation.	No
Cyst3	65	F	NA	Surveillance	MT-IPMN, PD 5mm, cyst 20mm, enhancing mural nodule.	No



Chapter 3

Size and concentration of extracellular vesicles in pancreatic juice from pancreatic ductal adenocarcinoma patients

K. Nesteruk, I. J. M. Levink, N. F. J. Dits, D. L. Cahen, M. P. Peppelenbosch, M. J. Bruno, G. M. Fuhler

Journal of Extracellular Vesicles. 2021 (submitted)

INTRODUCTION

Detection of pancreatic ductal adenocarcinoma (PDAC) at a curable stage is challenging due to the late presentation of symptoms and limited visibility of sub-centimeter lesions on imaging. Therefore, accurate biomarkers for early detection are urgently needed. In recent years, extracellular vesicles (EVs) have gained interest as potential disease biomarkers. EVs carry a unique molecular cargo to communicate between cells and are expected to represent a cell-specific signature [1]. Cancer cells release EVs to form a pre-metastatic niche [1]. Thus, detection of cancer-derived EVs based on their content may predict the presence of disease. While blood-born EVs are most frequently studied in this context, for PDAC, pancreatic juice (PJ) may be a promising biomarker source, as it is in close contact with ductal cells from which PDAC arises. Indeed, detection of microRNA molecules from EVs derived from PJ was able to distinguish PDAC from controls [2]. Interestingly, the concentration of EVs determined by nanoparticles tracking analysis in bile discriminated patients with malignant (including PDAC) from non-malignant common bile duct stenosis with 100% accuracy [3]. These data suggest that cancer cells emit elevated numbers of EVs in the extracellular space [3], which may be exploited to detect the presence of cancer without studying specific EV content and would greatly simplify testing. Here, we characterize the size and concentration of EVs in PJ and serum of PDAC patients and controls, to establish whether this may present a promising biomarker for early detection of PDAC.

METHODS

Secretin-stimulated PJ was collected from the duodenum during endoscopic ultrasound (EUS) from 54 individuals with sporadic PDAC (cases) and 117 non-malignant controls undergoing pancreatic surveillance [4]. PJ was collected with the endoscope for up to 8 minutes starting immediately after intravenous administration of human synthetic secretin. Serum was available for 46 cases and 58 controls. PLA2G1B concentration, as a measure of pancreatic specific content, was determined by enzyme linked immunosorbent assay, and the total protein concentration by Lowry protein assay [5]. EVs were extracted from 200 µL of supernatant of PJ or serum and analyzed with NTA, transmission electron microscopy (TEM), Lowry assay and Western blot. Details are provided in the **Supplementary methods**.

RESULTS

A summary of clinical characteristics is provided in **Supplementary Table 1**. Pancreas-specific PLA2G1B was detected in PJ of all PDAC patients and 115/117 controls, and concentrations of PLA2G1B ($p=0.22$) and total protein content ($p=0.24$) did not differ between cases and controls, indicating similar PJ quality (**Supplementary Figure S1**).

For both PJ and serum, isolated EV fractions showed round double-membrane vesicle-like structures, typical of EVs (**Figure 1A),** which express membrane and cytoplasmic EV markers such as CD81, caveolin-1 and GAPDH (**Figure 1B**). NTA analysis allowed visualization of heterogeneous populations

of spherical nanoparticles moving under Brownian motion (**Figure 1C**). The concentration of EVs was significantly higher in serum (median $3.28E^{12}$ particles/ml; 95% CI 2.85^{12}-3.68^{12}) than in PJ (median $8.42E^{11}$ particles/ml; 95% CI 7.53^{11}-9.49^{11}, p<0.001). When comparing the concentration of EVs between cases and controls, no difference was found for either biofluid as seen from NTA analysis and total protein concentration in EV isolates (**Figure 1D-E,** in pancreatic juice median 8.71^{11} particles/ml [95% CI 7.67^{11}-9.86^{11}] for controls vs 7.73^{11} [95% CI 5.92^{11}-1.145^{12}] for cases, *P* = 0.41; in serum median 3.28^{12} [95% CI 2.82^{12}-3.96^{12}] for controls vs 3.34^{12} [95% CI 2.71^{12}-3.94^{12}; p=0.84] for cases).

PJ-derived EVs appeared larger than serum EVs in TEM analysis, which was confirmed by NTA analysis (mode diameter of 116 nm [95% CI 114.2-120.2] for PJ and 82 nm [95% CI 80.2-84.1] for serum, *P* < 0.0001). No difference in mode diameter was observed between controls and cases for either PJ (**Figure 1F**) or serum (**Figure 1G**). However, when comparing the concentration of EVs according to their size distribution, a significant difference between cases and controls was seen in PJ, but not in serum. In PJ, particles with the sizes 102.5 nm, 355.5-385.5 nm, 534.5-629.5 nm, 631.5-642.5 nm and 645.5-647.5 nm reached significantly higher concentrations (p<0.05) in cases, as compared to controls (**Figure 1H**). When choosing a threshold of 350 nm, cases had a higher proportion of large EVs (size >350 nm) in PJ as compared to controls (*P* < 0.001, **Figure 1I**).

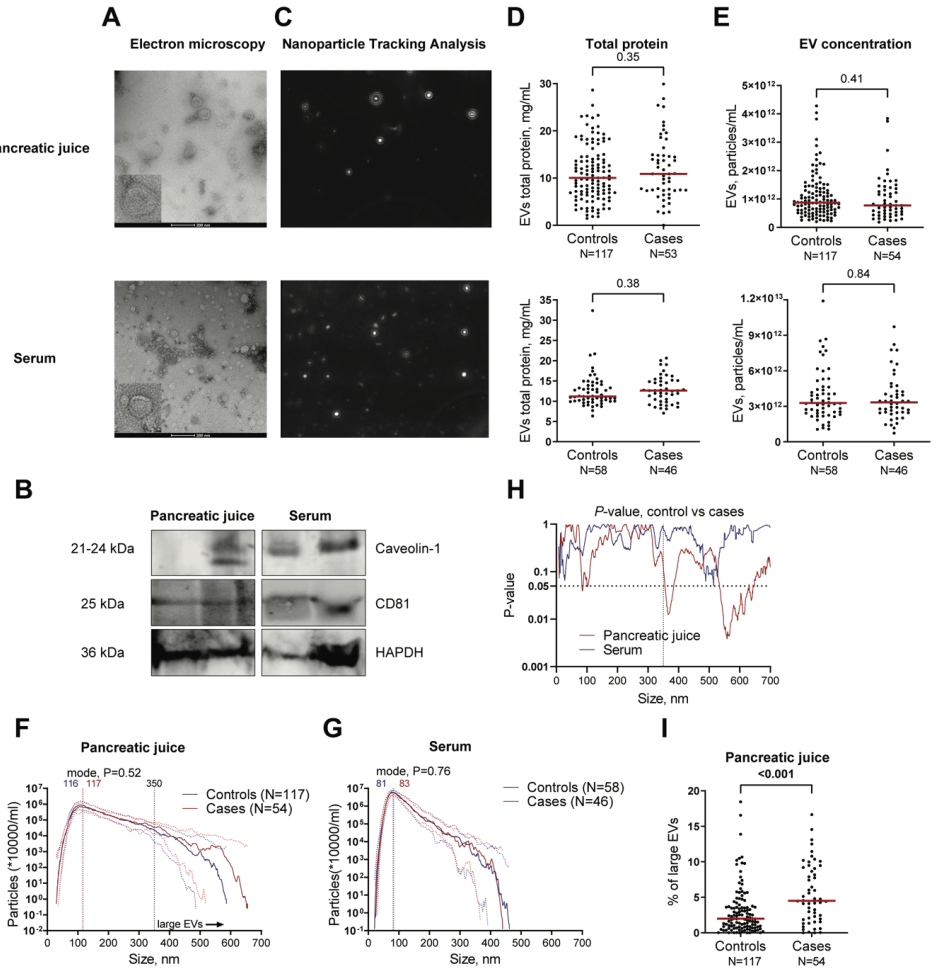


Figure 1. Analysis of EVs in PJ of controls and individuals with PDAC (cases) shows a different size distribution between these groups. A) Representative images of EVs extracted from pancreatic juice (PJ) and serum by transmission electron microscopy (TEM) showing the presence of double membrane vesicles. Notice the larger size of EVs in PJ compared to serum. B) Western blot analysis of typical EV markers commonly found in exosome subpopulations. C) Representative NTA images for pancreatic juice (PJ) and serum. (D-E) Total protein content in EVs isolated from PJ (upper panel) or serum (lower panel) are equal between controls and cases and NTA showed no difference in particle concentrations between controls and cases in PJ (E, median concentration of 8.71^{11} particles/ml [95% CI 7.67^{11}-9.86^{11}] for controls vs 7.73^{11} [95% CI 5.92^{11}-1.145^{12}] for cases, P = 0.41) or serum (E, median 3.28^{12} [95% CI 2.82^{12}-3.96^{12}] and 3.34^{12} [95% CI 2.71^{12}-3.94^{12}; p=0.84], respectively. F, G) Median concentration of EVs of different sizes (from 0 to 750 with stepwise increments of 0.5 nm) in PJ (F) and serum (G). For PJ, vertical lines indicate mode size of 116 and 117 nm for controls and cases respectively (P = 0.52). A threshold line of 350 nm indicating large size particles is indicated, with cases having more EVs with diameter > 350 nm then controls. For serum, vertical lines indicate mode diameter of 81 and 83 nm for controls and cases respectively (P = 0.76). Scattered lines indicate IQR. H) Comparison of the concentration of EVs of different sizes (from 0 to 750 with stepwise increments of 0.5 nm) between controls and cases. P-values for PJ are indicated in red, serum in blue. While the number of EVs in serum is similar between cases and controls across the size ranges, cases present significantly more EVs in the larger range as compared to controls in PJ. Dotted line indicates significance threshold level of p=0.05. I) Percentage of large EVs (>350 nm) in PJ is higher in cases vs controls.

DISCUSSION

This study shows that, as compared to serum, EVs from PJ are larger while their absolute concentration is lower, indicating a distinct proportional composition of vesicular subtypes in PJ. We, and others (Severino et al (2017) [3], did not find differences in EVs concentrations between PDAC cases and controls in serum, where the vast majority of EVs may be of non-tumor origin. In contrast to previous reports for bile [3], where elevated numbers of EVs were seen for cholangiocarcinoma and PDAC patients with biliary stenosis, we did not find differences in concentration of PJ-derived EVs concentrations between cases and controls. Interestingly, bile-derived EVs appeared to be larger in cancer patients compared to patients with chronic pancreatitis, although size distribution was not quantified [3]. In our study, the proportion of large EVs (>350 nm) in PJ of PDAC cases was significantly increased, suggesting a different prevalence of particular subtypes of EVs in these groups. The number of large EVs correlated with pancreas-specific PLA2G1B levels (not shown), implying that these larger-sized EVs are of pancreatic origin, and that EV size may be a promising tool to discriminate PDAC patients from controls.

EVs are classified based on size and their biogenesis: exosomes (<150 nm) are released through multivesicular bodies (MVBs) in the endosomal pathway, microvesicles (200-500 nm) are formed by budding from the plasma membrane, and apoptotic bodies of various sizes derive from programmed cell death. In addition, many other specialized EVs subtypes have been described [6]. Due to a significant overlap in size, similarities in composition and lack of specific markers, it is difficult to assign individual EVs to one of the biogenesis pathways, but the nature of the large EVs found in PJ of PDAC patients represents an interesting research question.

We show that EV extraction from PJ with isolation kits requiring microcentrifuges yields similar concentrations as reported for extraction with ultracentrifugation [2, 7]. As microcentrifuges are commonly available in laboratories, this finding facilitates the application of EVs as a clinical biomarker.

In summary, we characterized vesicular composition of pancreatic juice in PDAC and controls undergoing surveillance and found that PJ from individuals with PDAC harbor increased amounts of large EVs, which may be useful for future biomarker development.

ACKNOWLEDGMENTS

We would like to acknowledge E. de Vries for the help in sample preparation and nanoparticle tracking; I.J. Visser for her help with clinical characteristics dataset; Guido W. Jenster, Marten E. van Royen for the help with electron microscopy.

REFERENCES

1. Costa-Silva, B., et al., *Pancreatic cancer exosomes initiate pre-metastatic niche formation in the liver.* Nat Cell Biol, 2015. 17(6): p. 816-26.
2. Nakamura, S., et al., *Pancreatic Juice Exosomal MicroRNAs as Biomarkers for Detection of Pancreatic Ductal Adenocarcinoma.* Ann Surg Oncol, 2019. 26(7): p. 2104-2111.
3. Severino, V., et al., *Extracellular Vesicles in Bile as Markers of Malignant Biliary Stenoses.* Gastroenterology, 2017. 153(2): p. 495-504.e8.
4. Overbeek, K.A., et al., *Long-term yield of pancreatic cancer surveillance in high-risk individuals.* Gut, 2021.
5. Lowry, O.H., et al., *Protein measurement with the Folin phenol reagent.* J Biol Chem, 1951. 193(1): p. 265-75.
6. Maas, S.L.N., X.O. Breakefield, and A.M. Weaver, *Extracellular Vesicles: Unique Intercellular Delivery Vehicles.* Trends in Cell Biology, 2017. 27(3): p. 172-188.
7. Zheng, J., et al., *Extracellular matrix proteins and carcinoembryonic antigen-related cell adhesion molecules characterize pancreatic duct fluid exosomes in patients with pancreatic cancer.* HPB, 2018. 20(7): p. 597-604.
8. Levink, I.J.M., et al., *Optimization of Pancreatic Juice Collection: A First Step Toward Biomarker Discovery and Early Detection of Pancreatic Cancer.* Am J Gastroenterol, 2020. 115(12): p. 2103-2108.
9. European Study Group on Cystic Tumours of the, P., *European evidence-based guidelines on pancreatic cystic neoplasms.* Gut, 2018. 67(5): p. 789-804.

SUPPLEMENTARY DATA

Supplementary methods

Selection of subjects

This study was executed at the Erasmus University Medical Center and analyzed PJ and serum was collected between August 2018 and May 2020 in patients who participate in the following prospective study cohorts: 1) Patients with suspected (sporadic) PDAC (KRASPanc study, MEC-2018-038); 2) high-risk individuals under surveillance for a hereditary predisposition or familiarly history of PDAC (CAPS study, MEC-2012-448, www.caps-registry.com) [4]; 3) individuals under surveillance for neoplastic pancreatic cysts (PACYFIC study, MEC-2014-021, www.pacyfic.net). The Erasmus Medical Center ethical review board approved the studies, and the included individuals gave written consent before enrolment. The study was carried out according to the ethical principles for medical research involving human subjects from the World Medical Association Declaration of Helsinki.

PJ collection

PJ was collected as described before [8]. In short, during EUS, PJ collection was performed after visualization of the ampullary orifice. To reduce duodenal contamination, duodenal fluid was aspirated prior to juice collection. Next, a wash-out of PJ was stimulated by intravenous administration of human synthetic secretin (ChriRhoStim, Burtonsville, MD, 16µg/patient). PJ was collected for up to 8 minutes starting immediately after injection with the endoscope (Pentax Medical, Tokyo, Japan) and assembled in a mucus extractor (Pennine Healthcare, Derby, United Kingdom, 15 mL) attached to the proximal end of the endoscopic channel. PJ was aliquoted, snap frozen within 10 minutes after collection and stored at -80°C until further use.

EV analysis

PJ was centrifuged for 10 min at 4000 RPM 4°C to remove debris. Then, 100 µL of Total Exosome Isolation Reagent (Thermo Fisher Scientific, Waltham, MA, #4478359) was added to 200 µL of PJ supernatant and kept on a rollerbank overnight at 4°C. After this, samples were centrifuged for 1 h at 14000 RPM and the pellet was resuspended in 400 µL of PBS (pre-filtered with 0.2 µM filter).

Serum was centrifuged during 30 min at 2000 g at 4°C. 40 µL of Total Exosome Isolation (Thermo Fisher Scientific, Waltham, MA, #4478360) were added to 200 µL of serum supernatant and incubated 30 min at 4°C. Then, samples were centrifuged 10 min at 10000 g and pellet was resuspended in 200 µL of filtered PBS. EVs were stored at -80°C until further analysis.

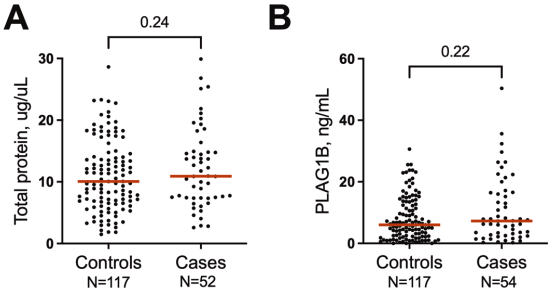
Total protein concentration was determined by Lowry assay (Bio-Rad, Hercules, CA, USA)[5]. For Nanoparticle Tracking Analysis (NTA), samples were diluted 1:1000 in filtered PBS. The size and concentration were detected by NanoSight NS300 (NTA 3.4 Build 3.4.003 software). Two

measurements of each sample were performed. EVs were visualized with Transmission electron microscopy (TEM). For this, 10 µL droplets were deposited on formvar/carbon coated 400 Mesh Cu grids and incubated for 10 min. Thereafter, remaining liquid was drained with filter paper, samples were stained with a drop of Uranyless stain for 1 minute. Remaining liquid was drained, and grids allowed to air dry. Grids were observed under the electron microscope Talos L120C TEM from Thermofisher Scientific at 120 kV. For Western blot, total proteins were extracted in 300 µl Laemmli Buffer [SDS 4%, glycerol 20%, Tris-Cl (pH 6.8) 120 mM, bromophenol blue 0.02% (w/v) and DTT 0.1 M]. Proteins were resolved by SDS-PAGE and blotted onto Immobilon FL PVDF membranes (Millipore, Bedford, MA, USA). Membranes were blocked in Odyssey Blocking Buffer (Thermo Fisher Scientific) and incubated overnight at 4 °C with primary antibody (Caveolin-1 [Cell Signaling Technology #3238], CD81, GAPDH [Santa Cruz Biotechnology #sc-51906]), followed by the appropriate Alexa-linked secondary antibodies, at 1:5000 dilution, in Odyssey Blocking Buffer for 1 h. The fluorescent bands were detected using fluorescent Odyssey Imaging System and densitometric analysis was performed with Image Studio Lite Ver.5.2.

Statistical analysis

Graphpad Prism 9 (GraphPad Software Inc.) and USA IBM SPSS statistic 25 (SAS INSTITUTE INC., Cary, NC, USA) software were used for the generation of graphs and statistical analyses. The Shapiro–Wilk test was used to determine data distribution; the Mann-Whitney U test was performed to compare 2 groups.

Supplementary tables and figures

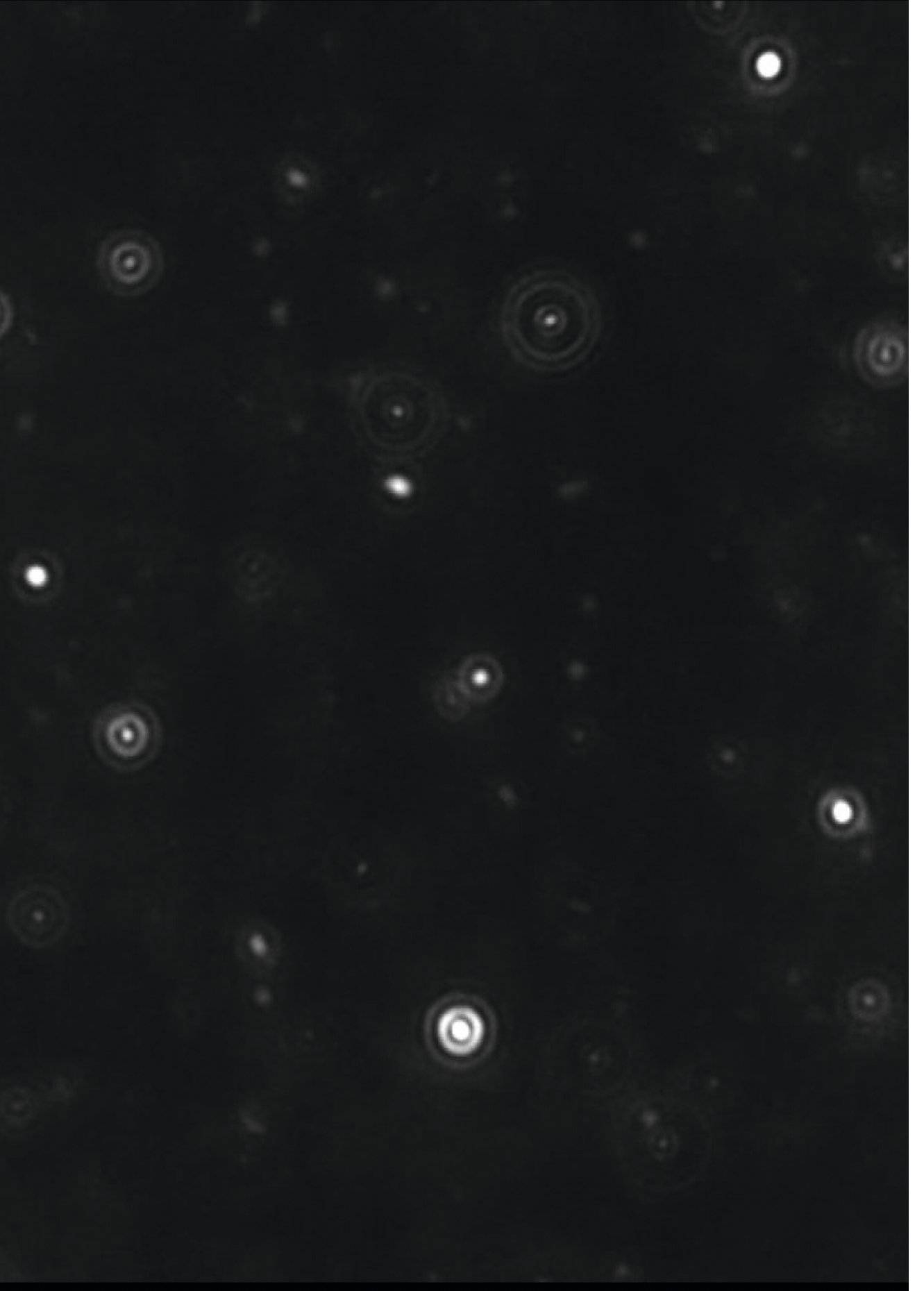


Supplementary figure 1. No difference was seen between controls and cases in pancreatic juice concentration of total protein (A) and phospholipase A2 group IB (PLA2G1B) (B), as a measure of pancreatic origin, indicating similar composition of pancreatic juice.

Supplementary table 1: Clinical characteristics at time of pancreatic juice collection

	Cases (N=54)	Controls (N=117)	*P*-value
Age in years, median (IQR)	67.5 (10.3)	62.0 (6.0)	0.001
Male gender, n (%)	34 (63.0)	40 (34.2)	<0.001
BMI in kg/m^2, median (IQR)	23.7 (3.7)	25.7 (5.0)	0.001
Familial/genetic predisposition, n (%)	0 (0.0)	66 (56.4)	<0.001
Member of FPC family [b]	.	31 (26.5)	.
CDKN2A p16	.	24 (20.5)	.
BRCA2 + 2 blood relatives with PDAC	.	5 (4.3)	.
BRCA1 + 2 blood relatives with PDAC	.	1 (0.9)	.
PALB2 + 2 blood relatives with PDAC	.	1 (0.9)	
BRCA2 + *CDKN2A* p16	.	1 (0.9)	.
STK11/LKB1	.	2 (1.7)	.
Diabetes mellitus, n (%)	21 (38.9)	16 (13.7)	0.001
Indication EUS, n (%)			<0.001
Suspected PDAC	35 (64.8)	4 (3.4)	.
Fiducial placement[c]	18 (33.3)	0 (0.0)	.
Surveillance	1 (1.9)	114 (96.6)	.
CBD stent in situ, n (%)			<0.001
CBD stent in situ	9 (16.7)	0 (0.0)	.
No CBD stent and CBD dilation	14 (25.9)	3 (2.5)	.
No CBD stent and no CBD dilation	31 (57.4)	115 (97.5)	.
Relative or absolute indications for surgery[d,][9], n (%)	54 (100.0)	26 (22.2)	<0.001
Enhancing mural nodule or hypodense lesion	54 (100.0)	4 (3.4)	
Caliber change	41 (75.9)	0 (0.0)	
Diffuse PD dilation > 5mm	0 (0.0)	14 (12.0)	
CA19.9 ≥37 kU/L	34 (63.0)	7 (6.0)	
Cyst size > 40mm	0 (0.0)	2 (1.7)	
New-onset diabetes[e]	9 (16.6)	2 (1.7)	
Recent acute pancreatitis[f,1]	2 (3.7)	6 (5.1)	
Lymphadenopathy	23 (42.6)	0 (0.0)	
Working diagnosis, n (%)			<0.001
No abnormalities	.	41 (35.0)	.
Unspecified cyst	.	9 (7.7)	.
SB-IPMN	1 (1.9)	50 (41.9)	.
MD-IPMN or MT-IPMN	.	14 (12.0)	.
MCN	.	1 (0.9)	
NET	.	1 (0.9)	.
Indeterminate lesion, not suspect for malignancy	.	2 (1.7)	.
Resectable PDAC	10 (18.5)	0 (0.0)	.
Locally advanced PDAC	43 (79.6)	0 (0.0)	.
Distal metastases (on imaging), n (%)	8 (14.8)	0 (0.0)	<0.001

[a]no germline mutation known and not fitting criteria of FPC family
[b] 2 first-degree relatives or 3 relatives (either first or second degree) or 2 second-degree relatives of which 1 with age <50 years at time of diagnosis.
[c] Received previous chemotherapy
[d]One can have developed multiple worrisome features
[e]Development of diabetes mellitus in two years before biomaterial collection
[f]Acute pancreatitis in 2 years before biomaterial collection (not related to performed ERCP).
[1]3 extra post-ERCP pancreatitis



Chapter 4

Extracellular vesicles derived microRNAs in pancreatic juice as biomarkers for detection of pancreatic ductal adenocarcinoma

K. Nesteruk*, I.J.M. Levink*, E. de Vries, I. J. Visser, M. P. Peppelenbosch, D. L. Cahen, G. M. Fuhler, M. J. Bruno

(in preparation)

INTRODUCTION

Pancreatic ductal adenocarcinoma (PDAC) has an overall 5-year survival rate of less than 10% [1]. While the 5-year survival rate improves to more than 30% for individuals diagnosed at an operable stage [2], the majority of patients are diagnosed with inoperable advanced disease due to the late presentation of symptoms [1, 3-5]. Additionally, in surveillance programs for individuals with an increased risk of developing PDAC, present imaging techniques (endoscopic ultrasound [EUS] and magnetic resonance imaging [MRI]) struggle to detect sub-centimeter cancer mass, even when both modalities are performed concurrently [6]. Conversely, the false-positivity rate of worrisome features is high [7-9], resulting in resection of lesions that on histological examination appear to be benign [10, 11]. Thus, a novel tool that diagnoses PDAC with a high specificity at an early stage is urgently needed.

One approach is to support clinical diagnosis with molecular markers. Currently, the main biomarker for (recurrence of) PDAC is serum levels of carbohydrate antigen 19-9 (CA19-9), which has a sensitivity of 70-90% [12-16]. However, this marker is not likely to be positive until PDAC reaches an advanced stage [17]. Furthermore, the false-positivity rate of CA19-9 is relatively high at a specificity of 70-90% [12-16], and its use in a surveillance setting is disputed as high levels (above the clinically used cut-off of 37kU/L) can be detected in patients with no or low-grade dysplasia [18]. Also, pancreatic tumors are highly heterogeneous, both within a tumor and between individuals [19]. Consequently, it is conceivable that the diagnostic performance of a single biomarker will not be sufficient, but that a robust panel of biomarkers is required to accurately detect PDAC at an early stage.

Currently, evidence shows that serum levels of cell-free miRNA [20-22] and miRNA isolated from extracellular vesicles present in serum (EV-miRNA) [23-25], can differentiate between patients with PDAC and healthy controls. Expression of several of these miRNAs are altered upon proliferation, angiogenesis, epithelial to mesenchymal transition, and metastasis of several human malignancies including PDAC [25-35]. However, none of these serum miRNAs have made it to clinical practice yet and it is possible that other biomaterial sources may be more relevant for biomarker detection. Pancreatic juice (PJ), which can be safely collected from the duodenal lumen during EUS, conceivably contains markers that are more pancreas-specific as compared to blood. Being in direct contact with the potential tumor cells, PJ is also expected to contain information from all tumor clones present. On the other hand, PJ harbors high concentrations of digestive enzymes and represents an abrasive environment which may result in the degradation of promising PDAC biomarkers. Interestingly, a recent study showed that compared to total cell-free miRNAs in PJ, the diagnostic performance improved when miRNA was isolated from EVs present in PJ [36], suggesting that biomarkers in PJ may be protected in EVs. EVs are a group of cellular particles which can be classified based on size and biogenesis. They include: 1. exosomes (<150 nm), which are released from cells through multivesicular bodies in the endosomal pathway; 2. microvesicles (200-500 nm), which arise through budding of the plasma membrane; and 3. apoptxfotic bodies (various sizes). On top of these major classes, many specialized EVs subtypes have been described [37]. Cancer cells, specifically, release EVs to create a pre-metastatic niche [38]. Their vesicular contents, which include proteins, DNA, RNA, and microRNAs, are cell-specific and expected to represent a signature of cellular pathology. A previous study showed that

the levels of EV-derived miR-21 and miR-155 in PJ discriminated PDAC cases from controls with an accuracy of 83% and 89% respectively [36]. However, limitations included the low number of subjects involved, and EV-miR expression in PJ and blood has never been directly compared to our best knowledge. Thus, our study aimed to evaluate the diagnostic performance of PJ and serum-derived EV-miRNA for the detection of PDAC in a larger cohort of patients and controls. Based on the most often described promising microRNAs for PDAC detection, we selected EV-miR-16, EV-miR-21, EV-miR-25, EV-miR-155, and EV-miR-210 for analysis [39]. We compared the expression of these EV-miRNAs and their diagnostic performance between PJ and serum and contrasted this against the currently available serum biomarker CA19.9.

MATERIAL AND METHODS

Selection of subjects

This study was executed at the Erasmus University Medical Center. PJ and serum were collected between August 2018 and May 2020 in patients who participate in the following prospective study cohorts: 1) Patients with suspected (sporadic) PDAC (KRASPanc study, MEC-2018-038); 2) high-risk individuals under surveillance for a hereditary predisposition or familiarly history of PDAC (CAPS study, MEC-2012-448, www.caps-registry.com); 3) individuals under surveillance for neoplastic pancreatic cysts (PACYFIC study, MEC-2014-021, www.pacyfic.net). The Erasmus Medical Center ethical review board approved the studies, and the included individuals gave written informed consent before enrolment. The study was carried out according to the ethical principles for medical research involving human subjects from the World Medical Association Declaration of Helsinki.

All inclusion criteria of the prospective cohort studies can be found in **Supplemental Table 1**. Samples were excluded if diagnostic work-up eventually revealed other pancreatic diseases (e.g., pancreatitis). For endpoint analysis, this cohort was divided into a case group (patients with pathologically proven high-grade dysplasia [HGD] or PDAC) and controls (individuals without HGD or PDAC). For subgroup analysis, the control group was subdivided into individuals with high-risk morphology who presented with worrisome features or indications for surgery (as described by the EU evidence-based guidelines [9]) and those with low-risk morphology and no indications for surgery [9].

Pancreatic Juice and Serum Collection

PJ was collected from the duodenum as described before [40]. During EUS, PJ collection was performed after visualization of the ampullary orifice. To reduce duodenal contamination, duodenal fluid was aspirated prior to juice collection. Next, a wash-out of PJ was stimulated by intravenous administration of human synthetic secretin (ChriRhoStim, Burtonsville, MD, 16 µg/patient). PJ was collected for up to 8 minutes starting immediately after injection with the endoscope (Pentax Medical, Tokyo, Japan) and assembled in a mucus extractor (Pennine Healthcare, Derby, United Kingdom, 15 mL) attached to the proximal end of the endoscopic channel. PJ was aliquoted to avoid freeze-thaw

cycles, snap-frozen within 10 minutes after collection, and stored at -80°C until further use. All serum samples were collected within a 3-week window of PJ collection. CA19.9 data were retrieved from patient records and only measurements performed within 3 weeks preceding or following PJ collection were scored.

Pancreatic juice quality

As a measure of PJ quality, we assessed the concentration of PLA2G1B (pancreas standard marker, MyBiosource, San Diego, USA, MBS703283) by enzyme linked immunosorbent assay (ELISA). Pre-coated, pre-blocked plates were used according to the manufacturer's protocol. Briefly, plates were blocked with ELISA diluent for 1 hour at room temperature. PJ (75 µL per well) was incubated at 4 °C overnight, after which biotin conjugated detection antibody was added at room temperature for 1 hour. Following incubation of avidin-HRP for 30 minutes at room temperature, TMB substrate was added. Reactions were stopped by addition of sulfuric and absorbance was read at 450 nm (Tecan Infinite200 pro plate reader). Cases and controls were equally distributed among batches. In addition, we measured the total protein concentration in PJ and EV isolates by Lowry protein assay (Bio-Rad, Hercules, CA, USA) [41].

EV isolation and analysis

Prior to EV isolation, PJ was centrifuged for 10 min at 4000 RPM 4°C to remove debris. Then, 100 µL of total exosome isolation reagent (Thermo Fisher Scientific, Waltham, MA, #4478359) was added to 200 µL of supernatant and kept on a rollerbank at 4°C overnight. After this, samples were centrifuged for 1h at 14000 RPM and the pellet was resuspended in 400 µL of PBS (filtered with 0.2µM filter). Serum was centrifuged during 30 minutes at 2000g at 4°C to remove debris. 40 µL of Total exosome isolation reagent (for serum; Thermo Fisher Scientific, Waltham, MA, #4478360) was added to 200 µL of supernatant and incubated for 30 minutes at 4°C. Then, samples were centrifuged again at 10 000 g for 10 min at room temperature. Supernatant was discarded and EVs were resuspended in 200 µL of filtered PBS. Concentrated EVs were stored at -80°C until further analysis. The size and concentration of particles were confirmed by NanoSight NS300 (NTA 3.4 Build 3.4.003 software). For nanoparticle tracking analysis samples were diluted 1:1000 in filtered PBS. Quality of EVs was confirmed by electron microscopy and western blot (data not shown).

EVs miRNA analysis

MicroRNA was isolated from 200 µL of concentrated EVs with QIAzol Lysis Reagent (Qiagen, Hilden, Germany, #79306) and miRNeasy Mini kit (Qiagen, Hilden, Germany, #217004) according to manufacturer's recommendations. MicroRNA-specific cDNA was prepared using the Taqman microRNA Reverse Transcription Kit (Thermo Fisher Scientific, Waltham, MA, #217004; miR-16, miR-

21, miR-205, miR-155) as described before [42, 43]. Every cDNA reaction consisted of 0.4 µl dNTP mix, 1.35 µl Multiscribe RT enzyme (500U/µL), 2.0 µl 10x RT Buffer, 0.25 µl RNase inhibitor, 1.0 µl of each RT primer, and 5 µl of diluted template RNA. The total reaction volume was adjusted to 20 µl with nuclease-free water. Each qPCR reaction consisted of 6 µl TaqMan Universal PCR Master Mix (Thermo Fisher Scientific, Waltham, MA, #4324018), 0.5 µl microRNA-specific PCR primer and 5.0 µl of the previously 1:5 diluted cDNA. The final volume of every PCR reaction was adjusted to 12 µl with nuclease-free water. All qPCR reactions were performed according to the manufacturer's instructions and carried out in duplicate. MiRNA expression changes were calculated relative to plate average using the $2^{-\Delta\Delta Ct}$ method [44] and presented as log2 fold change. When expression miRNA was not detected (only for miR-25 in 6.9% of samples), CT input value of 45 was imputed for quantification.

Statistical analysis

Shapiro–Wilk's test was used to determine data distribution. For normally distributed data, an unpaired Student's t-test was performed to compare two groups. For non-parametric data, a Mann-Whitney U (2 groups) test or a Kruskal-Wallis H test (>2 groups) was performed. Spearman's rank correlation coefficient was used to investigate correlations between biomarkers and continuous variables. A χ^2 test or Fisher's exact probability test was used to evaluate the association between categorical variables. Receiver operating characteristic (ROC) curves and their area under the curve (AUC) were used to assess the diagnostic performance of the biomarkers. The first optimal cut-off values in ROC curves were set to the value that maximizes the Youden index. The Youden index was defined as sensitivity + specificity − 1. For each biomarker, a second cut-off point on the ROC curve was chosen with a specificity of at least 90%, aiming for high specificity to minimize harm due to unnecessary biopsy or surgery in the surveillance population. For Ca19.9, the clinically used cut-off of 37kU/L was used. Multiple logistic regression models were created to test the performance and interaction between a combination of biomarkers. Confidence intervals for sensitivity, specificity, and accuracy are "exact" Clopper-Pearson confidence intervals. All tests were two-sided, and the significance level was set at $p < 0.05$. Excel, Graphpad Prism 5 (GraphPad Software Inc.), IBM SPSS Statistics (for Windows, Version 25.0. Armonk, NY: IBM Corp.) software were used for the analyses**.**

RESULTS

Patient's characteristics

In total, 54 cases and 118 controls were recruited for PJ collection. A summary of subject characteristics is provided in **Table 1.** Controls tended to be younger (62.1 vs 67.5, p=0.001) with a lower proportion of males (63% vs 34%, p=0.001) and higher BMI (25.7 kg/m^2 vs 23.7 kg/m^2, p=0.001). Cases more often suffered from diabetes mellitus (38.9% vs 13.4%, p=0.001).

Serum samples were available for 46 cases and 58 controls. Characteristics of this subpopulation are summarized in **Table 2.** Again, controls were younger (60 vs 68, p=0.001), with a lower proportion of

males (25.9% *vs* 63.0%, p<0.001). Fewer controls had diabetes (15.5% vs 41.3%, p=0.004). Cases had a lower BMI (23.2 kg/m^2) than controls (25.7 kg/m^2, p=0.003).

Table 1: Clinical characteristics at time of pancreatic juice collection

	Cases (N=54)	Controls (N=118)	*P*-value
Age in years, median (IQR)	67.5 (10.3)	62.1 (6.0)	0.001
Male gender, n (%)	34 (63.0)	41 (34.7)	0.001
BMI in kg/m^2, median (IQR)	23.7 (3.7)	25.7 (5.1)	0.001
Familial/genetic predisposition, n (%)	0 (0.0)	66 (55.9)	<0.001
Member of FPC family	.	32 (27.1)	.
CDKN2A p16	.	24 (20.3)	.
BRCA2 + 2 blood relatives with PDAC	.	5 (4.2)	.
BRCA1 + 2 blood relatives with PDAC	.	1 (0.8)	.
PALB2 + 2 blood relatives with PDAC	.	1 (0.8)	
BRCA2 + *CDKN2A* p16	.	1 (0.8)	.
STK11/LKB1	.	2 (1.7)	.
Diabetes mellitus, n (%)	21 (38.9)	16 (13.6)	0.001
Indication EUS, n (%)			<0.001
Suspected PDAC	35 (64.8)	4 (3.4)	.
Fiducial placement	18 (33.3)	0 (0.0)	.
Surveillance	1 (1.9)	114 (96.6)	.
CBD stent in situ, n (%)			<0.001
CBD stent in situ	9 (16.7)	0 (0.0)	.
No CBD stent and CBD dilation	14 (25.9)	3 (2.5)	.
No CBD stent and no CBD dilation	31 (57.4)	115 (97.5)	.
Relative or absolute indications for surgery, [1] n (%)	54 (100.0)	26 (22.0)	<0.001
Enhancing mural nodule or hypodense lesion	54 (100.0)	4 (3.4)	
Caliber change	41 (75.9)	0 (0.0)	
Diffuse PD dilation > 5mm	0 (0.0)	14 (11.9)	
CA19.9 ≥37 kU/L	34 (63.0)	7 (5.9)	
Cyst size > 40mm	0 (0.0)	2 (1.7)	
New-onset diabetes[1]	9 (16.6)	2 (1.7)	
Recent acute pancreatitis[2]	2 (3.7)[c]	6 (5.1)	
Lymphadenopathy	23 (42.6)	0 (0.0)	
Working diagnosis, n (%)			<0.001
No abnormalities	.	41 (34.7)	.
Unspecified cyst	.	9 (7.6)	.
SB-IPMN	1 (1.9)	50 (42.4)	.
MD-IPMN or MT-IPMN	.	14 (11.9)	.
MCN	.	1 (0.8)	
NET	.	1 (0.8)	.
Indeterminate lesion, not suspect for malignancy	.	2 (1.7)	.
Resectable PDAC	10 (18.5)	0 (0.0)	.
Locally advanced PDAC	43 (79.6)	0 (0.0)	.
Distal metastases (on imaging), n (%)	8 (14.8)	0 (0.0)	<0.001

[a]no germline mutation known and not fitting criteria of FPC family
[b] 2 first-degree relatives or 3 relatives (either first or second degree) or 2 second-degree relatives of which 1 with age <50 years at time of diagnosis.
[C] 3 extra post-ERCP pancreatitis

Table 2: Clinical characteristics (serum)

	Cases (N=46)	**Controls (N=58)**	***P*-value**
Age in years, median (IQR)	68 (10.5)	60 (7.3)	0.001
Male gender, n (%)	29 (63.0)	15 (25.9)	<0.001
BMI in kg/m^2, median (IQR)	23.2 (3.2)	25.7 (5.2)	0.003
Familial/genetic predisposition, n (%)	0 (0.0)	29 (50.0)	<0.001
Member of FPC family	.	13 (22.4)	
CDKN2A p16	.	10 (17.2)	
BRCA2 + 2 blood relatives with PDAC	.	3 (5.2)	
BRCA1 + 2 blood relatives with PDAC	.	0 (0.0)	
PALB2 + 2 blood relatives with PDAC	.	1 (1.7)	
BRCA2 + *CDKN2A* p16	.	1 (1.7)	
STK11	.	1 (1.7)	
Diabetes mellitus, n (%)	19 (41.3)	9 (15.5)	0.004
Indication EUS, n (%)			<0.001
Suspected PDAC	28 (60.9)	3 (5.2)	
Fiducial placement	18 (39.1)	0 (0.0)	
Surveillance	0 (0.0)	55 (94.8)	
CBD stent in situ, n (%)			<0.001
CBD stent in situ	8 (17.4)	0 (0.0)	
No CBD stent and CBD dilation	10 (21.7)	0 (0.0)	
No CBD stent and no CBD dilation	28 (60.9)	58 (100.0)	
Relative or absolute indications for surgery [1], n (%)	46 (100.0)	11 (19.0)	<0.001
Enhancing mural nodule or hypodense lesion	46 (100.0)	0 (0.0)	
Caliber change	35 (76.1)	0 (0.0)	
Diffuse PD dilation > 5mm	0 (0.0)	7 (12.1)	
CA19.9 ≥37 kU/L	27 (58.7)	4 (6.9)	
Cyst size > 40mm	0 (0.0)	2 (3.4)	
New-onset diabetes[1]	7 (15.2)	1 (1.7)	
Recent acute pancreatitis[2]	2 (4.3)	4 (6.9)	
Lymphadenopathy	21 (45.7)	0 (0.0)	
Working diagnosis, n (%)			<0.001
No abnormalities	.	18 (31.0)	
Unspecified cyst	.	10 (17.2)	
SB-IPMN	.	23 (39.7)	
MD-IPMN		2 (3.4)	
MT-IPMN	.	5 (8.6)	
MCN	.	0 (0.0)	
NET	.	0 (0.0)	
Indeterminate, not suspect for malignancy	.	0 (0.0	
Resectable PDAC	11 (23.9)	.	
Locally advanced PDAC	35 (76.1)	.	
Distal metastases (on imaging), n (%)	6 (13.0)	0 (0.0)	<0.001
Time in days between serum and PJ sample collection, median (IQR)	0 (0.0)	0 (9.3)	0.003

[c] Received previous chemotherapy
[d]One can have developeTad multiple worrisome features
[e]Development of diabetes mellitus in two years before biomaterial collection
[f]Acute pancreatitis in 2 years before biomaterial collection (not related to performed ERCP).

Quality of pancreatic juice

As a measure of PJ quality, pancreas-specific PLA2G1B and total protein concentration were determined. Both were similar between controls and cases, indicating a similar pancreatic component in PJ as a result of collection (P = 0.24 for total protein and P = 0.19 for PLA2G1B; **Figure 1**).

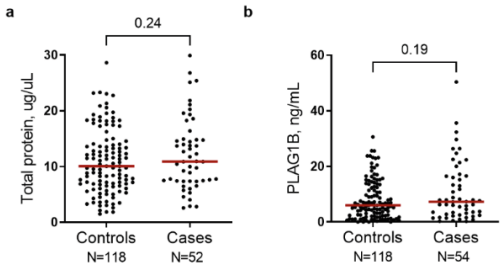


Figure 1. No difference was found between controls and cases in concentration of total protein (A) and phospholipase A2 group IB (PLA2G1B; as a measure of pancreatic origin) (B), indicating similar composition of PJ.

EV-miRNA expression

Nanoparticle tracking analysis revealed abundant small vesicles in PJ with a mode size of 116 nm and an overall concentration of 8.42^{11} particles/mL (range 1.85^{11} to 4.28^{12}, data not shown).

EV-miR-16, EV-miR-21, EV-miR-155, and EV-miR-210 were detectable in all subjects for both PJ and serum, while EV-miR-25 was detectable in 161/172 (93.6%) PJ samples and all serum samples. When comparing cases with controls, EV-miR-21 ($P = 0.002$), EV-miR-25 ($P = 0.005$), EV-miR-210 ($P = 0.02$), and EV-miR-16 ($P = 0.004$) were significantly overexpressed in PJ, while no difference was found for EV-miR-155 ($P = 0.11$) (**Figure 2 a-e**). In serum, as expected, CA19.9 level was higher in cases than controls ($P < 0.001$), yet only EV-miR-210 was significantly increased (**Figure 2 f-k**). Expression of EV-miRNAs was highly correlated with each other in both PJ and serum but no correlation was found between PJ and serum (**Supplementary tables 2 and 3**).

After correction for age, gender, BMI and diabetes mellitus, EV-miR-21, EV-miR-25, and EV-miR-16 remained significantly overexpressed in cases vs controls in PJ (**Table 3,** *Ors (95% CIs) and p-value determined by multivariate logistic regression analysis, adjusted for age, gender, BMI and diabetes) and EV-miR-210 in serum, thus showing an independent association with PDAC.

In subgroup analysis, expression of EV-miR-21 ($P = 0.07$) and EV-miR-210 ($P = 0.04$) in PJ tend to be higher in high-risk controls, who harbor worrisome features, as compared to low-risk controls, while no difference was found between high-risk controls and PDAC (**Figure 3**). CA19.9 was increased in PDAC compared to both control groups (**Figure 3**).

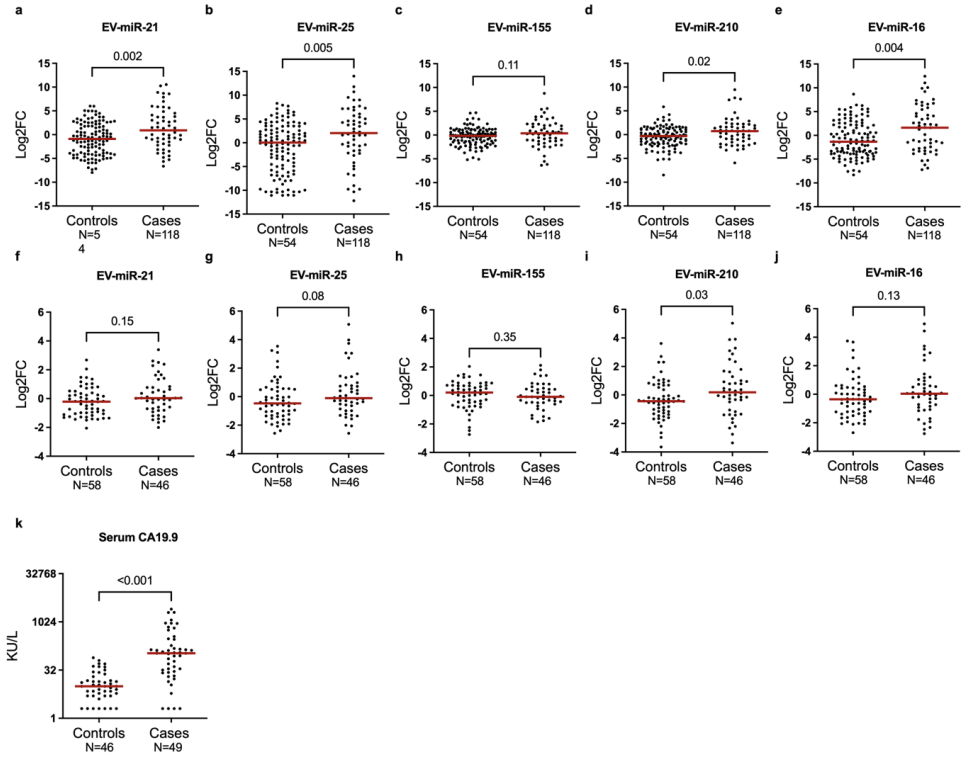


Figure 2. Relative expression of miR-21, miR-155, miR-210, miR-25 and miR-16 in pancreatic juice (a-e) and serum (f-j) of patients with pancreatic cancer (cases) and controls. k) Level of serum CA19.9 is increased in individuals with pancreatic cancer (cases).

Table 3. Binary logistic regression Crude OR and adjusted OR for controls and cases.

Parameter	Crude OR	*P*-value	Adjusted OR * (age, gender, BMI, diabetes)	*P*-value
Pancreatic juice				
Age	1.070	0.001	1.064	0.028
Gender	3.193	0.001	3.523	0.009
BMI	0.827	0.002	0.778	<0001
Diabetes mellitus	0.246	<0001	0.332	0.52
PJ_miR-21	1.173	0.001	1.152	0.03
PJ_miR-25	1.095	0.005	1.117	0.016
PJ_miR-210	1.224	0.006	1.183	0.094
PJ_miR-16	1.141	0.001	1.146	0.017
Serum				
Age	1.080	0.01	1.070	0.042
Gender	4.890	<0001	7.071	0.003
BMI	0.814	0.0083	0.757	0.006
Diabetes mellitus	0.261	0.0043	0.464	0.323
Serum_miR-210	1.317	0.0359	1.571	0.033

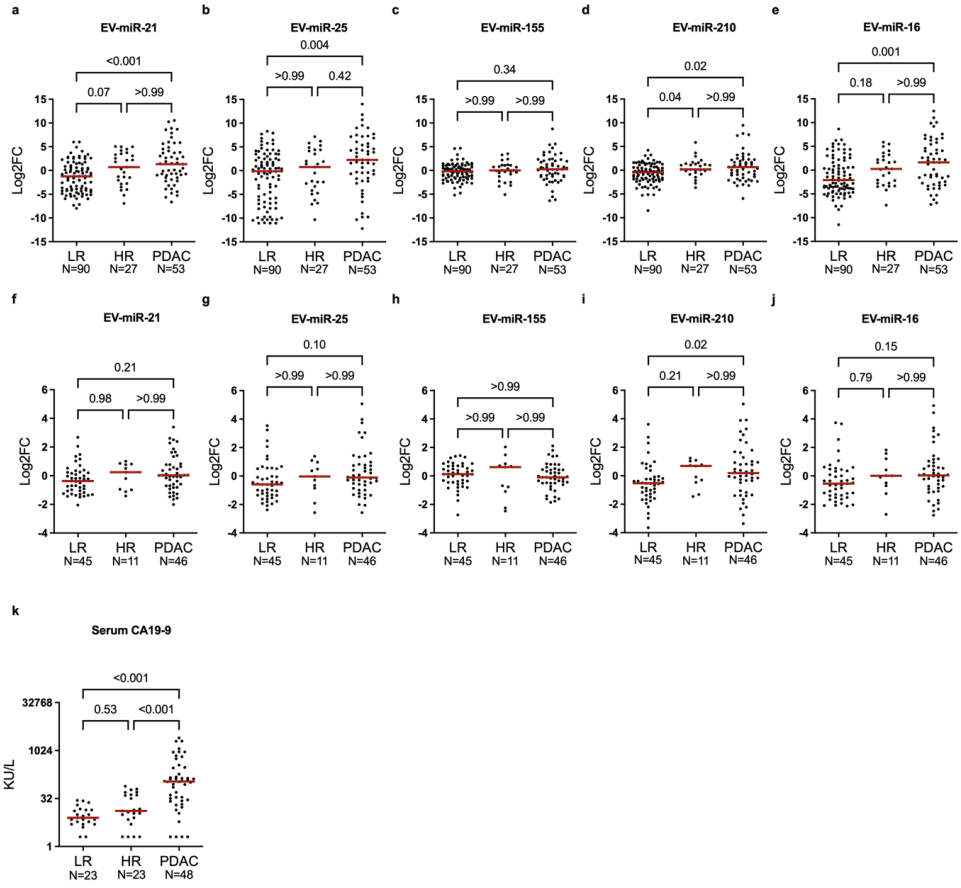


Figure 3. Relative expression of miR-21, miR-155, miR-210, miR-25 and miR-16 in pancreatic juice (a-e) and serum (f-j) of individuals with pancreatic cancer (PDAC), low-risk controls (LR) and high-risk controls (HR). k) Level of serum CA19.9 is increased in individuals with PDAC.

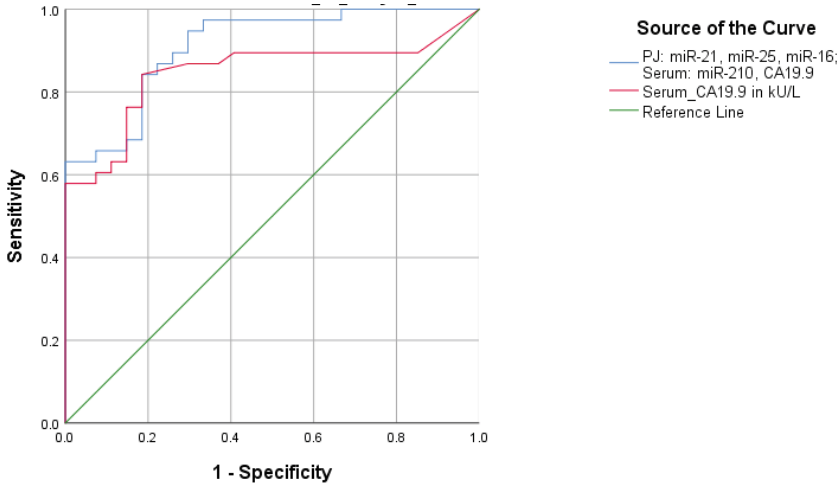
Diagnostic performance of analyzed miRNAs

ROC curves were constructed for independently associated EV-miRNAs in PJ and serum to compare their diagnostic value for PDAC detection. AUC values for individual overexpressed EV-miRNAs in PJ ranged from 0.61 to 0.64. For serum, EV-miR-210 achieved an AUC of 0.62 for PDAC prediction, while for CA19.9 an AUC of 0.85 was reached (**Table 4**). While performance of individual EV-miRNAs was poor, when combining PJ EV-miR-21, EV-miR-25, and EV-miR-16 with serum EV-miR-210 with CA19.9, the sensitivity and specificity were 84.2% and 81.5% (AUC = 0.91) respectively, compared to 85.7% and 73.3% for CA19.9 alone (**Figure 4**).

Table 4. Diagnostic performance of analyzed miRNAs

			Cut-off 1 (SP= 100%)		Cut-off 2 (the highest Youden Index)	
Individual performance	**N controls/ cases**	**ROC-AUC (95% CI)**	**SE, % (95% CI)**	**SP, % (95% CI)**	**SE, % (95% CI)**	**SP,% (95% CI)**
PJ_miR-21	118/54	0.64 (0.55-0.73)				
PJ_miR-25	118/54	0.63 (0.54-0.73)				
PJ_miR-16	118/54	0.64 (0.55-0.73)				
Serum_miR-210	58/49	0.62 (0.51-0.74)				
Serum_CA19.9	45/49	0.85 (0.77-0.93)	59.2 (44.2-73.0)	100.0 (92.1-100)	85.7 (72.8-94.1)	73.3 (58.1-85.4)
Logistic regression models						
PJ_miR-21+miR-25+miR-16+Serum_CA19.9	45/49	0.89 (0.82-0.95)	63.3 (48.3-76.6)	100.0 (92.1-100)	75.5 (61.1-86.7)	86.7 (73.2-95.0)
PJ_miR-21+miR-25+miR-16+Serum_CA19.9+Serum_miR-210	27/38	0.91 (0.84-0.98)	63.2 (46.0-78.2)	100 (87.2-100)	84.2 (68.6-94.0)	81.5 (61.9-93.7)

Figure 4. Diagnostic performance of CA19.9 level and logistic model (PJ: miR-21, miR-25, miR-16 and serum: miR-210 and CA19.9)



DISCUSSION

In the present study, we extracted EVs from serum and PJ and investigated EV-miRNA expression of 54 malignant cases and 118 non-malignant controls. Five EV-miRNAs were selected based on their promise inferred from literature (see **Supplementary Table 4**): EV-miR-21, EV-miR-25, EV-miR-210, EV-miR-155, EV-miR-16. Of these, EV-miR-21, EV-miR-25, EV-miR-16 were overexpressed in PJ from cases compared to controls, independently from other clinical characteristics. EV-miR-210 was the only miRNA to be overexpressed in serum from cases compared to controls. A combined panel of PJ EV-miR-21, EV-miR-25, EV-miR-16, and serum EV-miR-210 and CA19-9 is able to distinguish cases from controls undergoing surveillance with a specificity of 81.5% and sensitivity of 84.2%.

MiRNAs are short non-coding RNAs composed of 18–25 nucleotides which are functional regulators of gene expression. In PDAC cells, miRNA-21 regulates gene expression of *MMP2*, *MMP9*, and *VEGF* to enhance cellular proliferation and invasion [45]. MiR-21 expression in PDAC tumor-associated fibroblasts is also linked with decreased overall survival [46]. MiR-25 expression promotes cell proliferation by targeting the regulator of actin cytoskeleton dynamics *ABI2* in PDAC [47]. Zhang et al. have found that cigarette smoke-induced miR-25-3p excessive maturation promotes the development and progression of pancreatic cancer [48]. MiR-210 is a hypoxia marker in PDAC [49] and has been shown to mediate epithelial-mesenchymal transition induced by HIF-1α under hypoxic conditions by inhibition of HOXA9 in PDAC cell line [50]. Luciferase reporter assays suggested that miR-16 post-transcriptionally regulates *Bcl-2* expression in PDAC cells by targeting sites of the 3′ untranslated region of this gene [51]. Thus, the overexpressed miRNAs are clearly involved in pathogenesis and progression of PDAC, explaining why these molecules have been targeted for investigation as potential biomarkers.

Cell-free miRNA can either be isolated from whole biofluid, or from EVs obtained from these biobluids. To our best knowledge, only one study investigated EV-miRNA in PJ before [36]. Nakamure et al. demonstrated overexpression of EV-miR-21 and EV-miR-155 in PDAC cases (N=27) compared to chronic pancreatitis controls (N=8) [36]. While we obtained similar results for EV-miR-21, no difference in expression was found for EV-miR-155, likely due to different expressions in control groups (cystic lesions and pancreas without abnormalities in our study). Technical differences may also account for different study results, including 1) the method of PJ collection (endoscopic retrograde pancreatography vs less invasive EUS in our study), 2) the method of EV extraction (ultracentrifugation vs more clinically applicable method with Invitrogen kit in our study), and 3) normalization method for RNA expression (miR-16 vs PJ volume in our study). Other studies mostly concentrated on total miRNA [20, 52-54] or EV-miRNA [23, 55-57] in blood, but some also investigated total [58], EV-miRNA [36] or cellular miRNA [59] in PJ (**Supplementary table 4**). The expression profile of miRNAs in EVs is not the same as the corresponding cell-free total miRNAs, indicating that miRNAs have a strict sorting mechanism [60]. When side by side comparison of EV-miRNA with miRNA from the whole biofluid was performed, EV-miRNA was superior as a biomarker of PDAC in both serum [55] and PJ [36], and as such EV-miRNAs may be more useful and stable as biomarkers for PDAC detection compared to total miRNA in body fluids, including PJ. However, EV-miRNA in PJ is clearly under-investigated in this respect.

Here we found that EV-miRNA expression patterns do not correlate between serum and PJ. Thus, PJ may contain biomarkers that are not present in serum and *vice versa*. For instance, EV-miR-210 was overexpressed in serum of cases, but not in PJ after adjusting for clinical parameters. Aberrant high expression of miR-210 has been detected in many tumors [61] making it in theory a less specific PDAC biomarker when detected in blood. EVs in PJ are larger in size when compared to serum-derived EVs (unpublished data) and as size is linked to biogenesis this might also explain the different molecular cargo of these EVs. Thus, differences may be related to different subtypes of EVs present in these bodily fluids or the fact that most EVs in serum are not pancreas derived. Selection of candidate miRNAs for our study were based on available data which predominantly exists for blood. It has been estimated that each exosome can accommodate 70–25,000 small RNA or protein molecules [62]. Taking into account our findings on different expression of EV-miRNAs in blood and PJ, a more systematic approach is needed to determine miRNA composition of EVs in PJ and identify good candidate miRNAs for PDAC diagnosis in PJ.

The ideal biomarker (or panel) should be able to discriminate PDAC patients not only from heathy controls but more importantly also from other non-malignant pancreatic and other organs conditions. We previously found that a minority (38%) of individuals with a solid lesion on imaging and a minority (35%) of individuals undergoing surgery for a suspicious lesion had malignancy or high-grade dysplasia in the resected specimen [6]. To prevent unnecessary surgeries, it is important to test a potential biomarker on a relevant control group with non-malignant pancreatic masses. For this reason, we included individuals under surveillance for familiar or genetic predisposition to pancreatic cancer and for neoplastic pancreatic cysts. 59% of controls had non-malignant pancreatic abnormalities. This is also agreement with the current consensus [63, 64] that screening for pancreatic cancer should not be aimed at the general population but at individuals at increased risk of developing pancreatic cancer. In subgroup analysis, splitting low-risk controls and high-risk controls (based on the presence of worrisome features and indications for surgery [9]), a gradual increase in the expression of EV-miR-21, EV-miR-25 and EV-miR-16 was seen from low-risk controls to high-risk controls and PDAC in PJ, but not serum. Although high-risk controls were not statistically different from either low-risk controls or PDAC, low-risk controls were significantly different from PDAC. Our data in line with studies on total miRNA in PJ, demonstrating a decrease in expression levels from PDAC to chronic pancreatitis to non-pancreatic disease controls, with differences between PDAC and chronic pancreatitis patients not being significant [58]. Including only healthy individuals in a control group may lead to the overestimation of diagnostic performance of candidate biomarker. Indeed, it has been reported that the diagnostic performance of serum Ca19.9 for discriminating PDAC is higher when including healthy controls compared to benign pancreatic disease cases [65]. A drawback of our approach is that high-risk controls may theoretically harbor as yet undetected cancer, thus resulting in an underestimated performance of candidate biomarkers. However, the probability of this occurring is low, with a follow-up of a minimum of one year for all controls in our cohort. Most published studies (**Supplementary table 4**) include heathy controls and chronic pancreatitis patients, with some adding other cancers or non-healthy controls without pancreatic disease. High diagnostic values have been reported for heathy subjects vs PDAC, but at the same time increased miRNA levels were seen in cases with IPMN [20] or chronic pancreatitis [52].

Chronic pancreatitis is a risk factor for PDAC as they share many clinical symptoms, making a clear distinction between the two diseases difficult, particularly during the early stages of pancreatic cancer development. Thus, there is clinical need in distinguishing malignant and non-malignant abnormalities of the pancreas, which should be reflected in design of studies aiming to estimate diagnostic values of candidate biomarkers for detection of pancreatic cancer.

With biomarker panel we also aim to distinguish high-risk from low-risk controls and detect in high-risk individuals with sub centimeter lesions, but larger cohort is needed for this study. In addition, the panel of PJ-derived EV-miR-21, EV-miR-25, EV-miR-16 and serum miR-210 and CA19-9 only modestly increased the diagnostic performance of CA19-9 alone. Overlap between the samples does not allow to discriminate patients who do not have PDAC with high specificity. Thus, panel is not ready yet to be implemented in the clinical practice. Nevertheless, our study demonstrates differences between pancreatic juice and serum in terms of EV-miRNA expression for the first time to our best knowledge. An independent confirmation of our findings would be our recommended the next step.

CONCLUSIONS

A combined panel of PJ EV-miR-21, EV-miR-25, EV-miR-16, and serum EV-miR-210 and CA19-9 distinguishes cases with PDAC from controls undergoing surveillance with a specificity of 81.5% and sensitivity of 84.2%.

ACKNOWLEDGEMENTS

We would like to acknowledge Henk P. Roest his advises on qPCR; Guido W. Jenster, Marten E. van Royen for the help with electron microscopy.

REFERENCES

1. Siegel, R.L., K.D. Miller, and A. Jemal, *Cancer statistics, 2019.* CA Cancer J Clin, 2019. **69**(1): p. 7-34.
2. Siegel, R.L., K.D. Miller, and A. Jemal, *Cancer statistics, 2020.* CA: A Cancer Journal for Clinicians, 2020. **70**(1): p. 7-30.
3. Konings, I., et al., *Surveillance for pancreatic cancer in high-risk individuals.* BJS Open, 2019. **3**(5): p. 656-665.
4. Konings, I., et al., *Evolution of features of chronic pancreatitis during endoscopic ultrasound-based surveillance of individuals at high risk for pancreatic cancer.* Endosc Int Open, 2018. **6**(5): p. E541-E548.
5. Konings, I.C., et al., *Prevalence and Progression of Pancreatic Cystic Precursor Lesions Differ Between Groups at High Risk of Developing Pancreatic Cancer.* Pancreas, 2017. **46**(1): p. 28-34.
6. Overbeek, K.A., et al., *Long-term yield of pancreatic cancer surveillance in high-risk individuals.* Gut, 2021.
7. Yu, S., et al., *Validation of the 2012 Fukuoka Consensus Guideline for Intraductal Papillary Mucinous Neoplasm of the Pancreas From a Single Institution Experience.* Pancreas, 2017. **46**(7): p. 936-942.
8. Canto, M.I., et al., *International Cancer of the Pancreas Screening (CAPS) Consortium summit on the management of patients with increased risk for familial pancreatic cancer.* Gut, 2013. **62**(3): p. 339-47.
9. European Study Group on Cystic Tumours of the, P., *European evidence-based guidelines on pancreatic cystic neoplasms.* Gut, 2018. **67**(5): p. 789-804.
10. Pereira, S.P., et al., *Early detection of pancreatic cancer.* Lancet Gastroenterol Hepatol, 2020. **5**(7): p. 698-710.
11. Valsangkar, N.P., et al., *851 resected cystic tumors of the pancreas: a 33-year experience at the Massachusetts General Hospital.* Surgery, 2012. **152**(3 Suppl 1): p. S4-12.
12. Skulimowski, A., et al., *Comparison of clinical usefulness of serum Ca125 and CA19-9 in pancreatic adenocarcinoma diagnosis: meta-analysis and systematic review of literature.* Biomarkers, 2021: p. 1-9.
13. Goonetilleke, K.S. and A.K. Siriwardena, *Systematic review of carbohydrate antigen (CA 19-9) as a biochemical marker in the diagnosis of pancreatic cancer.* Eur J Surg Oncol, 2007. **33**(3): p. 266-70.
14. Xing, H., et al., *Diagnostic Value of CA 19-9 and Carcinoembryonic Antigen for Pancreatic Cancer: A Meta-Analysis.* Gastroenterol Res Pract, 2018. **2018**: p. 8704751.
15. Satake, K., et al., *A clinical evaluation of various tumor markers for the diagnosis of pancreatic cancer.* Int J Pancreatol, 1990. **7**(1-3): p. 25-36.
16. Nazli, O., et al., *The diagnostic importance of CEA and CA 19-9 for the early diagnosis of pancreatic carcinoma.* Hepatogastroenterology, 2000. **47**(36): p. 1750-2.
17. Sawabu, N., et al., *Serum Tumor Markers and Molecular Biological Diagnosis in Pancreatic Cancer.* Pancreas, 2004. **28**(3).
18. Kim, J.R., et al., *Clinical implication of serum carcinoembryonic antigen and carbohydrate antigen 19-9 for the prediction of malignancy in intraductal papillary mucinous neoplasm of pancreas.* J Hepatobiliary Pancreat Sci, 2015. **22**(9): p. 699-707.
19. Jenkinson, C., et al., *Biomarkers for early diagnosis of PC.* Expert Rev Gastroenterol Hepatol, 2015. **9**(3): p. 305-15.
20. Yu, Y., et al., *Identification of Serum microRNA-25 as a novel biomarker for pancreatic cancer.* Medicine (Baltimore), 2020. **99**(52): p. e23863.
21. Guz, M., et al., *Serum miR-210-3p can be used to differentiate between patients with pancreatic ductal adenocarcinoma and chronic pancreatitis.* Biomed Rep, 2021. **14**(1): p. 10.
22. Zhu, Y., et al., *Differential MicroRNA Expression Profiles as Potential Biomarkers for Pancreatic Ductal Adenocarcinoma.* Biochemistry (Mosc), 2019. **84**(5): p. 575-582.
23. Que, R., et al., *Analysis of serum exosomal microRNAs and clinicopathologic features of patients with pancreatic adenocarcinoma.* World J Surg Oncol, 2013. **11**: p. 219.
24. Madhavan, B., et al., *Combined evaluation of a panel of protein and miRNA serum-exosome biomarkers for pancreatic cancer diagnosis increases sensitivity and specificity.* Int J Cancer, 2015. **136**(11): p. 2616-27.
25. Gabriel, A.N., et al., *The involvement of exosomes in the diagnosis and treatment of PC.* Mol Cancer, 2020. **19**(1): p. 132.
26. Peng, Y. et al., *The role of MicroRNAs in human cancer.* Signal Transduction and Targeted Therapy, 2016. **1**(1): p. 15004.
27. Rawat, M., et al., *MicroRNA in Pancreatic Cancer: From Biology to Therapeutic Potential.* Genes, 2019. **10**(10): p. 752.
28. Daoud, A.Z., et al., *MiRNA in PC: biomarkers, prognostic, and therapeutic modulators.* BMC Cancer, 2019. **19**(1): p. 1130.
29. Elewaily, M.I. and A.R. Elsergany, *Emerging role of exosomes and exosomal microRNA in cancer: pathophysiology and clinical potential.* J Cancer Res Clin Oncol, 2021. **147**(3): p. 637-648.
30. Wang, X., et al., *Hypoxic Tumor-Derived Exosomal miR-301a Mediates M2 Macrophage Polarization via PTEN/PI3Kγ to Promote Pancreatic Cancer Metastasis.* Cancer Res, 2018. **78**(16): p. 4586-4598.
31. Chang, J., et al., *microRNA-21-5p from M2 macrophage-derived extracellular vesicles promotes the differentiation and activity of pancreatic cancer stem cells by mediating KLF3.* Cell Biol Toxicol, 2021.
32. Uysal-Onganer, P., et al., *Peptidylarginine Deiminase Inhibitor Application, Using Cl-Amidine, PAD2, PAD3 and PAD4 Isozyme-Specific Inhibitors in Pancreatic Cancer Cells, Reveals Roles for PAD2 and PAD3 in Cancer Invasion and Modulation of Extracellular Vesicle Signatures.* Int J Mol Sci, 2021. **22**(3).
33. Li, M., et al., *Pancreatic stellate cells derived exosomal miR-5703 promotes pancreatic cancer by downregulating CMTM4 and activating PI3K/Akt pathway.* Cancer Lett, 2020. **490**: p. 20-30.
34. Jiang, M.J., et al., *Dying tumor cell-derived exosomal miR-194-5p potentiates survival and repopulation of tumor repopulating cells upon radiotherapy in pancreatic cancer.* Mol Cancer, 2020. **19**(1): p. 68.
35. Sonohara, F., et al., *Exploration of Exosomal Micro RNA Biomarkers Related to Epithelial-to-Mesenchymal Transition in Pancreatic Cancer.* Anticancer Res, 2020. **40**(4): p. 1843-1853.
36. Nakamura, S., et al., *Pancreatic Juice Exosomal MicroRNAs as Biomarkers for Detection of Pancreatic Ductal Adenocarcinoma.* Annals of Surgical Oncology, 2019.
37. Maas, S.L.N., X.O. Breakefield, and A.M. Weaver, *Extracellular Vesicles: Unique Intercellular Delivery Vehicles.* Trends in Cell Biology, 2017. **27**(3): p. 172-188.
38. Costa-Silva, B., et al., *Pancreatic cancer exosomes initiate pre-metastatic niche formation in the liver.* Nat Cell Biol, 2015. **17**(6): p. 816-26.
39. Xue, J., et al., *Circulating microRNAs as promising diagnostic biomarkers for pancreatic cancer: a systematic review.* Onco Targets Ther, 2019. **12**: p. 6665-6684.
40. Levink, I.J.M., et al., *Optimization of Pancreatic Juice Collection: A First Step Toward Biomarker Discovery and Early Detection of Pancreatic Cancer.* Am J Gastroenterol, 2020. **115**(12): p. 2103-2108.
41. Lowry, O.H., et al., *Protein measurement with the Folin phenol reagent.* J Biol Chem, 1951. **193**(1): p. 265-75.
42. Verhoeven, C.J., et al., *MicroRNA profiles in graft preservation solution are predictive of ischemic-type biliary lesions after liver transplantation.* Journal of Hepatology, 2013. **59**(6): p. 1231-1238.

43. Farid, W.R.R., et al., *Hepatocyte-derived microRNAs as serum biomarkers of hepatic injury and rejection after liver transplantation.* Liver Transplantation, 2012. **18**(3): p. 290-297.
44. Livak, K.J. et al., *Analysis of relative gene expression data using RT-PCR and the 2(- Delta Delta C(T)) Method.* Methods, 2001. **25**(4): p. 402-8.
45. Moriyama, T., et al., *MicroRNA-21 modulates biological functions of pancreatic cancer cells including their proliferation, invasion, and chemoresistance.* Mol Cancer Ther, 2009. **8**(5): p. 1067-74.
46. Kadera, B.E., et al., *MicroRNA-21 in pancreatic ductal adenocarcinoma tumor-associated fibroblasts promotes metastasis.* PLoS One, 2013. **8**(8): p. e71978.
47. Lu, H., et al., *miR-25 expression is upregulated in pancreatic ductal adenocarcinoma and promotes cell proliferation by targeting ABI2.* Exp Ther Med, 2020. **19**(5): p. 3384-3390.
48. Zhang, J., et al., *Excessive miR-25-3p maturation via N.* Nat Commun, 2019. **10**(1): p. 1858.
49. Ho, A.S., et al., *Circulating miR-210 as a Novel Hypoxia Marker in Pancreatic Cancer.* Transl Oncol, 2010. **3**(2): p. 109-13.
50. Ni, J., et al., *Mechanism of miR-210 involved in epithelial-mesenchymal transition of pancreatic cancer cells under hypoxia.* J Recept Signal Transduct Res, 2019. **39**(5-6): p. 399-406.
51. Basu, A., et al., *MicroRNA-mediated regulation of pancreatic cancer cell proliferation.* Oncol Lett, 2010. **1**(3): p. 565-568.
52. Liu, J., et al., *Combination of plasma microRNAs with serum CA19-9 for early detection of pancreatic cancer.* Int J Cancer, 2012. **131**(3): p. 683-91.
53. Liu, G., et al., *Diagnostic Value of Plasma miR-181b, miR-196a, and miR-210 Combination in Pancreatic Cancer.* Gastroenterol Res Pract, 2020. **2020**: p. 6073150.
54. Deng, T., et al., *Identification of Circulating MiR-25 as a Potential Biomarker for Pancreatic Cancer Diagnosis.* Cell Physiol Biochem, 2016. **39**(5): p. 1716-1722.
55. Goto, T., et al., *An elevated expression of serum exosomal microRNA-191, - 21, -451a of pancreatic neoplasm is considered to be efficient diagnostic marker.* BMC Cancer, 2018. **18**(1): p. 116.
56. Kawamura, S., et al., *Exosome-encapsulated microRNA-4525, microRNA-451a and microRNA-21 in portal vein blood is a high-sensitive liquid biomarker for the selection of high-risk pancreatic ductal adenocarcinoma patients.* Journal of Hepato-Biliary-Pancreatic Sciences, 2019. **26**(2): p. 63-72.
57. Wu, L., et al., *Circulating exosomal microRNAs as novel potential detection biomarkers in pancreatic cancer.* Oncol Lett, 2020. **20**(2): p. 1432-1440.
58. Wang, J., et al., *Circulating microRNAs in Pancreatic Juice as Candidate Biomarkers of Pancreatic Cancer.* Journal of Cancer, 2014. **5**(8): p. 696-705.
59. Sadakari, Y., et al., *MicroRNA expression analyses in preoperative pancreatic juice samples of pancreatic ductal adenocarcinoma.* Jop, 2010. **11**(6): p. 587-92.
60. Squadrito, M.L., et al., *Endogenous RNAs modulate microRNA sorting to exosomes and transfer to acceptor cells.* Cell Rep, 2014. **8**(5): p. 1432-46.
61. Feng, S., et al., *Diagnostic significance of miR-210 as a potential tumor biomarker of human cancer detection: an updated pooled analysis of 30 articles.* Onco Targets Ther, 2019. **12**: p. 479-493.
62. Li, M., et al., *Analysis of the RNA content of the exosomes derived from blood serum and urine and its potential as biomarkers.* Philos Trans R Soc Lond B Biol Sci, 2014. **369**(1652).
63. Goggins, M., et al., *Management of patients with increased risk for familial pancreatic cancer: updated recommendations from the International Cancer of the Pancreas Screening (CAPS) Consortium.* Gut, 2020. **69**(1): p. 7.
64. Poruk, K.E., et al., *Screening for pancreatic cancer: why, how, and who?* Annals of surgery, 2013. **257**(1): p. 17-26.
65. Poruk, K.E., et al., *The clinical utility of CA 19-9 in PDAC: diagnostic and prognostic updates.* Curr Mol Med, 2013. **13**(3): p. 340-51.

SUPPLEMENTARY DATA

Supplementary tables

Supplementary Table 1: Inclusion criteria prospective cohort studies

KRASPanc study	
Inclusion	Patients (≥18 years of age) who undergo an EUS for (suspected) PDAC either as part of a diagnostic process or fiducial placement prior to radiotherapeutic treatment.
CAPS study	
Inclusion	Individuals (≥18 years of age) who, after evaluation by a clinical geneticist, have an estimated 10-fold increased risk of developing PDAC, this includes: (1) Carriers of a gene mutation in *CDKN2A* or *STK11*, regardless of the family history of pancreatic cancer (2) Carriers of a gene mutation in *BRCA1*, *BRCA2*, *p53*, or Mismatch Repair Gene with a family history of PDAC in ≥ 2 family members. (3) Familial PDAC (FPC) kindreds, defined as individuals with at least (1) 2 first-degree relatives (FDR) with PDAC, (2) 3 relatives with pancreatic cancer, either FDR or second degree relative (SDR), or (3) 2 SDR relatives with pancreatic cancer of which at ≥1 was <50 years at the time of diagnosis.
PACYFIC study	
Inclusion	Individuals (≥18 years of age) with a neoplastic pancreatic cyst (either newly or previously diagnosed, or previously operated upon) for which cyst surveillance is warranted, according to the treating physician.

Supplementary table 2. miRNA expression levels in cases (N=42) correlate with each other in serum and PJ, but not between these two biomarker sources. (*P<0.05)

Spearman R					
	Serum_miR-21	Serum_miR-25	Serum_miR-155	Serum_miR-210	Serum_miR-16
Serum_miR-21		0.81 (*)	0.40 (*)	0.85 (*)	0.75 (*)
Serum_miR-25			0.16	0.83 (*)	0.96 (*)
Serum_miR-155				0.23	0.18
Serum_miR-210					0.80 (*)
Serum_miR-16					
	PJ_miR-21	PJ_miR-25	PJ_miR-155	PJ_miR-210	PJ_miR-16
PJ_miR-21		0.90 (*)	0.83 (*)	0.86 (*)	0.93 (*)
PJ_miR-25			0.78 (*)	0.76 (*)	0.93 (*)
PJ_miR-155				0.85 (*)	0.86 (*)
PJ_miR-210					0.84 (*)
PJ_miR-16					
	PJ_miR-21	PJ_miR-25	PJ_miR-155	PJ_miR-210	PJ_miR-16
Serum_miR-21	-0.11	-0.03	0.00	0.02	-0.16
Serum_miR-25	0.06	0.20	0.05	0.13	0.02
Serum_miR-155	-0.03	-0.03	0.11	0.03	-0.01
Serum_miR-210	0.00	0.09	0.06	0.08	-0.03
Serum_miR-16	0.02	0.19	0.03	0.08	0.00
p-value					
	Serum_miR-21	Serum_miR-25	Serum_miR-155	Serum_miR-210	Serum_miR-16
Serum_miR-21		1.09E-10	0.01	1.67E-12	8.09E-09
Serum_miR-25			0.31	6.33E-12	1.76E-23
Serum_miR-155				1.48E-01	2.54E-01
Serum_miR-210					2.57E-10
Serum_miR-16					
	PJ_miR-21	PJ_miR-25	PJ_miR-155	PJ_miR-210	PJ_miR-16
PJ_miR-21		3.01E-16	1.15E-11	1.73E-13	5.72E-19
PJ_miR-25			1.47E-09	4.84E-09	8.90E-19
PJ_miR-155				1.29E-12	5.07E-13
PJ_miR-210					3.26E-12
PJ_miR-16					
	PJ_miR-21	PJ_miR-25	PJ_miR-155	PJ_miR-210	PJ_miR-16
Serum_miR-21	0.48	0.83	0.98	0.88	0.30
Serum_miR-25	0.73	0.20	0.76	0.40	0.91
Serum_miR-155	0.83	0.84	0.48	0.85	0.97
Serum_miR-210	1.00	0.57	0.72	0.61	0.85
Serum_miR-16	0.91	0.22	0.86	0.63	0.99

Supplementary table 3. miRNA expression levels in cases (N=55) correlate with each other in serum and PJ, but not between these two biomarker sources. (*P<0.05)

Spearman R					
	Serum_miR-21	Serum_miR-25	Serum_miR-155	Serum_miR-210	Serum_miR-16
Serum_miR-21		0.81 (*)	0.47 (*)	0.83 (*)	0.84 (*)
Serum_miR-25			0.37 (*)	0.88 (*)	0.97 (*)
Serum_miR-155				0.43 (*)	0.44 (*)
Serum_miR-210					0.89 (*)
Serum_miR-16					
	PJ_miR-21	PJ_miR-25	PJ_miR-155	PJ_miR-210	PJ_miR-16
PJ_miR-21		0.83 (*)	0.62 (*)	0.58 (*)	0.94 (*)
PJ_miR-25			0.67 (*)	0.53 (*)	0.88 (*)
PJ_miR-155				0.49 (*)	0.61 (*)
PJ_miR-210					0.63 (*)
PJ_miR-16					
	PJ_miR-21	PJ_miR-25	PJ_miR-155	PJ_miR-210	PJ_miR-16
Serum_miR-21	-0.06	-0.09	-0.25	-0.03	-0.05
Serum_miR-25	0.03	-0.06	-0.22	0.01	0.04
Serum_miR-155	0.19	0.12	0.26 (*)	0.29	0.16
Serum_miR-210	0.08	0.00	-0.18	0.12	0.09
Serum_miR-16	0.02	-0.07	-0.20	0.04	0.03
p-value					
	Serum_miR-21	Serum_miR-25	Serum_miR-155	Serum_miR-210	Serum_miR-16
Serum_miR-21		4.92E-14	2.61E-04	6.51E-15	1.14E-15
Serum_miR-25			0.01	3.27E-19	4.15E-33
Serum_miR-155				1.18E-03	9.00E-04
Serum_miR-210					2.10E-19
Serum_miR-16					
	PJ_miR-21	PJ_miR-25	PJ_miR-155	PJ_miR-210	PJ_miR-16
PJ_miR-21		5.30E-15	3.69E-07	3.32E-06	1.56E-25
PJ_miR-25			2.52E-08	3.77E-05	1.31E-18
PJ_miR-155				1.30E-04	7.11E-07
PJ_miR-210					2.03E-07
PJ_miR-16					
	PJ_miR-21	PJ_miR-25	PJ_miR-155	PJ_miR-210	PJ_miR-16
Serum_miR-21	0.66	0.52	0.06	0.82	0.71
Serum_miR-25	0.85	0.69	0.11	0.91	0.78
Serum_miR-155	0.17	0.40	0.05	0.03	0.24
Serum_miR-210	0.55	0.97	0.19	0.37	0.49
Serum_miR-16	0.87	0.63	0.14	0.78	0.81

Supplementary table 4. Summary of studies with selected miRNA, NG – not given, Ref-reference, SE – sensitivity, SP – specificity

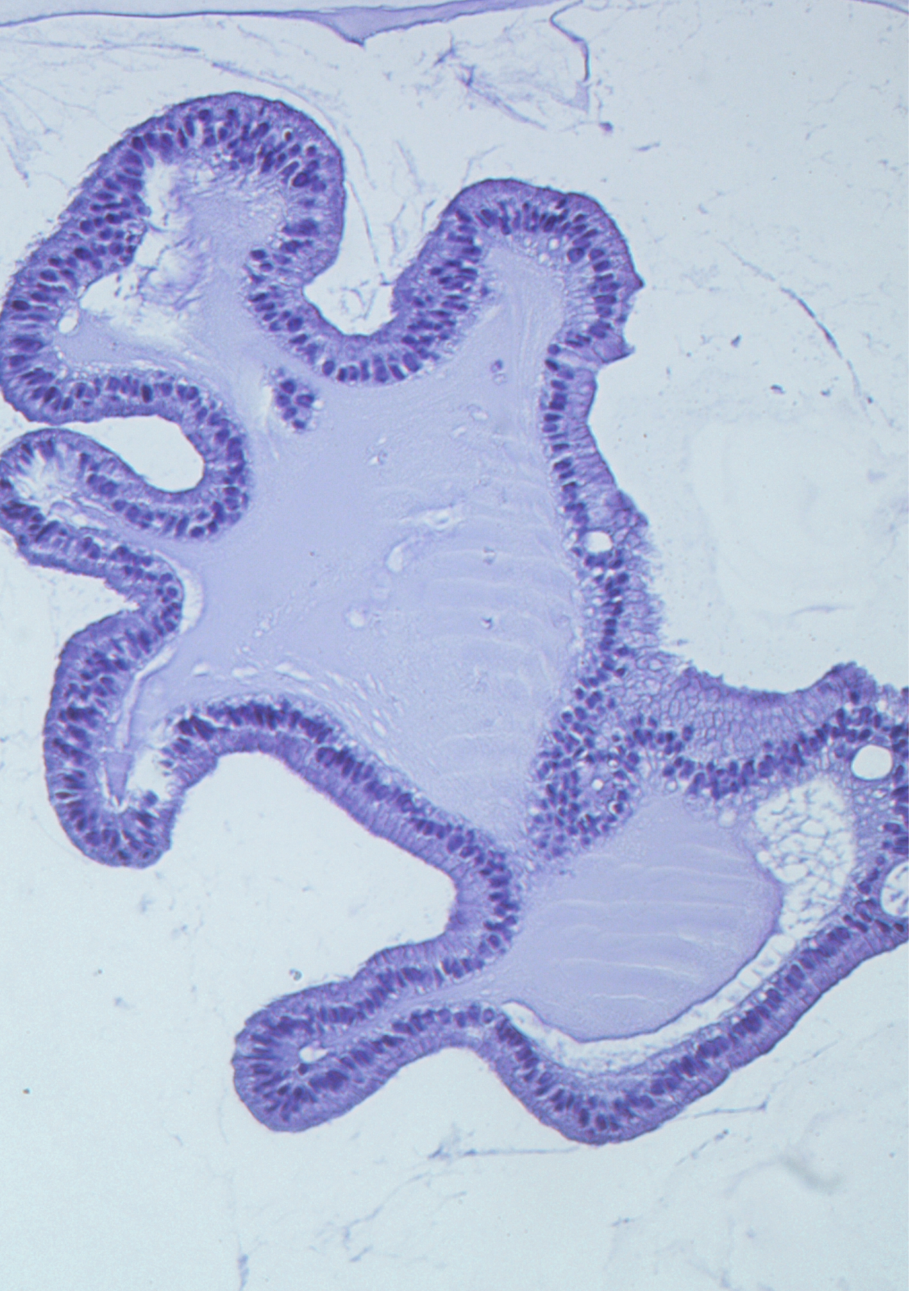
Ref	miR	Biofluid	Isolated EVs or total biofluid	N Cases/ controls	Control group	SE%	SP%	Normalization method
[2]	miR-25	Serum	Total	80/91	Healthy without malignancy	82.5%	93.6%	absolute quantification
[3]	miR-16	Serum	Total	140/ 111+68	Chronic pancreatitis + Healthy	NG, AUC=0.75	NG	cel-miR-39
[3]	miR-21	Serum	Total	140/ 111+68	Chronic pancreatitis + Healthy	NG, AUC= 0.78	NG	cel-miR-39
[3]	miR-155	Serum	Total	140/ 111+68	Chronic pancreatitis + Healthy	NG, AUC for 0.70	NG	cel-miR-39
[3]	miR-210	Serum	Total	140/ 111+68	Chronic pancreatitis + Healthy	NG, AUC=0.76	NG	cel-miR-39
[4]	miR-25	Serum	Total	303/ 600+ 160	Non-malignant controls + 40 with chronic pancreatitis, 20 with gastric cancer, 20 with lung cancer, 20 with esophageal cancer, 20 with colorectal cancer, 20 with liver cancer, and 20 with breast cancer	75.6%	93.0%	serum volume
[5]	miR-21	Serum	EV and total	32/22	Healthy without malignancy + other GI tract conditions (e.g. gastritis)	80.7% (total: 53.5%	81.0 (for total: 76.2%)	NG
[6]	miR-21	Serum	EVs	22/27	6 benign pancreatic tumors, 7 with ampullary carcinomas, 6 with chronic pancreatitis and 8 healthy	81.5%	95.5%	RNU6B

[6]	miR-155	Serum	EVs	22/27	6 benign pancreatic tumors, 7 ampullary carcinomas, 6 chronic pancreatitis, 8 healthy	Low expression		RNU6B
[8]	miR-21	Serum	EVs and Total	30/10	Chronic pancreatitis	80% (total 73%)	90% (For total 70%)	cel-miR-39
[8]	miR-210	Serum	EVs and Total	30/10	Chronic pancreatitis	83% (total 76%)	90% (for total 70%)	cel-miR-39
[9]	miR-21	Plasma	Total	31/28	Healthy	NG, AUC 0.85		absolute quantification and miR-39
[9]	miR-210	Plasma	Total	31/28	Healthy	NG, AUC 0.69		absolute quantification and miR-39
[9]	miR-155	Plasma	Total	31/28	Healthy	NG, AUC 0.82		absolute quantification and miR-39
[9]	miR-25	Plasma	Total	31/28	Healthy	NG, AUC 0.76		absolute quantification and miR-39
[7]	miR-210	Plasma	Total	40/40	Healthy	82.5%	80.0%	U6
[10]	miR-21	Plasma	EVs	55/20	Healthy	72.7% (peripheral blood 54.5%)	72.7% (peripheral blood 63.6%)	RNU6B
[11]	miR-210	PJ	Total	50/19+19	Non-pancreatic, non-healthy controls + chronic pancreatitis	76%	95%	RNU6B
[13]	miR-21	PJ	EVs and Total	27/8	Chronic pancreatitis	81% (total AUC=0.71)	88%	miR-16

[13]	miR-155	PJ	EVs and Total	27/8	Chronic pancreatitis	89% (total miR-155 AUC 0.56)	88%	miR-16
[12]	miR-21	PJ	Cell pellet	16/5	Chronic pancreatitis	NG	NG	RNU6B and miR-191
[12]	miR-155	PJ	Cell pellet	16/5	Chronic pancreatitis	NG	NG	RNU6B and miR-191

Supplementary references

1. European Study Group on Cystic Tumours of the, P., European evidence-based guidelines on pancreatic cystic neoplasms. Gut, 2018. 67(5): p. 789-804.
2. Yu, Y., et al., Identification of Serum microRNA-25 as a novel biomarker for pancreatic cancer. Medicine (Baltimore), 2020. 99(52): p. e23863.
3. Liu, J., et al., Combination of plasma microRNAs with serum CA19-9 for early detection of pancreatic cancer. Int J Cancer, 2012. 131(3): p. 683-91.
4. Deng, T., et al., Identification of Circulating MiR-25 as a Potential Biomarker for Pancreatic Cancer Diagnosis. Cell Physiol Biochem, 2016. 39(5): p. 1716-1722.
5. Goto, T., et al., An elevated expression of serum exosomal microRNA-191, - 21, -451a of pancreatic neoplasm is considered to be efficient diagnostic marker. BMC Cancer, 2018. 18(1): p. 116.
6. Que, R., et al., Analysis of serum exosomal microRNAs and clinicopathologic features of patients with pancreatic adenocarcinoma. World J Surg Oncol, 2013. 11: p. 219.
7. Liu, G., et al., Diagnostic Value of Plasma miR-181b, miR-196a, and miR-210 Combination in Pancreatic Cancer. Gastroenterol Res Pract, 2020. 2020: p. 6073150.
8. Wu, L., et al., Circulating exosomal microRNAs as novel potential detection biomarkers in pancreatic cancer. Oncol Lett, 2020. 20(2): p. 1432-1440.
9. Yu, Q., et al., Evaluation of Plasma MicroRNAs as Diagnostic and Prognostic Biomarkers in Pancreatic Adenocarcinoma: miR-196a and miR-210 Could Be Negative and Positive Prognostic Markers, Respectively. Biomed Res Int, 2017. 2017: p. 6495867.
10. Kawamura, S., et al., Exosome-encapsulated microRNA-4525, microRNA-451a and microRNA-21 in portal vein blood is a high-sensitive liquid biomarker for the selection of high-risk pancreatic ductal adenocarcinoma patients. Journal of Hepato-Biliary-Pancreatic Sciences, 2019. 26(2): p. 63-72.
11. Wang, J., et al., Circulating microRNAs in Pancreatic Juice as Candidate Biomarkers of Pancreatic Cancer. Journal of Cancer, 2014. 5(8): p. 696-705.
12. Sadakari, Y., et al., MicroRNA expression analyses in preoperative pancreatic juice samples of pancreatic ductal adenocarcinoma. Jop, 2010. 11(6): p. 587-92.
13. Nakamura, S., et al., Pancreatic Juice Exosomal MicroRNAs as Biomarkers for Detection of Pancreatic Ductal Adenocarcinoma. Annals of Surgical Oncology, 2019.



Chapter 5

Upregulated β-catenin signaling does not affect survival of pancreatic cancer cells during dual inhibition of GSK3B and HDAC

K. Nesteruk, R. Smits, M. Bruno, M. P. Peppelenbosch, G. M. Fuhler

Pancreatology. 2019

Chapter 5. Upregulated β-catenin signaling does not affect the survival of PDAC cells during dual inhibition of GSK3B and HDAC

INTRODUCTION

The newly synthesized molecule Metavert was recently introduced as a promising new agent for treatment of pancreatic ductal adenocarcinoma (PDAC) [1]. Metavert slows tumor growth and metastasis by inhibiting both glycogen synthase kinase 3 beta (GSK3B) and histone deacetylases (HDACs). Edderkaoui *et al* recently demonstrated that dual targeting of these pathways induces synergistic PDAC killing [1], and showed that Metavert decreases expression of cancer stemness markers associated with epithelial-to-mesenchymal transition and metastasis, which can still occur under inhibition of GSK3B alone [2]. However, an unexpected increase in β-catenin protein levels was seen in Metavert-treated PDAC cells, suggesting activation of Wnt/β-catenin signaling. Wnt/β-catenin signaling is complex and was shown to enhance PDAC development and malignancy [3-6]. However, it has also been suggested that β-catenin partly mediates killing effects of GSK3B inhibitors in KRAS-dependent tumors [7]. Furthermore, a specific dosage of β-catenin signaling is needed for tumor formation as an excessive accumulation of β-catenin leads to apoptosis in normal and carcinoma cells [8-10]. Thus, to what extent Wnt/β-catenin signaling plays a role in Metavert-mediated PDAC killing remains unclear. We therefore investigated the activity of this pathway upon inhibition of GSK3 and/or HDAC and determined its role in PDAC cell cytotoxicity.

METHODS

MTT test BxPC-3, Panc-1, and MIAPaCa-2 cell lines were treated with GSK3B inhibitors CHIR99021 and TWS119, HDAC inhibitor Vorinostat or Wnt3a conditioned medium. MTT test was performed after 72h [11].

β-Catenin Reporter Assays were performed as described [12]. After transfection with Wnt Responsive Element (WRE) or Mutant Responsive Element (MRE) vectors and TK-Renilla, luciferase activity was measured and normalized for transfection efficiency using the Dual Luciferase Reporter Assay system (Promega). WRE/MRE ratios are shown.

qPCR for Axin2 was performed as described [13]. In short, after 24h of treatment total RNA was isolated for cDNA preparation. Primers used: forward TATCCAGTGATGCGCTGAC, reverse TTACTGCCCACACGATAAGG.

siRNA-mediated gene knock-down. Smartpool ON-TARGETplus siRNAs targeting *CTNNB1* and nontargeting siRNA control #2 were introduced into cells using DharmaFECT. Successful knockdown was confirmed by Western blot analysis [14] using a β-catenin antibody (#610154, BD Transduction Laboratories) and IRdye-linked secondary antibodies. β-actin served as loading control.

Statistics. One-way ANOVA was used for statistical analysis of MTT, two-way ANOVA and student t-test was used for reporter assay and knockdown experiment. For dose-response curves, best fit sigmoidal dose-response (variable slope) curves are presented.

RESULTS

The β-catenin reporter assay showed that while the HDAC inhibitor Vorinostat alone does not change β-catenin signaling in PDAC cells (**Figure 1A, 1B, 1C**), the GSK3B inhibitor CHIR99021 activates this pathway in a dose dependent fashion, far exceeding β-catenin signaling induced by Wnt-3a conditioned medium. Importantly, dual targeting of GSK3B and HDAC causes synergistic β-catenin activation as compared to CHIR99021 alone. Furthermore, in line with previous reports [15-17], the less potent GSK3B inhibitor TWS119 showed lower activation of β-catenin signaling (**Figure 1, supplementary Figure 1**), but nevertheless also displayed synergistic β-catenin activation in Panc-1 and BxPC-3 cells in combination with HDAC inhibition. Vorinostat also strengthens this signal in combination with Wnt-3a conditioned medium. We verified these findings by investigation of mRNA levels of *AXIN2*, a downstream target gene of β-catenin **(Figure 1D, 1E, 1F**). Similar to CHIR99021, albeit at lower levels, TWS119 increases expression of *AXIN2* alone and in combination with Vorinosat in all three cell lines.

We next investigated whether this synergistic β-catenin signaling mediates PDAC cell killing. First we confirmed the effect of combined inhibition of GSK3B and HDAC on PDAC growth inhibition (**Figure 2**). Subsequently, we tested whether β-catenin signaling affects PDAC growth by addition of Wnt-3a conditioned medium, as activation of Wnt3a receptors Frizzled and LRP5/6 leads to stabilization of cytoplasmic β-catenin [18]. However, Wnt3a stimulation did not affect PDAC cell viability (**Figure 3**).

We then studied the direct effect of β-catenin through verified siRNA mediated knockdown (**Figure 4**). Although knockdown of β-catenin in itself decreased the growth of pancreatic cancer cell lines by 22% (Panc-1), 33% (MIAPaCa-2) and 20% (BxPC-3)(not shown, $p<0.05$), knockdown did not affect susceptibility of Panc-1 and MIAPaCa-2 to treatment of GSK3B and HDAC inhibitor, and only slightly increased survival of BxPC-3 compared to controls when treated with CHIR99021 and Vorinostat (Figure 4D).

DISCUSSION

Our data suggest that while β-catenin knockdown in itself may reduce PDAC viability to some extent, activation of this signaling pathway does not contribute to the cytotoxic effects induced by combined GSK3/HDAC inhibition. It is conceivable that the synergistic β-catenin signaling triggered by these inhibitors seen here mediates other anti-tumorigenic effects of this dual treatment i.e. EMT or metastasis. The exact role of β-catenin signaling during Metavert treatment of PDAC remains elusive and requires further investigation.

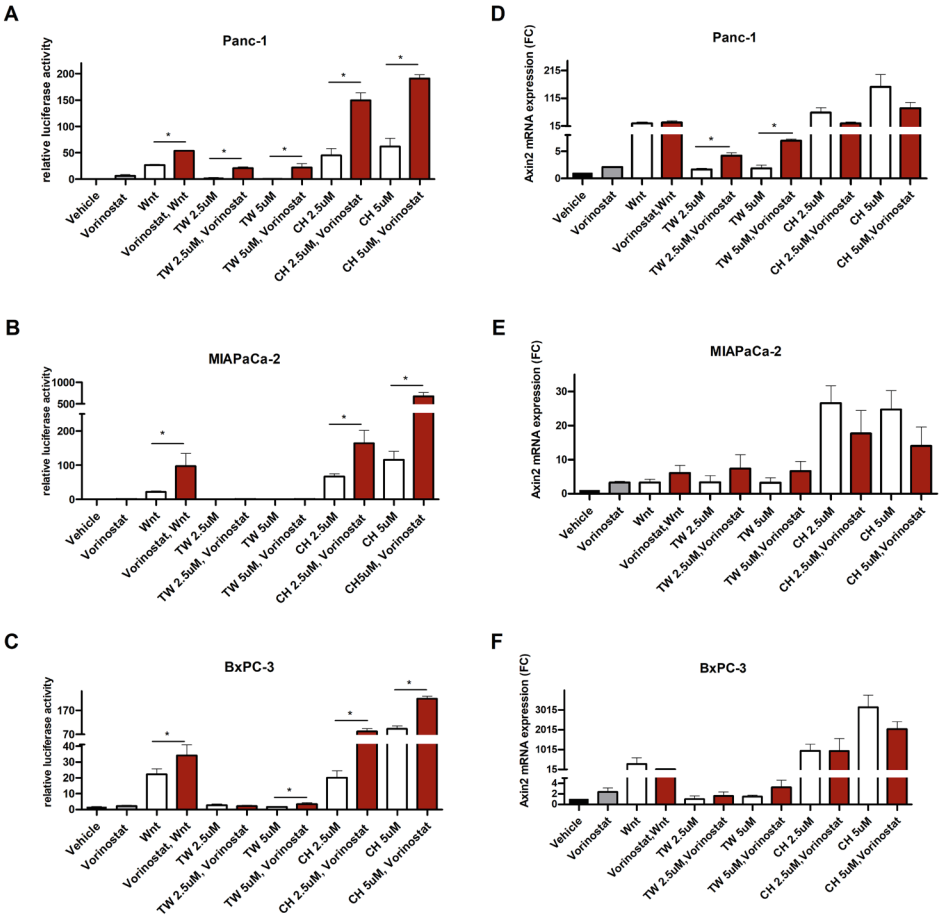


Figure 1. Treatment of Panc-1 (A, D), MIAPaCa-2 (B, E), BxPC3 (C, F) cells with GSK3B inhibitors CHIR99021 (CH) and TWS119 (TW) or their combination with HDAC inhibitor Vorinostat shows synergistic effect of these inhibitors on β-catenin signaling as determined by reporter assays or qPCR for AXIN2. Mean ± SEM, * $p<0.05$.

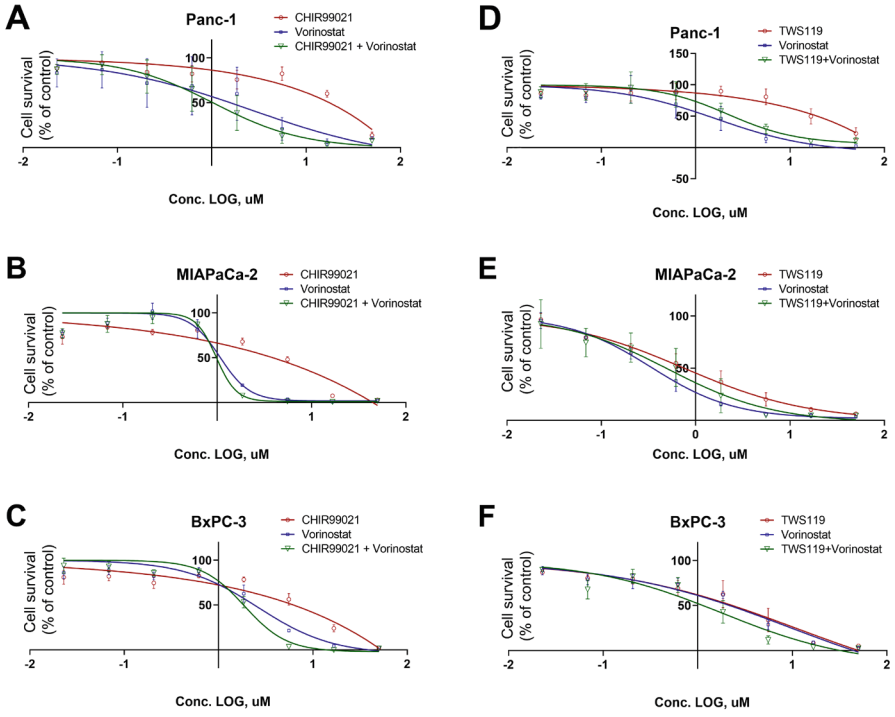


Figure 2. Treatment with CHIR99021, TWS119, Vorinostat or their combination induces killing of Panc-1 (A, D), MIAPaCa-2 (B, E), BxPC3 (C, F) cells as determined by MTT assay (Mean ± SEM) (B).

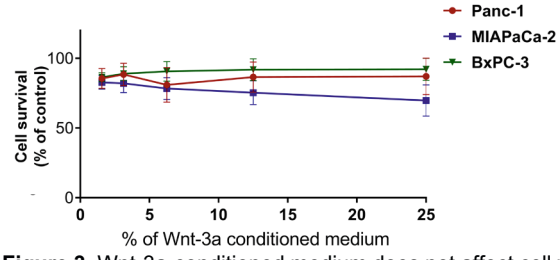


Figure 3. Wnt-3a-conditioned medium does not affect cell viability as determined by MTT assay (Mean ± SEM).

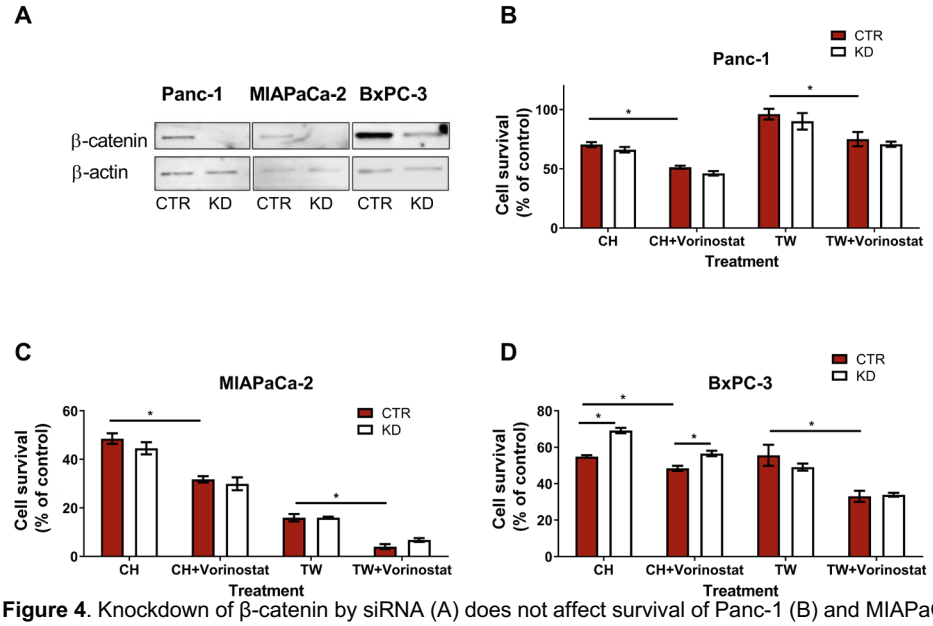


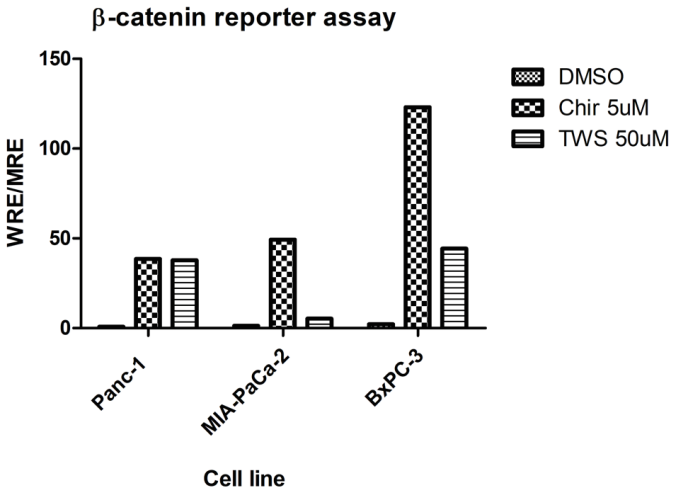
Figure 4. Knockdown of β-catenin by siRNA (A) does not affect survival of Panc-1 (B) and MIAPaCa-2 (C) cells after treatment with CHIR99021, TWS119 and Vorinostat and moderately inhibits cell killing of BxPC-3 (D) cells by CH and CH+Vorinostat (Mean ± SEM, * $p<0.05$).

REFERENCES

1. Edderkaoui M, Chheda C, Soufi B, et al. An Inhibitor of GSK3B and HDACs Kills Pancreatic Cancer Cells and Slows Pancreatic Tumor Growth and Metastasis in Mice. Gastroenterology 2018;155:1985-1998.e5.
2. Zheng H, Li W, Wang Y, et al. Glycogen synthase kinase-3 beta regulates Snail and β-catenin expression during Fas-induced epithelial–mesenchymal transition in gastrointestinal cancer. European Journal of Cancer 2013;49:2734-2746.
3. Sano M, Driscoll DR, DeJesus-Monge WE, et al. Activation of WNT/beta-Catenin Signaling Enhances Pancreatic Cancer Development and the Malignant Potential Via Up-regulation of Cyr61. Neoplasia 2016;18:785-794.
4. Wang L, Heidt DG, Lee CJ, et al. Oncogenic function of ATDC in pancreatic cancer through Wnt pathway activation and beta-catenin stabilization. Cancer Cell 2009;15:207-19.
5. Wang L, Yang H, Abel EV, et al. ATDC induces an invasive switch in KRAS-induced pancreatic tumorigenesis. Genes Dev 2015;29:171-83.
6. Yu M, Ting DT, Stott SL, et al. RNA seq of pancreatic CTC implicates WNT signalling in metastasis. Nature 2012;487:510-3.
7. Kazi A, Xiang S, Yang H, et al. GSK3 suppression upregulates β-catenin and c-Myc to abrogate KRas-dependent tumors. Nature Communications 2018;9:5154.
8. Albuquerque C, Breukel C, van der Luijt R, et al. The 'just-right' signaling model: APC somatic mutations are selected based on a specific level of activation of the β-catenin signaling cascade. Human Molecular Genetics 2002;11:1549-1560.
9. Kim K, Pang KM, Evans M, et al. Overexpression of beta-catenin induces apoptosis independent of its transactivation function with LEF-1 or the involvement of major G1 cell cycle regulators. Mol Biol Cell 2000;11:3509-23.
10. Albuquerque C, Bakker ERM, van Veelen W, et al. Colorectal cancers choosing sides. Biochimica et Biophysica Acta (BBA) - Reviews on Cancer 2011;1816:219-231.
11. Queiroz KC, Milani R, Ruela-de-Sousa RR, et al. Violacein induces death of resistant leukaemia cells via kinome reprogramming, endoplasmic reticulum stress and Golgi apparatus collapse. PLoS One 2012;7:e45362.
12. Bakker ER, Das AM, Helvensteijn W, et al. Wnt5a promotes human colon cancer cell migration and invasion but does not augment intestinal tumorigenesis in Apc1638N mice. Carcinogenesis 2013;34:2629-38.
13. Janmaat VT, Liu H, da Silva RA, et al. HOXA9 mediates and marks premalignant compartment size expansion in colonic adenomas. Carcinogenesis 2019.
14. Somasundaram R, Fernandes S, Deuring JJ, et al. Analysis of SHIP1 expression and activity in Crohn's disease patients. PLoS One 2017;12:e0182308.
15. Billin AN, Thirlwell H, Ayer DE. Beta-catenin-histone deacetylase interactions regulate the transition of LEF1 from a transcriptional repressor to an activator. Molecular and cellular biology 2000;20:6882-6890.
16. Bordonaro M, Lazarova DL, Sartorelli AC. The activation of beta-catenin by Wnt signaling mediates the effects of histone deacetylase inhibitors. Experimental cell research 2007;313:1652-1666.
17. Debeb BG, Lacerda L, Xu W, et al. Histone deacetylase inhibitors stimulate dedifferentiation of human breast cancer cells through WNT/β-catenin signaling. Stem cells (Dayton, Ohio) 2012;30:2366-2377.
18. MacDonald BT, He X. Frizzled and LRP5/6 receptors for Wnt/β-catenin signaling. Cold Spring Harbor perspectives in biology;4:a007880.

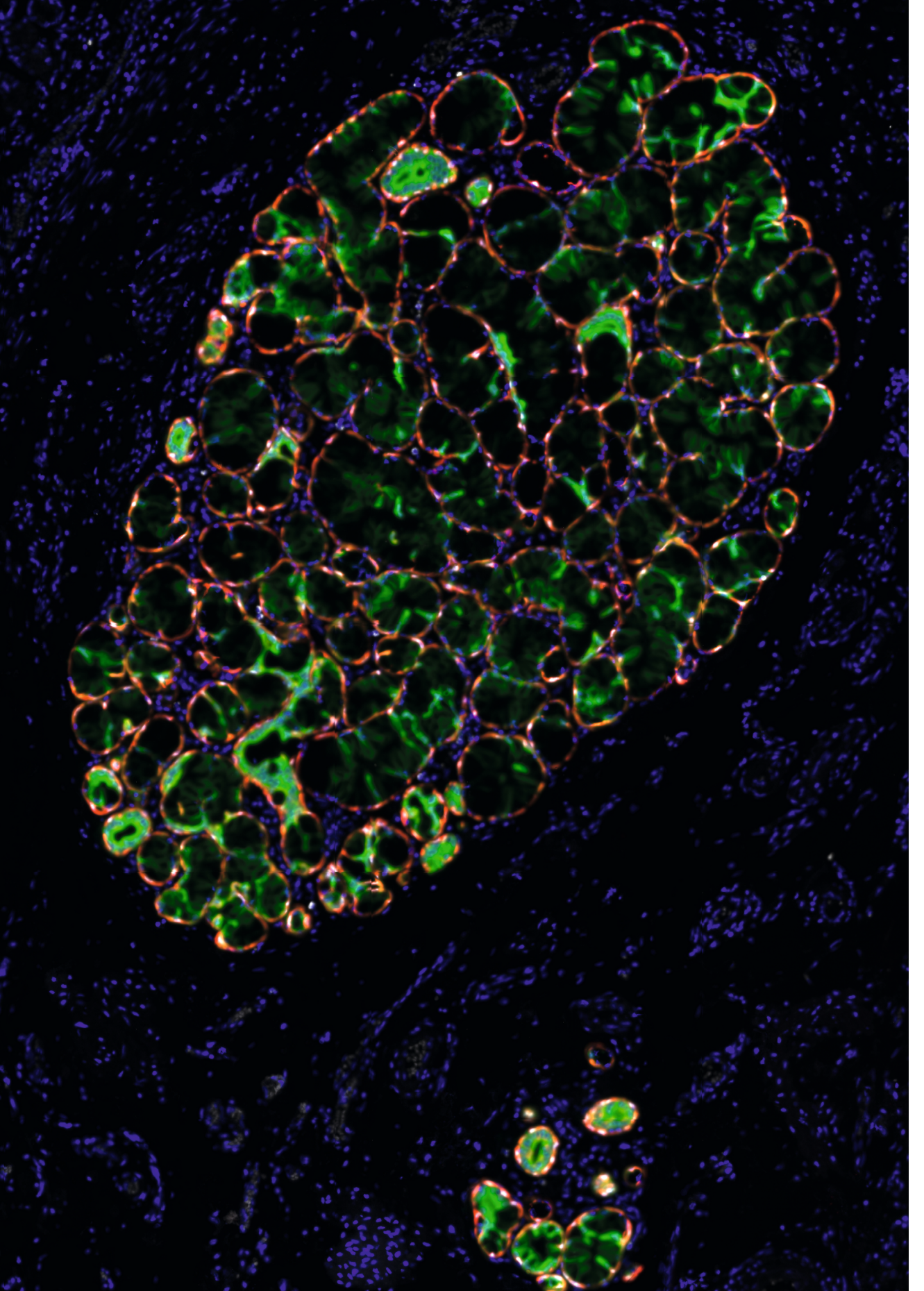
SUPPLEMENTARY DATA

Supplementary figures



Supplementary Figure 1. TWS119 (TW) at dosage of 50uM shows the same or lower level activation of β-catenin signaling as CHIR99021 (CR) at 5uM in three pancreatic cancer cell lines.

Esophageal lesions



Chapter 6

HOXA13 in etiology and oncogenic potential of Barrett's esophagus

V. T. Janmaat, K. Nesteruk, M. C.W. Spaander, A. P. Verhaar, B. Yu, R. A. Silva, W. A. Phillips, M. Magierowski, A. van de Winkel, H. S. Stadler, T. Sandoval-Guzmán, L.J.W. van der Laan, E. J. Kuipers, R. Smits, M. J. Bruno1, G. M. Fuhler#, N. J. Clemons#, & M. P. Peppelenbosch1#.

Nature Communications. 2021

ABSTRACT

Barrett's esophagus in gastrointestinal reflux patients constitutes a columnar epithelium with distal characteristics, prone to progress to esophageal adenocarcinoma. *HOX* genes are known mediators of position-dependent morphology. Here we show *HOX* collinearity in the adult gut while Barrett's esophagus shows high *HOXA13* expression in stem cells and their progeny. *HOXA13* overexpression appears sufficient to explain both the phenotype (through downregulation of the epidermal differentiation complex) and the oncogenic potential of Barrett's esophagus. Intriguingly, employing a mouse model that contains a reporter coupled to the *HOXA13* promotor we identify single HOXA13-positive cells distally from the physiological esophagus, which is mirrored in human physiology, but increased in Barrett's esophagus. Additionally, we observe that *HOXA13* expression confers a competitive advantage to cells. We thus propose that Barrett's esophagus and associated esophageal adenocarcinoma is the consequence of expansion of this gastro-esophageal *HOXA13*-expressing compartment following epithelial injury.

HIGHLIGHTS

Barrett's esophagus is a pro-oncogenic lesion in the proximal GI tract, but with a distal morphology
Barrett's esophagus is characterized by expression of the distal Homeobox gene *HOXA13*
HOXA13 is expressed in single cells of the physiological esophagus and distal GI tract
HOXA13 conveys phenotypic aspects of metaplasia and increases proliferation.

INTRODUCTION

Barrett's esophagus (BE) and gastric intestinal metaplasia (IM) are important risk factors for adenocarcinoma of the esophagus and stomach. In the esophagus, the chronic inflammation associated with gastroesophageal reflux disease (GERD) is believed to lead to Barrett's esophagus (BE), a crypt-structured columnar epithelium with distal gastrointestinal (GI)-tract characteristics, located just above the gastro-esophageal junction (GEJ). BE is a precursor lesion for esophageal adenocarcinoma (EAC) [1, 2], a disease which has shown a strong increase in incidence in the past decades. Analogously, *H. pylori*-infection can degenerate into atrophic gastritis and gastric IM, which in turn can progress into gastric cancer, the third leading cause of cancer-related death [3]. Similarly, while absolute risk is low, heterotopic tissues in Meckel's diverticula and gastric inlet patches of the proximal esophagus represent relatively high-risk regions for adenocarcinoma comparatively to other sites of the ileum and proximal esophagus, respectively [4, 5]. Therefore, a deeper understanding of the biology of BE and gastric IM is necessary for designing rational avenues for the prevention and treatment of GI cancers.

BE is characterized by the presence of cells with a caudal intestinal phenotype at a rostral location. Therefore, dysregulation of positional specification is likely involved in the etiology of BE. Regulation of rostral-caudal patterning of specialized tissue in embryology and adulthood is to a large extent

dependent on the concerted action of two evolutionary highly conserved gene systems, the Caudal-related Homeobox (*CDX*) transcription factor gene family and the genes of the Homeobox (*HOX*) cluster. A substantial research effort has been invested in investigating the role of *CDX* genes in positional misspecification in BE [6]. However, these efforts have not yielded convincing evidence that these genes are the principal mediators of the distal phenotype in this disease [7, 8]. Intriguingly, however, a microarray-based gene expression study of BE suggested potential misregulation of the *HOX* gene family in BE [9]. *HOX* genes are linked to morphological transformations and neoplasia [10, 11]. Four clusters of *HOX* genes, *HOXA* to *HOXD,* have been defined. The 3' to 5' sequence of *HOX* gene paralogues corresponds to the sequence in which they act along the rostrocaudal axis. This property is termed collinearity and links clustering to function. Previously, a *Hox* expression gradient was found along the murine embryonic gut [12]. Ectopic *Hox* expression in mice can alter intestinal differentiation[8]. A *HOX* gradient along the adult human gut has also been reported [13], but that study involved pooling full thickness gut specimens, limiting data interpretation. Nevertheless, we feel that there is sufficient evidence to prompt exploring the function of *HOX* gene expression with respect to positional identity in physiology and pathology of the GI-tract in general and in BE in particular.

Here, we show that single cells of the upper GI tract express the distal gene *HOXA13,* that their number is upregulated in BE and that HOXA13 conveys phenotypic metaplasia and increases proliferation.

METHODS

Collection of human material

All human tissues used in this study were obtained at the Erasmus University Medical Center, department of Gastroenterology & Hepatology. The use of these samples was approved by the Erasmus MC medical ethical committee (MEC-2015-208, MEC-2015-209, MEC-2015-199, MEC-2010-093; tissues were handled according to the FEDERA code of conduct and informed consent was obtained for all participants[60]. "The study was designed and carried out according to the ethical principles for medical research involving human subjects from the World Medical Association Declaration of Helsinki". Biopsy specimens to investigate *HOX* collinearity were obtained by double balloon enteroscopy. Nine biopsy specimens were obtained from each patient (n=3) at different locations along the GI-tract. Sequentially these locations were: esophagus, stomach, duodenum, jejunum, proximal ileum, distal ileum, ascending colon, descending colon and sigmoid/rectum (Supplementary fig. 2). Included patients had unexplained symptoms, mostly anemia, while inflammatory bowel disease patients were excluded. All biopsies for RNA isolation were stored in RNA-later at -80˚C. Squamous esophageal biopsies (n=13) originated 5 cm above the squamocolumnar junction (SCJ). Barrett's (BE) biopsies (n=13) originated caudal of the SCJ and cranial of the gastric folds (all patients were on PPI therapy), stomach biopsies (n=12) were from the corpus. All three types of biopsy specimens were derived from the same patients, (one stomach biopsy specimen was not obtained due to patient agitation during the gastroscopy). The squamous esophageal, BE, and stomach biopsies were taken in a paired fashion from 13 patients. Where the number of samples is indicated

below, this indicates the number of individual patients. Forceps biopsy specimens of EACs (n=12) were obtained. Pathological examination of simultaneously taken forceps biopsies around the study specimens had to be positive for EAC. Gastric inlet patch were sampled from proximal esophagus. To determine the proximal colonic HOXA13 border, biopsies were taken from the cecum at the appendix base, the ileocecal valve, 5 cm distal to the ileocecal valve, and from the transverse colon in each patient (n=5).

Collection of archival pathology specimens

FFPE material was collected from gastric IM (from the antrum, angulus, and corpus, i.e. not from the cardia; n=12), the gastric inlet patch (n=5), CLE (from the proximal esophagus; n=14), and Meckel's diverticula (n=14). For RNAscope RNA-ISH, one FFPE specimen of each of these origins was used. Depending on the extent of metaplasia, remainder of tissues was used whole, macroscopically separated, or processed with the Photo-Activated Localization Microscopy with the Laser Caption Microdissection (PALM LCM) for mRNA isolation and subsequent qPCR. Nuclease-free membrane slides treated with UV light at 254 nm for 30 minutes were used to mount 10µm sections, dried overnight at 56˚C, deparaffinized, stained with hematoxylin and eosin, and dehydrated. AdhesiveCap microtubes obtained from Zeiss (Oberkochen, Germany) were used to collect the tissue of interest after cutting and pulsing of the PALM LCM. Additionally, FFPE materials or fresh pinch biopsies were collected from the squamous esophagus of a patient without BE, the squamous esophagus of a BE patient, BE, EAC, stomach (the corpus), and the ileum. Colon was used as a positive control. FFPE materials that were collected only for RNA-ISH were pyloric metaplasia (from the colon; n=5), Paneth cell metaplasia (from the colon; n=5), fetal GEJ tissue (n=2 of 17 weeks, and n=1 of 20 weeks; this material originated from spontaneous abortions), and adult GEJ tissue consisting out of continuous strips of tissue containing squamous esophageal epithelium, GEJ, and oxyntic stomach epithelium (n=3). Two strips came from surgical specimens without evidence of BE, with a neuroendocrine tumor and decompensated achalasia (male of 71 and female of 56 years old). The third patient had surgery to remove an EAC (male of 63 years old). All tissues were obtained from the gastroenterology and pathology departments of the Erasmus MC according to the FEDERA code of conduct [60]. The use of archival pathology specimens was authorised by an institutional review board (METC – Erasmus MC). Informed consent is not required for leftover diagnostic material from the Erasmus MC Tissue Bank.

Animal studies

For the *Hoxa13* mRNA expression analysis throughout the murine gastrointestinal (GI) tract, four *C57BL/6J* wildtype mice were used between three and five months of age. The GI-tract was divided into 1: esophagus; 2: stomach; 3: duodenum; 4: jejunum; 5: proximal ileum; 6: distal ileum; 7: cecum; 8: proximal colon; 9: distal colon, of which sections were opened and rinsed in PBS followed by storage in RNAlater at -80˚C (Supplementary fig. 1b). For determining which cells express *Hoxa13* in the GI-tract, tissues from a *C57BL/6J-Hoxa13-GFP* heterozygous mutant mouse model were employed, in

which GFP expression is driven by the endogenous mouse *Hoxa13* promotor through the creation of a fusion protein[61]. Mice were generally kept with 12:12 hours light - dark, the animal room temperature is between 20 and 24oC and the relative humidity is 55±10%.These tissues were taken out and embedded in O.C.T. Compound bought from Qiagen Inc. (Hilden, Germany) and frozen at -80 °C. Cryosections were made which were mounted in fluoroshield mounting medium with DAPI obtained from Abcam (Cambridge, UK). Subsequently, the GEJ and the distal GI-tract were analyzed directly for GFP expression using the Zeiss confocal laser scan microscope LSM 510. Additionally, immunohistochemistry staining was performed with anti-GFP antibody (#AB3080, Bio-Connect BV) (see below). These murine experiments were approved by the Ethical Committee for Animal Experiments of the Erasmus MC and were performed according to the guidelines of the same institution.

Immunohistochemistry

For immunohistochemistry, slides were blocked in 10% of normal goat serum, antigens were retrieved by boiling samples in citrate buffer (pH6), and samples were incubated overnight at 4°C with primary antibody. Dilutions and manufactures of primary antibodies are presented in the Supplementary table 4. After incubation with HRP-conjugated secondary antibody (Dako EnVision+System-HRR labeled Polymer Anti Mouse, Dako) endogenous peroxidase was blocked in 3% H_2O_2 and antibody binding was visualized by DAB staining. IHC analysis for HOXA13 was tried using antibodies ab106503 and ab26084, however these failed to show specificity and have since been discontinued by the companies offering them. H&E staining was performed by [62]. For H&E stainings de-parafinized 4 µM slides were incubated during 3 minutes in hematoxylin solution, followed by tap water washes and 15sec of incubation with eosin. For PAS staining, de-parafinized slides were incubated with 0.5% Periodic Acid solution for 10 min, followed by two ddH2O washes and incubation in Schiff's reagent (Sigma Aldrich) for 15 minutes and hematoxylin for 3 mins.

Multiplex Immunofluorescent Staining

Triplex staining for keratin 5, keratin7 and p63 was done by automated multiplex IF using the Ventana Benchmark Discovery (Ventana Medical Systems Inc.). In brief, following deparaffinization and heat-induced antigen retrieval with CC1 (#950-500, Ventana) for 64 minutes at 97°C, the tissue samples were incubated firstly with Keratin 5 antibodies for 32 minutes at 37˚C followed by detection with Ultramap anti-rabbit HRP (#760-4315, Ventana) for 12 minutes followed by visualization with Red610 for 8 minutes (#760-245, Ventana). Antibody denaturing was performed using CC2 (#950-123, Ventana) for 20 minutes at 100˚C. Secondly, Keratin 7 antibodies were incubated for 32 minutes at 37˚C followed by detection with Ultramap anti-rabbit HRP (#760-4315, Ventana) followed by visualization with FAM (#760-243, Ventana) for 4 minutes. Antibody denaturing was performed using CC2 (#950-123, Ventana) for 8 minutes at 100˚C. Thirdly, P63 antibodies were incubated for 32 minutes at 37 ˚C followed by detection with Ultramap anti-mouse HRP (#760-4313, Ventana) for 12 minutes

followed by visualization with Cy5 for 12 minutes (#760-238, Ventana). Slides were incubated in PBS with DAPI for 15 minutes and covered with anti-fading medium (DAKO, S3023).

RNA isolation

RNA was isolated using the NucleoSpin RNA isolation kit (Macherey Nagel, Düren, Germany). Biopsies and animal tissues were homogenized by the TissueRuptor obtained from Qiagen Inc. RNA concentrations were measured using a Nanodrop spectrophotometer and samples were stored in RNA storage solution (Sodium Citrate pH 6.4), bought from Ambion (Foster City, USA) and kept at -80 °C. RNA integrity was checked with 1% agarose gel-electrophoresis. FFPE material was deparaffinized with xylene and ethanol, lysed, digested with proteinase K, and RNA was isolated with the High Pure FFPET RNA isolation kit obtained from Roche (Basel, Switzerland). RNA isolation from de-differentiated KH2 mouse embryonic stem cells (mESCs) was done using a picopure RNA isolation kit (Thermo-Fisher Scientific, Waltham, USA). After RNA isolation all samples for RNA-Sequencing were tested on the Agilent 2100 Bioanalyzer to determine RNA integrity and quantity.

cDNA and qPCR

cDNA was made from 1µg RNA using Primescript RT Master Mix according to manufacturer's instructions (Takara, Otsu, Japan), for 15 minutes at 37 °C and 5 sec at 95 °C, and stored at -20 °C. qPCR was performed for 40 cycles in the iQ5 Real-Time PCR detection system that was obtained from BioRad Laboratories (Veenendaal, The Netherlands). For each reaction 10 µL cDNA template, 12.5 µL SYBR GreenER purchased from Invitrogen (Carlsbad, CA), and 2.5 µL 10 pM/µl primer were used. Reactions were performed in duplicate. Primers used are shown in Supplementary table 2 and were ordered at Sigma-Aldrich (Darmstadt, Germany). qPCR data were analyzed with Microsoft Excel using the ΔΔCt method. Reference genes used for PCRs on human materials were *RP2*, *β-ACTIN*, and *GAPDH*. Reference genes used for PCRs on mice materials were *Eef2*, *Rpl37*, and *Leng8*. Differences in expression were analyzed with a two sided Student's t-test using Prism 5.01, obtained from GraphPad Software (San Diego, USA). Values from individual samples were excluded if they deviated more than 2SD from the mean. Correlations between *HOTTIP* expression in the squamous esophagus and BE, and correlations between *HOTTIP* and *HOXA13* expression levels in the squamous esophagus and BE were tested using nonparametric Spearman correlations. This is depicted in graphs by connecting lines between datapoints, also indicating the paired nature of the specimens, *i.e*. they are derived from the same patient, used for this analysis.

In situ hybridization by RNAscope

RNAscope was performed according to the instructions of the manufacturer of the probes and the reagent kit (VS Reagent Kit 320600; Advanced Cell Diagnostics), on proteinase K (0.1%, 10 minutes

at 37 °C) treated paraffin sections (5 µm). Subsequently, slides were hybridized with the RNA probe from RNAscopeVS Hs-*HOXA13*, (art. #ACDA 400226), or the control probe also from RNAscopeVS Hs-*PPIB* (art. #ACDA 313901) [63]. *PPIB* (peptidylprolyl isomerase B) is a ubiquitously expressed gene. The RNAscope probe Hybridization in situ Multiplex was bought from Advanced Cell Diagnostics (Newark, USA). Pyloric metaplasia and Paneth cell metaplasia of the colon were quantified using FIJI, for which a macro was made (Supplementary method 1) [64]. For illustrations of RNA-Scope slides in the paper, background grey signal reduction was performed using Photoshop.

Analysis of GSE datasets

Expression profiles from clonogenic human gastro-intestinal stem cell cultures were obtained from Gene Expression Omnibus datasets GSE57584 [15] and GSE65013 [18]. In silico analyses were performed using the NCBI Gene Expression Omnibus (GEO) database. Analyses in the GEO database were performed by using the GEO2R tool (www.ncbi.nlm.nih.gov/geo/geo2r/), R 3.2.3., Biobase 2.30.0, GEOquery 2.40.0, limma 3.26.8[65]. The results were represented as a ^{2}log-fold change (^{2}log-FC). In Microsoft Excel, this ^{2}log-FC was converted to fold change (FC). For each ^{2}log-FC an empirical Bayes moderated t-statistic was calculated. *p*-values were corrected for multiple testing using the Benjamini & Hochberg false discovery rate method.

Analysis of single cell RNA seq datasets

BE and ESMG Single Cell Experiment Matrix from supplementary Data files 6 ll three Experiment Matrixes have been mapped to hg38 standard human genome ('*TxDb.Hsapiens.UCSC.hg38.knownGene*' R-package), normalized as Reads Per Kilobase per Million mapped reads (RPKM). Genes expressed in less than in 0.5% cells were filtered out. Low-quality cells were excluded based on: (1) the number of expressed genes - for 10x Single-Cell sequence data, cells expressing less than 400 or more than 7000 genes, for smartSeq data cells expressing less than 1000 and more than 7000 were removed. Different numbers were chosen due to the different sequencing depth. (2) Boxplot representation of all cells – outliers, i.e. cells mapping higher or lower than 1.5x the first or third quartile were removed. (3) Based on % of reads - cells were removed if there were more than 20% of reads mapping to mitochondrial or ribosomal genes. *HOXA13-related genes query: HOXA13*-positive cells from normal esophagus were selected with R. Genes that were expressed in at least 70% of these *HOXA13*-positive cells (20445) were analyzed for their expression in *HOXA13*-negative cells of normal esophagus as well as *HOXA13* negative and positive cells in BE tissue. For *T-SNE* plot, 638 cells were included (388 cells from Barrett's tissue, 250 cells from normal esophagus) and plotted based on their location of origin (colour) as well as *HOXA13* expression (open vs closed symbols).

Cell culture

All cells were cultured with penicillin (100u/mL) and streptomycin (100u/mL) and were regularly STR-verified and checked for mycoplasma by handing in samples prepared according to instructions at GATC Biotech (Konstanz, Germany). Primary human esophageal epithelial cells transformed with hTERT (EPC2-hTERT) (gift of K.K. Krishnadath) [66], were cultured with Keratinocyte SFM medium, supplemented with bovine pituitary extract at 50 µg/ml and EGF at 1 ng/ml (Thermo-Fisher Scientific). HET1A, the primary immortalized human squamous esophageal cell line Het-1A was a gift of J.W.P.M. van Baal (University Utrecht, The Netherlands). These cells were grown in EPM2 medium obtained from AthenaES, (Baltimore, Maryland, USA). The primary immortalized human BE cell line (BAR-T) was a gift of dr. J.W.P.M. van Baal who had, in turn, received them from dr. R.F. Souza (University of Texas Southwestern Medical Center, USA). These cells were grown in supplemented keratinocyte basal medium (KBM2), bought from Lonza (Basel, Switzerland), according to the method of Jaiswal *et al.*[67]. KH2 mESCs were a gift of J. Gribnau and maintained in DMEM with 10% FCS, Non-Essential Amino Acids, sodium pyruvate, LIF, and β-mercapto-ethanol (embryonic stem cell medium; Supplementary table 3). Dishes were coated with attachment factor protein solution (Thermo-Fisher Scientific). Irradiated mouse embryonic fibroblasts (3T3-Swiss albino cells (gift of J.W.P.M. van Baal), cultured in DMEM with 10% FCS, were used as feeder cells. HEK293T cells were cultured in DMEM with 10% FCS.

Generation of EPC2-hTERT *HOXA13* overexpression model

The human *HOXA13* gene including its single intron was amplified using Q5 polymerase from gDNA using primers (AgeI HoxA13 F; GGTGGTACCGGTGCCACCATGACAGCCTCCGTGCTCCT, and XbaI HoxA13 R; ACCACCTCTAGATTAACTAGTGGTTTTCAGTT) and cloned into pEN_TmiRc3 using AgeI and XbaI restriction sites, a gift from Iain Fraser (Addgene (Cambridge, USA) #25748) [68]. Subsequently, the *HOXA13* insert was transferred into pSLIK-Venus, using a Gateway reaction [69]. pSLIK-Venus was a gift from Iain Fraser (Addgene #25734) [68]. A similar plasmid but without the *HOXA13* insert served as control. Both plasmids were sequenced by LGC Genomics (Teddington, UK). Next, plasmids were packaged into lentiviral particles following transfection in HEK293T cells with third generation packaging plasmids. The supernatant was collected and ultracentrifuged. EPC2-hTERT cells were transduced with the virus and Fluorescence-Activated Cell Sorted (FACS) for YFP (pSLIK-Venus) positive cells on the BD FACSCantoTM II that was bought from BD Biosciences (San Jose, USA). These cells were grown and analyzed as a cell pool. *HOXA13* was induced by the addition of 1,25 µg/ml doxycycline to the culture medium. Overexpression was determined by qPCR according to scientific standards [50].

Generation of KH2 embryonic stem cells *HOXA13* overexpression model

The human *HOXA13* gene including its single intron was amplified using Q5 polymerase from gDNA using primers with an added N-terminal FLAG-tag sequence (GACTACAAAGACGATGACGACAAG) and Kozak sequence (GCCGCCACC; Supplementary table 3). Next, this PCR product was ligated into EcoRI digested pgk-ATG-frt (Addgene #20734) using Gibson Cloning (New England BioLabs Inc., Ipswich, USA). pgk-ATG-frt was a gift from Rudolf Jaenisch [70]. KH2 mESCs were passaged the day before the electroporation and four hours before electroporation medium was replaced. Approximately 1.5 10^7 KH2 cells were electroporated with 50 µg of pgk-ATG-frt-*HOXA13* and 25 µg of pCAGGS-FLPe-puro (Addgene #20733) [71]. Cells were electroporated in 4 mm cuvettes, with two consecutive pulses (400V/250 µF) using a Gene PulserXcell (Bio-Rad Laboratories). The next day 140 µg/ml Hygromycin B (Thermo-Fisher Scientific) was added for the selection of correctly targeted colonies. DNA from resistant colonies was isolated with the Kleargene XL blood DNA extraction kit (LGC, Teddington, UK) and analyzed by Q5 PCR using the following primers: PGK-F1 or PGK-F2 and T1E2-HygroR6 and T1E2-HygroR7 (Supplementary table 3). Correctly-targeted clones were checked for proper *HOXA13* induction by the addition of 1.25 µg/ml doxycycline to the culture medium for 3 days. Three *HOXA13* overexpression versus three control biological replicates were selected and used for experiments.

Differentiation of KH2 mouse embryonic stem cells

An optimized version of the Ogaki protocol was used [72]. Cells were plated on 50% confluent pre-cultured M15 cells, a mesoderm-derived feeder cell line [73] (gift of N. Hastie, University of Edinburgh, UK). Cells grew six days in differentiation medium consisting of ESC medium without LIF, with the addition of Activin-A, basic Fibroblast Growth Factor, CHIR, and Noggin (Supplementary table 3). *HOXA13*-expression was induced on day four using doxycycline at 1.25 µg/ml. On day six, cells were analyzed by FACS by double staining with 0.8 µg PE Rat Anti-Mouse CD184 (CXCR4) and 2.0 µg Anti-CD324 Alexa Fluor® 488 (E-Cadherin) at 4 °C for 45 minutes (Supplementary table 3). The cells were analysed with a BD FACSCantoTM II (BD Biosciences, USA). Data were analyzed with BD FACSDiva v8.0.1 software, which was obtained from BD Biosciences, and processed using Microsoft Excel. Example of analysis is provided in the Supplementary figure 11. Double-positive cells were sorted and cultured for another day with doxycycline at 1.25 µg/ml before harvesting and RNA isolation took place, using the picopure RNA isolation kit (Thermo-Fisher Scientific).

Generation of the BAR-T *HOXA13* knock-out model

Functional *HOXA13* was removed from BAR-T cells using CRISPR/Cas9-mediated gene editing. A *HOXA13* sgRNA targeting exon 1 was cloned into pTLCV2, by ligating two annealed oligonucleotides, *i.e.* Guide1sgRNA F and R (Supplementary table 4). TLCV2 was a gift from Adam Karpf (Addgene #87360)[74]. Following sequence verification, the pTLCV2-*HOXA13*sgRNA plasmid was packaged into lentiviral particles by cotransfection into HEK293T cells with pSPAX2 and pMD2.G, gifts from Didier

Trono (Addgene #12260 and #12259). The supernatant was harvested and ultracentrifuged after which BAR-T cells were transduced. Mixed populations of transduced cells were plated at very low confluence, single cell clones could subsequently be isolated using glass cloning cylinders and low melting point agarose from Sigma-Aldrich, followed by DNA isolation using the Kleargene kit, followed by sequence verification with primers TIL*HOXA13*R3 and Pre-*HOXA13*-FW2 flanking the sgRNA-site (Supplementary table 4). Three cell lines in which both alleles were affected by unique out-of-frame deletions were selected along with three control cell lines.

RNA-Sequencing

The EPC2-hTERT samples (n=8) were treated with the TruSeq Stranded mRNA Library Prep Kit. Sequencing took place according to the Illumina TruSeq v3 protocol on an Illumina HiSeq2500 sequencer. Sample preparation and sequencing was performed at the Erasmus MC. Reads of 50 base-pairs were generated and mapped against reference genome hg19 with Tophat (version 2.0.10). Expression was quantified using HTseq-count (0.6.1). Stranded libraries of the BAR-T (n=6), and both non-differentiated and differentiated KH2 mESCs (n=6 each) were prepared with the NEBNext RNA Ultra sample prep kit. Sequencing took place according to the Illumina NestSeq 500 protocol on an Illumina HiSeq2500 sequencer. Sample preparation and sequencing was performed at GenomeScan in Leiden, The Netherlands. Reads of 75 base-pairs were generated, mapped against reference genome hg19 or mm9 with Tophat (version 2.1.0), and quantified using HTSeq (version 0.6.1p1). Data were processed using R. version 3.2.5, [75] in combination with the module DeSeq2[76]. Generated FCs and p-values adjusted for multiple testing, i.e. *q*-values, were analyzed using Ingenuity Pathway Analysis (IPA) version 42012434, obtained from Qiagen Inc. (Hilden, Germany)[77]. We limited the number of genes analyzed to a maximum of 1000 by eliminating genes with a (relatively) low fold change if differentially expressed genes number was above 1000. The dataset cut-offs used were always a *q* value of 0.05, the fold change cut-off was set at: nondifferentiated KH2-mESCs, FC 2, 888 genes; differentiation of KH2-mESCs, FC 5, 924 genes; differentiated KH2-mESCs, FC not restricted, 665 genes; BAR-T, FC not restricted, 146 genes; EPC2-hTERT, FC 1.3, 990 genes. Activity scores are known in IPA as "*z*-scores" which represents the number of standard deviations from the mean of a normal distribution. For analysis and visualization of gene expression in the epidermal differentiation complex the raw counts from both models normalized to total reads were used. Genes for which one of both cell models had less than ten reads in the control or experimental samples were excluded. Overlap in multiple testing corrected differentially expressed genes in the BAR-T and EPC2-hTERT datasets was calculated as follows; the proportion of overexpressed genes in the EPC2-hTERT dataset was determined. Half of the differentially expressed genes in the BAR-T dataset would be expected to be regulated in the same direction if regulation would be random. This expected overlap if regulation was random, and the observed overlap, were used as input for an X^2 test. Information included in Supplementary Data 2 and 3 in the "known function" and "Detailed description" columns was obtained through non-systematic review and should not be considered as an exhaustive overview of the

literature. Association of expression of molecules in the distal GI-tract with their regulation by *HOXA13* expression was reviewed using the human proteome atlas and depicted in Supplementary Data 2 [32].

Acid and bile exposure

For assessment of *HOXA13* mRNA expression upon acid/bile exposure, EPC2-hTERT and HET-1A cells were treated for 30 minutes with cell culture medium adjusted to a pH of 7.0 or 4.0 using HCl. Cells were subsequently washed using PBS and given standard medium. Acid experiments were performed four times in duplo. Cells were separately exposed to medium with a bile acid mixture in concentrations of 0, 200, (and 400 for EPC2-hTERT) µmole/L for 30 minutes at a pH of 7.0. The bile acid mixture consisted of 25% deoxycholic acid, 45% glycocholic acid and 30% taurochenodeoxycholic acid. Cells were subsequently washed using PBS and given normal medium. Bile experiments were performed twice in duplicate. After 24 hours, the cells were harvested and RNA was isolated. Methods were derived from Bus *et al.*[78]. To assess the effect of bile/acid on expansion of cells, EPC2-hTERT cells transduced with *HOXA13* or control vector as described above were seeded in 96-well plate with at least 2 wells per condition. Next day, medium was replaced with 100 µL of bile/acid mixture in cell culture medium (50 µM of sodium glycocholatenhydrate, 50 µM taurochenodeoxycholic acid, pH=4.95). After incubation for 4 days, MTT test was performed as described below. Experiment was performed at least five times.

BAR-T spatial distribution experiments

These were performed with three biological replicate cell lines containing *HOXA13* knock-out and three control cell lines. 40.000 BAR-T cells were seeded in a 6 well plate and pictures were taken the second day after seeding. Per well three pictures were taken. These pictures were analyzed using FIJI, using the multipoint tool, an X and Y (pixel) coordinate table was generated [64]. The distance between each cell and its three closest neighbors was quantified using Microsoft Excel and analyzed by two sided student's t-test. The experiment was performed in three independent cell lines and repeated three times.

MTT assay

For assessment of cell growth of EPC2 and BAR-T cells, we performed a 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) assay [79]. We seeded 1000 cells per well in 96 well plates for each of the three wild-type and three *HOXA13* knock-out cell lines. Per condition at least 2 wells were used. On days one, three, five, and seven 10µl MTT at 5 µg/ml was added and incubated for three hours, the medium was removed, and the precipitate was dissolved in 100µl DMSO, which was incubated for five minutes under continuous shaking. For BAR-T cells, absorption was measured in a BioRad microplate reader Model 680 XR at 490 and 595 nm, the average absorption was used to process the data. For EPC2 cells it was measured with Tecan microplate reader Model Infinite 200 pro

at 565 nm with reference wavelength 670 nm. The experiment was repeated three times and a two sided Student's t-test was used to test for statistical significance.

3D culture EPC2-hTERT cells

3D culturing of EPC2-hTERT cells was performed as previously described [48]. 4000 EPC2-hTERT cells in culture medium were mixed 1:1 with ice-cold Matrigel basement membrane matrix (Corning BV), seeded in 50 µL drops in a 24 well plate for cell suspension, and incubated at 37℃ for 30 minutes. After solidification, 500 µL of culturing medium supplemented with 0.6 mM CaCl2 was added. Y27632 (10 µM) was included in medium only the first 24 h after seeding. Medium was refreshed and pictures were taken every three days. The morphology of spheroids (based on number of extrusions, or 'invadosomes') was counted on day 5. The area of the spheroids was measured with FIJI[62]. For H&E staining and IHC analysis of involucrin (see above), spheroids were fixed in 4% formaldehyde for 7 minutes on day 11, washed with PBS, put in 2% agarose, and embedded in paraffin, then 4 µM slices were sectioned. Quantification was based on the percentage of positive cells and the intensity of the staining (scores ranged from 0, 2 to 9).

Organotypic air-liquid interface culture

Plate inserts (Sigma-Aldrich, Germany) were covered with bovine collagen I (Thermo Fisher Scientific, USA). The fibroblast (3T3-Swiss albino) feeder layer was embedded within a collagen matrix and was allowed to mature for 7 days, after which time BAR-T *HOXA13* knock-out and control epithelial cells were seeded on top and allowed to grow to confluence for an additional 3 days as described [80.] Then the culture media level of the upper well was reduced, exposing the apical side of keratinocytes to the air, while maintaining liquid levels at the basolateral side. On day 15, cultures were harvested for histologic examination. 4 µM paraffin-embedded sections were deparaffinized, and staining with hematoxylin and eosin, PAS staining and immunohistochemistry for involucrin were performed.

Rat trachea *in vivo* tissue reconstitution model

500,000 parental BAR-T cells (derived from six independent clones) or *HOXA13* knock-out clones (three independent clones) in 30 µl of medium were sealed in the lumen of devitalized and denuded rat tracheas and implanted under the dorsal skin of NOD SCID gamma mice as described by Croagh et al.[81]. Mice were housed in microisolator cages with a 14 hr light/10hr dark cycle, standard chow and water ad libitum, and temperature and humidity maintained at 21±1°C and 50±10%, respectively. Mice were sacrificed after four or six weeks. Harvested rat tracheas were formalin fixed, decalcified, embedded in paraffin and sectioned. Staining with hematoxylin and eosin, alcian blue, PAS staining and immunohistochemistry using antibodies against human mitochondria, CK7, TFF3, CDX2, p63, CK5, and involucrin (a gift from Prof. Pritinder Kaur, Curtin University, Australia or #I9018-100UL from

Sigma-Aldrich) were performed (Supplementary fig. 9 and table 4). These murine experiments were approved by the Peter MacCallum Cancer Centre Animal Experimentation Ethics Committee and were performed according to the guidelines of the same institution.

Data availability statement

The RNA sequence data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus[82] and are accessible through GEO Series accession number GSE173170 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE173170) [83]. There are no restrictions regarding data availability. Supplemental figures, tables, and a method are included. Source data are provided with this paper and all relevant data are available from the authors.

Datasets used in the manuscript: RNA seq data GSE57584 [15], GSE65013 [18], single cell RNA data seq [21], GSE134520 [58], GSE81861 [59].

RESULTS

HOX cluster gene expression in the GI tract is collinear in men and mice

Investigating *HOX* gene mRNA expression in the murine and human gastrointestinal tract, we observed collinearity that is similar in adult humans and mice (Fig. 1a for human *HOXA*, Supplementary fig. 1 for all *HOX* genes and Supplementary fig. 2 for graphical presentation of the studied *HOX* clusters and the locations of biopsies taken along the human [n=3] and mouse [n=4] GI tract). The highest *HOX* gene cluster expression was observed in the colon, except for the *HOXC* cluster. For individual paralogues, there is a higher expression of 5' *HOXA/B* genes in the distal GI-tract from *HOXA5/B5* onward. Of all *HOXA* paralogues, expression of *HOXA13* was highest and restricted to the colon (Fig. 1a). *HOXA13* expression is regulated by LncRNA *HOTTIP*, which is located 5' to *HOXA13* [14]. Accordingly, *HOTTIP* and *HOXA13* share a similar expression pattern (Supplementary fig. 1a**)**. For *HOXD,* all paralogue genes have increased expression in the distal colon, while *HOXC* expression is mainly localized in the proximal and ileal regions. Thus, *HOX* gene expression is linked to positional identity in the mammalian gut, and collinearity is particularly strong for the *HOXA/B* paralogues.

Subsequently, we addressed the question as to whether GI HOX coding is already present at the GI stem cell stage, or is established only upon the formation of differentiated derivatives. For this, we used publicly available data published by Wang *et al*. which contains the mRNA expression of human stem cells isolated from the GI-tract and either cultured as stem cells or differentiated in an air-liquid interface (ALI) [15]. Analysis of this data shows that *HOX* gene expression patterns in stem cell and ALI cultures are similar. *HOXA* and *B* cluster genes have a significantly higher expression in the large intestine as compared to the small intestine, in particular 5' *HOXA* genes including *HOXA13* (Fig. 1d for *HOXA* genes, for clusters *HOXB*, *C*, and *D* see Supplementary fig. 3). No clear regulation of the *HOXC* or *D* clusters is seen in this in dataset, with exception of an upregulation of *HOXC10* in the large intestine.

Hence, *HOX* coding is an inherent feature of the location-specific stem cell and is maintained in its derivatives.

HOXA13 in BE, GI heterotopias and GI cancers

As positional phenotype is linked to HOX status in physiology, we subsequently characterized *HOX* mRNA expression in several metaplastic tissues known to assume the morphological phenotype of other intestinal locations, as well as their sequelae. BE shows upregulation of *HOXA10*, *11*, and *13*, and *HOXB 6*, *7*, *9*, and *13* mRNA by qPCR when compared to the normal squamous esophagus (Fig. 1b and Supplementary fig. 4), which closely resembles colonic *HOXA* and *B* expression patterns. High 5' *HOXA* gene expression is also present in columnar-lined esophagus without goblet cells (CLE; a BE-related condition), esophageal adenocarcinoma (EAC), and IM of the stomach (Fig. 1b, c). In accordance with a regulatory role for *HOTTIP* on *HOXA13* expression, we find that *HOTTIP* is also overexpressed in BE, and correlates with *HOXA13* expression patterns (Supplementary fig. 5a, b, c). *HOTAIR*, a lncRNA located in the *HOXC* cluster and associated with chromatin reprogramming in cancer progression[16] is upregulated as well (Supplementary fig. 5d, e, f) [14]. We concluded that BE, EAC and various metaplasias with caudal histo-morphological characteristics have *HOXA* and *HOXB* expression patterns typical of the caudal GI-tract, with upregulation of *HOXA13* expression being the prominent feature. Heterotopias, namely the gastric inlet patch in the proximal esophagus and heterotopia of the Meckel's diverticulum, are tissues which have a physiological appearance, but are normally found in a different location. Both these heterotopias are characterized by abundant *HOXA13* mRNA expression (Fig. 1c), although intriguingly the direction of epithelial metaplasia for Meckel's diverticulum is of an anterior rather than posterior phenotype, indicating an exception from the pattern in case of Meckel's diverticulum. One of the existing hypotheses on the cell of origin of BE states that BE may arise from cells with progenitor properties that are able to give rise to a variety of cell types [17]. To investigate whether aberrant HOX gene expression in BE is established at the level of the epithelium-specific stem cell, we interrogated the publically available data of Yamamoto *et al* [15, 18]. *HOX* gene expression patterns in squamous esophageal and BE stem cells as well as their respective ALI-differentiated derivatives were retrieved. HOX gene expression in stem cell cultures from these locations is similar to their ALI differentiated counterparts (Fig. 1e, Supplementary Fig. 3). In BE stem cells, an upregulation of 5' *HOXA* genes (Fig. 1e) as well as *HOXB6*, *7*, *13*, and *HOXC10* is seen (Supplementary fig. 3), reaching levels similar to those observed in the colon. Thus, alternative HOX coding associated with BE is established at the epithelium-specific stem cell level and is maintained in derivatives of the stem cells involved.

According to the collinearity theory, a paralogue group 13 member is more likely to confer the distal characteristics seen in BE as compared to more anterior paralogue group members [19]. Of the paralogue group 13 members, *HOXA13* and *HOXB13* are overexpressed in BE, with *HOXA13* showing much higher expression compared to *HOXB13* in BE, EAC, and IM of the stomach (Supplementary fig. 4). Therefore, while HOX genes such as *HOXA11, B6, B9*, and *B13* are also potentially interesting

candidates, here we chose to focus on the *HOXA13* gene for further in-depth analysis of different metaplastic tissues. As immunohistochemistry for *HOXA13* was unsuccessful, (two anti-HOXA13 antibodies were tested, but lacked specificity) we resorted to *in situ* hybridization (ISH) for *HOXA13* to further confirm the observed atypical expression of this gene in different tissues (examples shown in Fig. 1f. Metaplasia is found throughout the GI-tract. While BE and IM acquire a more distal phenotype, distally located colonic pyloric and Paneth cell metaplasia, related to inflammatory bowel disease, acquire a more rostral phenotype [20]. Accordingly, downregulation of *HOXA13* expression (corrected for *PPIB* expression as a reference gene) relative to adjacent non-metaplastic tissue, was seen for these tissues (Fig. 1g, again supporting a role for *HOXA13* in positional identity.

Binary regulation of *HOXA13* expression

To study in more detail which of the cells in the healthy GI-tract express Hoxa13, we employed a murine model in which the endogenous mouse *Hoxa13* promoter drives the expression of a Hoxa13-GFP fusion protein. Within the epithelial compartment, the proximal expression border is located at the transition from the distal to the proximal colon as can be seen from fluorescent images and images of anti-GFP IHC staining (Fig. 2a, b and Supplementary fig. 6 a-d for bigger overview images). This proximal expression border seems to be crypt-clonal, with some crypts expressing Hoxa13 and others not (see arrows in Fig. 2b and close-up in Fig. 2c). Functional consequences of this clonality are unknown and, while beyond the scope of the present manuscript, present an interesting biological question. The distal Hoxa13-GFP expression is limited by the anal squamocolumnar junction (SCJ; Supplementary fig. 6e, please note this cannot be appreciated in Fig. 2a, as this part was damaged for this mouse). To investigate whether these local gradients of Hoxa13 expression are also present in humans, *HOXA13* mRNA expression was assessed by qPCR in an additional set of biopsies taken from different colonic locations. Cecal biopsies are *HOXA13* negative, while *HOXA13* expression increases from the ileocecal valve to the distal transverse colon, demonstrating a similar expression pattern as observed in the mouse (Fig. 2d).

In addition to a Hoxa13 gradient along the GI tract, epithelial Hoxa13-GFP expression is also tightly regulated along the baso-luminal axis of individual crypts. Proximally, only apical expression is seen, while distally Hoxa13-GFP is expressed along the entire baso-luminal axis of the crypts (Fig. 2e). In addition, mesenchymal expression is observed in the cells just beneath the epithelium in the proximal colon (Fig. 2e). Within the cell, the strongest signal is co-localized with nuclei, as expected, but cytoplasmic staining is also seen which can be explained by ribosomal synthesis (Fig. 2e).

We concluded that spatial regulation of HOXA13 expression is very precise, robust and colon-specific, raising questions as to the cellular origin of the HOXA13 expression observed in BE.

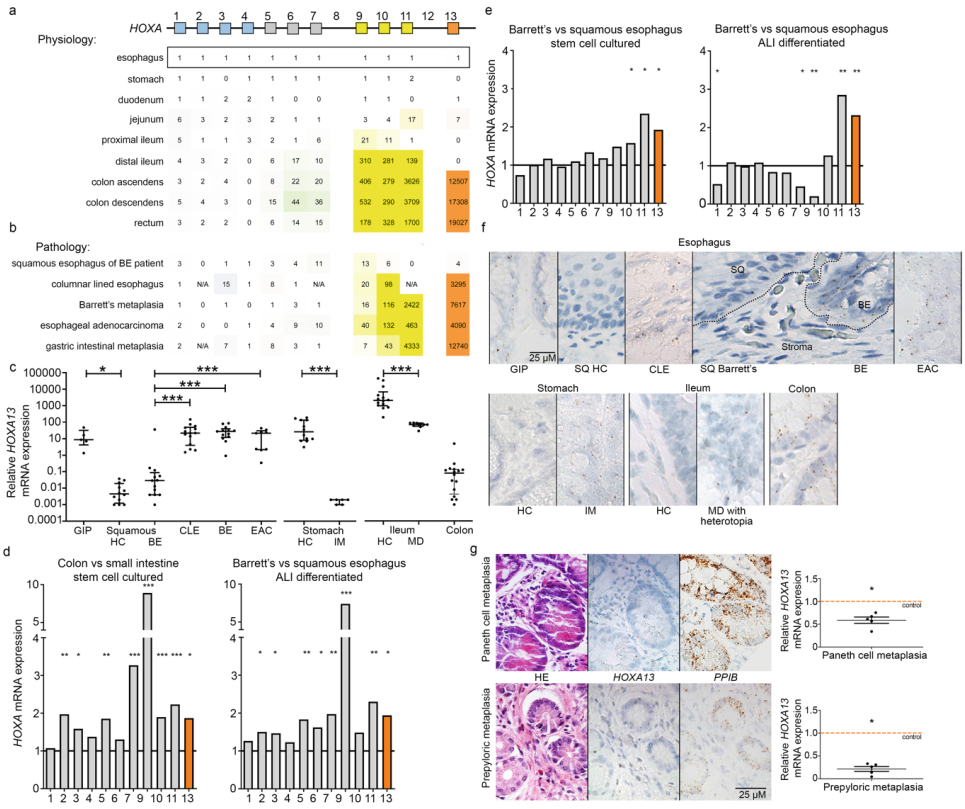


Fig. 1 | HOXA cluster gene expression shows collinearity along the adult gastrointestinal tract but is deregulated in Barrett's esophagus (BE), various metaplasias and esophageal adenocarcinoma (EAC). a) HOXA cluster genes are collinearly expressed along the gastro-intestinal (GI)-tract of adult humans (n=3). Numbers represent mRNA fold changes relative to the esophagus and thus can be compared within each HOXA paralogue member but not between them. b) HOXA cluster gene expression in the squamous esophagus of BE patients (n=13), columnar lined esophagus (CLE) (n=14), BE (n=13), EAC (n=12), and gastric intestinal metaplasia (IM) (n=12) is characterized by an upregulation of 5' HOXA genes. Numbers represent mRNA fold changes relative to the esophagus of healthy individuals. c) HOXA13 expression quantified by qPCR in BE, CLE, IM of the stomach, and heterotopias along the GI tract with their corresponding physiological epithelia. Squamous epithelium (SQ) Barrett's and BE are derived from the same person (n=13). Gastric inlet patch (GIP; n=5); healthy control (HC) squamous esophagus (n=12); CLE (n=14); BE (n=13); EAC (n=12); stomach (n=14); gastric IM (n=12); ileum (n=6); Meckel's diverticulum (MD) with gastric heterotopia (n=14), and colon (n=9). Median±IQR, *p<0.05; **p<0.01; ***p<0.001. For esophagus, Kruskal-Wallis test with Dunn's multiple comparisons test (SQ healthy vs. GIP, p=0.015; SQ healthy vs. CLE, p<0.0001; SQ healthy vs. BE, p<0.0001; SQ healthy vs. EAC, p=0.0009). For stomach and ileum, Mann-Whitney test (two-tailed), p <0.0001. d) HOXA cluster genes, in particular 5' HOXA genes including HOXA13, have a higher expression in the large intestine (n=3 in technical duplicate) compared to the small intestine (n=3 in technical duplicate), in both stem cells (left panel) and differentiated cells (right panel). Normalization was performed by setting mRNA expression to 1 for the small intestine. *p<0.05; **p<0.01; ***p<0.001. This figure includes no estimate of variance as the empirical Bayes-moderated two-sided t-statistic was used which does not generate a standard error. e) 5' HOXA cluster gene expression in BE is higher compared to the squamous esophagus in stem cell and air-liquid interface (ALI) differentiated cultures. n=12 (BE) versus n=2 (squamous esophagus) in technical duplicates are depicted for stem cell cultures and n=1 each for ALI differentiated samples in technical duplicates. Normalization was performed by setting mRNA expression to squamous esophagus. *p<0.05; **p<0.01; ***p<0.001. This figure includes no estimate of variance as the empirical Bayes-moderated two-sided t-statistic was used which does not generate

a standard error. f) Deregulation of HOXA13 expression in gastrointestinal tract pathology as evaluated with RNA in situ hybridization in clinical samples. HOXA13 is upregulated in IM and heterotopia and downregulated in pyloric and Paneth cell metaplasia in the colon. One sample of each tissue type was analyzed. g) Downregulation of HOXA13 expression (corrected for peptidylprolyl isomerase B - PPIB expression) relative to adjacent non-metaplastic tissue, was observed for Paneth cell metaplasia (n=5; FC 0.59; p=0.0003) and pyloric metaplasia (n=5; FC 0.22; p=0.0001) (lower panels). Unpaired t-test (two-tailed). Representative images of hematoxylin and eosin (HE) staining, HOXA13 RNA-scope, and PPIB reference gene RNA-scope of Paneth cell metaplasia (from the colon) present in two glands to the bottom right (upper panels) and pyloric metaplasia (from the colon) in the top left two glands (middle panels) are shown.

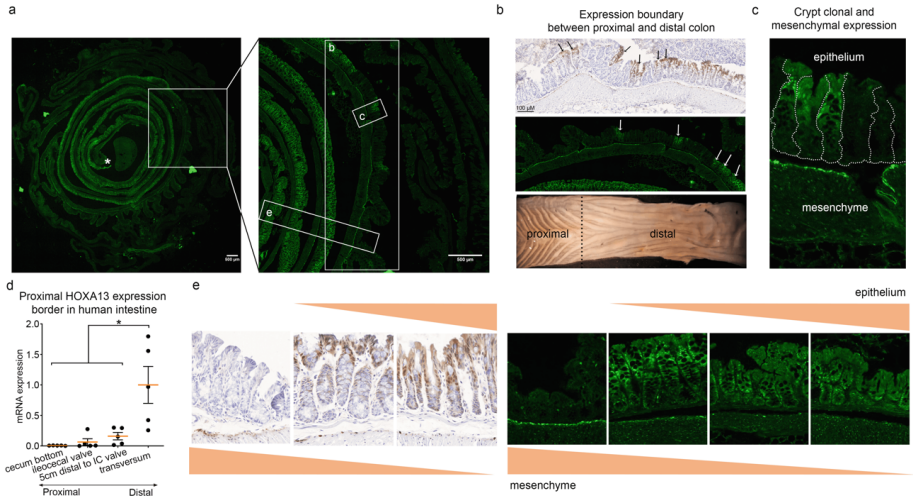


Fig. 2 | Murine and human *HOXA13* expression is subject to strict spatial control in the colon. a) A representative example from 3 mice of a "Swiss roll" configuration of the large intestine of the Hoxa13-GFP heterozygous mouse model. An asterisk indicates the most distal portion of the epithelium. Magnification of the insets are shown in panels b, c and e. b) The proximal border of physiological Hoxa13 expression in the adult mouse is patchy and located between the proximal and distal colon, indicated by a black dashed line in the bottom panel (macroscopic image of an opened mouse colon). Representative images of anti-GFP IHC and confocal microscopy are shown. Arrows indicate crypts that are positive for Hoxa13 among Hoxa13-negative crypts. c) The Hoxa13 expression is crypt clonal. This is observed for n=1. d) In adult humans the cecum bottom is negative for *HOXA13* while positivity increases distally (n=5 independent sampes). Mean±SEM, *p<0.05, repeated measures ANOVA with Holm-Šídák's multiple comparisons test, p=0.001. *HOXA13* mRNA levels were normalized to levels in the transverse colon. e) Hoxa13 expression is tightly regulated along the baso-luminal axis. Distally, Hoxa13 is expressed along the entire baso-luminal axis of the colonic crypts, proximally only expression at the luminal side is seen. In addition, a mesenchymal expression is observed in the cells just beneath the epithelium, predominantly in the proximal colon. Anti-GFP IHC and confocal images are shown.

Individual Hoxa13/*HOXA13*-positive cells in the upper GI tract

No significant expression of *HOXA13* mRNA was seen in the squamous esophagus of BE patients by qPCR (Fig. 1c, e), suggesting that GERD does not provoke *HOXA13* expression *per se.* Indeed, when two primary immortalized squamous esophageal cell lines (EPC2-hTERT and HET-1A) were exposed to either bile or acid, only minor effects on *HOXA13* expression were observed (two to fourfold from a low baseline expression; Fig. 3a, b), more in agreement with cells having a relatively high *HOXA13*

expression showing better survival of the treatment rather than upregulation of expression *per se*. This was confirmed by analysis of the publicly available single cell RNAseq database recently published by Owen *et al*[21]. Results at single cell level demonstrate the presence of a small population of *HOXA13*-positive cells in the normal squamous esophagus of BE patients (8%). In BE tissue, the percentage of these HOXA1ositive cells increase to 30%, but their individual *HOXA13* mRNA levels are not increased as compared to HOXA13-expressing cells of the normal esophagus (Fig. 3c, d). Similarly, the number of *HOXA13*-positive cells, but not *HOXA13* expression *per* cell, is increased in IM of the stomach, early gastric cancer, and colorectal cancer (Fig. 3c, d). Thus, we further investigated Hoxa13 at the cellular level in our samples. Although *Hoxa13* mRNA expression was detectable in only one of four mice in the upper GI-tract by q-PCR (Supplementary fig. 2), detailed inspection of specimens involved did identify single Hoxa13-positive cells in the stomach of Hoxa13-GFP mice by immunohistochemistry. Such signal was present at the basolateral side along the stomach starting from the GEJ, but not seen in the squamous cells along the esophagus, nor the stroma (Fig. 4a and Supplementary fig. 7). This is of particular interest as the GEJ has been suggested as a place of origin of BE[17]. A littermate negative for Hoxa13-GFP showed no positivity (Supplemenary fig. 6d, Supplementary fig. 7d). Subsequently, we employed ISH for *HOXA13* on surgical samples from the human GEJ of three adult patients to analyze the presence of *HOXA13*-expressing cells in the human upper GI tract. In all three specimens, the GE junction area contained a clear positive signal for *HOXA13* mRNA, some signal was seen in cells of the proximal stomach, while signal was even lower in the squamous epithelium and stroma (Fig. 4b and Supplementary fig. 8 a, b). Esophageal submucosal glands (ESMG) [21] were present in one sample and were *HOXA13* positive (Fig. 4c). Interestingly, ESMG were also highly positive for *KRT7*$^+$*KR5*$^+$*TP63*$^+$ cells (previously postulated as the cell of origin BE origin in GEJ [22]) although unlike HOXA13$^+$ cells, *KRT7*$^+$*KR5*$^+$*TP63*$^+$ triple positive cells were not identified in the stomach (Fig. 4d). (Of note, this is showed for one sample and we were unable to assess possible *HOXA13* co-expression with these triple positive cells due to absence of specific HOXA13 antibodies). We also studied *HOXA13* expression in the GEJ of three spontaneously aborted human fetuses of 17-20 weeks of age, a gestation period characterized by transition of the esophageal epithelium from columnar to a squamous phenotype. We observed high and specific *HOXA13* expression at the gastric cardia, while more distal stomach and esophageal epithelium were less positive (Fig. 4e, Supplementary fig. 8c, d). These data imply that *HOXA13*-positive cells are present in the human embryonic esophagus during the epithelial transition period, reduced in adult squamous esophagus, and increase again in BE. Thus, the epithelium of both the human and mouse adult upper GI tract, in particular the GEJ and ESMGs for human, is characterized by the presence of a subpopulation of *HOXA13*/Hoxa13-positive cells in an otherwise *HOXA13*/Hoxa13 negative surrounding.

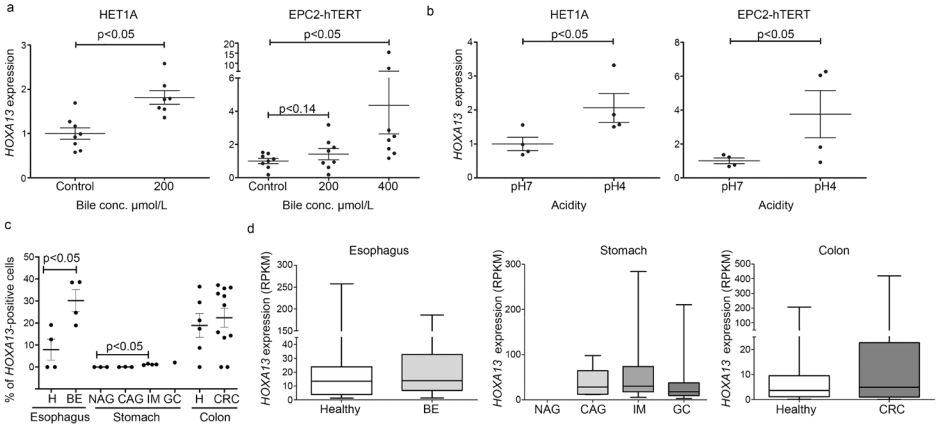


Fig. 3 | Number of *HOXA13+* cells rather than cellular expression levels are associated with metaplasia. Exposure to bile (at pH 7) (a) or to acid (b), in two *in vitro* model systems of gastroesophageal reflux disease (GERD), marginally induces the expression of *HOXA13* from low baseline expression levels in two primary immortalized squamous esophageal cell lines. Error bars represent the 95% CI of the mean. n=4 independent experiments. For Het1a, two-tailed t-test was used (p=0.0096 in a, p=0.0322 in b), for EPC2-hTERT Dunn's multiple comparisons test in a (control vs. 200 µM, p>0.99; control vs. 400 µM, p=0.0112; 200 µM vs. 400 µM, p=0.14), one-tailed t-test in b, p=0.0486. (c) The number of *HOXA13*-postive cells are increased in Barrett's esophagus (BE) and intestinal metaplasia (IM) as compared to normal esophagus or stomach tissue. Healthy and BE esophageal samples are derived from the same patients. Mean±SEM are shown. n=4 individuals for esophagus, n=3 for NAG, CAG, n=4 for IM, n=1 for GC, n=6 for healthy colon, n=11 for CRC. Graphs are based on the analysis of single cell RNA data seq [21], GSE134520 [58], GSE81861[59]. Two-tailed t-test was used for esophagus (p=0.0443) and colon (p=0.6808), one-way analysis of variance for stomach with Dunnett's Multiple Comparison Test (p< 0.0001). (d) *HOXA13* expression level per cell is unchanged. Expression level presented in *HOXA13*$^+$ cells only. H – healthy, BE - Barret's esophagus, NAG – non-atrophic gastritis, CAG - chronic atrophic gastritis, IM - intestinal metaplasia of stomach, GC - early gastric cancer, and CRC - colorectal cancer. n=37 for H, n=132 for BE, NA for NAG, n=5 for CAG, n=163 for IM, n=69 for GC, n=37 for healthy colon, n=87 for CRC of single cells from the individuals mentioned in (c). For GC, statistics are not presented as data per patient was not provided. Boxplots with middle line is the median, the lower and upper hinges correspond to 25th to 75th percentiles, and whiskers representing min-max values.

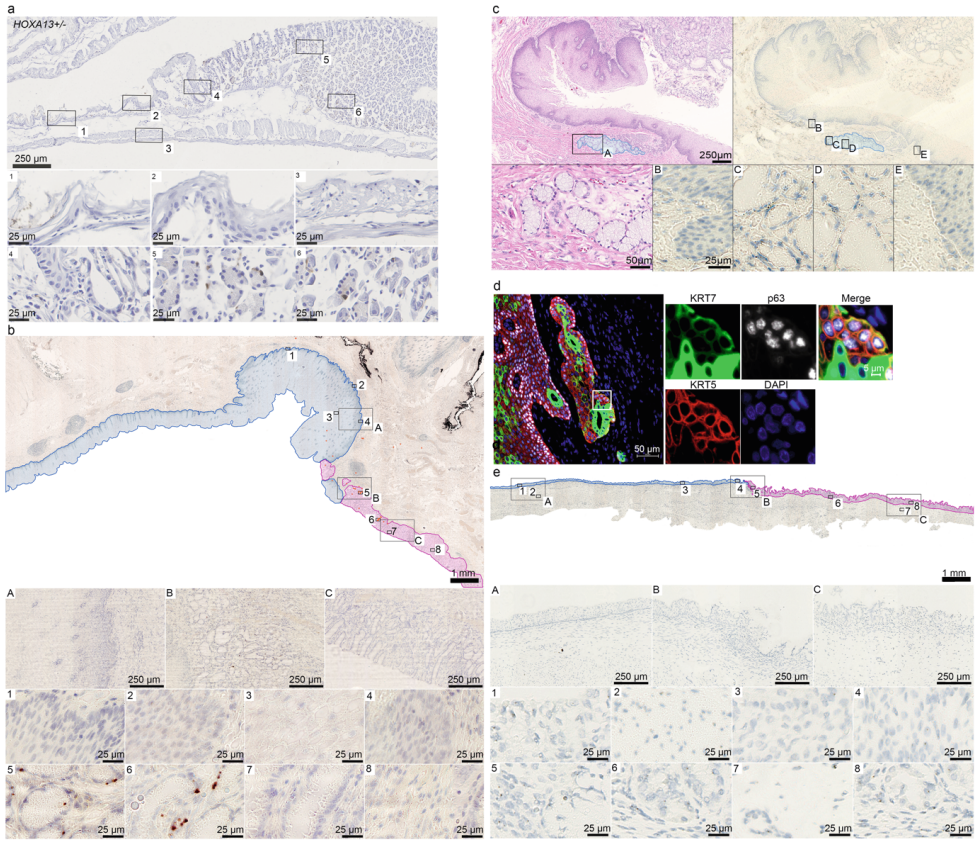


Fig. 4 | HOXA13 expression in the upper GI tract. a) Representative example of anti-GFP immunohistochemistry of a Hoxa13-GFP heterozygous mouse with gastroesophageal junctions (GEJ) (SQ; squamous epithelium, ST; stomach) (n=3). Hoxa13 is expressed in single cells of the stomach starting from the GEJ (5-6) and absent in the esophagus and stroma (1-4). b) HOXA13 expression as measured by RNA ISH in a representative example from n=3 with similar results of an adult human GEJ with magnification panel of: A – esophagus, B – GEJ area, C – proximal stomach. Orange circles indicate the positive signal in the overview image. c) HOXA13 expression as measured by RNA ISH in human esophageal submucosal gland (ESMG) with magnification panel of: A –H&E, C,D – ESMG, B, E – squamous esophagus, n=1. d) Keratin 7 (KRT7), keratin 5 (KRT5) and p63 triple positive cells are found in the ESMG, n=1 e) Overview of a representative example of a 17-week old fetus GEJ: A) Stratified esophageal epithelium of the distal esophagus (blue), B) GEJ area, C) gastric epithelium of the proximal stomach (pink), n=3.

HOXA13 affects differentiation potential and posteriorizes

Having established that individual HOXA13-positive cells reside in the physiological upper GI tract and are enhanced in BE tissue, we next set out to investigate the potential role of this population of cells in the etiology of BE. To this end, we further analyzed the single cell RNA-seq [21] data set mentioned above. In this study, the GEJ was not sampled for analysis. However, the 8% of cells of the normal esophagus that express *HOXA13* exhibit transcriptional overlap with cells derived from BE tissue as seen from the t-Distributed Stochastic Neighbour Embedding (t-SNE) plot (Fig. 5a). Gene expression analysis indicates that these cells are derived from ESMG (Fig. 5b and Supplementary Data 1) [21, 23] Specifically, and in contrast to the *HOXA13*-negative cell population, >70% of the *HOXA13*-positive cells from the normal squamous mucosa are positive for submucosal markers *LEFTY1* and *OLFM4*, designated ESMG markers, which have also been described as markers of BE progenitor cells [21]. Additionally, *HOXA13*-positive cells express mucosal markers *TFF3*, *Lyz* and *SOX9*, as well as columnar and BE markers *TFF1, KRT7, VIL1, MUC5B, MUC3A, MUC13, MUC1,* and *CEACAM5*, while being negative for keratinization marker *IVL* and basal epithelial cell marker *p63* (Fig. 5b and Supplementary Data 1 for the list of genes enriched in *HOXA13*-positive cells). In BE, the percentage of cells positive for these columnar and ductal markers increase also in the *HOXA13*-negative population, suggesting either that upon differentiation some of these cells might lose *HOXA13* expression, or that there is more than one population giving rise to BE tissue. This would be in line with mouse data, as the murine esophagus lacks ESMGs and Hoxa13-positive cells. Interestingly, although rare in this dataset, within the *TFF3*$^{+}$ population four cells were identified to be triple positive for *KRT14* (a gene pair with *KRT5*), *TP63* and *KRT7*[21] but these were not positive for *HOXA13.*

The cell of origin with respect to formation of the BE segment should be able to generate a variety of differentiated cell types that exhibit colonic, gastric, pancreatic acinar or other phenotypes [24, 25]. In chick embryos, *HOXA13* regulates regionalization after 1.5 days of development, showing the involvement of *HOXA13* in early differentiation, consistent with an effect of this gene on cellular phenotype in such pluripotent progenitor cells [26]. In an effort to experimentally test the influence of *HOXA13* on cell fate, we generated *HOXA13*-inducible pluripotent mouse embryonic stem cells (mESCs). These pluripotent mESCs can be efficiently differentiated to multipotent definitive endoderm, as determined by membrane expression of CXCR4 and E-cadherin (Fig. 5c). This was further confirmed by RNAseq, showing a strong upregulation of definitive endoderm markers such as *Sox17* and *Foxa1* in these differentiated cells, while pluripotency markers such as *Nanog* are downregulated (see Supplementary table 1). Using ingenuity pathway analysis (IPA) to further analyze differently expressed genes, a positive association was found with “differentiation of embryonic cells” (z=1.82, p=6.38·10^{-15}). Intriguingly, when *HOXA13* expression was induced, cells differentiated less effectively towards definitive endoderm as determined by CXCR4^{+}/E-cadherin^{+} expression and morphological assessment (Fig. 5c and d). Consistent with a reduced unilinear differentiation, clones expressing HOXA13 showed greater expansion (Fig. 5d).

We next contrasted the transcriptome of non-differentiated, pluripotent *HOXA13*-overexpressing and control cultures to identify potential molecular mediators of the *HOXA13* effects observed. Results of

IPA analysis of differential gene expression are broadly consistent with *HOXA13* conferring a pluripotent phenotype. Specifically, forced *HOXA13* expression results in upregulation of the "role of *Nanog* in mammalian embryonic cell pluripotency" category (z=1.34, p=2.32·10^{-3}), an effect that involves *Sox2*, *Nanog*, *Tbx3*, *Hesx-1*, and *Dppa*-1 amongst others [27, 28] (See Table 1 for more details/results, fold changes, and q-values with regard to this experiment). *HOXA13* expression also appears to downregulate Wnt signaling, possibly through BMP signaling [29]. Wnt signaling is known to promote mesoendodermal differentiation [30], these results are consistent with *HOXA13*-mediated downregulation of Wnt signaling during axial elongation [31]. Thus, the transcriptional profile provoked by *HOXA13* is consistent with maintaining a relatively pluripotent phenotype which in turn may increase compartment expansion.

HOXA13 expression does not block endodermal differentiation of mESC cells completely, suggesting that a role for *HOXA13* in this compartment is still relevant. Definitive endoderm is a feature of the entire GI tract epithelium, and does not distinguish upper and lower GI epithelium *per se*. To investigate the role of *HOXA13* in this cell compartment and test our prediction that *HOXA13* expression would predispose endoderm to acquire distal phenotypes, we sorted CXCR4^{+}/E-cadherin^{+} cells of *HOXA13* positive and negative cultures and contrasted their mRNA expression. *HOXA13* upregulates gene expression associated with determination of morphology in definitive endoderm cells. In IPA analysis, "actin cytoskeleton signaling" was most activated (z=3.00, p=3.74·10^{-2}). "RhoA signaling", which stimulates actin polymerization, (z=2.12, p=1.12·10^{-2}) was also stimulated. *HOXA13* supports distal epithelial functions with upregulation of microvillus-associated genes, *Ezr* and *Vill*, keratins, *Krt19* and *Krt20*, tetraspan network genes, *Igsf8*, and exocrine function associated genes such as *Gcnt3*, normally expressed in the distal GI-tract epithelium [32]. In addition, more transcripts of "Cell proliferation of carcinoma cell line" (z=1.13, p=1.5·10^{-6}) and "Neoplasia of cancer cells" (z=1.13, p=2.22·10^{-4}) categories, such as *Fgfr2* and *Nek2*, were detected. Thus, forced *HOXA13* expression during endodermal differentiation supports caudal epithelial functions and proliferative potential (see Table 1 for fold changes and q-values; see Supplementary Data 2 for additional relevant molecules).

Together, these data are in apparent agreement with *HOXA13*-expressing cells displaying a progenitor phenotype and having a competitive advantage, while simultaneously driving the acquisition of a more distal columnar phenotype once committed to differentiation (see Fig. 5e).

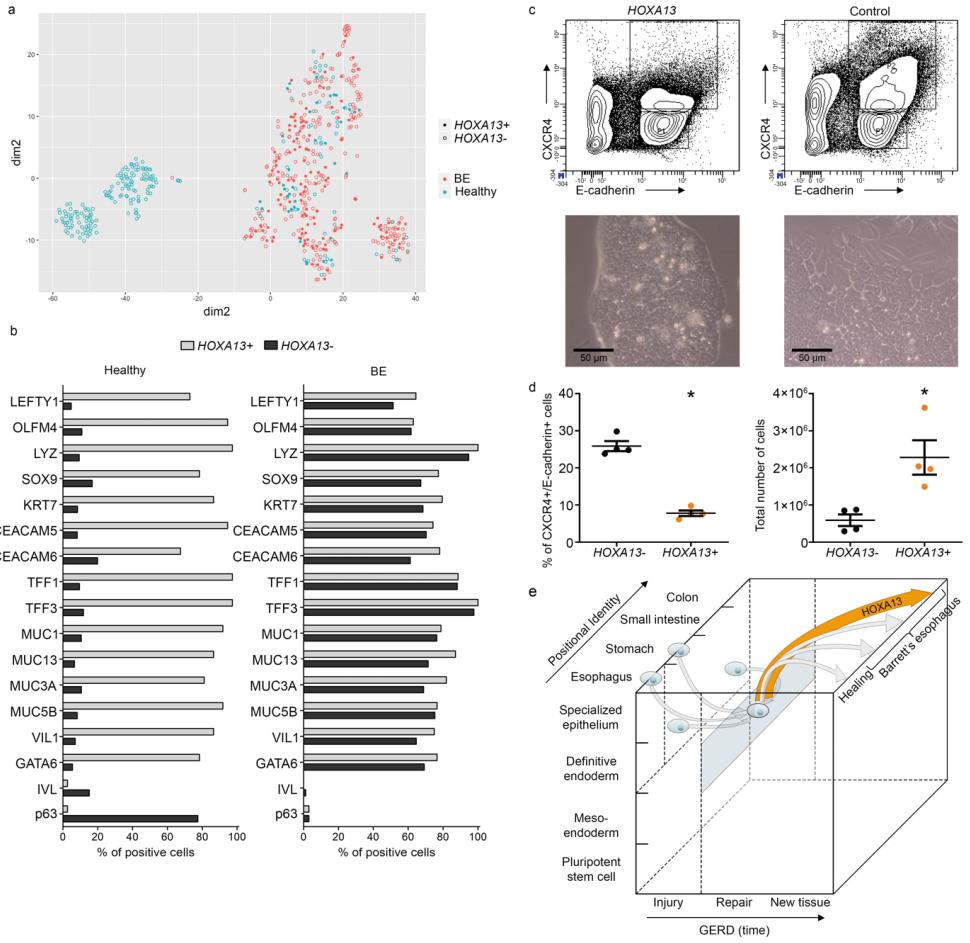


Fig. 5 | HOXA13 cellular expression modulates cell fate. a) *HOXA13*$^+$ cells of normal esophagus cluster together with BE cells in t-distributed Stochastic Neighbor Embedding (*T-SNE)* plot based on single cell RNA expression profiling[21] b) Analysis of single cell RNA seq data revealed that in contrast to *HOXA13*$^-$ cells, *HOXA13*$^+$ cells express submucosal gland markers, Barrett's esophagus (BE) markers and have decreased expression of squamous markers (p63, IVL) in healthy esophagus. This difference is not observed in BE. n=846 of HOXA13$^-$ cells in healthy esophagus, n=37 of *HOXA13*$^+$ in healthy esophagus, n=263 HOXA13$^-$ cells in BE, n=132 of *HOXA13*$^+$ cells in BE. c) *HOXA13*-overexpressing definitive endoderm is relatively resistant to terminal differentiation. Mouse embryonic stem cells (mESC) cells with and without forced *HOXA13* expression were differentiated from pluripotent stem cells to definitive endoderm. The percentage of differentiated definitive endoderm cells, defined as CXCR4$^+$/E-cadherin$^+$ cells, was analyzed by FACS analysis (upper panels). Lower panels (representative light microscopy images) show morphological differences in cultures of *HOXA13* overexpressing and wildtype mESCs upon differentiation to definitive endoderm, which induces a flattening of cell layers, with larger and irregular shaped cells. d) Quantification of FACS analysis results indicates that the percentage of CXCR4$^+$/E-cadherin$^+$ cells is decreased in *HOXA13*-overexpressing cell cultures under differentiation conditions (p<0.0001). *HOXA13*-expressing cells expand faster during the differentiation process compared to control cells (total number of cells increased) (p=0.0135). Mean±SEM, ***p<0.001, n=4 independent experiments, t-test (two-tailed).e) Model of cellular identity in BE development. The X-axis represents time (hypothetical units) following exposure to GERD-inducing agents. Y-axis shows differentiation during embryology and pathology. Z-axis indicates the positional identity of GI-tract tissues. Several theories exist regarding the cell of origin of BE: they may be fully differentiated esophageal or stomach cells, or less differentiated cells within these organs

(depicted by the 4 cells on the Y-Z plane). Irrespective of its location or differentiation state, this cell or origin might lose its correct positional identity or maintain its aberrant positional identity and resembles a definitive endoderm like cell. This is visualized by the blue rectangle harboring the cell with the thicker blue contour. For the model of cellular identity in BE, our data suggest that *HOXA13* expressing clones in the GEJ, depicted in orange, may outcompete clones with another positional identity, providing an explanation for the distal phenotype observed in BE.

Table 1 | Fold changes and q-values for the mRNAs mentioned in the results section of the main text pertaining to cell culture models analyzed by RNA-Seq.

Gene name	Fold change	*q*-value
Forced *HOXA13* in mESC confers a relative competitive advantage in multipotent cell cultures through upregulation of *Nanog* signaling and downregulation of *Wnt* signaling		
Sox2	1.43	0.01
Nanog	1.47	0.00
Tbx3	3.53	0.00
Hesx-1	20.56	0.00
Dppa-1	64.95	0.00
Igf2	3.61	0.00
Wnt3	0.43	0.00
Wnt4	0.36	0.00
Wnt6	0.36	0.00
Wnt8a	0.48	0.01
Sp8	0.12	0.00
Lef1	0.34	0.00
Tbxt	0.41	0.00
Axin2	0.55	0.04
Fgf8	0.10	0.00
Cdx1	2.47	0.00
Grhl3	5.00	0.00
Vill	1.87	0.00
Forced *HOXA13* expression supports caudal epithelial functions and promotes proliferation in DE		
Sox17	12.55	0.00
Lgr5	5.39	0.00
Nanog	0.20	0.02
Ezr	2.20	0.00
Vill	2.58	0.048
Krt19	2.34	0.00
Krt20	2.84	0.01
Igsf8	4.67	0.00
Gcnt3	3.51	0.00
Fgfr2	2.81	0.00
Nek2	2.35	0.01
HOXA13 downregulates the chromosome 1 epidermal differentiation complex, is pro-oncogenic, and conveys typical characteristics of the BE phenotype		
ANXA9	0.48	0.04
EVPL	0.61	0.03
SCEL	0.52	0.01
KLK7	0.42	0.01
EMP1	0.56	0.03
SERPINB13	0.38	0.00
DLL1	2.57	0.00
FURIN	1.49	0.03
JAG1	1.85	0.04

HOXA13 and the chromosome 1 epidermal differentiation complex

Further support for a role of *HOXA13* in the loss of the squamous phenotype and the appearance of caudal columnar phenotypes in the esophagus comes from experiments in which we investigated the effect of *HOXA13* directly on esophageal cell models. To this end, we used CRISPR-Cas9 technology to delete *HOXA13* from BAR-T, a primary monoclonal immortalized cell line derived from metaplastic tissue of a BE patient, with cells expressing both columnar and squamous markers [33]. Three separate *HOXA13* knock-out clones were selected to circumvent potential off target effects. Reversely, we provoked lentivirus-mediated *HOXA13* expression in EPC2-hTERT, an immortalized squamous esophageal cell line. For these latter experiments we used a mixed cell population of lentivirally transduced cells as to avoid clonal artifacts influencing results. Transcriptomes in these two models (Fig. 6a) were contrasted to their respective control lines. There was substantial overlap in the gene sets significantly affected by losing *HOXA13* in BAR-T compared with those significantly affected by gaining *HOXA13* in EPC2-hTERT, taking into account the direction of regulation (X^2 test: $p=4.74 \cdot 10^{-34}$) (see Supplementary Data 3). Investigation of this overlap across the two technically independently generated datasets limits the incidence of chance findings or single model system bias. Overlapping genes positively affected by *HOXA13* expression in esophageal cells are *IL7r*, *FAM196B*, *ADAMTS6*, *NRG1*, *LTBP1*, *JAG1*, *ELL2, SMAD7*, *C12ORF75*, *AXL*, *TIPARP*, *IKBIP*, *DUSP7*, and *GOLIM4*. Downregulated by *HOXA13* expression are *SERPINB13*, *MYO5C*, *KLK7*, *ANXA9*, *TMPRSS4*, *TTC9*, *MATN2*, *TNFAIP2*, *RAB27B*, *HCAR2, C6ORF132*, *EXPH5*, *MAP3K5*, and *FUCA1*. IPA analysis of the results predicts an increase in "(malignant) cell transformation" ($z=2.00$, $p=5.81 \cdot 10^{-3}$) and a decrease in "inflammation of an organ" ($z=-2.59$, $p=8.29 \cdot 10^{-3}$; gene function is described in Supplementary Data 3) in cells expressing *HOXA13*. Intriguingly, *HOXA13* downregulates the epidermal differentiation complex (EDC); Fig. 6b and c). The EDC, located on chromosome 1q21.3, contains clustered multigene families of genes associated with cornified envelope formation in stratified squamous epithelia, such as the S100 and the small proline-rich region (SPRR) genes [34]. Among the overlapping downregulated genes in both cell models, *ANXA9* is also associated with differentiating keratinocytes [35], and *EVPL*, *SCEL*, and *KLK7* are cornified envelope genes [36-38]. *EMP1* and *SERPINB13* downregulation is associated with increased disease severity in gastric cancer (*EMP1*) and head and neck squamous cell carcinoma (SCC) (*SERPINB13*) [39, 40]. See Table 1 for fold changes and *q*-values and Supplementary Data 3 for more differentially expressed molecules related to morphology. The downregulation of a gene region known to be essential for maintaining a squamous phenotype provides mechanistic support to the notion that altered *HOXA13* expression is cardinal for provoking the BE phenotype.

These experiments also provide mechanistic support for the notion that *HOXA13* expression may offer an explanation as to why BE is prone to progression to EAC. *HOXA13* mediates down-regulation of the EDC and many EDC and cornified envelope genes are progressively down-regulated in the BE to EAC cascade [36, 41]. In BE, EAC, and esophageal SCC, loss of heterozygosity (LOH) of the EDC is common [42-44]. Low EDC gene expression predicts chemotherapy non-response and LOH of the EDC is associated with reduced survival in curatively treated EAC patients [43, 45]. In our experimental models, we observed *HOXA13*-mediated upregulation of genes associated with Notch signaling,

specifically *DLL1*, *FURIN*, and *JAG1*. Notch signaling is associated with malignant transformation [46, 47]. In IPA analysis "Non-melanoma solid tumor" (z=2.03, p=9.28·10^{-8}) and "invasion of cells" (z=2.08, p=2.47·10^{-3}) were shown to be activated by *HOXA13*, whereas *HOXA13* expression negatively influenced the "Apoptosis" (z=-1.52, p=2.23·10^{-3}) and "killing of cells" (z=-2.03, p=2.03·10^{-3}) categories. Many individual genes showed differential regulation in a pro-oncogenic direction (see Supplementary Data 3).

In conclusion, using *HOXA13* knock-out and overexpression in a Barrett's and a squamous cell line, we show that *HOXA13* downregulates the epithelial differentiation complex and other cornified envelope genes which normally function to maintain squamous epithelial morphology and act as tumor suppressor genes. Additionally, Notch signaling is overexpressed and many individual genes show differential regulation in a pro-oncogenic direction.

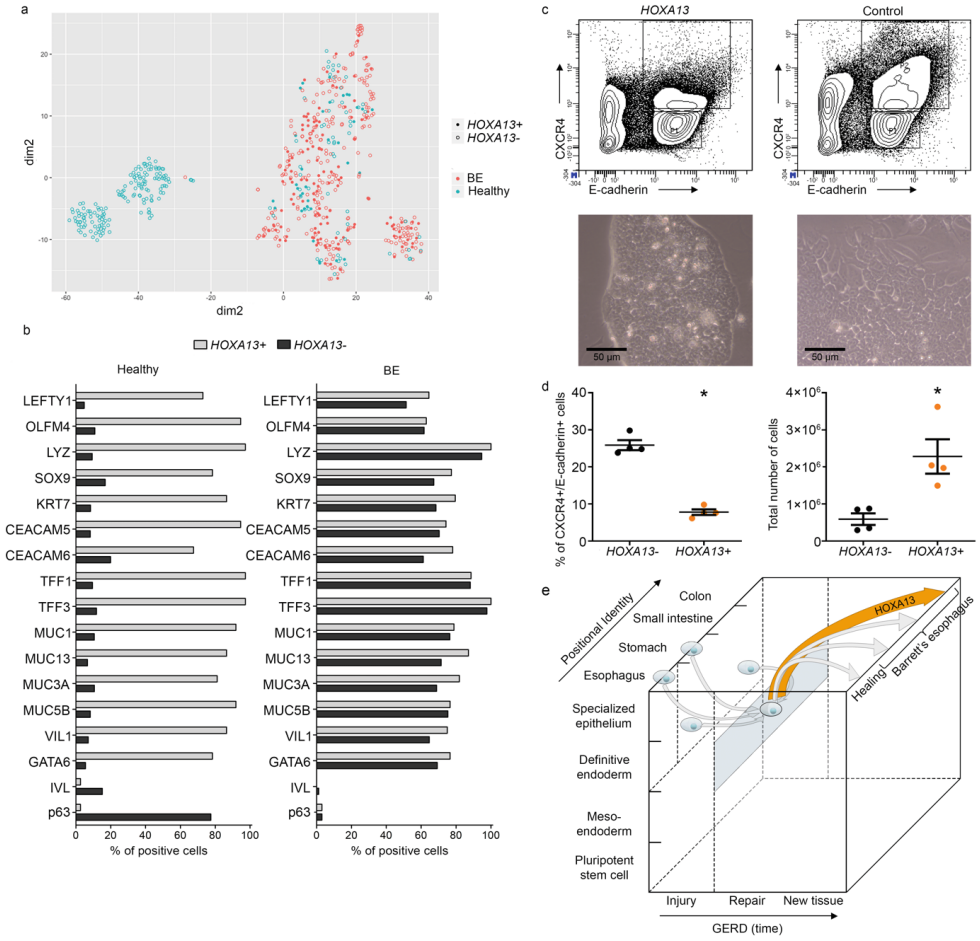


Fig. 6 | *HOXA13* counteracts squamous identity and increases growth of esophageal cells. a) Two models were constructed to investigate the function of *HOXA13* at the gastroesophageal junctions

(GEJ). One model used EPC2-hTERT, a primary immortalized human squamous esophageal cell line, characterized by low *HOXA13* expression, in which *HOXA13* was transduced. The second model employed BAR-T, a primary immortalized human Barrett's esophagus (BE) cell line, characterized by high *HOXA13* expression, in which *HOXA13* was knocked out. b) Hematoxylin and eosin (H&E) staining of the squamous esophagus of a patient without BE indicating the expected location of some of the products of the Ch1q21.3 epidermal differentiation complex along with other genes from the cornified envelope of the epidermis. c) *HOXA13* leads to a downregulation of genes in the Ch1q21.3 epidermal differentiation complex in both model systems. A cubic spline fit of *HOXA13* mRNA regulation is shown, with the BAR-T control transduced cell line presented compared to its *HOXA13* knock-out counterpart, and *HOXA13* overexpressing EPC2-hTERT cells presented compared to their parental line. FC – fold change. d) *HOXA13* knock-out in a BE cell line reduces the growth of the cell pool, as measured by MTT assay. Mean±SEM, **p*<0.05, exact *p*=0.0204, two-tailed t-test, n=9 independent experiments. e) *HOXA13* overexpression in a EPC2-hTERT cell line increases its growth in 3D culture (area of spheroids, mean±SEM, **p*<0.05, *p*= 0.0174, two-way ANOVA with Sidak's multiple comparisons test). f) EPC2-hTERT cells with *HOXA13* overexpression are less sensitive to bile/acid exposure (*p*=0.0343). MTT data presented as % of corresponding vehicle-treated controls. Mean±SEM, **p*<0.05, t-test (two-tailed).

HOXA13 supports columnar phenotype and provides proliferative advantage

Having established the transcriptional effect of *HOXA13* on BAR-T and EPC2-hTERT cells, we next investigated the functional consequences of *HOXA13* in these cells. As was seen for mESCs, *HOXA13* expression significantly enhances the growth-rate of esophageal cells. For BAR-T cells, a proliferative advantage of *HOXA13* expression was seen in 2D cultures (Fig. 6d), while for EPC2-hTERT the positive effect of *HOXA13* expression on cell growth was more noticeable under 3D culture conditions (Fig. 6e). Moreover, *HOXA13* expression decreases the sensitivity of keratinocytes to bile/acid exposure (Fig. 6f), consistent with the notion that *HOXA13* confers cellular protection under GERD-like conditions.

To gain further insight into the role of *HOXA13* in cell morphology and organization, we made use of the fact that EPC2-hTERT cells can be differentiated in 3D spheroid cultures, and become organized in layers with a more flattened cytological aspect in the middle of spheroids and high expression of keratinization markers such as involucrin, similar to esophageal stratified epithelium (see example in Fig. 7a top panel for differentiated morphology) [48]. Upon overexpression of *HOXA13*, EPC2-hTERT spheroids increase in size while maintaining a less differentiated phenotype (undifferentiated morphology, Fig. 7a bottom panel). Quantification of these morphological states indicates that in control cultures, 80% of spheroids attain a stratified epithelial phenotype, while overexpression of HOXA13 reduces this number to 28.6% (*p*<0.05, Fig. 7b left panel). This was further confirmed by staining for involucrin as a marker of keratinization, showing a decreased expression in spheroids derived from *HOXA13*-overexpressing EPC-hTERT cells (*p*<0.05, Fig. 7b right panel). Thus, in primary immortalized esophageal cells *HOXA13* overexpression reduces keratinization.

We further investigated the morphological role of HOXA13 in the BE-derived BAR-T cell line. In 2D cultures, an altered spatial distribution in growth pattern was observed, with cells growing more closely together in the absence of HOXA13 suggesting an effect on tissue morphology (Fig. 7c). The BAR-T cell model also allows testing the effect of *HOXA13* on columnar *versus* squamous differentiation in *in vitro* and *in vivo* settings. A 3D *in vivo* tissue reconstitution model was employed in which BAR-T cells were grafted in the lumina of devitalized and denuded rat tracheas and implanted in NOD SCID mice.

Under these conditions, parental non-transfected BAR-T cells produce both intestinal-type columnar epithelium and stratified squamous epithelium from the same clone. Hence this cell line has the potential to produce two types of morphological distinct epithelia [33, 49] (Fig. 7e and Supplementary fig. 9). Thus the epithelium in the model finds itself on a tipping point between both morphologies. This characteristic makes the *in vivo* tissue reconstitution model suited for studying the influence of modulators of morphology, *i.e.* to show if the modulator favours intestinal-type columnar epithelium or stratified squamous epithelium. Studying the effect of *HOXA13* knock-out, two important observations were made. Firstly, *HOXA13* knock-out decreases the length of columnar-like epithelium which contains PAS positive cells and is negative for involucrin (Fig. 7e, f). Thus, loss of *HOXA13* counteracts the proliferation of the intestinal-type columnar epithelium while the stratified squamous epithelial proliferation remains present. Secondly, *HOXA13* knock-out impairs epithelial proliferation in general, as inferred from the thickness of the epithelial layer (Fig. 7d; Supplementary fig. 9b). *In vitro* 2D organotypic ALI cultures of these cell lines confirm the *in vivo* findings, with *HOXA13* knockout reprogramming the BAR-T epithelial cells towards a squamous keratinized differentiated epithelium (Fig. 7f). In conclusion, *HOXA13* supports intestinal-type columnar epithelial differentiation and proliferation of the Barrett's epithelium confirming the notion that *HOXA13* expression can mediate both a competitive advantage as well as a predisposition to the formation of columnar phenotypes.

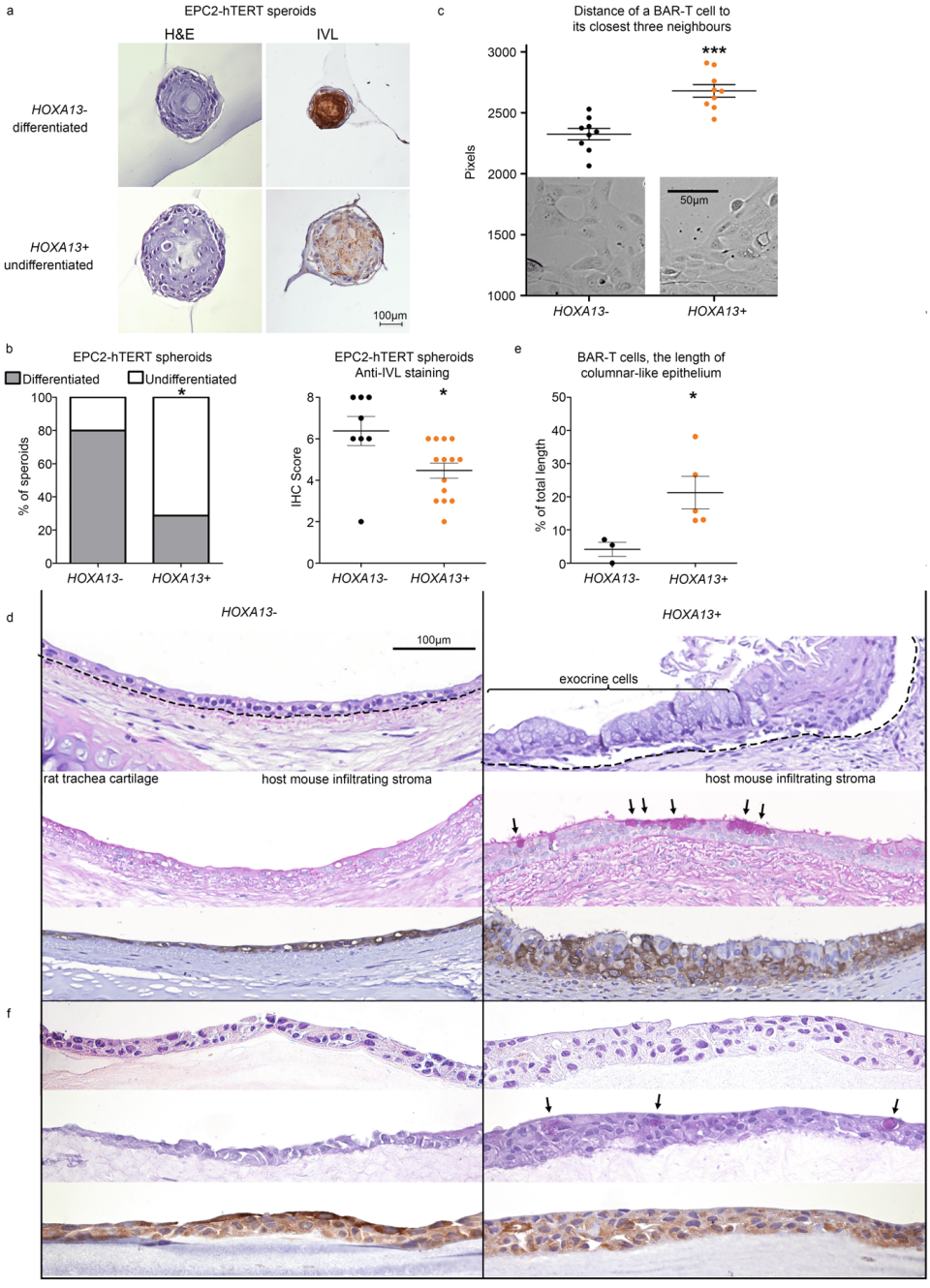


Fig. 7 | *HOXA13* supports intestinal-type columnar epithelial differentiation *HOXA13* overexpression impairs squamous differentiation of EPC2-hTERT spheroids as seen from the representative pictures (a) and quantitative assessment of morphologies based on hematoxylin and eosin (H&E) or anti-involucrin (IVL) immunohistochemistry (IHC) (p=0.0086). (b). Median with interquartile range, *p<0.05, Mann-Whitney test (two-tailed). c) *HOXA13* knockdown (KO) affects spatial distribution of BAR-T cells. Mean±SEM, ***p<0.001, p=0.0001, t-test (two-tailed). e) The length of columnar and mixed BAR-T

epithelium decreases upon *HOXA13* KO in the rat trachea *in vivo* tissue reconstitution model. Mean±SEM, **p*< 0.05, exact *p*=0.0439. d) Representative examples of H&E, PAS staining, anti-IVL IHC of BAR-T epithelium from the rat trachea *in vivo* tissue reconstitution model. H&E staining shows more layers of cells in animals transplanted with *HOXA13*$^{+}$ wild type cells. Periodic acid–Schiff (PAS) stains polysaccaride molecules, and positivity is indicative of goblet-like cells. The arrows point to PAS positivity, which is present in the right panel but not in the left panel, where the BAR-T *HOXA13*$^{-}$ cells are shown. IVL staining is strong in morphologically squamous cells in the left hand panel and weaker in the *HOXA13*$^{+}$ epithelium (n=3 for *HOXA*$^{-}$ and n=5 for *HOXA13*$^{+}$). f) *HOXA13*$^{-}$ and *HOXA13*$^{+}$ representative pictures of H&E staining, PAS staining and IVL IHC of the BAR-T organotypic cell culture system indicate that *HOXA13* KO reprograms the columnar epithelial phenotype towards squamous keratinized epithelium (n=1 independent experiments).

DISCUSSION

In this study, we characterized *HOX* gene expression and localization in mice and men, demonstrating a collinearity of these genes along the GI tract. Following analysis of one of these *HOX* genes, *HOXA13*, we observed single *HOXA13*+ cells in the upper GI tract, which present exceptions to the *HOX* gene collinearity theory. Specifically, in the normal physiology of the esophagus and proximal stomach, non-squamous structures such as the epithelium at the GEJ, glandular cells of ESMGs and glands of stomach contain single cells expressing *HOXA13*. The fact that these cells have not been described before may be a reflection of the fact that homogenization of tissues for qPCR masks this fraction, and that single cell analysis of the GI tract for this gene has not been performed before. We observe that GI pathology with distal phenotypes like intestinal metaplasia of the esophagus and stomach are characterized by an expansion of *HOXA13*-positive cells, while conversely, a relatively low expression of *HOXA13* is found in the phenotypically rostral Paneth cell metaplasia and pyloric metaplasia of the colon, compared to the surrounding physiological tissue.

It is clear that in normal physiology, *HOXA13* contributes to the distal phenotype of the caudal GI tract, begging the question as to the role and origin of the *HOXA13*-expressing compartment now observed in the upper GI tract. We demonstrate that esophageal *HOXA13*-positive cells express columnar and BE markers and show gene expression patterns overlapping with BE-derived cells. Functionally, *HOXA13* provides cells with several properties required for development of a BE segment. *HOXA13* maintains cells in a stem-like progenitor state, while conferring a proliferative advantage, promoting cellular migration [50] and resistance to bile and acid exposure. Furthermore, in cells that are lineage committed, *HOXA13* supports a phenotypically columnar phenotype, most likely partly driven by downregulation of the chromosome 1 epidermal differentiation complex. Thus, our data are consistent with the hypothesis that BE arises as a consequence of the expansion of resident *HOXA13*-positive cells under abrasive environments such as GERD. Several potential theories have been proposed as to the origin of BE: transdifferentiation of basal cells in the squamous epithelium, extension of a special population of cells from the GEJ, repopulation of the esophagus after injury with cells derived from progenitors ESMGs or ducts, resident embryonic stem cells or circulating bone marrow cells[51]. These potential sources of esophageal columnar epithelium are not mutually exclusive, and BE may have more than one precursor cell or location. Our study supports the previously proposed hypothesis that BE may originate from ESMGs and the GEJ as *HOXA13* is expressed in OLMF4^{+}, LEFTY1^{+} cells of

ESMGs, recently suggested as a cell of BE origin [21]. The fact that in addition to the GEJ, rare *HOXA13*$^+$ cells are found in the human esophagus and stomach, is consistent with the observation that after esophagogastrostomy BE can reoccur in patients, indicating that the involvement of the GEJ is not an absolute prerequisite for the development of BE [52]. Furthermore, our data show that *HOXA13* is already present at stem cell level, supportive of the notion that BE may arise from a cell with stem-cell like characteristics. While *HOXA13* expression overlaps greatly with *KRT7,* a columnar cytokeratin seen in Barrett's, we did not observe direct transcriptional overlap with the previously described *KRT7*$^+$*KRT14/5*$^+$*TP63*$^+$ cell of BE origin. However, *KRT7*$^+$*KR5*$^+$*TP63*$^+$ cells gave rise to BE-like epithelium only upon ectopic expression of *CDX2* [22]. Lineage-tracing studies are needed to further confirm whether one or more types of cells of origin might exist for BE. While we focused on HOXA13 here, it is conceivable that other *HOX* paralogues are involved in BE pathophysiology, in particular caudal genes such as *HOXA10*, *11*, *B13*, and *C10* are interesting candidates for further investigation, in particular as disruption of collinearity was reported for cluster B in BE [9] and in duodenum of murine embryos [12].

BE is considered as the precursor lesion for EAC, a dangerous form of cancer of which the incidence has substantially increased in recent decades. Increased insight into the pathogenesis of BE may aid development of prevention and treatment strategies for EAC. *HOXA13* is involved in ESCC [53] and other types of cancer [54-57]. Here we show that expression of *HOXA13* also increases in EAC and colorectal cancer, provides proliferative advantage to the cells and activates cancer-related gene transcription like Notch signaling. Hence, we speculate that *HOXA13* may play a role in BE progression towards EAC.

In toto, the present study identifies a importance of regional patterning by HOX genes in the gut epithelium. In Barrett's esophagus, gastric IM, and heterotopia of the upper GI-tract, a colon-like *HOX* gene expression is present, especially characterized by *HOXA13* upregulation. Single cells expressing the generally thought to be distally-restricted *HOXA13* gene are present in the physiological upper GI tract, in particular the GEJ, where it supports a columnar phenotype and may confer a relative competitive advantage. Thus, *HOXA13* mediates BE phenotype and proliferative potential and hence appears a rational target for strategies aimed at counteracting EAC development.

ACKNOWLEDGMENTS

We would like to acknowledge H.F.B.M. Sleddens, H. Stoop, M.H.W. van Dullemen, P. Vasic, I.T.A. Edelijn, M.J. van der Lee, P.J. Zwalua, W.W. van Dam, E. Zielhuis, J. Knoop, M. Doukas, A.L. Nigg and T.P.P. van den Bosch, Erasmus MC - University Medical Center Rotterdam, for their involvement in investigation, W.N.M. Dinjens for providing material, and F. McKeon, The Jackson Laboratory for Genomic Medicine, for conceptualization and providing resources. FAPESP n. 2016/01139-0; 2017/01046-5 for funding. M.Magierowski was supported by a grant from National Science Centre (Poland): UMO-2016/23/D/NZ4/01913.

REFERENCES

1. Grotenhuis, B.A. *et al.* Delay in diagnostic workup and treatment of esophageal cancer. *Journal of Gastrointestinal Surgery* 14, 476-483 (2010).
2. Thrift, A.P. & Whiteman, D.C. The incidence of esophageal adenocarcinoma continues to rise: analysis of period and birth cohort effects on recent trends. *Ann Oncol* 23, 3155-3162 (2012).
3. Ferlay, J. *et al.* Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer* 136, E359-386 (2015).
4. Thirunavukarasu, P. *et al.* Meckel's diverticulum--a high-risk region for malignancy in the ileum. Insights from a population-based epidemiological study and implications in surgical management. *Ann Surg* 253, 223-230 (2011).
5. Orosey, M., Amin, M. & Cappell, M.S. A 14-Year Study of 398 Esophageal Adenocarcinomas Diagnosed Among 156,256 EGDs Performed at Two Large Hospitals: An Inlet Patch Is Proposed as a Significant Risk Factor for Proximal Esophageal Adenocarcinoma. *Dig Dis Sci* (2017).
6. Barros, R. *et al.* Dynamics of SOX2 and CDX2 Expression in Barrett's Mucosa. *Dis Markers* 2016, 1532791 (2016).
7. Mari, L. *et al.* A pSMAD/CDX2 complex is essential for the intestinalization of metaplasia. *Cell Rep* 7, 1197-1210 (2014).
8. Silberg, D.G. *et al.* Cdx2 ectopic expression induces gastric intestinal metaplasia in transgenic mice. *Gastroenterology* 122, 689-696 (2002).
9. di Pietro, M. *et al.* Evidence for a functional role of epigenetically regulated midcluster HOXB genes in the development of Barrett esophagus. *Proceedings of the National Academy of Sciences of the United States of America* 109, 9077-9082 (2012).
10. Pearson, J.C., Lemons, D. & McGinnis, W. Modulating Hox gene functions during animal body patterning. *Nat Rev Genet* 6, 893-904 (2005).
11. Shah, N. & Sukumar, S. The Hox genes and their roles in oncogenesis. *Nat Rev Cancer* 10, 361-371 (2010).
12. Kawazoe, Y. *et al.* Region-specific gastrointestinal Hox code during murine embryonal gut development. *Dev Growth Differ* 44, 77-84 (2002).
13. Yahagi, N. *et al.* Position-specific expression of Hox genes along the Gl tract. *Congenit Anom (Kyoto)* 44, 18-26 (2004).
14. Wang, K.C. *et al.* A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* 472, 120-124 (2011).
15. Wang, X. *et al.* Cloning and variation of ground state intestinal stem cells. *Nature* 522, 173-178 (2015).
16. Gupta, R.a. *et al.* Long noncoding RNA HONTAIR reprograms chromatin state to promote cancer metastasis. *Nature* 464, 1071-1076 (2010).
17. Que, J. *et al.*. Pathogenesis and Cells of Origin of Barrett's Esophagus. *Gastroenterology* 157, 349-364.e341 (2019).
18. Yamamoto, Y. *et al.* Mutational spectrum of Barrett's stem cells suggests paths to initiation of a precancerous lesion. *Nature Communications* 7, 10380 (2016).
19. Gaunt, S.J. The significance of Hox gene collinearity. *Int J Dev Biol* 59, 159-170 (2015).
20. Symonds, D.A. Paneth cell metaplasia in diseases of the colon and rectum. *Arch Pathol* 97, 343-347 (1974).
21. Owen, R.P. *et al.* Single cell RNA-seq reveals profound transcriptional similarity between Barrett's oesophagus and oesophageal submucosal glands. *Nature Communications* 9, 4261 (2018).
22. Jiang, M. *et al.* Transitional basal cells at the squamous-columnar junction generate BE. *Nature* 550, 529-533 (2017).
23. Krüger, L. *et al.* Ductular and proliferative response of esophageal submucosal glands in a porcine model of esophageal injury and repair. *American Journal of Physiology-Gastrointestinal and Liver Physiology* 313, G180-G191 (2017).
24. Rotterdam, H. Pathology of the gastric cardia. *Verh Dtsch Ges Pathol* 83, 37-42 (1999).
25. Johansson, J. *et al.* Pancreatic acinar metaplasia in the distal oesophagus and the gastric cardia: prevalence, predictors and relation to GORD. *J Gastroenterol* 45, 291-299 (2010).
26. Matsushita, S., Ishii, Y., Scotting, P.J., Kuroiwa, A. & Yasugi, S. Pre-gut endoderm of chick embryos is regionalized by 1.5 days of development. *Dev Dyn* 223, 33-47 (2002).
27. Boyer, L.A. *et al.* Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* 122, 947-956 (2005).
28. Shin, M.R., Cui, X.S., Jun, J.H., Jeong, Y.J. & Kim, N.H. Identification of mouse blastocyst genes that are downregulated by double-stranded RNA-mediated knockdown of Oct-4 expression. *Mol Reprod Dev* 70, 390-396 (2005).
29. Voorneveld, P.W. *et al.* The BMP pathway either enhances or inhibits the Wnt pathway depending on the SMAD4 and p53 status in CRC. *Br J Cancer* 112, 122-130 (2015).
30. Bakre, M.M. *et al.* Generation of multipotential mesendodermal progenitors from mouse embryonic stem cells via sustained Wnt pathway activation. *J Biol Chem* 282, 31703-31712 (2007).
31. Denans, N., *et al.* Hox genes control vertebrate body elongation by collinear Wnt repression. *Elife* 4 (2015).
32. Uhlen, M. *et al.* Proteomics. Tissue-based map of the human proteome. *Science* 347, 1260419 (2015).
33. Jaiswal, K.R. *et al.* Characterization of telomerase-immortalized, non-neoplastic, human Barrett's cell line (BAR-T). *Dis Esophagus* 20, 256-264 (2007).
34. Mischke, D., Korge, B.P., Marenholz, I., Volz, A. & Ziegler, A. Genes encoding structural proteins of epidermal cornification and S100 calcium-binding proteins form a gene complex ("epidermal differentiation complex") on human chromosome 1q21. *J Invest Dermatol* 106, 989-992 (1996).
35. Boczonadi, V. & Maatta, A. Annexin A9 is a periplakin interacting partner in membrane-targeted cytoskeletal linker protein complexes. *FEBS Lett* 586, 3090-3096 (2012).
36. Dai, Y. *et al.* Genome-Wide Analysis of Barrett's Adenocarcinoma. A First Step Towards Identifying Patients at Risk and Developing Therapeutic Paths. *Transl Oncol* 11, 116-124 (2017).
37. Kalinin, A., *et al.* Assembly of the epidermal cornified cell envelope. *J Cell Sci* 114, 3069-3070 (2001).
38. Lundstrom, A. & Egelrud, T. Stratum corneum chymotryptic enzyme: a proteinase which may be generally present in the stratum corneum and with a possible involvement in desquamation. *Acta Derm Venereol* 71, 471-474 (1991).
39. Sun, G. *et al.* Association of EMP1 with gastric carcinoma invasion, survival and prognosis. *Int J Oncol* 45, 1091-1098 (2014).
40. de Koning, P.J. *et al.* Downregulation of SERPINB13 expression in head and neck squamous cell carcinomas associates with poor clinical outcome. *Int J Cancer* 125, 1542-1550 (2009).
41. Kimchi, E.T. *et al.* Progression of Barrett's metaplasia to adenocarcinoma is associated with the suppression of the transcriptional programs of epidermal differentiation. *Cancer Res* 65, 3146-3154 (2005).
42. Chaves, P. *et al.* Chromosomal analysis of Barrett's cells: demonstration of instability and detection of the metaplastic lineage involved. *Mod Pathol* 20, 788-796 (2007).
43. Maru, D.M. *et al.* Frequent loss of heterozygosity of chromosome 1q in esophageal adenocarcinoma: loss of chromosome 1q21.3 is associated with shorter overall survival. *Cancer* 115, 1576-1585 (2009).

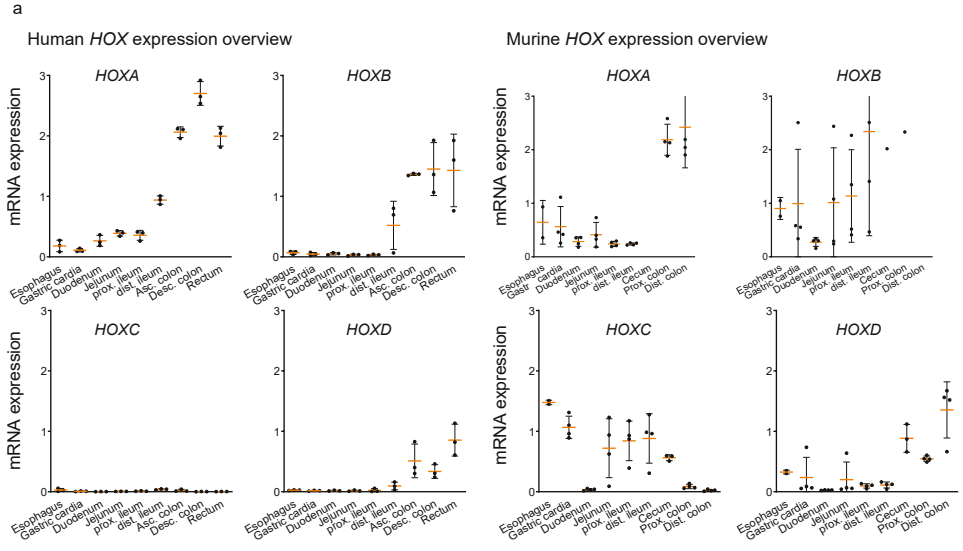
44. Li, J. *et al.* Allelic imbalance of chromosome 1q in esophageal squamous cell carcinomas from China: a novel region of allelic loss and significant association with differentiation. *Cancer Lett* 220, 221-230 (2005).
45. Luthra, M.G. *et al.* Decreased expression of gene cluster at chromosome 1q21 defines molecular subgroups of chemoradiotherapy response in esophageal cancers. *Clin Cancer Res* 13, 912-919 (2007).
46. Quante, M. *et al.* Bile acid and inflammation activate gastric cardia stem cells in a mouse model of Barrett-like metaplasia. *Cancer Cell* 21, 36-51 (2012).
47. Menke, V. *et al.* Conversion of metaplastic Barrett's epithelium into post-mitotic goblet cells by gamma-secretase inhibition. *Dis Model Mech* 3, 104-110 (2010).
48. Kasagi, Y. *et al.* The Esophageal Organoid System Reveals Functional Interplay Between Notch and Cytokines in Reactive Epithelial Changes. *Cell Mol Gastroenterol Hepatol* 5, 333-352 (2018).
49. Bajpai, M. *et al.* Repeated exposure to acid and bile selectively induces colonic phenotype expression in a heterogeneous Barrett's epithelial cell line. *Lab Invest* 88, 643-651 (2008).
50. Nesteruk, K. *et al.* Forced expression of HOXA13 confers oncogenic hallmarks to esophageal keratinocytes. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* 1866, 165776 (2020).
51. Nesteruk, K., Spaander, M.C.W., Leeuwenburgh, I., Peppelenbosch, M.P. & Fuhler, G.M. Achalasia and associated esophageal cancer risk: What lessons can we learn from the molecular analysis of Barrett's-associated adenocarcinoma? *Biochim Biophys Acta Rev Cancer* 1872 (2019).
52. Hamilton, S.R. & Yardley, J.H. Regnerative of cardiac type mucosa and acquisition of Barrett mucosa after esophagogastrostomy. *Gastroenterology* 72, 669-675 (1977).
53. Gu, Z.D. *et al.* HOXA13 promotes cancer cell growth and predicts poor survival of patients with esophageal squamous cell carcinoma. *Cancer Res* 69, 4969-4973 (2009).
54. Qin, Z. *et al.* Elevated HOXA13 expression promotes the proliferation and metastasis of gastric cancer partly via activating Erk1/2. *OncoTargets and therapy* 12, 1803-1813 (2019).
55. Deng, Y. *et al.* The expression of HOXA13 in lung adenocarcinoma and its clinical significance: A study based on The Cancer Genome Atlas, Oncomine and reverse transcription-quantitative polymerase chain reaction. *Oncol Lett* 15, 8556-8572 (2018).
56. Quagliata, L. *et al.* High expression of HOXA13 correlates with poorly differentiated hepatocellular carcinomas and modulates sorafenib response in in vitro models. *Lab Invest* 98, 95-105 (2018).
57. Dong, Y. *et al.* HOXA13 is associated with unfavorable survival and acts as a novel oncogene in prostate carcinoma. *Future Oncol* 13, 1505-1516 (2017).
58. Zhang, P. *et al.* Dissecting the Single-Cell Transcriptome Network Underlying Gastric Premalignant Lesions and Early Gastric Cancer. *Cell Rep* 27, 1934-1947 e1935 (2019).
59. Li, H. *et al.* Author Correction: Reference component analysis of single-cell transcriptomes elucidates cellular heterogeneity in human colorectal tumors. *Nat Genet* 50, 1754 (2018).
60. Federa (2011).
61. Perez, W.D., Weller, C.R., Shou, S. & Stadler, H.S. Survival of Hoxa13 homozygous mutants reveals a novel role in digit patterning and appendicular skeletal development. *Dev Dyn* 239, 446-457 (2010).
62. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat Methods* 9, 676-682 (2012).
63. Wang, F. *et al.* RNAscope: a novel in situ RNA analysis platform for formalin-fixed, paraffin-embedded tissues. *J Mol Diagn* 14, 22-29 (2012).
64. Schindelin, J. *et al.* The ImageJ ecosystem: An open platform for biomedical image analysis. *Mol Reprod Dev* 82, 518-529 (2015).
65. Barrett, T. *et al.* NCBI GEO: archive for functional genomics data sets--update. *Nucleic Acids Res* 41, D991-995 (2013).
66. Harada, H. *et al.* Telomerase induces immortalization of human esophageal keratinocytes without p16INK4a inactivation. *Molecular cancer research : MCR* 1, 729-738 (2003).
67. Jaiswal, K.R. *et al.* Characterization of telomerase-immortalized, non-neoplastic, human Barrett's cell line (BAR-T). *Diseases of the esophagus : official journal of the International Society for Diseases of the Esophagus / I S D E* 20, 256-264 (2007).
68. Shin, K.J. *et al.* A single lentiviral vector platform for microRNA-based conditional RNA interference and coordinated transgene expression. *Proc Natl Acad Sci U S A* 103, 13759-13764 (2006).
69. Katzen, F. Gateway((R)) recombinational cloning: a biological operating system. *Expert Opin Drug Discov* 2, 571-589 (2007).
70. Beard, C., Hochedlinger, K., Plath, K., Wutz, A. & Jaenisch, R. Efficient method to generate single-copy transgenic mice by site-specific integration in embryonic stem cells. *Genesis* 44, 23-28 (2006).
71. Buchholz, F., Angrand, P.O. & Stewart, A.F. Improved properties of FLP recombinase evolved by cycling mutagenesis. *Nat Biotechnol* 16, 657-662 (1998).
72. Ogaki, S., Shiraki, N., Kume, K. & Kume, S. Wnt and Notch signals guide embryonic stem cell differentiation into the intestinal lineages. *Stem Cells* 31, 1086-1096 (2013).
73. Larsson, S.H. *et al.* Subnuclear localization of WT1 in splicing or transcription factor domains is regulated by alternative splicing. *Cell* 81, 391-401 (1995).
74. Sanjana, N.E., Shalem, O. & Zhang, F. Improved vectors and genome-wide libraries for CRISPR screening. *Nat Methods* 11, 783-784 (2014).
75. R Development Core Team *R: A language and environment for statistical computing.* . (Vienna, Austria. ; 2008).
76. Love, M.I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15, 550 (2014).
77. Kramer, A., Green, J., Pollard, J., Jr. & Tugendreich, S. Causal analysis approaches in Ingenuity Pathway Analysis. *Bioinformatics* 30, 523-530 (2014).
78. Bus, P. *et al.*. Upregulation of miRNA-143, -145, -192, and -194 in esophageal epithelial cells upon acidic bile salt stimulation. *Diseases of the esophagus : official journal of the International Society for Diseases of the Esophagus / I S D E* (2013).
79. Queiroz, K.C. *et al.* Violacein induces death of resistant leukaemia cells via kinome reprogramming, endoplasmic reticulum stress and Golgi apparatus collapse. *PLoS One* 7, e45362 (2012).
80. Kalabis, J. *et al.* Isolation and characterization of mouse and human esophageal epithelial cells in 3D organotypic culture. *Nature Protocols* 7, 235 (2012).
81. Croagh, D. *et al.* Reconstitution of stratified murine and human oesophageal epithelia in an in vivo transplant culture system. *Scand J Gastroenterol* 43, 1158-1168 (2008).
82. Edgar, R., Domrachev, M., Lash, A.E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res 30 (1), 207-210* (2002).
83. Janmaat V.T. *et al.* HOXA13 in etiology and oncogenic potential of Barrett's esophagus. NCBI GEO database (GSE173170.

SUPPLEMENTARY DATA

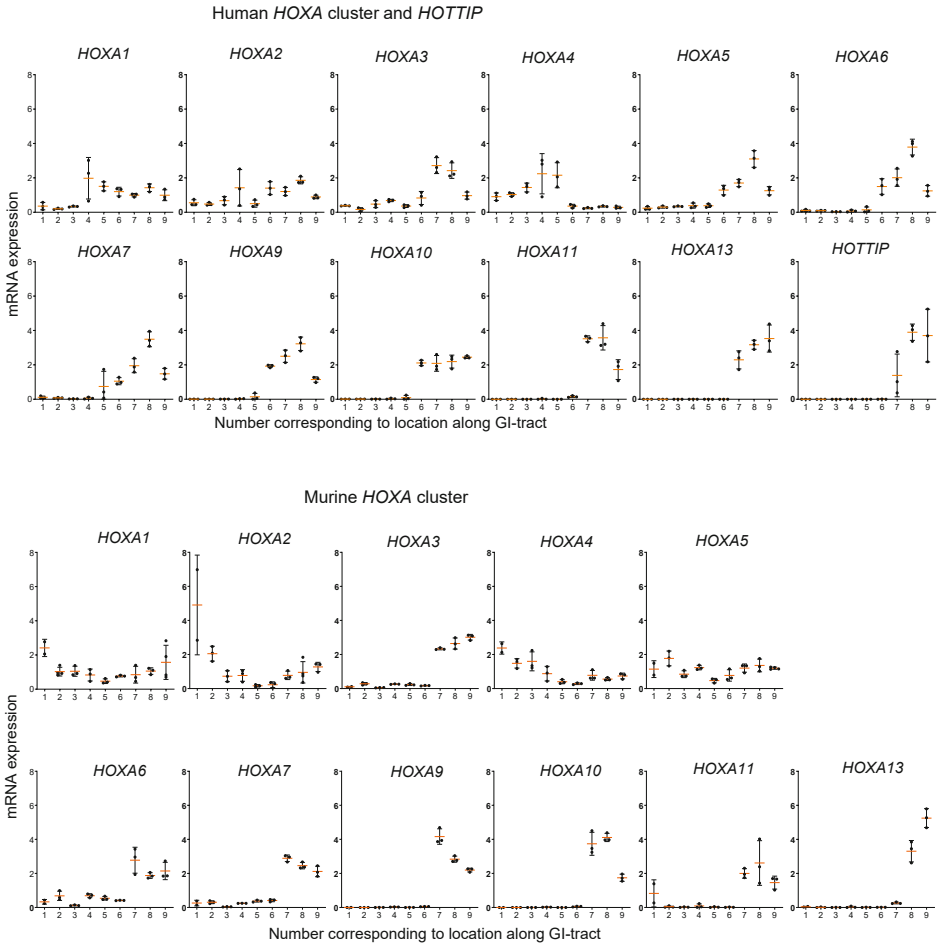
Supplementary method 1: macro used to quantify HOXA13-ISH in FIJI

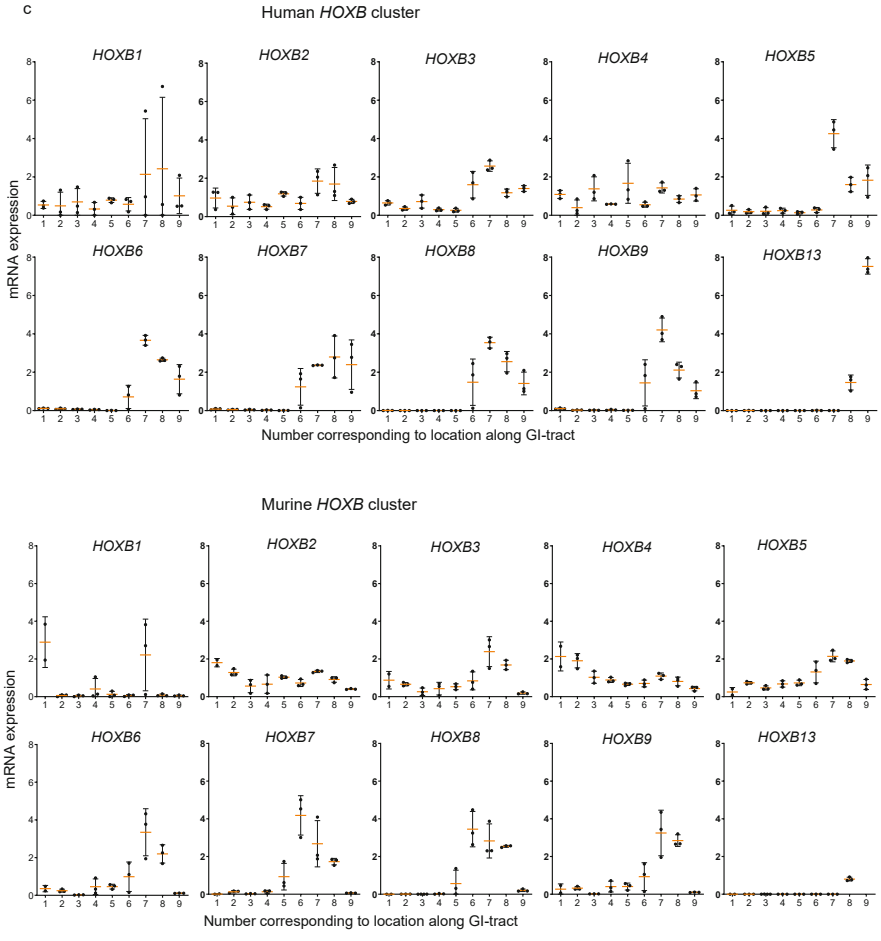
```
//setTool("freehand");
run("Cut");
run("Internal Clipboard");
selectWindow("Clipboard");
run("Colour Deconvolution", "vectors=[H&E DAB]");
selectWindow("Clipboard-(Colour_2)");
close();
selectWindow("Clipboard-(Colour_1)");
close();
selectWindow("Colour Deconvolution");
close();
selectWindow("Clipboard");
close();
selectWindow("Clipboard-(Colour_3)");
run("Measure");
run("Duplicate...", " ");
selectWindow("Clipboard-(Colour_3)");
setAutoThreshold("Default");
//run("Threshold...");
setThreshold(0, 10);
//setThreshold(0, 10);
run("Convert to Mask");
run("Measure");
close();
selectWindow("Clipboard-(Colour_3)-1");
setThreshold(3, 150);
//setThreshold(3, 150);
run("Convert to Mask");
run("Measure");
close();
String.copyResults();
IJ.deleteRows(0, 4);
```

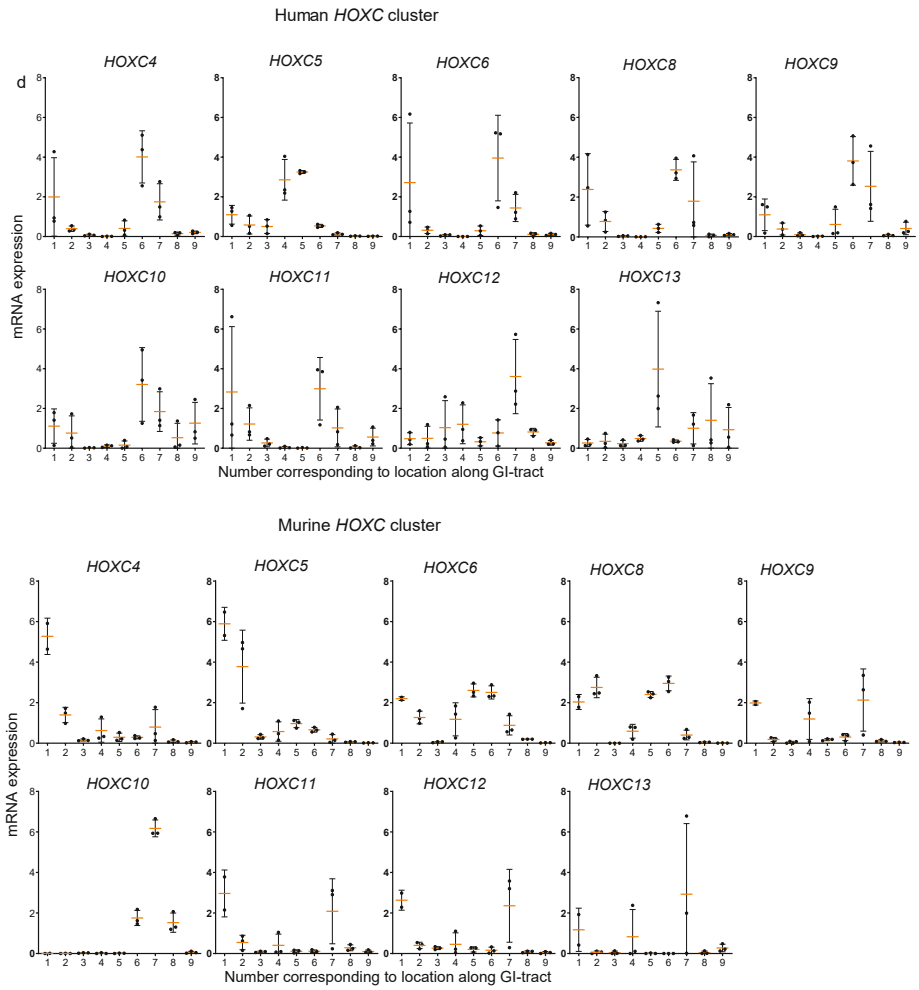
Supplementary figures

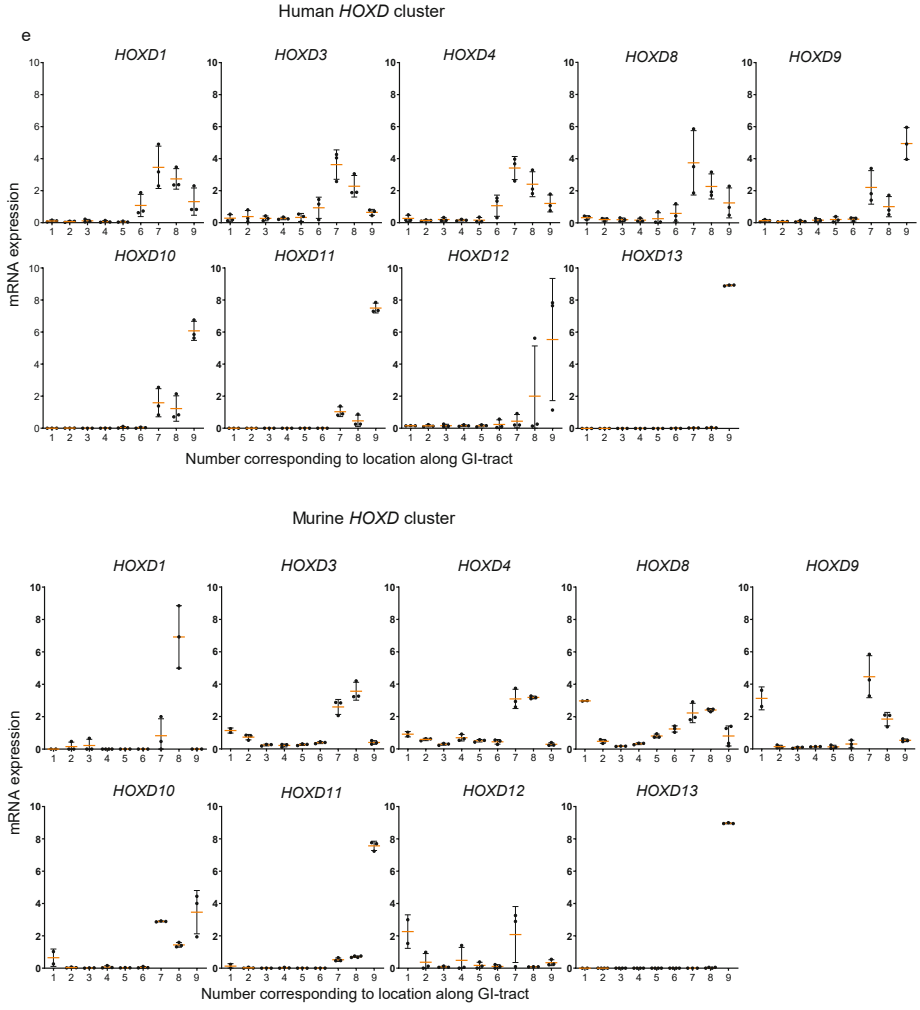


b

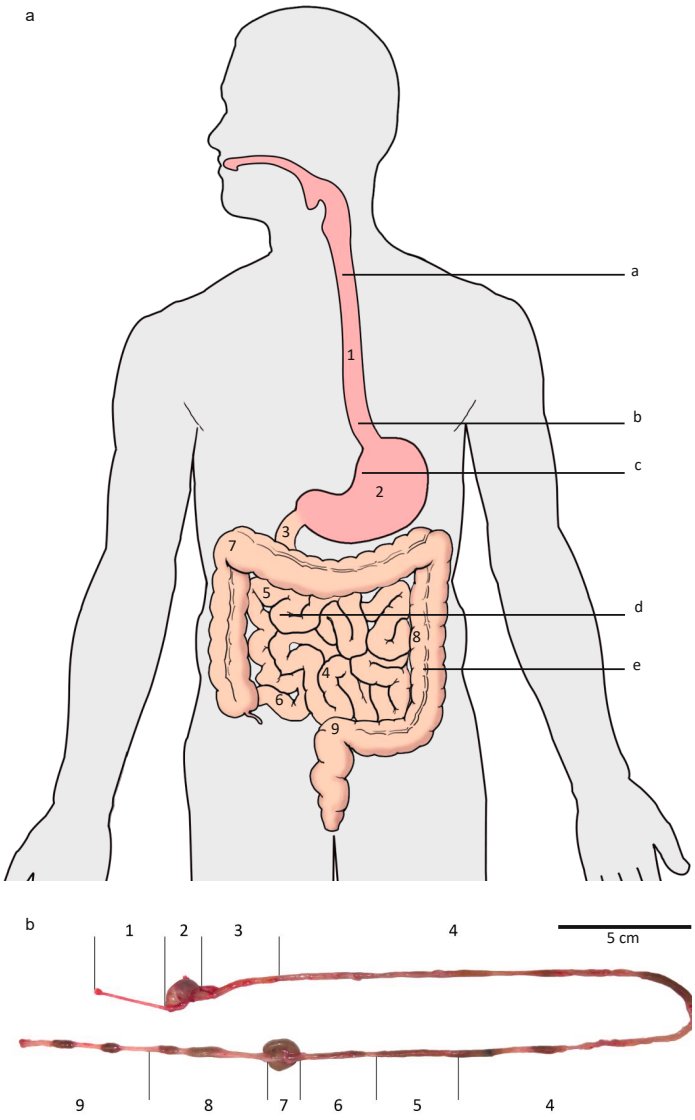

Human HOXA cluster and HOTTIP
HOXA1
HOXA2
HOXA3
HOXA4
HOXA5
HOXA6
HOXA7
HOXA9
HOXA10
HOXA11
HOXA13
HOTTIP
mRNA expression
Number corresponding to location along GI-tract
Murine HOXA cluster
HOXA1
HOXA2
HOXA3
HOXA4
HOXA5
HOXA6
HOXA7
HOXA9
HOXA10
HOXA11
HOXA13
mRNA expression
Number corresponding to location along GI-tract


c
Human HOXB cluster
HOXB1
HOXB2
HOXB3
HOXB4
HOXB5
HOXB6
HOXB7
HOXB8
HOXB9
HOXB13
mRNA expression
Number corresponding to location along GI-tract
Murine HOXB cluster
HOXB1
HOXB2
HOXB3
HOXB4
HOXB5
HOXB6
HOXB7
HOXB8
HOXB9
HOXB13
mRNA expression
Number corresponding to location along GI-tract

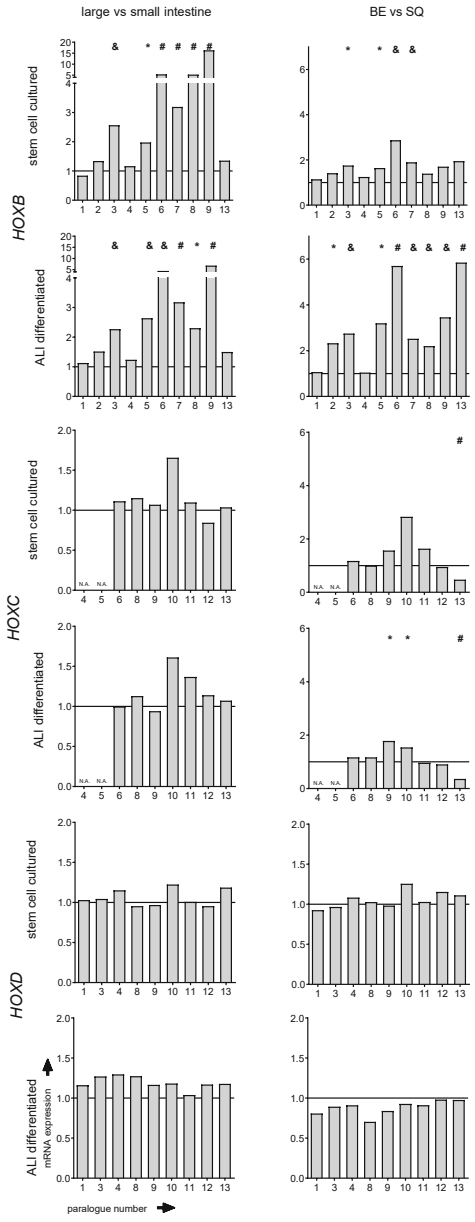

Human HOXC cluster
d
HOXC4
HOXC5
HOXC6
HOXC8
HOXC9
HOXC10
HOXC11
HOXC12
HOXC13
mRNA expression
Number corresponding to location along GI-tract
Murine HOXC cluster
HOXC4
HOXC5
HOXC6
HOXC8
HOXC9
HOXC10
HOXC11
HOXC12
HOXC13
mRNA expression
Number corresponding to location along GI-tract



Supplementary fig. 1: HOX cluster gene expression along the adult human and mouse gut. a) Overview of HOX cluster gene expression in the different epithelial regions along the human and mouse adult GI tract. The Y-axis represents the fold changes of mRNA expression relative to the average mRNA expression of all HOX genes of a given cluster. b)-e) Individual HOX gene expression of the HOXA, B, C, and D clusters is depicted. The Y-axis represents the fold changes of mRNA expression relative to the average mRNA expression of the depicted HOX gene. Mean with SD. Human data, n=3. Mouse data, n=4.

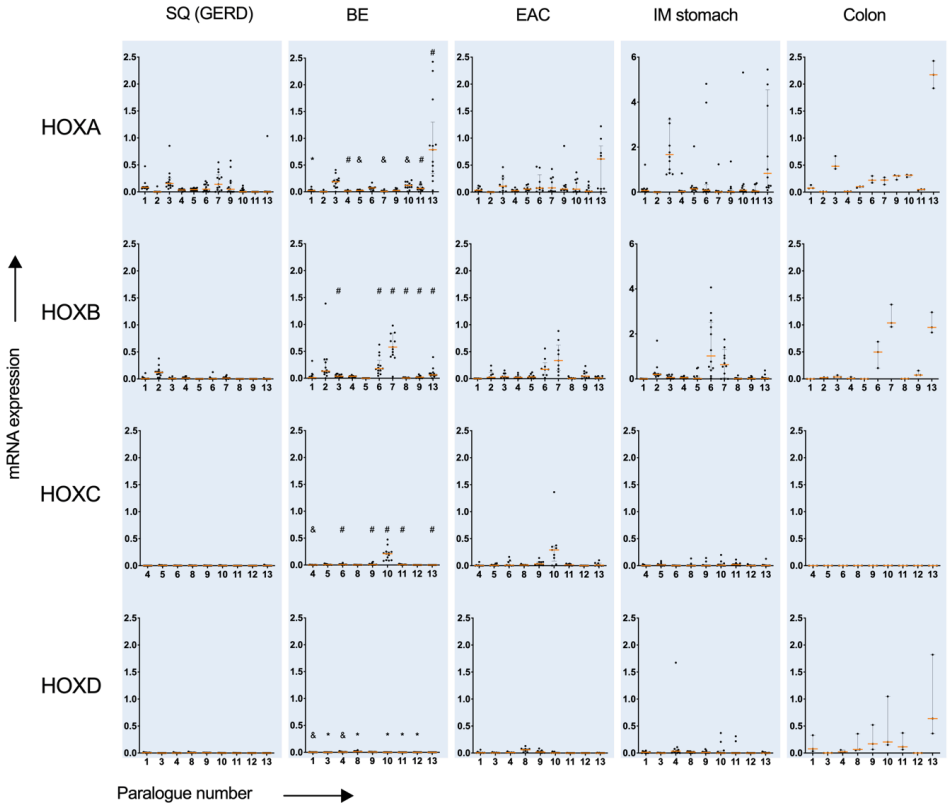


Supplementary fig. 2: locations of forceps biopsies taken along the human GI-tract and sections of murine GI-tract analyzed. a) Location of the forceps biopsies taken along the GI-tract are indicated by number in the schematic illustration of the GI-tract: 1) esophagus, 2) stomach, 3) duodenum, 4) jejunum, 5) proximal ileum, 6) distal ileum, 7) ascending colon, 8) descending colon and 9) sigmoid/rectum. Lesions studied: a) gastric inlet patch; b) CLE, BE, EAC; c) gastric IM; d) Meckel's diverticulum; e) pyloric and Paneth cell metaplasia (from the colon). b) Sections of mouse GI-tract used: 1) esophagus, 2) stomach, 3) duodenum, 4) jejunum, 5) proximal ileum, 6) distal ileum, 7) cecum and proximal colon, 8) proximal colon and distal colon and 9) distal colon.

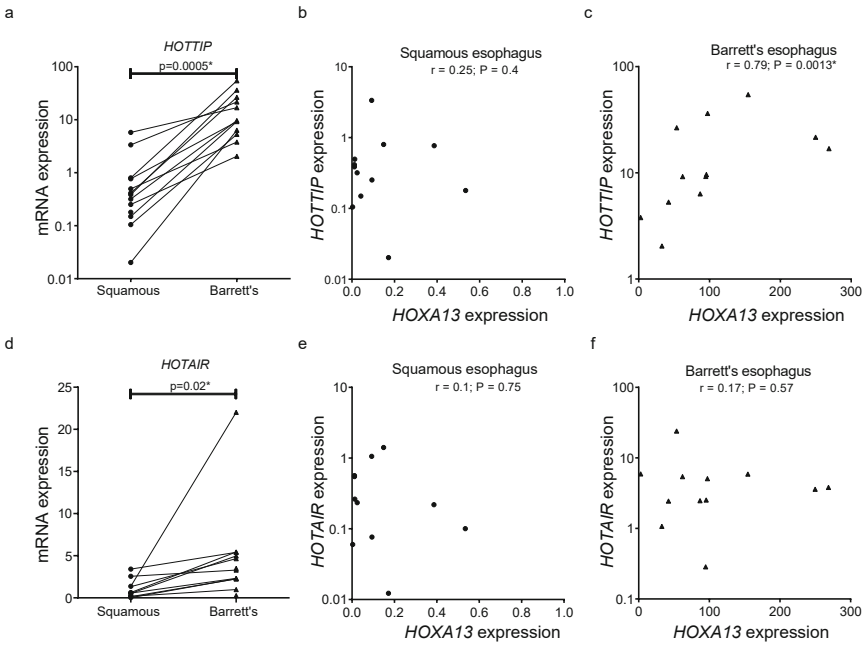


Supplementary fig. 3: HOX coding is established at the level of the stem cell. Public data of Wang et al. 1 contains the mRNA expression of human stem cells isolated from the GI-tract and either cultured as stem cells or differentiated in an air-liquid interface (ALI). Right column: HOXB, C, and D cluster gene expression in the large (n=3 in technical duplicate) compared to the small intestine (n=3 in technical duplicate). Left panels: HOXB, C, and D cluster gene expression in Barret's esophagus (BE, n=12) vs squamous esophagus (n=2) in technical duplicates are depicted for stem cell cultures and n=1 each for ALI differentiated samples in technical duplicates. Normalization was performed by setting mRNA expression to 1 for the small intestine or squamous esophagus. HOX gene expression in stem cell and ALI cultures are similar, which is not seen in the dataset in general. HOXB cluster genes have a higher expression in the large versus the small intestine (left column). No clear regulation of the HOXC or D clusters is seen, with exception of an upregulation of HOXC10 in the ascending colon. HOXB

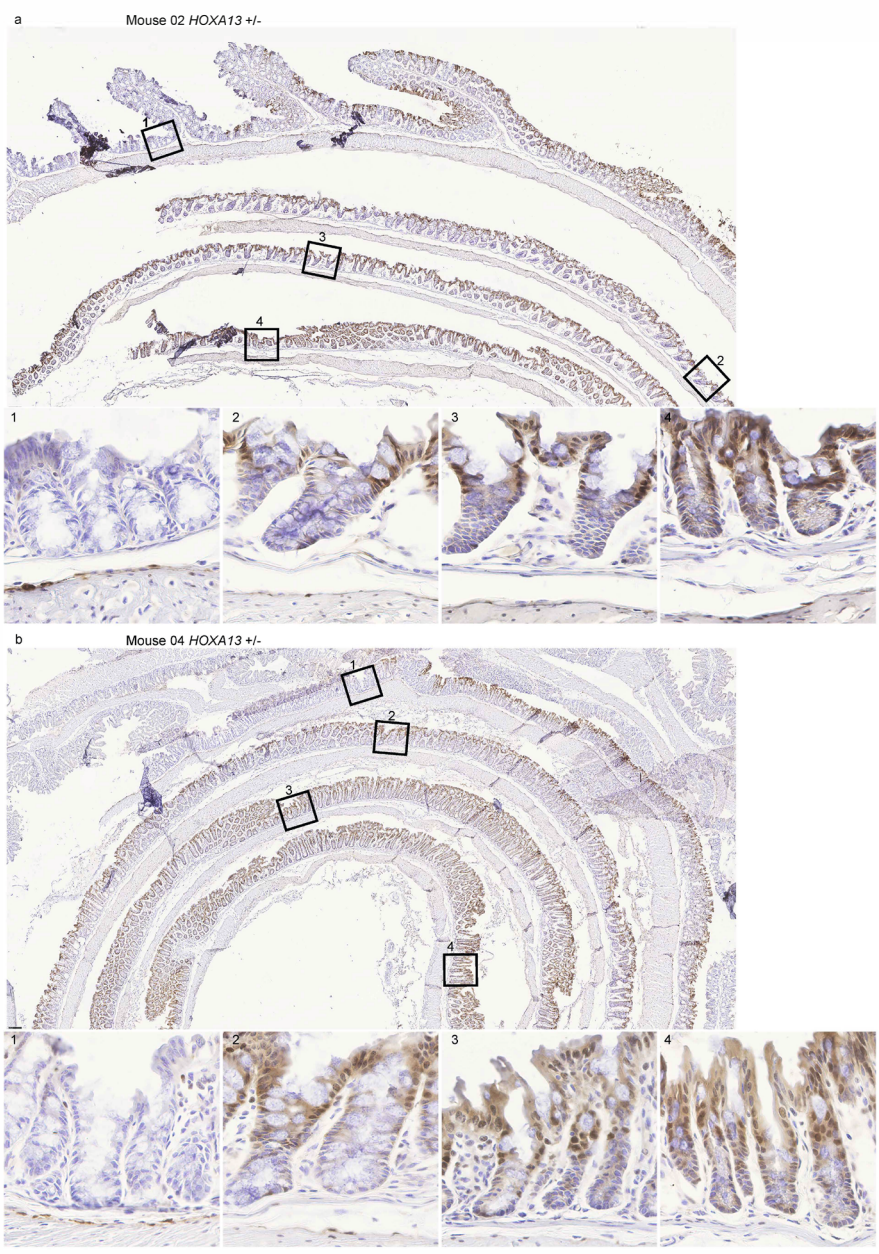
cluster genes are upregulated in BE stem cells vs squamous esophagus, including mid cluster and 5' HOXB genes. HOXC10 is the most pronounced HOXC gene upregulated in BE stem cells. *p<0.05; **p<0.01; ***p<0.001; NA: not available.

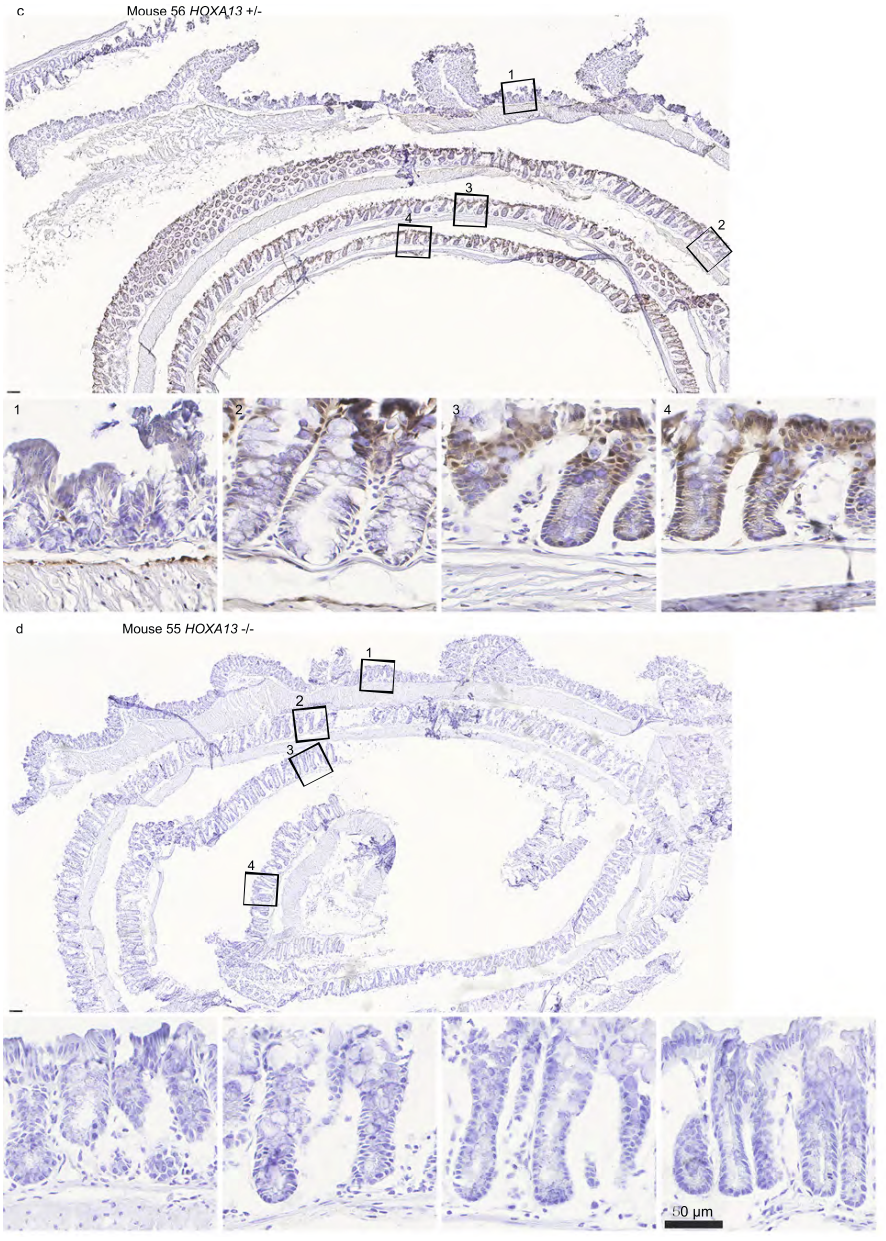


Supplementary fig. 4: HOX cluster gene expression in BE, esophageal adenocarcinoma and gastric IM, and colon. Comparison of the mRNA expression levels of the squamous esophagus from GERD patients (n=13), BE (n=13), EAC (n=9), IM of the stomach (n=12), and material from the colon (n=3). Expression in BE was compared to expression in the squamous esophagus of matched samples from the same patients (SQ, GERD) by two sided Student's t-tests. *p<0.05; **p<0.01; ***p<0.001. Y-axis values represent the fold change in mRNA expression in relation to the average mRNA expression of all HOX genes in all samples. Median with IQR.



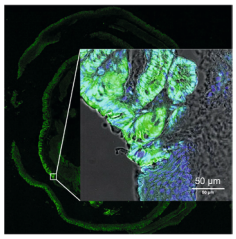
Supplementary fig. 5: Gene expression of lncRNA *HOTTIP* and lncRNA *HOTAIR*, in the normal squamous esophagus and BE tissue from BE patients. a) *HOTTIP* expression, normalized to mean expression in the squamous esophagus, is overexpressed in BE as tested with a Wilcoxon signed rank test. Matched squamous esophageal and BE were taken from the same patient (n=13). b, c) Correlations between expression levels of *HOTTIP* and *HOTAIR* and *HOXA13* in squamous and BE tissues were tested with a non-parametric Spearman test. *HOTTIP* expression does not correlate with *HOXA13* in normal squamous esophagus (b), but does correlate to *HOXA13* in BE tissue (c). d) *HOTAIR* is overexpressed in BE tissue (Wilcoxon signed rank test) and does not correlate with *HOXA13* expression (e, f). *p<0.05; &p<0.01;# p<0.001. Error bars represent the SEM.


a
Mouse 02 HOXA13 +/-
1
2
3
4
b
Mouse 04 HOXA13 +/-
1
2
3
4

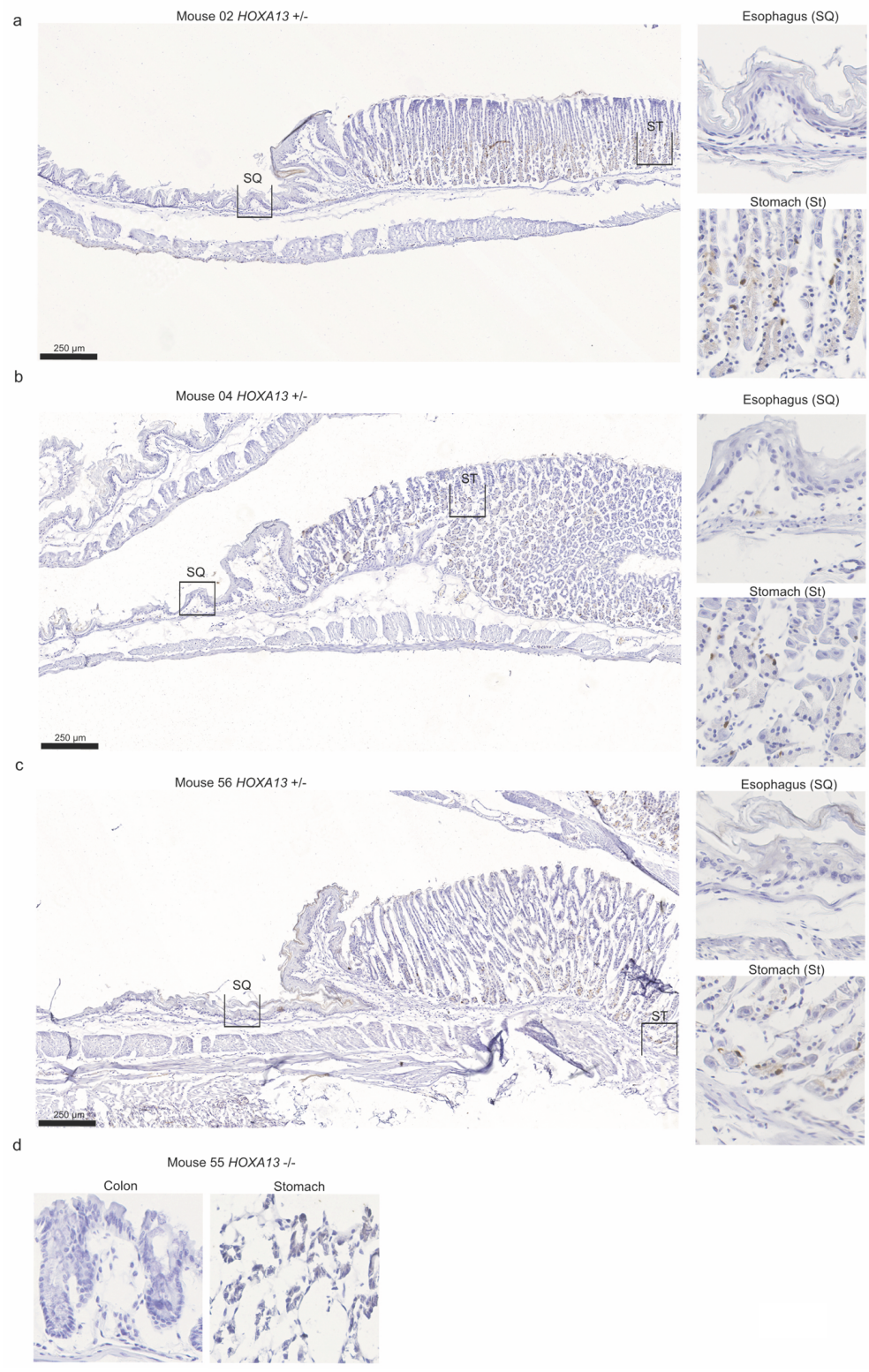

c
Mouse 56 HOXA13 +/-
1
2
3
4
1
2
3
4
d
Mouse 55 HOXA13 -/-
1
2
3
4
50 µm

e

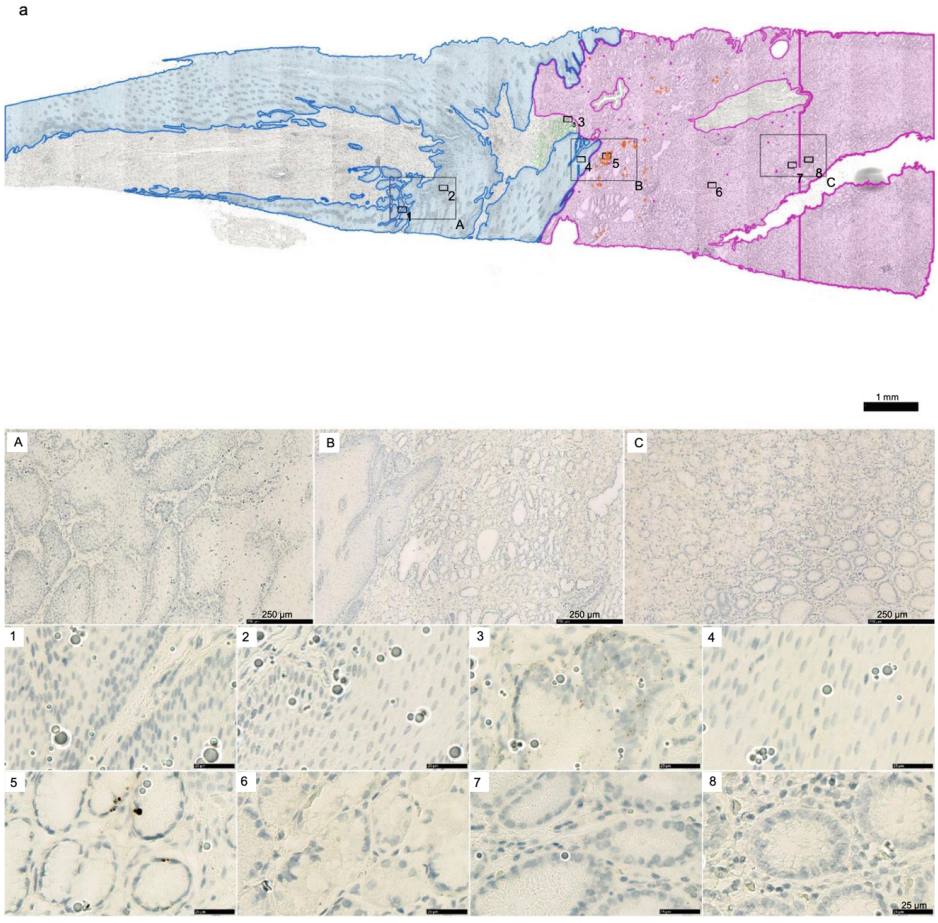
Distal expression boundary at anal squamous columnar junction

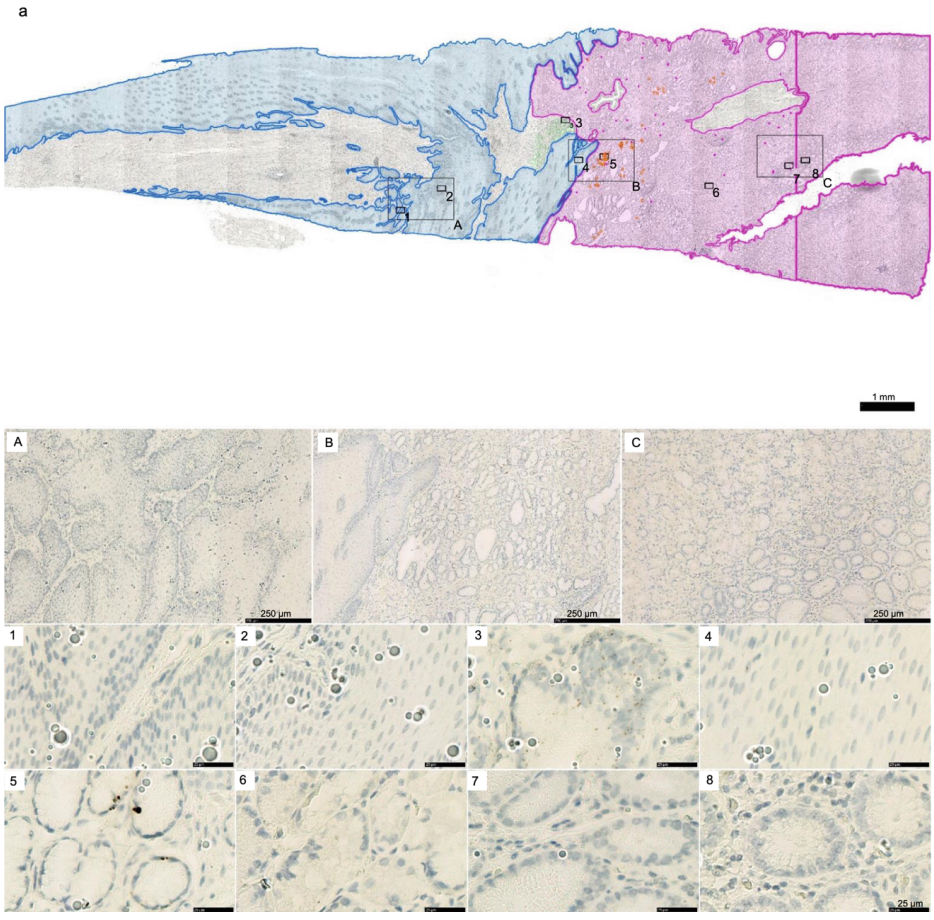


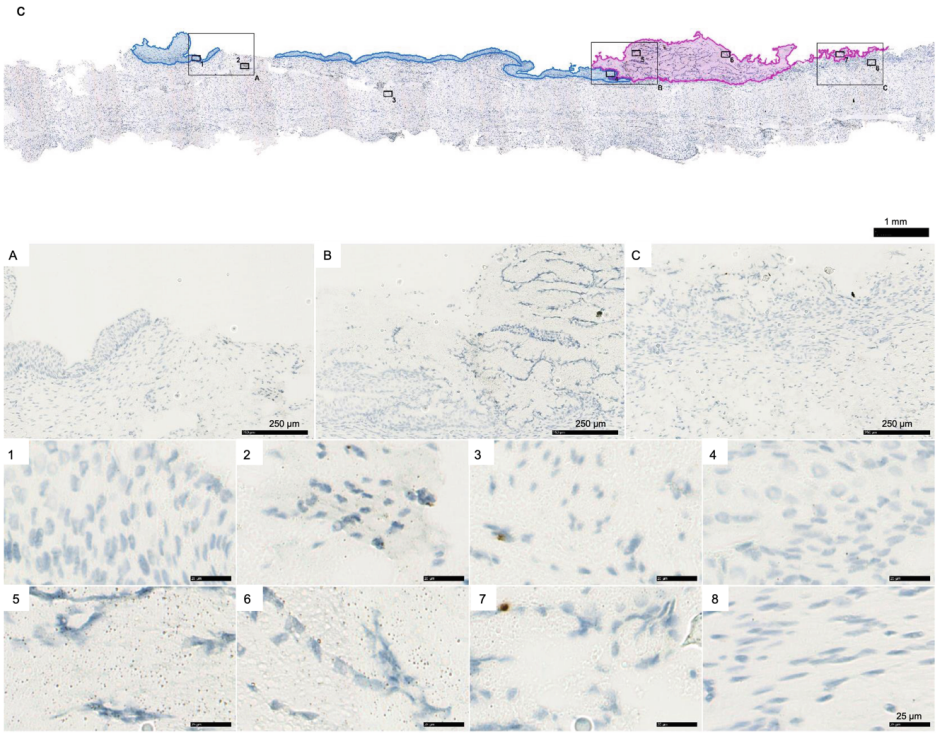
Supplementary fig. 6: Expression of HOXA13 in murine GI tract. Anti-GFP IHC of Hoxa13-GFP mouse model was performed on swiss roles of bowels isolated from three HOXA13+/- mice (a-c) and one HOXA13-/- negative control mouse (d). Overall presentation of swiss role and close ups of proximal (1) and distal (2-4) colon are shown. e) The distal Hoxa13 expression border is the anal SCJ (confocal images).

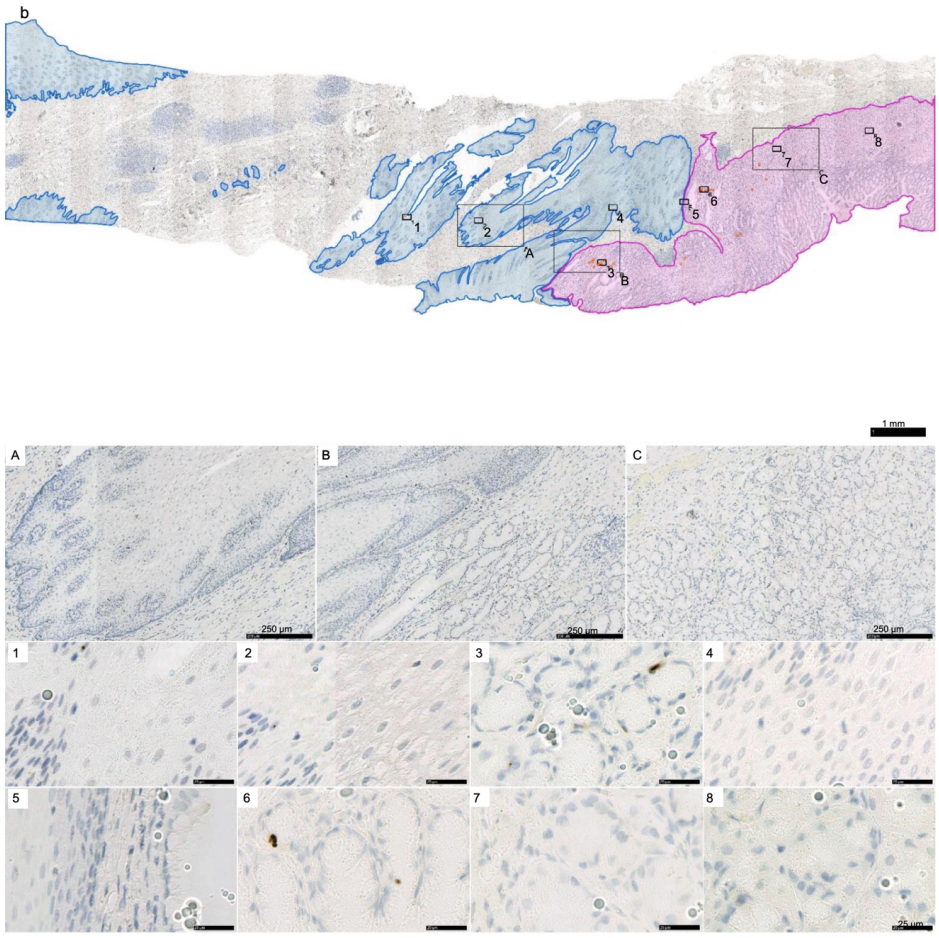

a
Mouse 02 HOXA13 +/-
Esophagus (SQ)
ST
SQ
Stomach (St)
250 µm
b
Mouse 04 HOXA13 +/-
Esophagus (SQ)
ST
SQ
Stomach (St)
250 µm
c
Mouse 56 HOXA13 +/-
Esophagus (SQ)
SQ
ST
Stomach (St)
250 µm
d
Mouse 55 HOXA13 -/-
Colon
Stomach

Supplementary fig. 7. Expression of HOXA13 in murine upper GI tract. Anti-GFP IHC of Hoxa13-GFP mouse model was performed on upper gastrointestinal tract isolated from three HOXA13+/- mice (a-c) and one HOXA13-/- negative control mouse (d). Magnification of squamous esophagus and stomach indicated in the overview image are presented on the right.

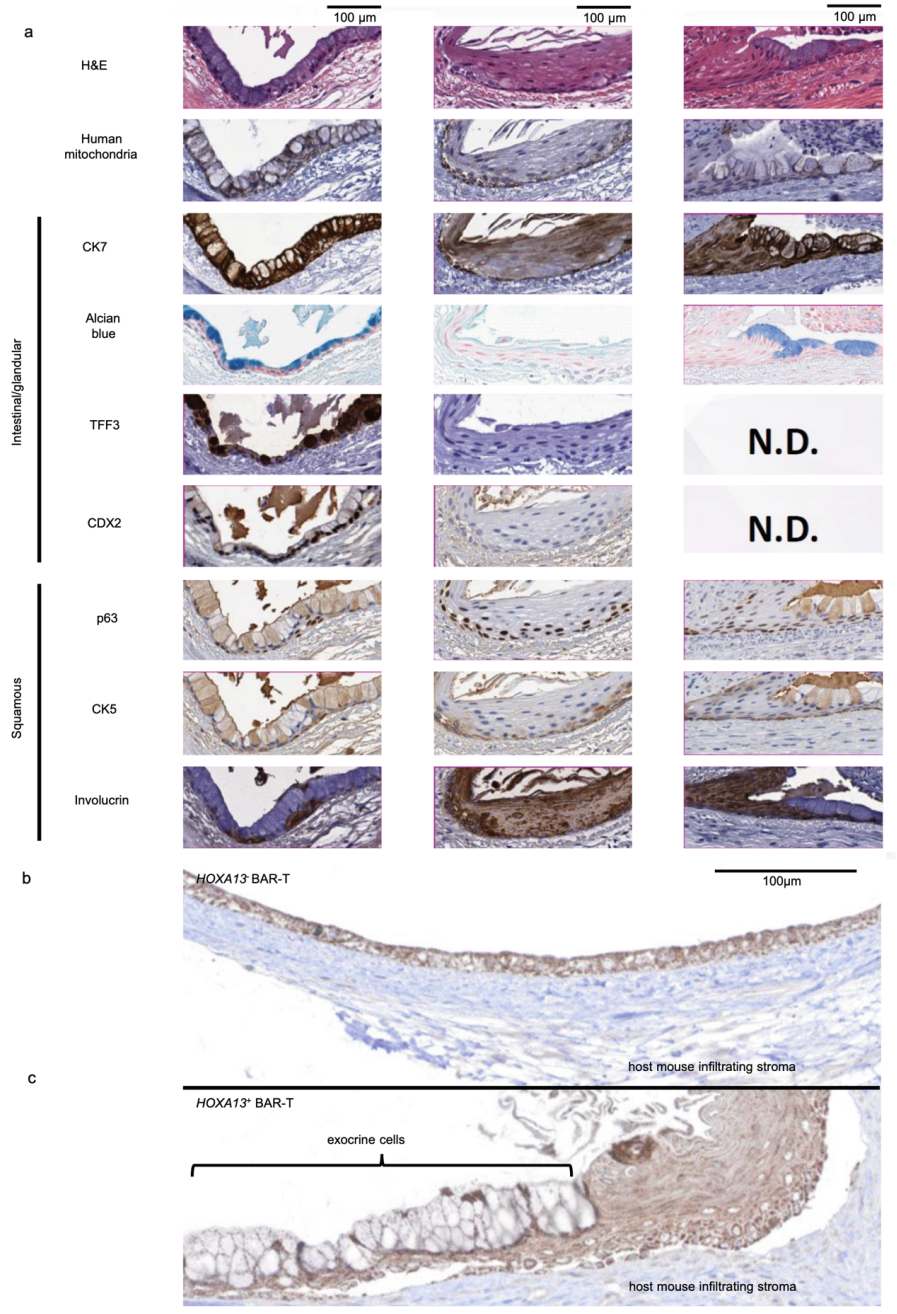



a
1
2
3
4
5
6
7
8
A
B
C
1 mm
A
B
C
250 µm
250 µm
250 µm
1
2
3
4
5
6
7
8
25 µm

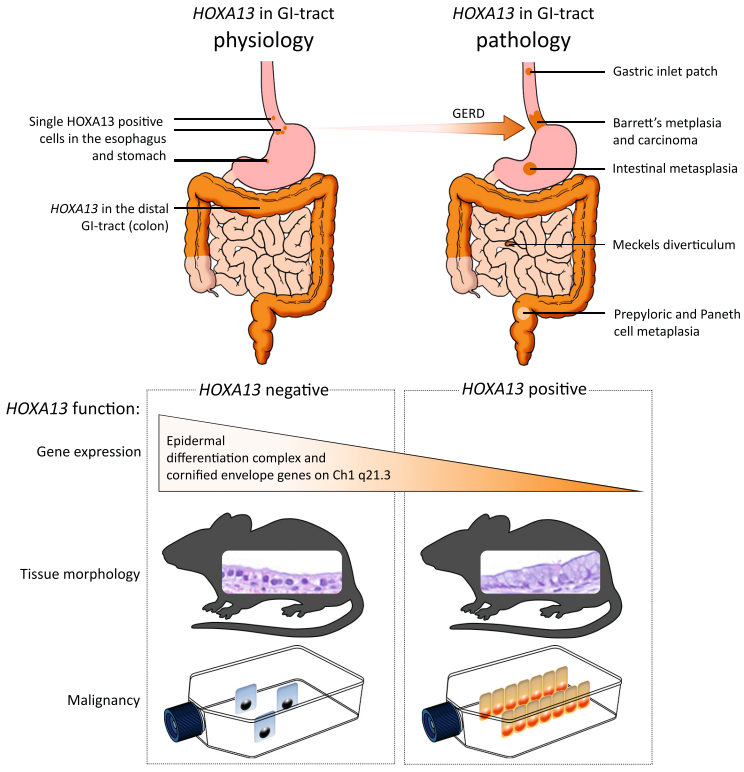

C
1 mm
A
B
C
250 µm
250 µm
250 µm
1
2
3
4
5
6
7
8
25 µm



Supplementary fig 8. a, b) HOXA13 expression as measured by RNA ISH in representative examples of an adult human GEJ with magnification panel of: A – esophagus, B – GEJ area, C – proximal stomach. c, d) HOXA13 expression as measured by RNA ISH in a representative examples of an fetus human GEJ with magnification panel of: A – esophagus, B – GEJ area, C – proximal stomach

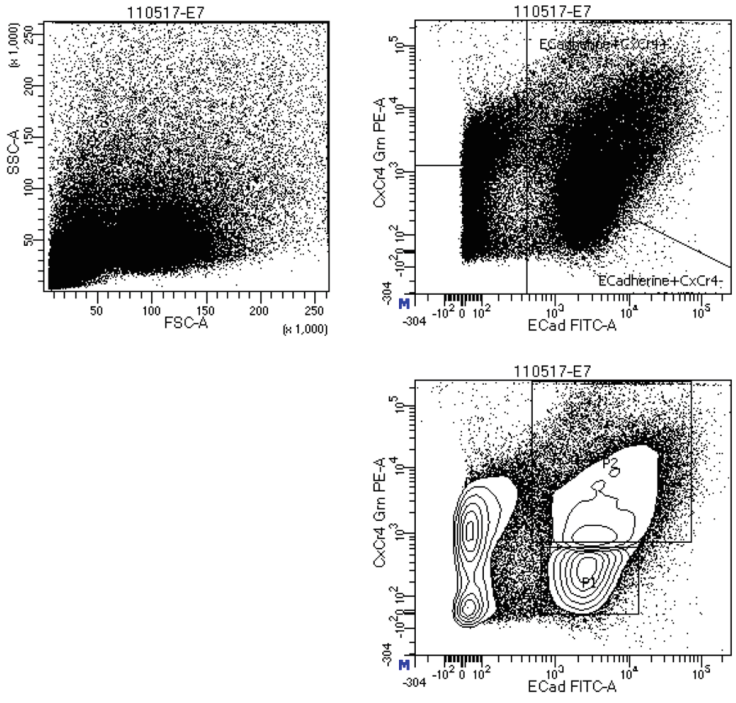


Supplementary fig. 9: Proof of the human origin of epithelium in the rat trachea in vivo tissue reconstitution model. a) Staining of the in vivo tissue reconstitution model with intestinal/glandular and squamous markers shows that parental BAR-T cells form intestinal-type columnar epithelium (left), squamous epithelium (middle), and multi layered epithelium with mixed phenotype (right) from the same clone. Human origin of the epithelium was confirmed by staining for human mitochondria. b) HOXA13- and c) HOXA13+ representative examples of the BAR-T epithelium stained for human mitochondria.



Supplementary fig. 10: summary figure.

BD FACSDiva 8.0.1



Tube: E7

Population	#Events	%Parent	%Total
All Events	100,000	####	100.0
Q1	13,259	13.3	13.3
ECadherine+CxCr4+	39,401	39.4	39.4
Q3	22,456	22.5	22.5
ECadherine+CxCr4-	24,884	24.9	24.9
P1	30,841	30.8	30.8
P2	29,293	29.3	29.3

BD FACSDiva 8.0.1

Supplementary fig. 11: Panel 1: forward / sideward scatter plot. Panel 2: Dotplot showing cells stained for ECadherin- FITC and CXCR4-PE. Panel 3: contour plots showing gating for ECadherin+ / CXCR4- cells (P1) and ECadherin+ / CXCR4+ (P2) cells, with gates positioned based on negative staining controls.

Supplementary tables

Supplementary table 1. differentially regulated genes in definitive endoderm versus non-differentiated KH2 mESCs.

	Fold change and q-value of CXCR4$^+$/E-cadherin$^+$ vs non-differentiated mESCs cells		**reference**
definitive endoderm markers:			
Sox17	12.55	0.00	2-5
Foxa1	4.30	0.00	3,5
Gata4	6.86	0.00	2
Lgr5	5.39	0.00	6
Pax3	25.91	0.00	3
Bmp2	16.32	0.00	4
Tacstd2	17.55	0.00	3
Bmp4	4.73	0.00	4
pluripotency markers:			
Nanog	0.20	0.02	4
Tcl1	0.01	0.00	7
Dppa3	0.02	0.00	8

Pluripotent mESC gene expression was compared to CXCR4$^+$/E-cadherin$^+$ FACS selected definitive endoderm cell gene expression. Both samples did not express *HOXA13*. Common definitive endoderm or pluripotency markers are included.

Supplementary table 2. All primers used in this study

primer name	sequence (5' to 3')
huHOXA1 L	TCTTCTCCAGCGCAGACTTT
huHOXA1 R	TTGACCCAGGTAGCCGTACT
huHOXA2 L	CCAAGAAAACCGCACTTCTG
huHOXA2 R	CATCGGCGATTTCCAGG
huHOXA3 L	ATGCAAAAAGCGACCTACTACG
huHOXA3 R	TACGGCTGCTGATTGGCATTA
huHOXA4 L	GAAGAAGATCCATGTCAGCG
huHOXA4 R	GGAACTCCTTCTCCAGCTCC
huHOXA5 L	GCGCAAGCTGCACATAAGTC
huHOXA5 R	GAACTCCTTCTCCAGCTCCA
huHOXA6 L	AAAGCACTCCATGACGAAGG
huHOXA6 R	CATGGCTCCCATACACAGC
huHOXA7 L	CAATTTCCGCATCTACCCCT
huHOXA7 R	GGAACTCCTTCTCCAGCTCC
huHOXA9 L	AATGCTGAGAATGAGAGCGG
huHOXA9 R	GTATAGGGGCACCGCTTTTT
huHOXA10 L	CCGGAGAAGGATTCCCTG
huHOXA10 R	CAGTGTCTGGTGCTTCGTGT
huHOXA11 L	ACACTGAGGACAAGGCCG
huHOXA11 R	GAAGAAGAACTCCCGTTCCA
huHOXA13 L	CCTCTGGAAGTCCACTCTGC
huHOXA13 R	GCACCTTGGTATAAGGCACG
huHOXB1 L	AGGAGACGGAGGCTATTTTCA
huHOXB1 R	GTCTGCTCGTTCCCATAAGGG
huHOXB2 L	CGCCAGGATTCACCTTTCCTT
huHOXB2 R	CCCTGTAGGCTAGGGGAGAG
huHOXB3 L	ATATTCACATCGAGCCCCAG
huHOXB3 R	CGTCATGAATGGGATCTGC
huHOXB4 L	CTTCTCCAGCTCCAAGACCT
huHOXB4 R	CTGGATGCGCAAAGTTCAC
huHOXB5 L	GGAACTCCTTTTCCAGCTCC

huHOXB5 R	GGAAGCTTCACATCAGCCAT
huHOXB6 L	GGGGACATGGACAAAATGAG
huHOXB6 R	GTGAGAACTGAGGAGCGGAC
huHOXB7 L	CTTTCTCCAGCTCCAGGGTC
huHOXB7 R	AACTTCCGGATCTACCCCTG
huHOXB8 L	GAACTCCTTCTCCAGCTCCA
huHOXB8 R	ACACAGCTCTTCCCCTGGAT
huHOXB9 L	TCCAGCGTCTGGTATTTGGT
huHOXB9 R	GAAGCGAGGACAAAGAGAGG
huHOXB13 L	GCTGTACGGAATGCGTTTCT
huHOXC4 L	GAGGTCTGGGGGTTGAGC
huHOXC4 R	GGAGCTGAGACAGGCTCG
huHOXC5 L	GAGTCTGGTAGCGCGTGTAAC
huHOXC5 R	CCACAGATTTACCCGTGGAT
huHOXC6 L	GATCATAGGCGGTGGAATTG
huHOXC6 R	GGCACAGAATGAGGGAAGAC
huHOXC8 L	CAAGGTCTGATACCGGCTGT
huHOXC8 R	ATCAAAACTCGTCTCCCAGC
huHOXC9 L	GTACTTGGTGTAGGGGCAGC
huHOXC9 R	ACAAAGAGGAGAAGGCCGAC
huHOXC10 L	ACCTCTTCTTCCTTCCGCTC
huHOXC10 R	GACACCTCGGATAACGAAGC
huHOXC11 L	ATAAGGGCAGCGCTTCTTG
huHOXC11 R	GAACACAAATCCCAGCTCGT
huHOXC12 L	GCAACTTCGAATAGGGCTTG
huHOXC12 R	AGCTTGGTATCGCCGTTG
huHOXC13 L	GCTGCACCTTAGTGTAGGGC
huHOXC13 R	CCACCTCTGGAAGTCTCCCT
huHOXD1 L	TTCTGTCAGTTGCTTGGTGC
huHOXD1 R	GGATGAAAGTGAAGAGGAATGC
huHOXD3 L	CACCTCCAATGTCTGCTGAA
huHOXD3 R	CAAAATTCAAGAAAACACACACA
huHOXD4 L	AGTTCTAGGACTTGCTGCCG
huHOXD4 R	CTACCCCTGGATGAAGAAGG
huHOXD8 L	TCTTCCTCTTCGTCTACCAGG
huHOXD8 R	TAATATTGGCGAGGACCCAG
huHOXD9 L	CTTTCTCCAGCTCAAGCGTC
huHOXD9 R	CAGCAGCAACTTGACCCAA
huHOXD10 L	TTCTGCCACTCTTTGCAGTG
huHOXD10 R	CTGAGGTCTCCGTGTCCAGT
huHOXD11 L	AAAGAAAAACTCGCGTTCCA
huHOXD11 R	CGAGAAGAGCAGCAGCG
huHOXD12 L	GCTGCTTCGTGTAGGGTTTC
huHOXD12 R	TGAACATGACAGTGCAGGC
huHOXD13 L	CCTCTTCGGTAGACGCACAT
huHOXD13 R	CAGGTGTACTGCACCAAGGA
huRP2 L	AAGCTGAGGATGCTCAAAGG
huRP2 R	CCCATTAAACTCCAAGGCAA
huβ-ACTIN L	GCACAGAGCCTCGCCTT
huβ-ACTIN R	GTTGTCGACGACGAGCG
huGAPDH L	AAGGTCGGAGTCAACGGATTT
huGAPDH R	ACCAGAGTTAAAAGCAGCCCT
moHoxA1 L	AAAAGAAACCCTCCCAAAACA
moHoxA1 R	AGCTCTGTGAGCTGCTTGGT
moHoxA2 L	GATGAAGGAGAAGAAGGCGG
moHoxA2 R	TGCCATCAGCTATTTCCAGG
moHoxA3 L	GCTGCCTGGTCATTCAAAGT
moHoxA3 R	GTCTCCAGTTCCAGGTGCTC
moHoxA4 L	ACCTTGATGGTAGGTGTGGC
moHoxA4 R	ACGCTGTGCCCCAGTATAAG
moHoxA5 L	CTCAGCCCCAGATCTACCC

moHoxA5 R	CAGGGTCTGGTAGCGAGTGT
moHoxA6 L	CCCTGTTTACCCCTGGATG
moHoxA6 R	GTCTGGTAGCGCGTGTAGGT
moHoxA7 L	AAGCCAGTTTCCGCATCTAC
moHoxA7 R	CTTCTCCAGTTCCAGCGTCT
moHoxA9 L	ACAATGCCGAGAATGAGAGC
moHoxA9 R	GTAAGGGCATCGCTTCTTCC
moHoxA10 L	CTCCAGCCCCTTCAGAAAAC
moHoxA10 R	TCTTTGCTGTGAGCCAGTTG
moHoxA11 L	AGGCTCCAGCCTACTGGAAT
moHoxA11 R	CCTTTTCCAAGTCGCAATGT
moHoxA13 L	GCTGCCCTACGGCTACTTC
moHoxA13 R	GCGGTGTCCATGTACTTGTC
moHoxB1 F	GGTGAAGTTTGTGCGGAGAC
moHoxB1 R	TTCGACTGGATGAAGGTCAA
moHoxB2 F	GAACCAGACTTTGACCTGCC
moHoxB2 R	GAGCTGGAGAAGGAGTTCCA
moHoxB3 F	ATCTGTTTGGTGAGGGTGGA
moHoxB3 R	CCGCACCTACCAGTACCACT
moHoxB4 F	GACCTGCTGGCGAGTGTAG
moHoxB4 R	CTGGATGCGCAAAGTTCAC
moHoxB5 F	CTGGTAGCGAGTATAGGCGG
moHoxB5 R	AGGGGCAGACTCCACAGATA
moHoxB6 F	TCCTATTTCGTGAACTCCACCT
moHoxB6 R	GCATAGCCAGACGAGTAGAGC
moHoxB7 F	GAGCAGAGGGACTCGGACTT
moHoxB7 R	GTCTGGTAGCGCGTGTAGGT
moHoxB8 F	CCTGCGCCCCAATTATTATGA
moHoxB8 R	AACTCCTGGATTTGCGAAGGG
moHoxB9 F	TCCAGCGTCTGGTATTTGGT
moHoxB9 R	GAAGCGAGGACAAAGAGAGG
moHoxB13 F	TGCCCCTTGCTATAGGGAAT
moHoxB13 R	ATTCTGGAAAGCAGCGTTTG
moHoxC4 F	CTACCCTGAGCGTCAGTATAGC
moHoxC4 R	CGCAGAGCGACTGTGATTTCT
moHoxC5 F	TTCTCGAGTTCCAGGGTCTG
moHoxC5 R	ATTTACCCGTGGATGACCAA
moHoxC6 F	CAGGGTCTGGTACCGAGAGTA
moHoxC6 R	TCCAGATTTACCCCTGGATG
moHoxC8 F	CAAGGTCTGATACCGGCTGT
moHoxC8 R	ATCAGAACTCGTCTCCCAGC
moHoxC9 F	ACTCGCTCATCTCTCACGACA
moHoxC9 R	GGACGGAAAATCGCTACAGTC
moHoxC10 F	ACCTCTTCTTCCTTCCGCTC
moHoxC10 R	ACTCCAGTCCAGACACCTCG
moHoxC11 F	TCCAACCTCTATCTGCCCAGT
moHoxC11 R	CAAGACGAGTAGCTGTTCCGA
moHoxC12 F	AATACGGCTTGCGCTTCTT
moHoxC12 R	GACCCTGGCTCTCTGGTTTC
moHoxC13 F	GGGCTATGGTTACCCATTTGG
moHoxC13 R	CTGGAGGACAGGTCGTCAC
moHoxD1 F	CAGCACTTTCGAGTGGATGA
moHoxD1 R	GCTCTGTCAGTTGCTTGGTG
moHoxD3 F	ACCAGCTGAGCACTCGTGTA
moHoxD3 R	AGAACAGCTGTGCCACTTCA
moHoxD4 F	CTCCCTGGGCTGAGACTGT
moHoxD4 R	CCCTGGGAACCACTGTTCT
moHoxD8 F	GTAATATTGGCGAGGACCCA
moHoxD8 R	CTACCAGGAGCTTGTGGTC
moHoxD9 F	GCTGAAGGAGGAGGAGAAGC
moHoxD9 R	GTGTAGGGACAGCGCTTTTT

moHoxD10 F	TCTCCTGCACTTCGGGAC
moHoxD10 R	GGAGCCCACTAAAGTCTCCC
moHoxD11 F	GAAAAAGCGCTGTCCCTACA
moHoxD11 R	AGGTTGAGCATCCGAGAGAG
moHoxD12 F	TGCTTTGTGTAGGGTTTCCTCT
moHoxD12 R	CTTCACTGCCCGACGGTA
moHoxD13 F	TGGTGTAAGGCACCCTTTTC
moHoxD13 R	CCCATTTTTGGAAATCATCC
Eef2 F	GCTTCCCTGTTCACCTCTGA
Eef2 R	CGGATGTTGGCTTTCTTGTC
Rpl37 F	GTCGGATGAGGCACCTAAAG
Rpl37 R	GAAGAACTGGATGCTGCGAC
Leng8 F	GGTTGTCTTGAAGCTGCCTT
Leng8 R	GACCTTGGGGTGTAGGGAAT
HOTTIP F	CCTAAAGCCACGCTTCTTTG
HOTTIP R	TGCAGGCTGGAGATCCTACT
HOTAIR F	GGTAGAAAAAGCAACCACGAAGC
HOTAIR R	ACATAAACCTCTGTCTGTGAGTGCC
MEIS1 F	GGGCATGGATGGAGTAGGC
MEIS1 R	GGGTACTGATGCGAGTGCAG
AgeI HoxA13 F	GGTGGTACCGGTGCCACCATGACAGCCTCCGTGCTCCT
XbaI HoxA13 R	ACCACCTCTAGATTAACTAGTGGTTTTCAGTT
HOXA13gibson F	ctccgcggccccgaagccgccaccatggactacaaagacgatgacgacaagATGACAGCCTCCG TGCTC
HOXA13gibson R	cgaagcggccatgaaTTAACTAGTGGTTTTCAGTTTGTTGATG
HOXA13colonyPCR F	CCTCTGGAAGTCCACTCTGC
HOXA13colonyPCR R	GCACCTTGGTATAAGGCACG
pBS31-TetO-F	CCATCCACGCTGTTTTGAC
MF13-R	AGCGGATAACAATTTCACACAGGA
T1E2 HygroR6	TGTATTGACCGATTCCTTGC
T1E2 HygroR7	AGGACATTGTTGGAGCCGAA
PGK-F1	AACAGCTATGACCATG
PGK-F2	GGGCCTTTCGACCTGCATCCATC
Guide1sgRNA F	CACCGTTTCTCTACGACAACGGCGG
Guide1sgRNA R	AAACCCGCCGTTGTCGTAGAGAAAC
px330-F	GATACAAGGCTGTTAGAGAG
TILHOXA13R3	CGAGCAGGGGCTGCATTG
Pre HOXA13 FW2	GCTTTGCATACGCCGTGG
Rat HoxA13 1 F	GGGCTATGACAGCCTCCGT
Rat HoxA13 1 R	ATGTTCTTGTTGAGCTCGTCGG
Rat HoxA13 2 F	GTCGTCTCCCATCCTTCAGA
Rat HoxA13 2 R	TATCCTCCTCCGTTTGTCCTT
Rat HoxA13 3 F	CTGGAACGGCCAAATGTACT
Rat HoxA13 3 R	CCTCCGTTTGTCCTTGGTAA
Rat Hmbs F	TCCTGGCTTTACCATTGGAG
Rat Hmbs R	TGAATTCCAGGTGAGGGAAC
Rat Hprt F	AGGCCAGACTTTGTTGGATT
Rat Hprt R	GCTTTTCCACTTTCGCTGAT
Rat Sdha F	TCCTTCCCACTGTGCATTACAA
Rat Sdha R	CGTACAGACCAGGCACAATCTG
Rat Mapk6 F	TAAAGCCATTGACATGTGGG
Rat Mapk6 R	TCGTGCACAACAGGGATAGA
Rat Rps18 F	AAGTTTCAGCACATCCTGCGAGTA
Rat Rps18 R	TTGGTGAGGTCAATGTCTGCTTTC
HOXA7methF	GACTGCGCCTACCTGAAGAC
HOXA7methR	CAACAGCCCCCTTTATCAGA
HOXA9methF	TGTAGGTCCCCACAGCTACC
HOXA9methR	AATCCTGATTGCCAGCTGAT
HOXA10methF	GGTGTCCTCGTCCCTAGTCA

HOXA10methR	CAGACAGGCAGACACAAGGA
HOXA11methF	TCGAAAAACTGGTCGAAAGC
HOXA11methR	CAATCTGGCCCACTGCTACT
HOXA13methF	AGTACATTTGGCCGTTCCAG
HOXA13methR	CTTCTACCACCAGGGCTACG
HOTTIPmethF	CTTCGAGCGTTTGAAGGAAG
HOTTIPmethR	GTCGCGTTGTGCATTAAGAA

Supplementary table 3. Mouse ESC culture medium and differentiation medium components

Details of used products to culture mESCs.			
Product	**Product details**		**Manufacturer**
Dulbecco's Modified Eagle Medium (DMEM)	DMEM 4.5 g/L Glucose with L-Glutamine	82%	Lonza
Foetal Bovine Serum (FBS)		15%	Biowest (Nuaillé, France)
Penicillin/Streptomycin	10,000 Units/mL Penicillin, 10,000 µg/mL Streptomycin	1%	Thermo-Fisher Scientific
MEM Non-Essential Amino Acids		1%	Thermo-Fisher Scientific
Leukaemia Inhibitory Factor (LIF)		0.01%	Department of Developmental Biology, Erasmus MC
2-Mercaptoethanol / β-Mercaptoethanol	55 mM in DPBS	0.1%	Thermo-Fisher Scientific
Sodium Pyruvate	100 mM	1%	Thermo-Fisher Scientific

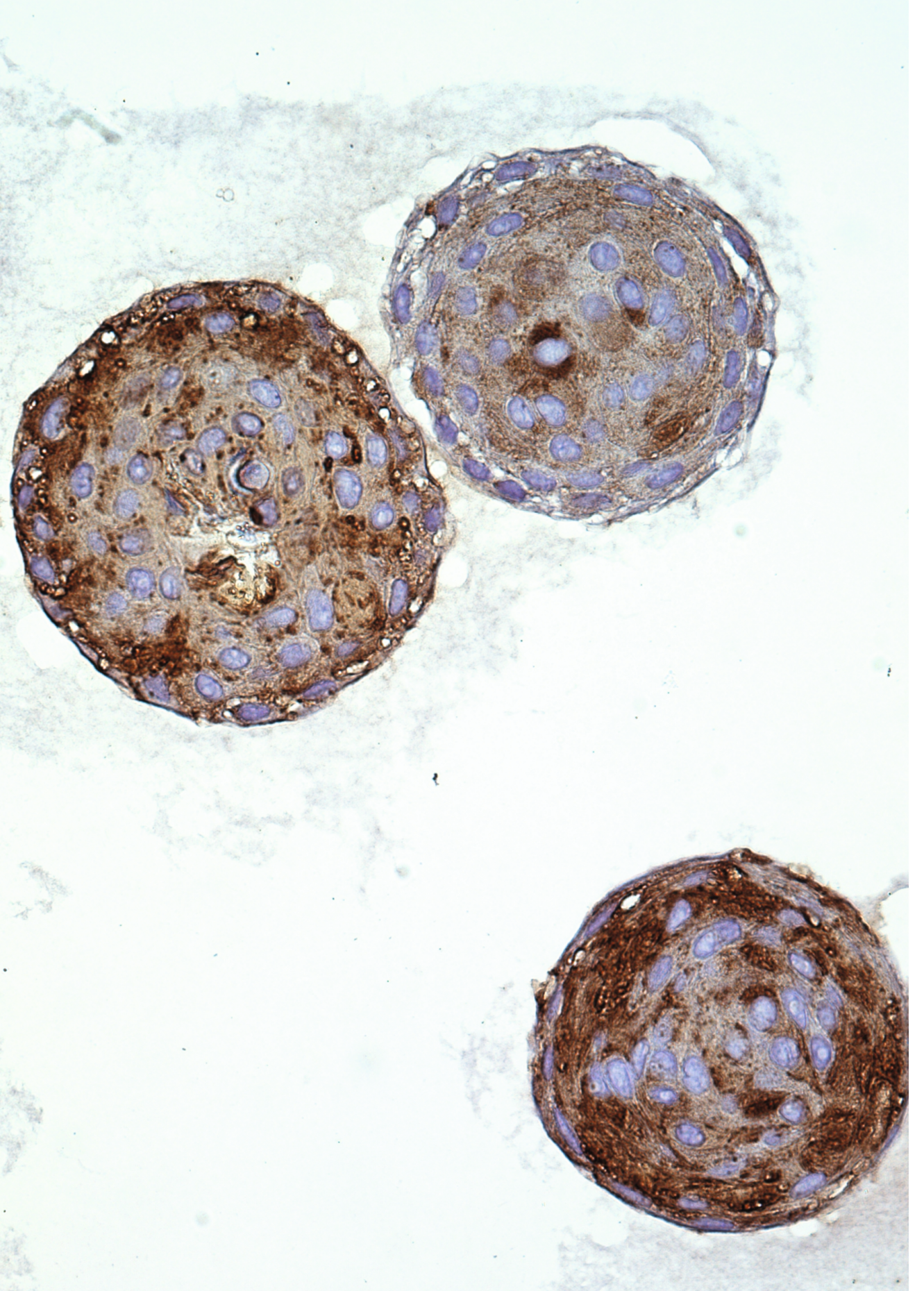
Details of the components used to differentiate the mESCs to definitive endoderm cells.		
Product	**Product details**	**Manufacturer**
Activin A	50 ng/mL	Thermo-Fisher Scientific
(Recombinant Human β-Fibroblast Growth Factor, 154 a.a.)	50 ng/mL	PeproTech EC Ltd. (London, UK)
CHIR	5 µM	Cayman Chemical (Ann Arbor, USA)

Supplementary table 4. All antibodies used in this study

Details of antibodies.				
Antibody	**Conc**	**Manufacturer**	**Product #**	**RRID**
PE rat anti-mouse (clone 2B11) CD184 (CXCR4)	1:250	BD Pharmingen	551966	AB_394305
Alexa Fluor 488 rat anti-mouse (clone DECMA-1) anti-CD324 (E-Cadherin)	1:250	Thermo-Fisher Scientific	53-3249-80	AB_10671270
Anti-human mitochondria (clone 113-1)	1:500	Merck Millipore, Billerica, USA	MAB1273	AB_94052
Mouse anti-human monoclonal (clone OV-TL 12/30) CK7	1:100	Dako Cytomation, Glostrup, Denmark	M7018	AB_2134589
Mouse anti-human monoclonal (clone 415909) TFF3	1:50	R&D Systems, Minneapolis, USA	MAB4407	AB_2271768
Rabbit anti-human monoclonal (clone EPR2764Y) CDX2	1:100	Cell Marque, Rocklin, CA	235R-14	AB_1516797
Mouse anti-human monoclonal (clone DAK-p63) P63	1:100	Dako Cytomation, Glostrup, Denmark	M7317	NA
Rabbit anti-human monocolonal (clone EP1601Y) CK5	1:100	Cell Marque, Rocklin, CA	305R-16	AB_1159468
Rabbit anti-human Involucrin	1:100	gift from A/Prof. Pritinder Kaur, Curtin University, Australia	NA	NA
Mouse anti-human Involucrin	1:500	Sigma-Aldrich, Sigma-Aldrich, St. Louis, Missouri, USA	#I9018	AB_477129
Rabbit polyclonal anti-GFP	1:100	Merck Millipore, Billerica, USA	#AB3080	AB_91337
Rabbit anti-human KRT5 (clone SP27)	0.51 µg/ml	Ventana, USA	760-4935	NA
Rabbit anti-human KR7 (clone SP52	0.536 µg/ml	Ventana, USA	790-4462	NA
Mouse anti-human P63 (cone 4AU)	0.140 µg/ml	Ventana, USA	790-4509	NA

Supplementary references

1. Wang, X. *et al.* Cloning and variation of ground state intestinal stem cells. *Nature* 522, 173-178 (2015).
2. Li, F. *et al.* Combined Activin A/LiCl/Noggin treatment improves production of mouse embryonic stem cell-derived definitive endoderm cells. *Journal of Cellular Biochemistry* 112, 1022-1034 (2011).
3. Sherwood, R.I. *et al.* Prospective isolation and global gene expression analysis of definitive and visceral endoderm. *Dev Biol* 304, 541-555 (2007).
4. Teo, A.K. *et al.* Activin and BMP4 synergistically promote formation of definitive endoderm in human embryonic stem cells. *Stem Cells* 30, 631-642 (2012).
5. Wang, P. *et al.* A molecular signature for purified definitive endoderm guides differentiation and isolation of endoderm from mouse and human embryonic stem cells. *Stem Cells Dev* 21, 2273-2287 (2012).
6. Tsai, Y.H. *et al.* LGR4 and LGR5 Function Redundantly During Human Endoderm Differentiation. *Cell Mol Gastroenterol Hepatol* 2, 648-662 e648 (2016).
7. Ivanova, N. *et al.* Dissecting self-renewal in stem cells with RNA interference. *Nature* 442, 533-538 (2006).
8. Waghray, A. *et al.* Tbx3 Controls Dppa3 Levels and Exit from Pluripotency toward Mesoderm. *Stem Cell Reports* 5, 97-110 (2015).
9. Polyak, K. & Weinberg, R.A. Transitions between epithelial and mesenchymal states: acquisition of malignant and stem cell traits. *Nat Rev Cancer* 9, 265-273 (2009).
10. Shapiro, L. *et al.* Structural basis of cell-cell adhesion by cadherins. *Nature* 374, 327-337 (1995).



Chapter 7

Forced expression of *HOXA13* confers oncogenic hallmarks to esophageal keratinocytes

K. Nesteruk*, V. T. Janmaat*, H. Liu, T. L.M. Ten Hagen, M. P. Peppelenbosch, G. M. Fuhler

Biochim Biophys Acta Mol Basis Dis. 2020

Chapter 7. Forced expression of *HOXA13* confers oncogenic hallmarks to esophageal keratinocytes

ABSTRACT

HOXA13 overexpression has been detected in human ESCC tissue and high HOXA13 protein expression is correlated with a shorter median survival time in ESCC patients. Although aberrant expression of *HOXA13* in ESCC has thus been established, little is known regarding the functional consequences thereof. The present study aimed to examine to what extent aberrant *HOXA13* might drive carcinogenesis in esophageal keratinocytes. To this end, we overexpressed *HOXA13* in a non-transformed human esophageal cell line EPC2-hTERT, performed gene expression profiling to identify key processes and functions, and performed functional experiments. We found that *HOXA13* expression confers oncogenic hallmarks to esophageal keratinocytes. It provides proliferation advantage to keratinocytes, reduces sensitivity to chemical agents, regulates MHC class I expression and differentiation status and promotes cellular migration. Our data indicate a crucial role of *HOXA13* at early stages of esophageal carcinogenesis.

HIGHLIGHTS

HOXA13 overexpression provides a proliferative advantage to esophageal keratinocytes

HOXA13 increases migration of esophageal keratinocytes

Overexpression of *HOXA13* decreases MHC class I in esophageal keratinocytes

HOXA13 overexpression inhibits differentiation of EPC2-hTERT-derived spheroids

INTRODUCTION

Esophageal cancer is the 8th most common cancer worldwide and the 6th cause of cancer related deaths [1-3]. Moreover, the prevalence of esophageal cancer has been growing; it rose by 44% from 1990 and reached 455,800 new cases per year in 2012. Approximately 85% of patients have a histological subtype called esophageal squamous cell carcinoma (ESCC), which is especially frequent in Eastern Asia, particularly in China [2, 4]. Contributing to ESCC development are environmental factors (alcohol consumption and tobacco use, a diet low in fruits and vegetables, ingestion of very hot food and beverages, etc.), genetic factors (e.g. aldehyde dehydrogenase [ALDH2] deficiency) and predisposing diseases (achalasia, tylosis) [3, 5]. ESCC arises from dysplastic precursor lesions: patches of squamous epithelial cells exhibiting nuclear atypia and abnormal maturation, but which do not invade through the basement membrane until disease progression to invasive carcinoma occurs [6]. ESCC is usually diagnosed at an advanced stage and prognosis is poor, with only 15% to 25% of patients diagnosed with ESCC surviving for 5 years after diagnosis [7].

While some studies have investigated the molecular pathways underlying ESCC development, disease etiology is still poorly understood. However, a possible role for *HOX* genes in ESCC development is now emerging. *HOX* genes are a highly conserved family of transcription factors which play a crucial role in the development of an embryo along the anterior-posterior axis [8, 9]. In humans, 39 *HOX* genes

are expressed with temporal and spatial collinearity [10, 11] which persists in adult tissues such as the skeleton and digestive system [12]. For example, the *HOX13* paralogues (*HOXA13, HOXB13*, and *HOXD13*) show high expression in the hindgut region and weak expression in the foregut including the esophagus [13]. As carcinogenesis can be seen as an aberrant form of organogenesis, these transcription factors may also regulate carcinogenic pathways [14-19]. Both tumor-promoting and tumor-suppressing properties have been ascribed to *HOX* genes [20]. *HOXA13* overexpression has been detected in human ESCC tissue [21], and in other types of cancer like gastric cancer, cervical cancer, ovarian cancer and prostate carcinoma [22-25]. High HOXA13 protein expression is correlated with a shorter median survival time in ESCC patients [26] and poor clinicopathological characteristics of patients [27]. The expression profile of HOXA13, ANXA2 and SOD2 was suggested as predictive marker of the postoperative outcome of patients with ESCC [28]. Expression of FGF2, the normal morphogen of HOXA13, also correlates with poor survival of patients with ESCC [29].

Although aberrant expression of *HOXA13* in ESCC has thus been established, little is known regarding the functional consequences thereof. One study investigated the molecular targets of *HOXA13* in a cancer cell model of ESCC by CHIP-DSL and identified 1938 gene promotors. The targeted genes mostly regulate cell proliferation, survival, and migration [30] and functional assays confirmed that knockdown of *HOXA13* decreased tumor growth *in vivo* and colony formation of ESCC cell lines *in vitro* [26]. Similarly, elevated HOXA13 expression promoted the proliferation and metastasis of gastric cancer partly via activating Erk1/2 [31] while downregulation of HOXA13 sensitizes human ESCC to chemotherapy [32].

Although *HOXA13* seems to play a prognostic role when esophageal cancer has already been established, it remains unknown if there is a causal relationship between *HOXA13* and ESCC and whether this factor can drive neoplastic transformation. Advancement of high-throughput genomic technologies has led to a better understanding of the molecular basis of ESCC development [33, 34]. ESCC and even its precursor lesion are highly mutated and heterogeneous diseases, but early events of ESCC are not completely clear. The present study aimed to examine to what extent aberrant *HOXA13* might drive oncogenic hallmarks in esophageal keratinocytes. To this end, we overexpressed *HOXA13* in a non-transformed human esophageal cell line, performed gene expression profiling to identify key processes and functions, and employed functional experiments to study the role of *HOXA13* in keratinocytes.

METHODS

Cell line

EPC2-hTERT cells [35] are normal hTERT immortalized human esophageal keratinocytes. Cells were routinely cultured in keratinocyte–serum-free medium (KSFM) without calcium chloride (CaCl2) (17005042, Gibco), supplemented with 50 µg/ml bovine pituitary extract (BPE)(129-5, Cell Applications), 1 ng/ml human recombinant epidermal growth factor (EGF) (E9644-.2 Sigma) and Penicillin-Streptomycin (100u/ml, Gibco). Cell line identity was confirmed with short tandem repeats

(STR) analysis by DSMZ and cells were routinely checked for Mycoplasma infection (Eurofins, Ebersberg, Germany).

Generation of EPC2-hTERT *HOXA13* overexpression model

Amplification of the human *HOXA13* gene including its single intron was performed with Q5 polymerase using primers (AgeI HoxA13 F; GTGGTACCGGTGCCACCATGACAGCCTCCGTGCTCCT, and XbaI HoxA13 R; ACCACCTCTAGATTAACTAGTGGTTTTCAGTT). The gene was cloned into pEN_TmiRc3 using AgeI and XbaI restriction sites, a gift from Iain Fraser (Addgene #25748, Cambridge, USA) [36]. Subsequently, two plasmids with and without the *HOXA13* insert were prepared. The *HOXA13* insert was transferred into pSLIK-Venus, using a Gateway reaction [37]. pSLIK-Venus was a gift from Iain Fraser (Addgene #25734) [36]. Both plasmids were sequenced by LGC Genomics (Teddington, UK). Next, they were packaged into lentiviral particles following transfection in HEK293T cells with third generation packaging plasmids. The supernatant was collected and ultracentrifuged. EPC2-hTERT cells were transduced with the virus and YFP (pSLIK-Venus) positive cells were sorted by Fluorescence-Activated Cell Sorting (FACS; BD FACSCantoTM II, BD Biosciences, San Jose, CA). These cells were grown and analyzed as a heterogeneous cell pool. Cells transduced with control vector are hereafter called 'control', while cells transduced with the HOXA13-containing plasmid are denoted as HOXA13^{+} cells. While *HOXA13* gene expression was supranormally induced by 1.25 µg/ml doxycycline in the culture medium, 'leakage' of the vector caused *HOXA13* overexpression even in absence of doxycycline (**Supplementary Figure S1A**) [38, 39]. Doxycyclin itself affected growth of EPC2-hTERT cells (**Supplementary Figure S1B**). For this reason, doxycycline was not added to functional assays with longer timepoints.

RNA isolation

RNA was isolated using the NucleoSpin RNA isolation kit (Macherey Nagel, Düren, Germany). RNA concentrations were measured using a Nanodrop spectrophotometer and samples were stored in RNA storage solution (Sodium Citrate pH 6.4), obtained from Ambion (Foster City, USA) and kept at -80 °C. RNA integrity and quantity were determined by the Agilent 2100 Bioanalyzer.

RNA-Seq

The EPC2-hTERT samples (N=8) were prepared with the TruSeq Stranded mRNA Library Prep Kit. Sequencing took place according to the Illumina TruSeq v3 protocol on an Illumina HiSeq2500 sequencer. 50 base-pairs reads were generated and mapped against reference genome hg19 with Tophat (version 2.0.10). Expression was quantified using HTseq-count (0.6.1). Data were processed using R. version 3.2.5, [40] module DeSeq2 [41]. Generated fold changes (FCs) and p values were

analyzed using ingenuity pathway analysis (IPA) (QIAGEN Inc., https://www.qiagenbioinformatics.com/products/ingenuitypathway-analysis) [42].

Only differentially expressed genes with a p-value <0.05 in RNA-Seq were used as input data. In IPA analysis, *p*-value (calculated using a Right-Tailed Fisher's Exact Test) reflects the likelihood that the association or overlap between a set of significant molecules from the experiment and a given process/pathway/transcription neighborhood is due to random chance. The smaller the p-value, the less likely that the association is random. The *p*-value does not consider the directional effect of one molecule on another or the direction of change of molecules in the dataset. Z-scores, a statistical measure of correlation between relationship direction and gene expression were considered significant when > 2 or < -2. Z score takes into account the directional effect of one molecule on another molecule or on a process and the direction of change of molecules in the dataset. Canonical pathway analysis identified the pathways most significant to the data set, based the ratio of the number of proteins from the data set that map to a pathway divided by the total number of proteins assigned to this canonical pathway.

FACS

EPC2-hTERT cells were stained with 5µl of Anti-human HLA-ABC (APC) antibody per 50µl (Clone W6/32, eBioscience, #17-9983-41) in 2% mouse serum, for 15 minutes at room temperature. Cells were analyzed on a FACSCanto II flow cytometer (BD Biosciences, San Jose, CA) and analyzed with FlowJo v10 (FLOWJO, LLC).

MTT assay

MTT assays were performed as previously described [43]. Transduced EPC2-hTERT cells were seeded in a 96-wells plate, 1000 cells/well. After 24 h, 3, 5 and 7 days 10µL of 5mg/mL MTT reagent (Sigma-Aldrich Chemie BV) was added to 100µL of culturing medium. After 3h of incubation at 37°C, medium was replaced by dimethyl sulfoxide (DMSO; Sigma-Aldrich). OD was measured in a Model 680 XR microplate reader (Bio-Rad, USA). This experiment was repeated three times.

Cell adhesion test

EPC2-hTERT cells were in seeded in 96 well plate (20000 cells per well). After 60 mins, 90 mins, 2h, 3h, 4h, 6h unattached cells were removed from the wells and counted by hemocytometer with Trypan Blue (Sigma-Aldrich Chemie BV). This experiment was repeated four times.

3D culture EPC2-hTERT cells

3D culturing of EPC2-hTERT cells was performed as previously described [44]. 4000 EPC2-hTERT cells were seeded in 50µL drop of ice-cold 1:1 mixture of Matrigel basement membrane matrix (Corning BV) with culture medium in a 24 well plate for cell suspension, and incubated at 37℃ for 30 minutes. After solidification, 500 µL of medium was added supplemented with 0.6mM CaCl2. Y27632 (10µM) was included in medium only first 24h after seeding. Medium was refreshed every 2-3 days. Pictures were made every three days. Morphological assessment was performed on day 12. Differentiated spheroids were characterized by at least three layers of prolonged cells and a nuclei-free mass in the middle, undifferentiated spheroids had round nuclei and lacked the cell-free area in the center. The area of the spheroids was measured with FIJI [45] on photographs taken on day 2, 5 and 8 of culture.

Histology and immunohistochemistry of 3D culture

EPC2-hTERT spheroids were fixed in 4% formaldehyde for 7 mins on day 11, washed with PBS, put in 2% agarose, and embedded in paraffin. Then 4µM slices were sectioned for H&E and immunohistochemistry staining.

For IHC, slides were deparaffinized and rehydrated followed by sodium citrate antigen retrieval (microwaved for 15 min at 200 Watt). Then they were blocked with Goat serum diluted 1:10 in PBS and incubated overnight at 4℃ with anti-IVL (mouse monoclonal anti-IVL I9018, 1:500; Sigma-Aldrich) or anti-CK19 (rabbit monoclonal anti-cytokeratin-19 EP72, 1:100, BSB 5382, ITK Diagnostic). After this, secondary antibodies (Dako EnVision+System-HRR labeled Polymer Anti Mouse, Dako) were applied for 30 mins at RT. Next, slides were counterstained with hematoxylin for 10s, dehydrated, and mounted with Pertex. Stained objects were captured and imaged with Axiovert 40 CFL Zeiss microscope (20x objective), Leica DFC400 digital camera and Leica Application Suite software (Leica Microsystems). Quantification was based on the percentage of positive cells and intensity of the staining (scores ranged from 0, 2 to 9).

2D Migration assays

2D migration assay were performed as previously described [46]. Sterile coverslips placed in an Attofluor incubation chamber were coated with gelatin (1 mg/ml) and incubated for 1 h at 37°C, prior to cell seeding. A removable circular sterile migration barrier was inserted into the chamber, which prevents cell growth in the center of the coverslip. 2.5×10^5 EPC2-hTERT *HOXA13* overexpression and control cells were seeded around this barrier and the rings were incubated at 37°C for 24 h. A confluent monolayer grew in the periphery and a cell-free area was present in the center of the coverslip. After removing the migration barrier, time-lapse imaging was conducted at 37°C under humidified 5% CO2 airflow for 24 h on an Axiovert 100M inverted microscope, equipped with an AxioCam MRC digital camera, using a 10X/0.30 Plan-Neofluar objective (Carl Zeiss B.V., Sliedrecht, Netherlands). ‘Total migration’ is the net track movement of cells in 24h, ‘effective migration’ is the directional movement of

cells to the cell-free center of the coverslip. Migration efficiency was determined as the percentage of directional movement over the total track distance. Velocity was defined as distance per hour. For each cell line, at least three independent migration assays were performed, data of one representative experiment are depicted.

3D-migration using cell dispersion assay

The procedure was performed as described before [47]. Cytodex-3 microcarrier beads (Sigma–Aldrich) were mixed with 5×10^5 EPC2-hTERT *HOXA13* overexpression and control cell suspensions, which constitutes a density of 40 cells per bead. These suspensions were incubated at 37°C for 6 h with gentle mixing. The bead suspension was transferred to a 25 cm^2 tissue culture flask and incubated for 48 h to ensure complete coating of beads and to remove unattached cells. Coated beads were embedded in 1.6 mg/ml collagen gel (collagen: modified Eagle's medium: 7.5% w/v $NaHCO_3$ in the ratio 11:8:1) in a 24-well plate such that each well had approximately 150 beads. Plates were incubated at 37°C for 2 h for the beads to settle in the gel and the polymerized gels were covered with 500 µl DMEM, 10% FBS, 1% p/s. Cell dispersion was measured as the maximum migrated distance from the surface of the bead into the collagen gel. All measurements were performed using AxioVision 4.5 software and assays were performed twice with ten beads per group. Two-way analysis of variance was performed to calculate *p*-values.

Phosphoprotein profiling

EPC2-hTERT control and *HOXA13* transduced cells were seeded in a 6 well plate. When they reached 80-90% of confluency, total proteins were extracted in 300 µL Laemmli Buffer [SDS 4%, glycerol 20%, Tris-Cl (pH 6.8) 120 mM, bromophenol blue 0.02% (w/v) and DTT 0.1 M] and the protein concentrations were measured using RC DC Protein Assay (Bio Rad). Western blotting was performed as described before [48, 49]. Briefly, proteins were resolved by SDS-PAGE and blotted onto Immobilon FL PVDF membranes (Millipore, Bedford, MA, USA). Membranes were blocked in Odyssey Blocking Buffer (Thermo Fisher Scientific) and incubated overnight at 4°C with primary antibody (See **Table S1** for details), followed by the appropriate Alexa-linked secondary antibodies, at 1:5 000 dilution, in Odyssey Blocking Buffer for 1 h. The fluorescent bands were detected using fluorescent Odyssey Imaging System and densitometric analysis was performed with Image Studio Lite Ver.5.2 [50]. All blots were reprobed for Actin to control for equal loading and normalized results are represented as ratios of protein of interest over Actin levels per lane. Three independent experiments were performed, run together on one blot, and heat maps of the phospho-protein profile in the 6 samples were constructed with CIMminer (Genomics and Bioinformatics Group, Laboratory of Molecular Pharmacology, Center for Cancer Research, National Cancer Institute) [51]. For some samples, more than one western blot was run for particular phospho-proteins – in this case the mean for that particular sample was used for heatmap preparation.

Drug sensitivity assay

10 000 EPC2-hTERT cells per well were seeded in 96 well plate. Next day, 10µL of chemical compounds were added to 90µL of cell culture medium and added to cells (see **Table S2** for the information on compounds and range of its serial dilution). The final concentration of solvents (DMSO, Ethanol or dH2O) was 1%. Appropriate controls for solvents were made. After incubation for 72 h, MTT test was performed as described above. Each concentration was tested in quadruplicates, and experiments were performed at least three times for each drug (**Table S2**).

Statistics

A two-way analysis of variance (ANOVA) was used to test for significant differences at each time point in the MTT assay, measurement of area of spheroids, and for 3D migration assay (Graphpad Prism 5; GraphPad Software Inc., USA). For the comparison of the level of MHC class I, IHC score, Western Blot data, and for the 2D migration assay a t-test or Mann-Whitney test was used based on the result of normality test (either a Komogorov-Smirnov test, the D'Agostino, Pearson omnibus normality test or the Shapiro-Wilk normality test). P-values <0.05 were considered to be statistically significant. Statistical analyses of proportions were performed with "N-1" Chi-squared test using MedCalc for Windows, version 18.11.3 (https://www.medcalc.org/calc/comparison_of_proportions.php.), MedCalc Software, Ostend, Belgium).

RESULTS

HOXA13 alters keratinocyte gene expression profiles

In order to investigate the tumor-initiating role of *HOXA13* in ESCC, we overexpressed this transcription factor in the primary immortalized esophageal squamous epithelial cell line EPC2-hTERT, which has low endogenous levels of *HOXA13*, and determined the ensuing molecular consequences by RNAseq profiling. A log2FC of 5.24 confirmed successful overexpression of *HOXA13* ($p<3.14$ E-215). This affected 2995 genes: 1745 (58.3%) were downregulated and 1250 upregulated ($p < 0.05$) (log2 fold change (FC) ranging from -5.28 to 4.97). The top 20 of *HOXA13*-induced differentially expressed genes and functions of their products are reported in **Table 1**. Upregulation of *ANPEP, MAGEA11, LCP1, CSAG1, CSAG1, ZNF486, MAGEA12, GPC4, CYP24A1, LRRC38* was observed, while *UBR1, PSMB8, UBR1, PSMB8, EPSTI1, SAMD9L, APOL6, TLR3, GBP1, SLC12A7, HS6ST2, SFRP1* are down-regulated. Subsequently, *in silico* functional enrichment analysis was performed for all differentially expressed genes. Canonical pathway analysis indicates that *HOXA13* influenced both metabolic and signaling pathways. The top canonical pathways affected include Antigen Presentation Pathway, Molecular Mechanisms Of Cancer, Epithelial Adherent Junction Signaling and 14-3-3 protein-Mediated Signaling. An extended list of pathways based on Z-scores > 2 and < -2 is shown in **Figure 1A and B**. A clear indication of altered cytoskeletal rearrangement was seen, as evidenced by the signaling by

Rho family GTPases, Rac GTPase signaling and its downstream PAK signaling. IPA prediction indicates that altered transcriptome upon *HOXA13* expression would affect the following molecular and cellular categorical functions: Cell Death And Survival, Cellular Movement, Cellular Assembly And Organization, Cellular Function And Maintenance, Cellular Development. On organismal level *HOXA13* overexpression affects Physiological System Development, Organismal Survival, Organismal Development, and Cardiovascular Development and Function.

Functional analysis identified the toxic functions and diseases that were most significant to the data set. The top three Disease and Disorders categories identified were Cancer (p-value = 7.98E-08 – 1.22E-83, #Molecules = 2739), Organismal Injury And Abnormalities (p-value = 8.08E-08 – 1.22E-83, #Molecules = 2768) and Gastrointestinal Disease (p-value = 8.08E-08 – 7.69E-08 – 9.19E-68, #Molecules = 2533). Z-score for cancerous processes in general was negative (Z-score = -2.507) indicating inhibition of such processes, however, for such categories as Upper Gastrointestinal Tract Tumor, Upper Gastrointestinal Tract Cancer and Squamous-cell Carcinoma, Z-scores were positive (Z-score = 2.451, Z-score = 2.236, Z-score = and 2.157 respectively) indicating activation of these processes (**Table 2, Figure 1C**).Upstream regulator analysis for the 2995 genes involved in the preceding processes was used to identify the potential upstream transcriptional regulators that can explain the observed gene expression changes our dataset [42]. *TP53* (log2FC = 0.032, activation Z-score = -1.723, P = 2,95 E-35), *TGFB1* (log2FC = 0.156, activation Z-score = 1.456, P = 5.88E-33), *TNF-α* (log2FC = 0.155, activation Z-score = -2.773, P= 3.78E-32), *IFNL1* (activation Z-score = -6.991, P= 1.93E-30) and *OSM* (activation Z-score = -2.167) were indicated as the most significant regulators, of which *TP53* [52], *TGFB1* [53] and *TNF*-α [54] have previously been implicated in ESCC pathogenesis. In total, these results suggest a specificity of *HOXA13* for gastrointestinal tumorigenesis and squamous cells carcinomas in particular.

Table 1. Top HOXA13-induced differentially expressed genes

Gene	log2FoldChange	*P* value	Protein function, biological processes
HOXA13	5.46	3.14E-215	
ANPEP	4.97	2.35E-255	membrane alanyl aminopeptidase
MAGEA11	2.55	3.91E-39	part of the androgen receptor signaling pathway, linked to cancer development
LCP1	2.38	3.37E-60	actin binding, actin filament network formation, cell migration
CSAG1	2.13	1.11E-26	unknown, tumor antigen
ZNF486	2.11	6.81E-37	DNA binding, regulation of transcription
MAGEA12	2.10	1.40E-26	protein binding, tumor antigen
GPC4	1.94	6.34E-22	transmembrane receptor, cell proliferation and differentiation
CYP24A1	1.82	3.75E-20	mitochondrial monooxygenase
LRRC38	1.81	6.58E-26	potassium channel regulator, ion transport
UBR1	-3.06	3.14E-293	ubiquitin-protein ligase activity, protein catabolic process
PSMB8	-3.11	3.00E-138	antigen presentation, interferon signaling, protein ubiquitination
EPSTI1	-3.15	2.04E-59	unknown
SAMD9L	-3.72	1.21E-123	protein binding, proliferation, cell division, differentiation
APOL6	-3.86	4.44E-161	lipid binding, lipid transport; lipoprotein metabolic process
TLR3	-3.94	2.06E-110	transmembrane receptor, pathogen recognition and activation of innate immunity
GBP1	-4.16	5.79E-135	guanylate binding, cell response to interferon
SLC12A7	-4.60	1.95E-150	electroneutral potassium-chloride cotransporter, cell volume homeostasis
HS6ST2	-4.64	1.03E-150	heparan sulfate 6-O-sulfotransferase, glycosaminoglycan biosynthesis
SFRP1	-5.28	2.40E-263	cysteine endopeptidase, soluble modulators of Wnt signaling

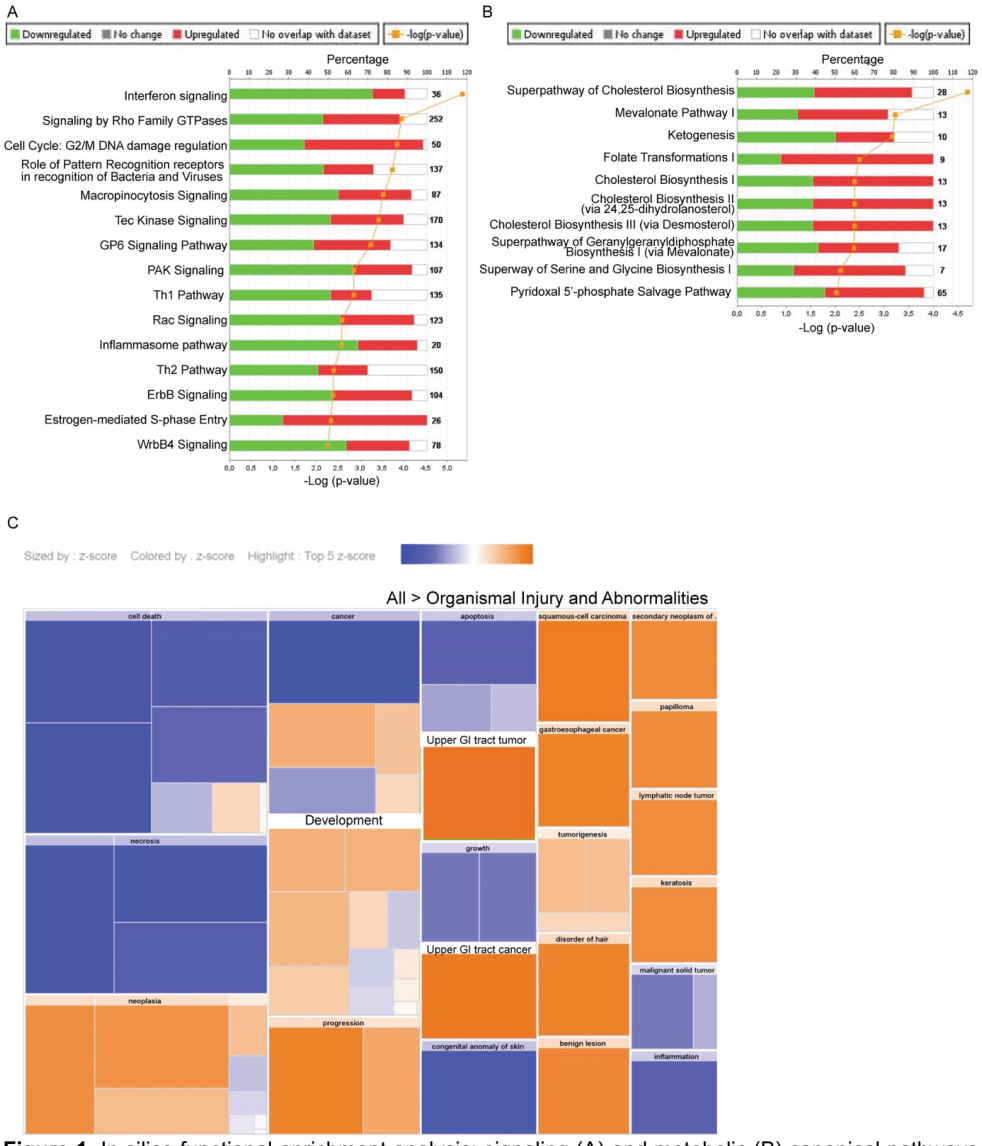


Figure 1. In silico functional enrichment analysis: signaling (A) and metabolic (B) canonical pathways regulated by HOXA13. Z-score > 2 or < -2. (C) Organismal injury and abnormalities caused by HOXA13 overexpression. Downregulated processes by HOXA13 are shown in blue, upregulated processes in orange. General cancerous processes are decreased ('cancer'), while "upper GI tract tumor", "upper GI tract cancer", "squamous cell carcinoma" are increased.

Table 2. IPA predicted toxic functions and diseases caused by HOXA13 overexpression

Categories	P-value	Predicted activation state	Activation Z-score	#Molecules
Cancer	1.80E-76	Decreased	-2.5	2690
Necrosis of tumor	1.83E-08	Decreased	-2.4	135
Cell death of tumor cells	4.35E-08	Decreased	-2.4	131
Necrosis of tumor	1.83E-08	Decreased	-2.4	135
Upper gastrointestinal tract tumor	2.20E-18	Increased	2.5	754
Upper gastrointestinal tract cancer	4.12E-17	Increased	2.2	599
Squamous-cell carcinoma	2.95E-14	Increased	2.2	805
Gastroesophageal cancer	6.94E-13	Increased	2.0	483

HOXA13 influences oncogenic cellular phosphoprofile

Next, we investigated to what extent *HOXA13*-induced transcriptomic changes are translated to altered signal transduction patterns. To this end, we performed phosphoprotein profiling to quantify the expression and activation status of several important signal transduction pathways and targeted some of these pathways with molecular inhibitors. A distinctly altered phosphoprofile was seen upon *HOXA13* overexpression, as evidenced by the clustering of control and overexpressing samples (**Figure 2A, Supplementary Figure S2 for individual western blot examples and quantification**). *HOXA13* overexpressing cells showed inactivation of the tumor suppressor lipid phosphatase PTEN as evidenced by increased inhibitory phosphorylation at Ser380. PTEN is known to inactivate the Akt survival pathway, which was consistent with a non-significant increase in Akt phosphorylation at Ser473, although surprisingly, phosphorylation at Thr308 was decreased. However, *HOXA13* overexpressing cells were significantly less sensitive to inhibition of Akt by treatment of cells with a SHIP2 lipid phosphatase inhibitor [55] (**Figure 2B-i**), suggesting overall enhanced Akt activity levels in these cells.

Significant activation of mitogen-activated growth signaling was also observed through enhanced activation of ERK1/2 (as evidenced by phosphorylation at Thr202/Tyr204) and its target substrate MEK1 (at Thr292). As MEK1 phosphorylation at Thr292 itself inhibits functionality of this protein, we also investigated cell growth in the presence of a MEK1/2 inhibitor **(Figure 2B-ii)**. Overexpression of *HOXA13* did not make cells more sensitive to targeting of MEK1, suggesting that activation of ERK1/2 upon *HOXA13* overexpression exerts its effects via other targets. Indeed, a significant upregulation of the ERK2 substrate 4e-BP1 (Thr70) in *HOXA13* overexpressing cells was seen, which is known to enhance mRNA translation.

In light of the effects seen on cytoskeletal regulation on mRNA level in our sequencing efforts, we also investigated several phospho-proteins involved in actin modulation and adhesion. In particular adhesive properties appeared to be affected by *HOXA13* overexpression, as evidenced by significantly enhanced phosphorylation of Integrin-β3 (Tyr 785) and focal adhesion kinase (FAK, Tyr391). While inhibition of FAK affected control cells more than *HOXA13*-overexpressing cells at low concentrations, higher concentrations indicated enhanced sensitivity upon *HOXA13* overexpression, suggesting that the effect of FAK activity present in the cell may be dichotomous (**Figure 2B-iii**). *HOXA13*-overexpressing EPC2-

hTERT cells were more resistant to inhibition of PAK, a target of the cytoskeletal GTPases Rac and CDC42 (Figure 3B-iv). While we did not observe a clear difference in phosphorylation of PAK1, PAK2 phosphorylation could not be visualized. Taken together, these data suggest that oncogenic signaling is activated upon *HOXA13* overexpression, with, in particular, ERK-induced translational control and cytoskeletal signaling playing important roles.

HOXA13 downregulates MHC class I in keratinocytes

Having demonstrated that *HOXA13* overexpression induces numerous molecular changes associated with tumorigenesis, we next investigated to what extent these changes translate to cellular consequences. Functional enrichment analysis predicted that the antigen presentation pathway is affected (**Figure 3A**), with *HOXA13* regulating expression of genes associated with the major histocompatibility complex (MHC) class I: HLA-A (FC = 0.22), HLA-B (FC = 0.21), HLA-C (FC = 0.57), B2M (FC = 0.55) and TAP1 (FC = 0.63) (**Figure 3B**). To validate this on protein level, EPC2-hTERT cells were stained with antibodies against MHC class I and analyzed by FACS. Confirming RNA sequencing data, *HOXA13* overexpression decreases MHC class I protein expression on EPC2-hTERT cells (FC = 0.17, p = 0.0286) **(Figure 3C).** This suggests activation of immune escape mechanisms upon *HOXA13* expression in esophageal keratinocytes.

HOXA13 overexpression provides a proliferative advantage to esophageal keratinocytes and decreases paclitaxel-induced cell death

Next, we investigated cellular proliferation potential of EPC2-hTERT cells, as a hallmark of tumorigenesis. As shown in **Figure 4A**, induction of *HOXA13* in esophageal keratinocytes does not affect 2D growth as determined by MTT assay. However, in line with RNA analysis predicted decrease in cell death upon *HOXA13* overexpression (**Table 2**), we observed a decreased sensitivity of *HOXA13* overexpressing keratinocytes to paclitaxel, a chemotherapeutic agent targeting the cytoskeleton and used for ESCC treatment (**Figure 4B**).

To gain further insight into the role of *HOXA13* in cell growth, we cultured EPC2-hTERT spheroids in 3D cultures, allowing assessment of growth of individual colonies in a more physiological setting. We found that spheroids derived from EPC2-hTERT cells with *HOXA13* overexpression are 1.5 times bigger in size upon 8 days of culture ($p < 0.05$) (**Figure 5A, B**), indicating a proliferative advantage of cells upon overexpression of *HOXA13*.

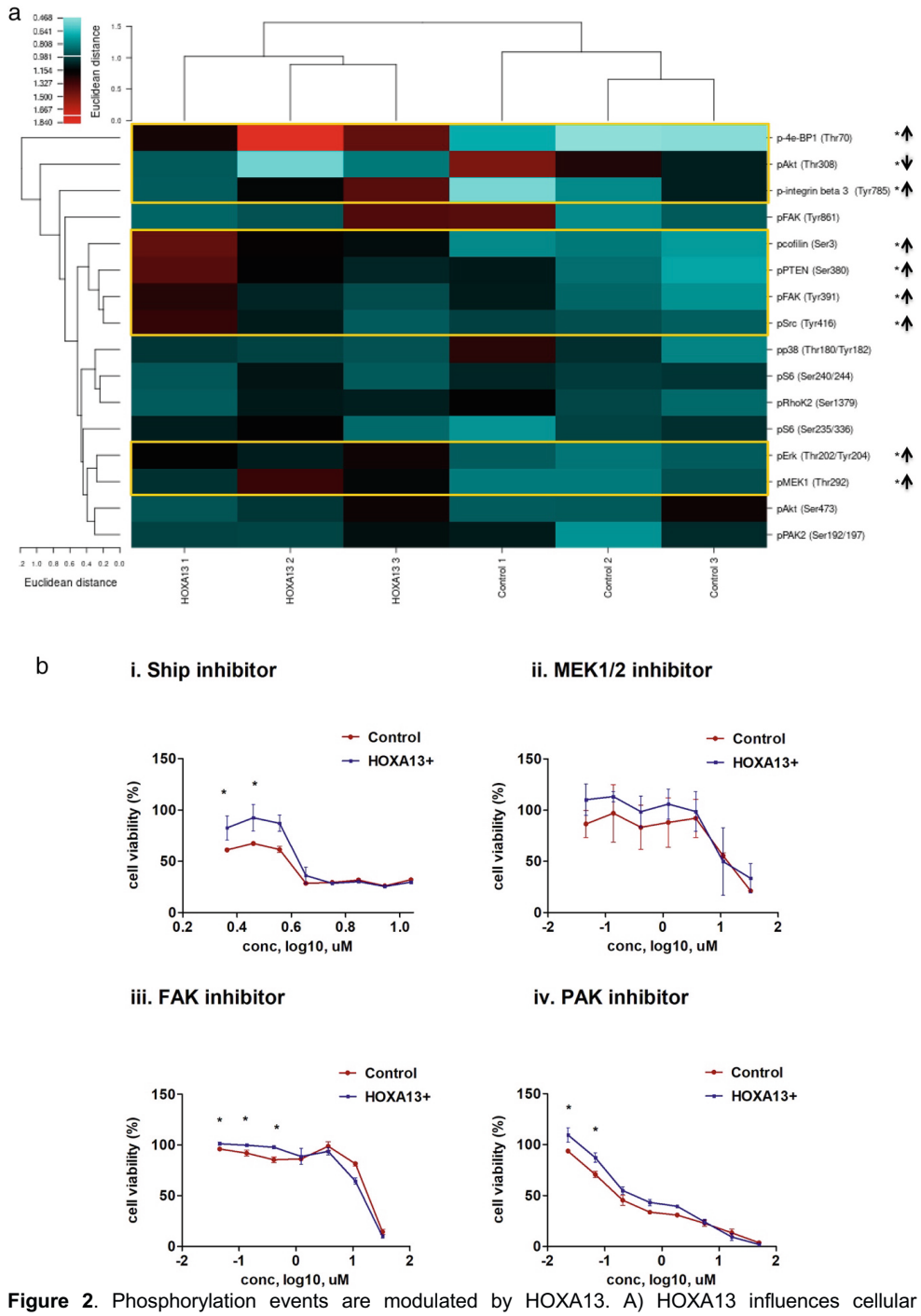
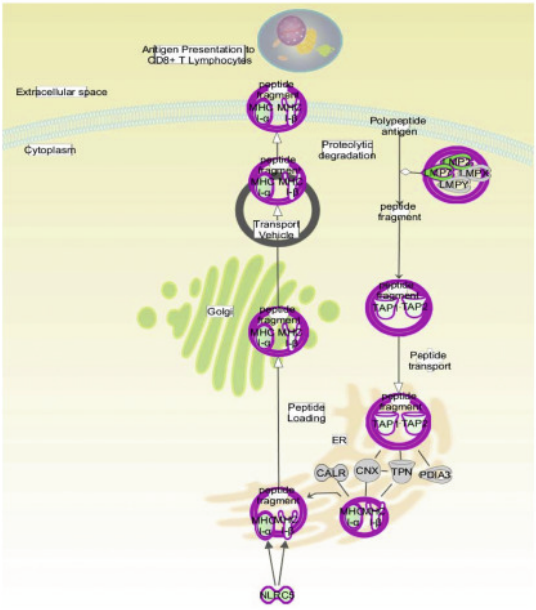


Figure 2. Phosphorylation events are modulated by HOXA13. A) HOXA13 influences cellular phosphoprofile as determined by western blot analysis. Heat map of the phospho-protein profile is shown. Increased phosphorylation is depicted in red, conversely decreased phosphorylation is depicted in blue. Magnitude of the phosphorylation differences is indicated by the scale bar in the top left corner. Statistical significance in phosphorylation status of individual proteins as calculated in Figure S2 is

indicated on the right side of the figure with an asterisk *P < 0.05. B) HOXA13 overexpression causes different sensitivity of keratinocytes to protein activity inhibitors (Ship, MEK, FAK and PAK). Representative figures are shown. *P < 0.05.

A



B

mRNA	Fold change HOXA13+/control	q value
HLA-A	0.22	2.5E-49
HLA-B	0.21	9.1E-34
HLA-C	0.57	2.1E-09
B2M	0.55	1.2E-08
TAP1	0.63	3.0E-05

C

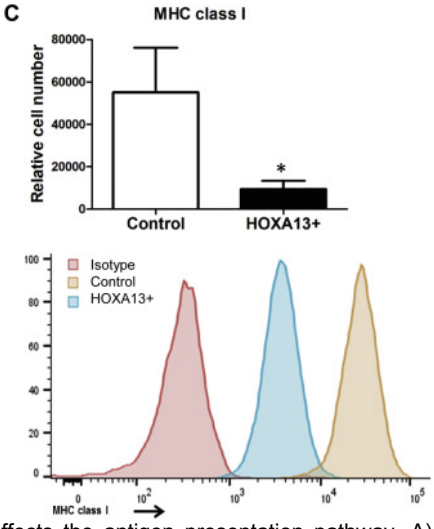


Figure 3. HOXA13 downregulates MHC class I and affects the antigen presentation pathway. A) Illustration of the antigen presentation pathway. Molecules with a decreased expression upon HOXA13 overexpression are indicated in green. IPA analysis based on RNAseq data. B) RNA expression data of genes associated with MHC class I as determined by RNAseq. C) HOXA13 decreases expression protein expression of MHC class I. Quantification of FACS data is shown in upper panel ($p < 0.05$), representative example is shown in lower panel

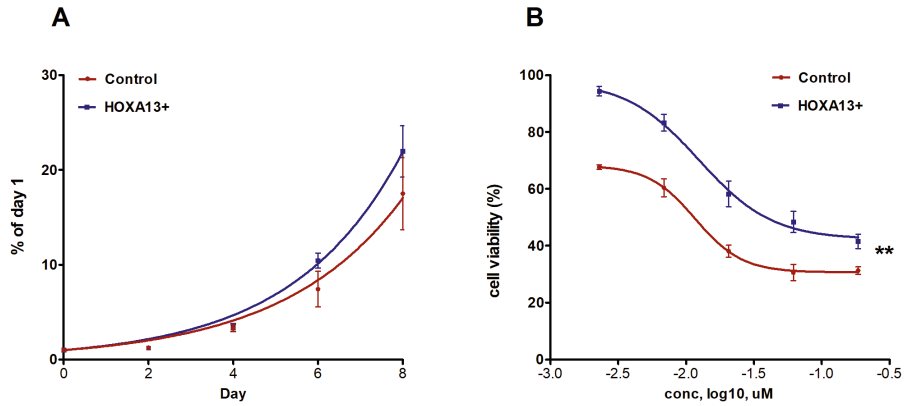


Figure 4. HOXA13 confers resistance to drug-induced cell death. HOXA13 overexpression does not affect 2D growth of esophageal cells (MTT data) (A) but makes them more resistant to paclitaxel-induced death (B). MTT was performed after 3 days of paclitaxel treatment (0.002-0.19μM). Data are presented as % of vehicle-treated corresponding control. Mean±SEM, *P < 0.05.

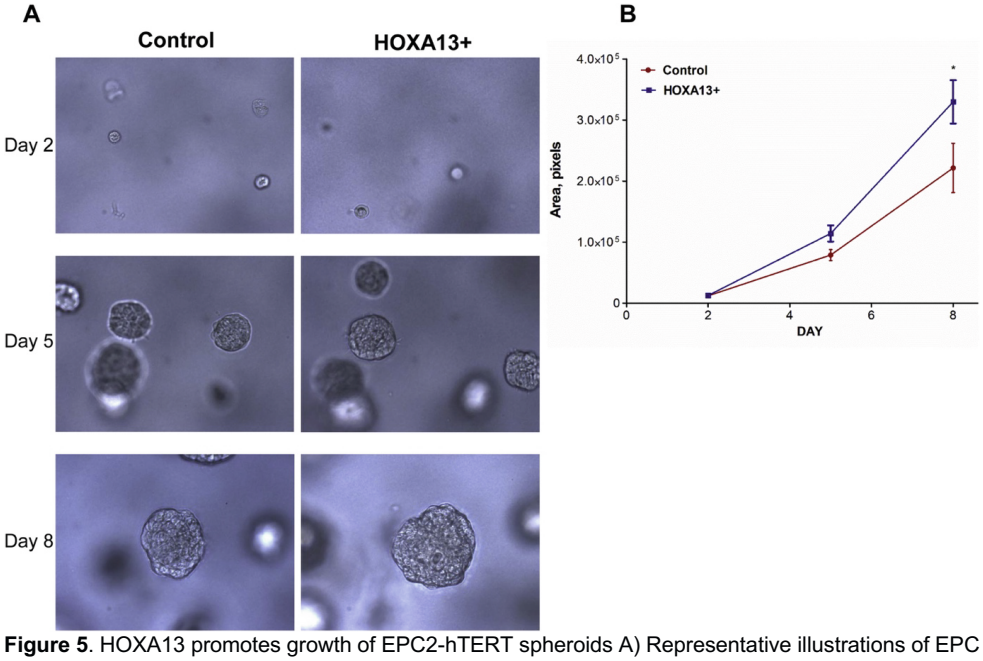


Figure 5. HOXA13 promotes growth of EPC2-hTERT spheroids A) Representative illustrations of EPC spheroids after 2-8 days in 3D culture B) Area of spheroids measured on day 2, 5 and 8, Mean±SEM, *P < 0.05.

HOXA13 overexpression changes the morphology and differentiation status of EPC2-hTERT-derived spheroids

EPC2-hTERT spheroids are characterized by a proliferation-differentiation gradient with Ki-67 staining seen in the basaloid cell layer and more differentiated cells toward the center of these structures [44]. As shown in **Figure 6A**, upon one week of culturing, EPC2-hTERT spheroids become organized in layers with a more flattened cytological aspect in the middle of spheroids, reminiscent of differentiated stratified epithelium. Upon overexpression of *HOXA13*, not only do spheroids become bigger (notice the increased number of nuclei, consistent with a proliferative advantage), they also are kept in a less differentiated state. Quantification of morphology indicates that 50% of spheroids obtain a differentiated morphology in control cells vs 22% upon forced *HOXA13* expression ($p < 0.05$, **Figure 6A, lower panel)**. We also investigated the expression of involucrin and cytokeratin 19 as a markers of differentiation and keratinization. Intensity of both stainings was significantly lower in *HOXA13* overexpressing EPC2-hTERT spheroids (IVL: median IHC score for control = 3, N=28, for *HOXA13*$^+$ cells = 2, N=43, $P < 0.05$, CK19: IHC score for control = 5, N=22, for *HOXA13*$^+$ cells = 0, N=22, $P < 0.05$). Thus, these results indicate that *HOXA13* overexpression prevents the differentiation of keratinocytes, which is consistent with the effect of *HOXA13* seen on cellular assembly and organization as well as on cellular development in IPA analyses.

HOXA13 promotes cellular migration

IPA analysis, as well as phosphoprotein profiling, indicated a direct effect of *HOXA13* on cytoskeletal rearrangement and cellular adhesion. Therefore, we first tested the adhesive strength of EPC2-hTERT cells upon replating, but no difference in time to adhesion was seen between control and overexpressing cells (**Figure 7A)**. More advanced adhesive properties are also required for migration and invasion of cells. Therefore, we investigated a migratory ability of these cells using 2D ring-barrier assays to track individual cell movement. Indeed, *HOXA13* significantly increases total migration, effective migration and velocity of EPC2-hTERT cells in 2D migration assay (**Figure 7**) ($P < 0.01$). Furthermore, a pronounced increase in migratory distance of keratinocytes was seen in 3D cultures upon *HOXA13* overexpression, indicating that *HOXA13* enhances invasive potential of esophageal keratinocytes ($p = 0.0038$) (**Figure 8 A,B**).

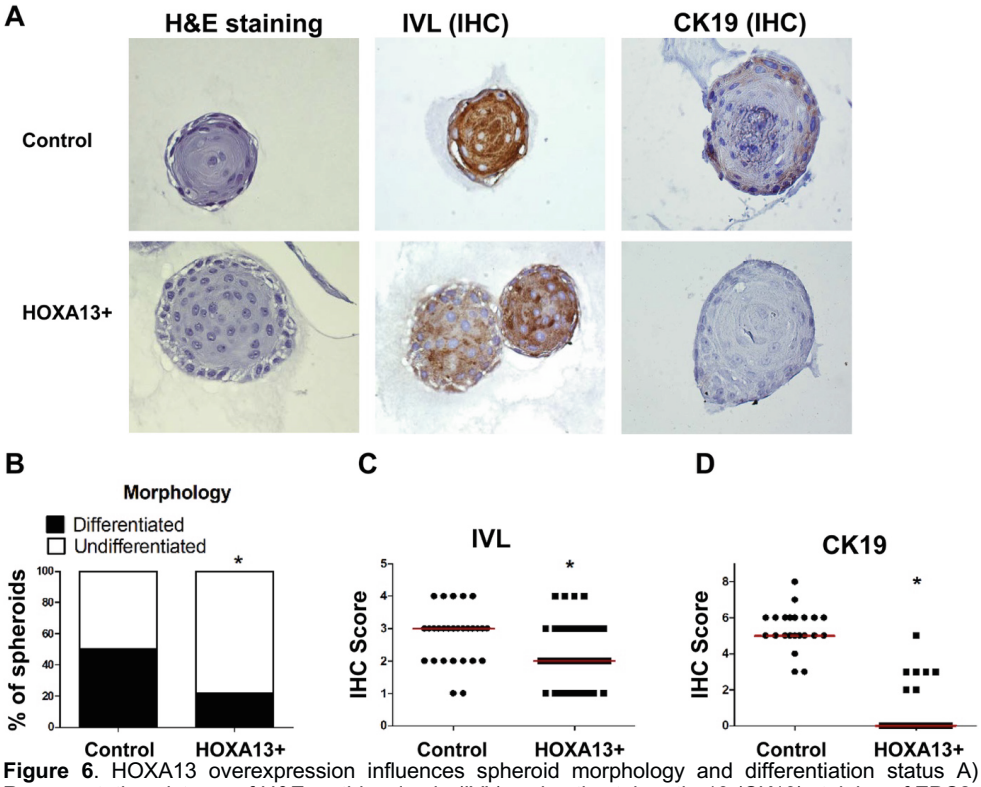


Figure 6. HOXA13 overexpression influences spheroid morphology and differentiation status A) Representative pictures of H&E, anti-involucrin (IVL) and anti-cytokeratin-19 (CK19) staining of EPC2-hTERT spheroids. B) Spheroids with a more differentiated aspect are more frequently observed for control EPC2-hTERT spheroids compared to HOXA13-overexpressing EPC2-hTERT spheroids (*P < 0.05) C) EPC2-hTERT spheroids with HOXA13 overexpression express less involucrin (IVL) and D) cytokeratin 19 (CK19) (Median with range,*P < 0.05).

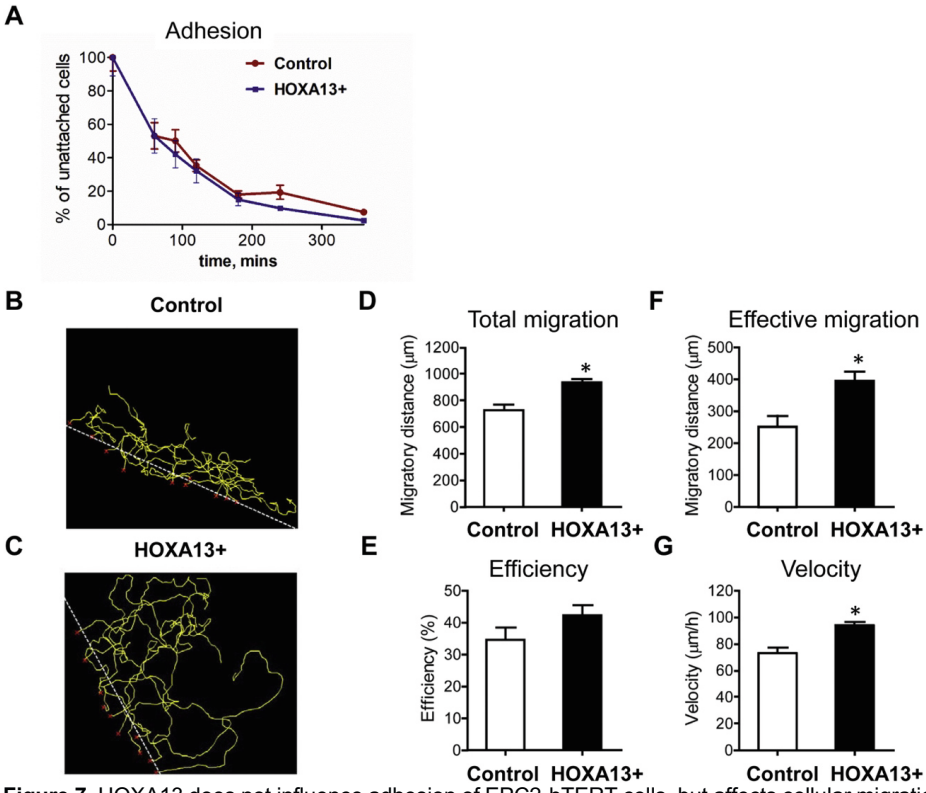


Figure 7. HOXA13 does not influence adhesion of EPC2-hTERT cells, but affects cellular migration. A) Adhesive properties of keratinocytes per se are not affected by HOXA13 overexpression. B-G) HOXA13 promotes cellular migration in 2D migration assay as determined by: D) total migration E) efficiency, F) effective migration and G) velocity of migration. Mean±SEM, **P < 0.01. Representative pictures of Individual cell tracking show cell migration in 24h time-lapse microscopy of control cells (B) and HOXA13-overexpressing cells (C).

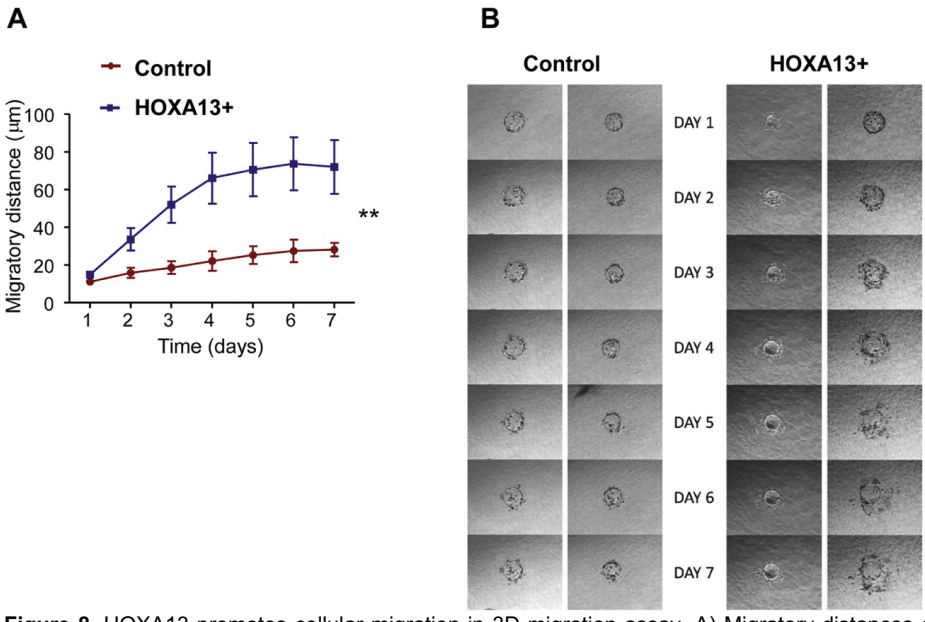


Figure 8. HOXA13 promotes cellular migration in 3D migration assay. A) Migratory distances of cell invading collagen matrix were measured for HOXA13-overexpressing cells and control cells over 7 days, Mean±SEM, **P < 0.01. B) representative photographs of cells migrating in collagen matrix from microcarrier beads in 7 days.

DISCUSSION

A better understanding of the molecular mechanisms behind ESCC development might reveal new targets for its treatment and early diagnosis. Homeobox genes were not only shown to be responsible for proper embryonic development and differentiation of stem cells but they are also associated with cancer development [56]. One of these genes, transcriptional factor HOXA13, has previously been investigated in human ESCC tissue and in cancerous cell lines [26, 27]. However, its role was not reported for early stages of ESCC or for squamous dysplasia. In this study, we overexpressed *HOXA13* in primary immortalized esophageal keratinocytes and compared them to empty vector-transduced controls in terms of hallmarks of cancer to investigate *HOXA13* as a driver of esophageal carcinogenesis. Initiation of cancer implies the cellular acquisition of several oncogenic characteristics, including selective growth and proliferative advantage, altered stress response favoring overall survival, vascularization, invasion and metastasis, metabolic rewiring, an abetting microenvironment, and immune modulation [57, 58]. Furthermore, cancerous cells are characterized by some level of dedifferentiation and heterogeneity [59].

First of all, we found that *HOXA13* downregulates MHC class I in keratinocytes and affects an antigen presentation pathway. Downregulation of MHC class I and escape from immune response is associated with the clinical course of ESCC [60]. Our results indicate that *HOXA13* expression may drive the immune escape of neoantigen-bearing transformed keratinocytes.

Second, we observe that overexpression of *HOXA13* provides a proliferative advantage to keratinocytes and decreases their sensitivity to paclitaxel-induced cell death. Upon overexpression of *HOXA13*, cells showed increased resistance to paclitaxel treatment and formed bigger spheroids in 3D culture. These data are in line with clinical data previously obtained [32] showing that downregulation of *HOXA13* sensitizes human ESCC to chemotherapy and with experiments done on cancerous cell lines showing that HOTTIP/HOXA13 enhances ESCC cell proliferation *in vitro* and *in vivo* [27]. However, while these earlier publications suggest that *HOXA13* plays a role in the maintenance of tumor cell characteristics, our data suggest that overexpression of this gene can drive tumorigenesis. Nevertheless, thus far, no activating *HOXA13* mutations have been reported for ESCC, suggesting alternative mechanisms for its over-expression in this tumor type. IPA analysis of our dataset revealed *TGF-β1* and *TP53* as the most significant regulators of signaling pathways affected by *HOXA13* overexpression. Mutation of *TP53* is an early and most common event in esophageal carcinogenesis and is typical also for squamous dysplasia and other types of esophageal cancer [34]. TGF-β1 also plays an important role in pathogenesis of ESCC [61] as it regulates epithelial-to-mesenchymal transition (EMT) of ESCC. As *HOXA13* was reported to induce the EMT cascade [27], it is conceivable that *in vivo*, TGF- β1 drives this effect.

Losing epithelial traits (dedifferentiation) is an important step during tumorigenesis which at early stages is required for local migration/invasion and at later stages contributes to macroscopic metastases [62]. Moreover, differentiation status of ESCC is associated with clinical outcome [63, 64]. The prognosis of patients with keratinizing ESCC has been reported to be significantly better than that of patients with non-keratinizing tumors [63]. In the present study, RNA expression data, morphology data on EPC2-hTERT spheroids and anti-IVL staining all indicate that *HOXA13* limits the differentiation of keratinocytes. Concomitantly, we observed that *HOXA13* promotes cellular migration, which is in line with data from cancerous cell lines [27]. Our study further suggests that Integrin-β3, FAK, PAK, GTPases Rac, and CDC42 are likely candidates to be involved and mediate this effect.

We acknowledge several limitations to our study. Tumors are heterogeneous, and HOXA13 may not play a similar role in all patients developing ESCC. Here, we employed a heterogeneous pool of EPC2-hTERT cells, but while cell lines are by nature heterogeneous, this does not fully reflect the heterogeneity of patients. Future studies are needed to confirm HOXA13 overexpression in esophageal premalignant lesions. Second, while our studies implicate a role for HOXA13 in driving oncogenic hallmarks, and previous publications have clearly shown the overexpression of HOXA13 in malignant tissues, the driving mechanisms for this overexpression remain to be elucidated.

CONCLUSIONS

In toto, we show here that *HOXA13* expression confers oncogenic hallmarks to esophageal keratinocytes. It provides proliferation advantage to keratinocytes, reduces sensitivity to chemical agents, regulates MHC class I expression and differentiation status and promote cellular migration. Our data indicate a crucial role of *HOXA13* at early stages of esophageal carcinogenesis.

ACKNOWLEDGEMENTS

This work was supported by the Stichting Proefdiervrij. Netherlands.

REFERENCES

1. Pennathur, A., et al., *Oesophageal carcinoma.* Lancet, 2013. 381(9864): p. 400-12.
2. Kamangar, F., G.M. Dores, and W.F. Anderson, *Patterns of cancer incidence, mortality, and prevalence across five continents: defining priorities to reduce cancer disparities in different geographic regions of the world.* J Clin Oncol, 2006. 24(14): p. 2137-50.
3. Ferlay, J., et al., *Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012.* Int J Cancer, 2015. 136(5): p. E359-86.
4. Pickens, A. and M.B. Orringer, *Geographical distribution and racial disparity in esophageal cancer.* Ann Thorac Surg, 2003. 76(4): p. S1367-9.
5. Domper Arnal, M.J., A. Ferrandez Arenas, and A. Lanas Arbeloa, *Esophageal cancer: Risk factors, screening and endoscopic treatment in Western and Eastern countries.* World J Gastroenterol, 2015. 21(26): p. 7933-43.
6. Taylor, P.R., C.C. Abnet, and S.M. Dawsey, *Squamous dysplasia--the precursor lesion for esophageal squamous cell carcinoma.* Cancer Epidemiol Biomarkers Prev, 2013. 22(4): p. 540-52.
7. Pennathur, A., et al., *Esophagectomy for T1 esophageal cancer: outcomes in 100 patients and implications for endoscopic therapy.* Ann Thorac Surg, 2009. 87(4): p. 1048-54; discussion 1054-5.
8. Copeland, N.G., N.A. Jenkins, and D.L. Court, *Recombineering: a powerful new tool for mouse functional genomics.* Nat Rev Genet, 2001. 2(10): p. 769-79.
9. Knosp, W.M., et al., *HOXA13 regulates the expression of bone morphogenetic proteins 2 and 7 to control distal limb morphogenesis.* Development, 2004. 131(18): p. 4581-92.
10. Lewis, E.B., *A gene complex controlling segmentation in Drosophila.* Nature, 1978. 276(5688): p. 565-70.
11. Noordermeer, D., et al., *The dynamic architecture of Hox gene clusters.* Science, 2011. 334(6053): p. 222-5.
12. Mallo, M., D.M. Wellik, and J. Deschamps, *Hox genes and regional patterning of the vertebrate body plan.* Dev Biol, 2010. 344(1): p. 7-15.
13. Yahagi, N., et al., *Position-specific expression of Hox genes along the gastrointestinal tract.* Congenit Anom (Kyoto), 2004. 44(1): p. 18-26.
14. McGinnis, W. and R. Krumlauf, *Homeobox genes and axial patterning.* Cell, 1992. 68(2): p. 283-302.
15. Pearson, J.C., D. Lemons, and W. McGinnis, *Modulating Hox gene functions during animal body patterning.* Nat Rev Genet, 2005. 6(12): p. 893-904.
16. Hueber, S.D., et al., *Analysis of central Hox protein types across bilaterian clades: on the diversification of central Hox proteins from an Antennapedia/Hox7-like protein.* Dev Biol, 2013. 383(2): p. 175-85.
17. Abate-Shen, C., *Deregulated homeobox gene expression in cancer: cause or consequence?* Nat Rev Cancer, 2002. 2(10): p. 777-85.
18. Reya, T. and H. Clevers, *Wnt signalling in stem cells and cancer.* Nature, 2005. 434(7035): p. 843-50.
19. Taipale, J. and P.A. Beachy, *The Hedgehog and Wnt signalling pathways in cancer.* Nature, 2001. 411(6835): p. 349-54.
20. Joo, M.K., J.-J. Park, and H.J. Chun, *Impact of homeobox genes in gastrointestinal cancer.* World Journal of Gastroenterology, 2016. 22(37): p. 8247-8256.
21. Chen, K.-N., et al., *Expression of 11 HOX Genes Is Deregulated in Esophageal Squamous Cell Carcinoma.* Clinical Cancer Research, 2005. 11(3): p. 1044.
22. Liu, C., et al., *Long Non-coding RNA DLEU1 Promotes Proliferation and Invasion by Interacting With miR-381 and Enhancing HOXA13 Expression in Cervical Cancer.* Front Genet, 2018. 9: p. 629.
23. Yu, H., et al., *Long noncoding RNA LUCAT1 promotes malignancy of ovarian cancer through regulation of miR-612/HOXA13 pathway.* Biochem Biophys Res Commun, 2018. 503(3): p. 2095-2100.
24. Han, Y., et al., *HOXA13 contributes to gastric carcinogenesis through DHRS2 interacting with MDM2 and confers 5-FU resistance by a p53-dependent pathway.* Mol Carcinog, 2018. 57(6): p. 722-734.
25. Dong, Y., et al., *HOXA13 is associated with unfavorable survival and acts as a novel oncogene in prostate carcinoma.* Future Oncol, 2017. 13(17): p. 1505-1516.
26. Gu, Z.-D., et al., *HOXA13 Promotes Cancer Cell Growth and Predicts Poor Survival of Patients with Esophageal Squamous Cell Carcinoma.* Cancer Research, 2009. 69(12): p. 4969.
27. Lin, C., et al., *Transcriptional and posttranscriptional regulation of HOXA13 by lncRNA HOTTIP facilitates tumorigenesis and metastasis in esophageal squamous carcinoma cells.* Oncogene, 2017. 36: p. 5392.
28. Ma, R.L., L.Y. Shen, and K.N. Chen, *Coexpression of ANXA2, SOD2 and HOXA13 predicts poor prognosis of esophageal squamous cell carcinoma.* Oncol Rep, 2014. 31(5): p. 2157-64.
29. Barclay, C., et al., *Basic fibroblast growth factor (FGF-2) overexpression is a risk factor for esophageal cancer recurrence and reduced survival, which is ameliorated by coexpression of the FGF-2 antisense gene.* Clin Cancer Res, 2005. 11(21): p. 7683-91.
30. Shen, L.Y. and K.N. Chen, *Exploration of target genes of HOXA13 in esophageal squamous cell carcinoma cell line.* Cancer Lett, 2011. 312(1): p. 18-23.
31. Qin, Z., et al., *Elevated HOXA13 expression promotes the proliferation and metastasis of gastric cancer partly via activating Erk1/2.* OncoTargets and therapy, 2019. 12: p. 1803-1813.
32. Shi, Q., et al., *Downregulation of HOXA13 sensitizes human esophageal squamous cell carcinoma to chemotherapy.* Thorac Cancer, 2018. 9(7): p. 836-846.
33. Contino, G., et al., *The Evolving Genomic Landscape of Barrett's Esophagus and Esophageal Adenocarcinoma.* Gastroenterology, 2017. 153(3): p. 657-673 e1.
34. Chen, X.-X., et al., *Genomic comparison of esophageal squamous cell carcinoma and its precursor lesions by multi-region whole-exome sequencing.* Nature Communications, 2017. 8(1): p. 524.
35. Harada, H., et al., *Telomerase induces immortalization of human esophageal keratinocytes without p16INK4a inactivation.* Mol Cancer Res, 2003. 1(10): p. 729-38.
36. Shin, K.J., et al., *A single lentiviral vector platform for microRNA-based conditional RNA interference and coordinated transgene expression.* Proc Natl Acad Sci U S A, 2006. 103(37): p. 13759-64.
37. Katzen, F., *Gateway((R)) recombinational cloning: a biological operating system.* Expert Opin Drug Discov, 2007. 2(4): p. 571-89.
38. Merten, O.-W., M. Hebben, and C. Bovolenta, *Production of lentiviral vectors.* Molecular therapy Methods & clinical development, 2016. 3: p. 16017-16017.

39. Kubo, S. and K. Mitani, *A New Hybrid System Capable of Efficient Lentiviral Vector Production and Stable Gene Transfer Mediated by a Single Helper-Dependent Adenoviral Vector.* Journal of Virology, 2003. 77(5): p. 2964.
40. R Development Core Team, *R: A language and environment for statistical computing.* . 2008, Vienna, Austria. .
41. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.* Genome Biol, 2014. 15(12): p. 550.
42. Kramer, A., et al., *Causal analysis approaches in Ingenuity Pathway Analysis.* Bioinformatics, 2014. 30(4): p. 523-30.
43. Queiroz, K.C., et al., *Violacein induces death of resistant leukaemia cells via kinome reprogramming, endoplasmic reticulum stress and Golgi apparatus collapse.* PLoS One, 2012. 7(10): p. e45362.
44. Kasagi, Y., et al., *The Esophageal Organoid System Reveals Functional Interplay Between Notch and Cytokines in Reactive Epithelial Changes.* Cell Mol Gastroenterol Hepatol, 2018. 5(3): p. 333-352.
45. Schindelin, J., et al., *Fiji: an open-source platform for biological-image analysis.* Nat Methods, 2012. 9(7): p. 676-82.
46. Das, A.M., A.M. Eggermont, and T.L. ten Hagen, *A ring barrier-based migration assay to assess cell migration in vitro.* Nat Protoc, 2015. 10(6): p. 904-15.
47. Liu, H., et al., *A microcarrier-based spheroid 3D invasion assay to monitor dynamic cell movement in extracellular matrix.* Biological Procedures Online, 2020. 22(1): p. 3.
48. Somasundaram, R., et al., *Analysis of SHIP1 expression and activity in Crohn's disease patients.* PLoS One, 2017. 12(8): p. e0182308.
49. de Sousa, R.R., et al., *Phosphoprotein levels, MAPK activities and NFkappaB expression are affected by fisetin.* J Enzyme Inhib Med Chem, 2007. 22(4): p. 439-44.
50. Schindelin, J., et al., *The ImageJ ecosystem: An open platform for biomedical image analysis.* Mol Reprod Dev, 2015. 82(7-8): p. 518-29.
51. Myers, T.G., et al., *A protein expression database for the molecular pharmacology of cancer.* Electrophoresis, 1997. 18(3-4): p. 647-53.
52. Shimada, H., *p53 molecular approach to diagnosis and treatment of esophageal squamous cell carcinoma.* Ann Gastroenterol Surg, 2018. 2(4): p. 266-273.
53. Jin, G., et al., *TGFB1 and TGFBR2 functional polymorphisms and risk of esophageal squamous cell carcinoma: a case-control analysis in a Chinese population.* J Cancer Res Clin Oncol, 2008. 134(3): p. 345-51.
54. Tu, Y., et al., *Pristimerin targeting NF-kappaB pathway inhibits proliferation, migration, and invasion in esophageal squamous cell carcinoma cells.* Cell Biochem Funct, 2018. 36(4): p. 228-240.
55. Fuhler, G.M., et al., *Therapeutic potential of SH2 domain-containing inositol-5'-phosphatase 1 (SHIP1) and SHIP2 inhibition in cancer.* Mol Med, 2012. 18: p. 65-75.
56. Bhatlekar, S., J.Z. Fields, and B.M. Boman, *Role of HOX Genes in Stem Cell Differentiation and Cancer.* Stem Cells Int, 2018. 2018: p. 3569493.
57. Fouad, Y.A. and C. Aanei, *Revisiting the hallmarks of cancer.* American journal of cancer research, 2017. 7(5): p. 1016-1036.
58. Hanahan, D. and Robert A. Weinberg, *Hallmarks of Cancer: The Next Generation.* Cell, 2011. 144(5): p. 646-674.
59. Friedmann-Morvinski, D. and I.M. Verma, *Dedifferentiation and reprogramming: origins of cancer stem cells.* EMBO reports, 2014. 15(3): p. 244-253.
60. Liu, Q., et al., *Down-regulation of HLA class I antigen-processing machinery components in esophageal squamous cell carcinomas: association with disease progression.* Scand J Gastroenterol, 2009. 44(8): p. 960-9.
61. Pang, L., et al., *TGF-beta1/Smad signaling pathway regulates epithelial-to-mesenchymal transition in esophageal squamous cell carcinoma: in vitro and clinical analyses of cell lines and nomadic Kazakh patients from northwest Xinjiang, China.* PLoS One, 2014. 9(12): p. e112300.
62. Pattabiraman, D.R. and R.A. Weinberg, *Targeting the Epithelial-to-Mesenchymal Transition: The Case for Differentiation-Based Therapy.* Cold Spring Harb Symp Quant Biol, 2016. 81: p. 11-19.
63. Ohbu, M., M. Saegusa, and I. Okayasu, *Apoptosis and cellular proliferation in oesophageal squamous cell carcinomas: differences between keratinizing and nonkeratinizing types.* Virchows Arch, 1995. 427(3): p. 271-6.
64. Nozoe, T., et al., *Significance of immunohistochemical expression of p27 and involucrin as the marker of cellular differentiation of squamous cell carcinoma of the esophagus.* Oncology, 2006. 71(5-6): p. 402-10.

SUPPLEMENTARY DATA

Supplementary tables

Supplementary Table S1. List of primary antibodies used for phosphoprotein [1]

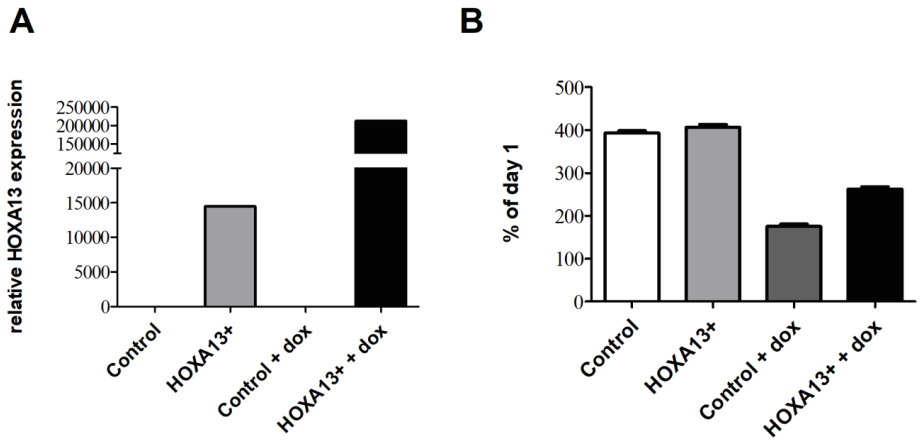
#	Antibody	Company	Cat #	Species	Dilution	Result of phosphorylation on protein activity
1	pErk (Thr202/Tyr204)	CST	4696	mouse	1:1000	+
2	pS6 (Ser235/336)	CST	4856	rabbit	1:1000	+
3	pS6 (Ser240/244)	CST	5364	rabbit	1:1000	+
4	pAkt (Thr308)	Signal way	11055-2	rabbit	1:1000	+
5	pAkt (Ser473)	CST	4060S	rabbit	1:1000	+
6	p-4e-BP1 (Thr70)	CST	9455	rabbit	1:1000	-
7	pp38 (Thr180/Tyr182)	CST	4511	rabbit	1:1000	+
8	pFAK (Tyr391)	Invitrogen	44-625G	rabbit	1:1000	+
9	pFAK (Tyr861)	ITK	21076-1	rabbit	1:1000	+
10	pcofilin (Ser3)	Signal Way	11139	rabbit	1:1000	-
11	pRhoK2 (Ser1379)	Signal Way	13005	rabbit	1:1000	-
12	pSrc (Tyr416)	CST	2113	rabbit	1:1000	+
13	p-integrin beta 3 (Tyr785)	Signal way	11060-1	rabbit	1:1000	-
14	pPAK2 (Ser192/197)	CST	2605	rabbit	1:1000	
15	pMEK1 (Thr292)	Merck	07-348	rabbit	1:1000	-
16	pPTEN (Ser380)	CST	9551	rabbit	1:1000	-
17	β-actin	SCBT - Santa Cruz Biotechnology	477778	mouse	1:1000	N/A

[1] Hornbeck, P.V., et al., PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. Nucleic Acids Res, 2015. 43(Database issue): p. D512-20.

Supplementary Table S2. List of chemicals for drug sensitivity test

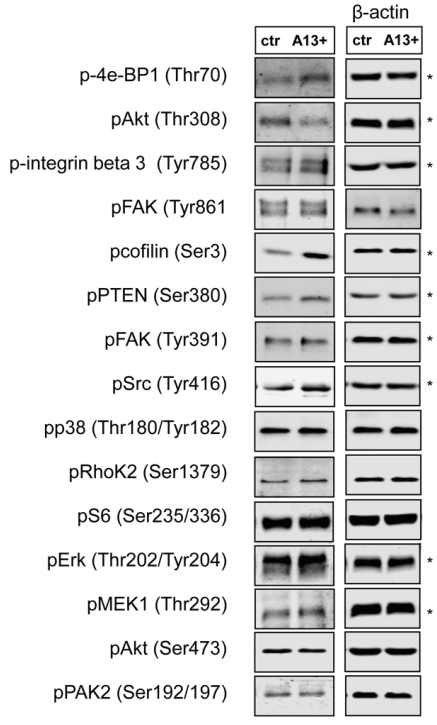
	Compound	Target	Manufacturer	Solvent	Concentration range
1	Paclitaxel	Tubulin	Sigma-Aldrich, 7191-1MG	DMSO	0.002-0.2 uM
2	FAK inhibitor 14	FAK	Sigma-Aldrich, SML0837-10MG	H2O	0.05-33 uM
3	FRAX1036	PAK-1	Selleckchem, S7989	Etanol	0.02-50 uM
4	U0126-EtOH	MEK1/2	MedChemExpress, HY-12031	DMSO	0.05-33 uM
5	2PIQ	SHIP2	Synthesized [2]	DMSO	2.3-11 uM

[2]F2uhler, G.M., et al., Therapeutic potential of SH2 domain-containing inositol-5'-phosphatase 1 (SHIP1) and SHIP2 inhibition in cancer. Mol Med, 2012. 18: p. 65-75.

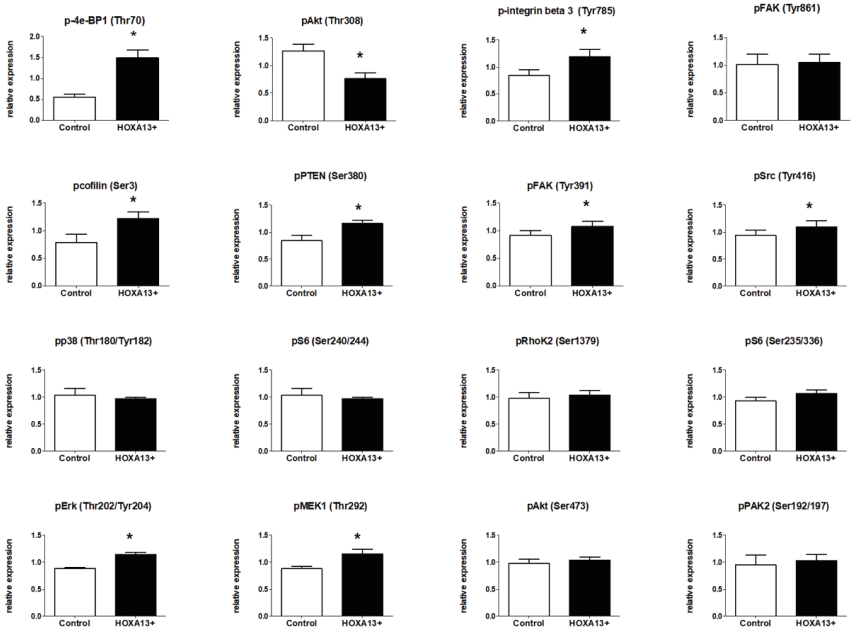


Supplementary Figure S1. A) HOXA13 is overexpressed in EPC2-hTERT HOXA13-transduced cells even without doxycycline (dox) treatment in contrast to control EPC2-hTERT cells (empty vector transduced). qPCR data are calculated relatively to corresponding control with or without dox treatment B) Effect of doxycycline on growth of EPC2-hTERT cells with and without HOXA13 overexpression (MTT). Mean±SEM, *P < 0.05.

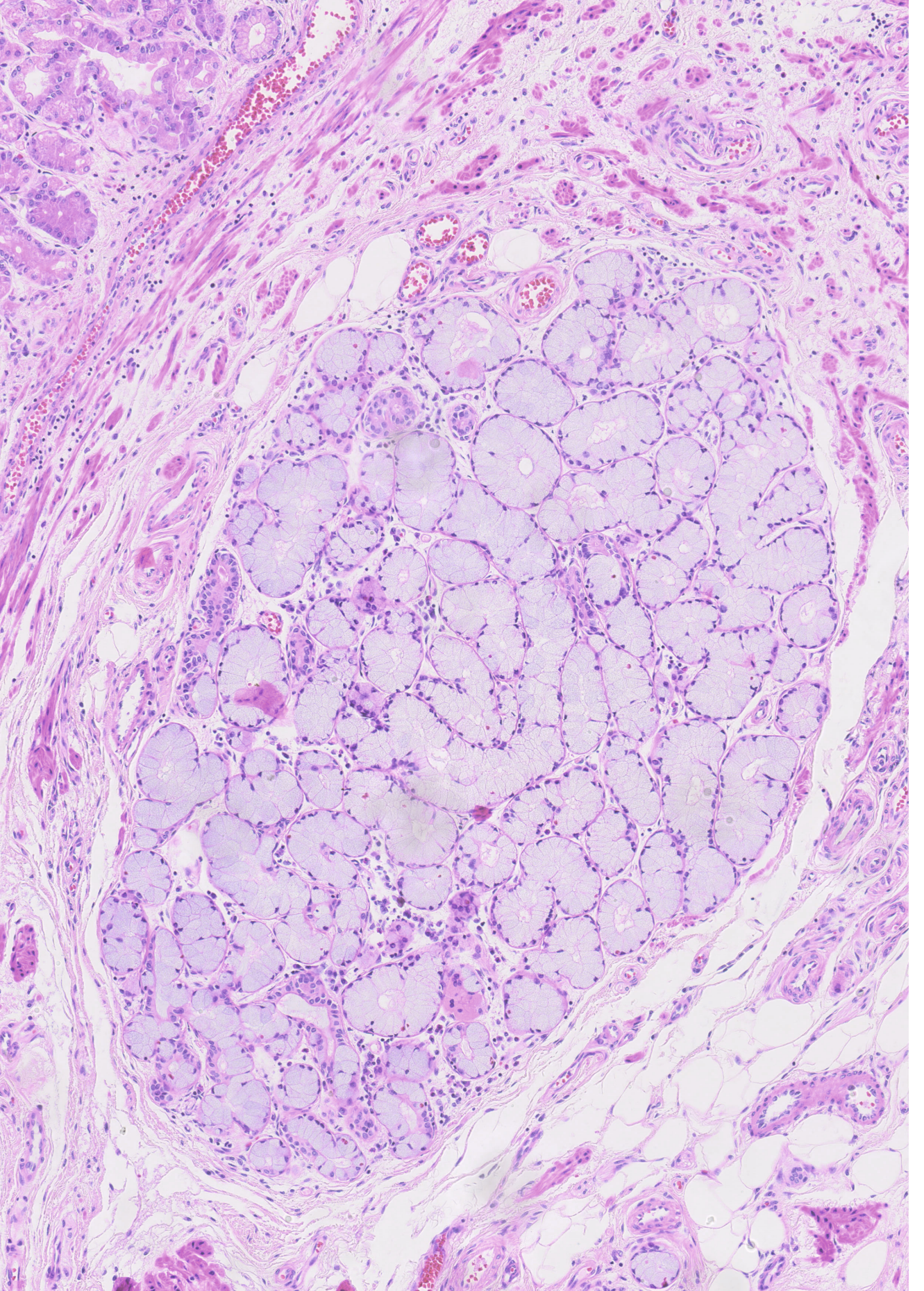
A



B



Supplementary Figure S2. HOXA13 influences cellular phosphoprofile as determined by western blot analysis. Three independent experiments were performed to obtain lysates of HOXA13+ and control cells. These lysates were subsequently blotted a minimum of 1 time, resultin in an N of at least 3 for control and HOXA13+ cells. A) Representative examples of control and HOXA13+ cell lysates are shown for the presentce of indicated phopho-proteins. B) Quantification of western blots.



Chapter 8

BARX1 is overexpressed in (pre-)malignant esophageal lesions, but does not act via Wnt signaling in adult human esophageal cells

K. Nesteruk, X. Guo, V. T. Janmaat, B. Yu, S. Arupillai, M. Magierowski, R. Smits, M.P. Peppelenbosch, G.M. Fuhler

(In preparation)

INTRODUCTION

While the incidence rate of esophageal adenocarcinoma (EAC) is increasing in many countries, overall survival remains poor with a 5-year survival rate of 20% [1]. Outcomes have not improved substantially over last 15 years and survival can be as low as 5% for late-stage disease [2, 3]. EAC commonly arises from its precursor lesion Barrett's esophagus (BE): replacement of normal squamous epithelium with metastatic columnar epithelium which, histologically and molecularly, greatly overlaps with colonic epithelium, including the presence of enterocytes, enteroendocrine cells, and goblet cells. BE occurs as a consequence of gastroesophageal reflux disease (GERD) when the normal squamous epithelium of the esophagus is injured by refluxate containing gastric acid, pepsin, bile, pancreatic enzymes, ingested foods and their metabolites [4]. The non-dysplastic BE can sequentially progress to low-grade dysplasia, high-grade dysplasia and eventually to invasive carcinoma. The annual EAC incidence rate in BE cohorts ranges from 0.12% to 3.55% in different studies [5, 6]. Contributing to BE and EAC development are several clinical risk factors such as male gender, aging, smoking, high-fat diet and obesity [7]. Risk increases progressively with the length of the Barrett's segment, and the presence and grade of dysplasia. While an understanding of the disease at a molecular level may improve strategies for early detection, prevention, and treatment of BE and EAC, the pathogenesis of BE is not completely clear.

In 2013, a large epidemiological genome-wide association study (GWAS) identified genetic variants associated with BE and EAC [8]. Later, a separate GWAS study provided support for some of these identified variants, including loci containing *BARX1 (*rs11789015), *FOXP1* (rs2687201) and *FOXF1 (*rs9936833, rs3111601, rs1728400, rs3950627, rs2178146, rs13332095) [9]. Interestingly, *BARX1, FOXP1* and *FOXF1* all encode transcription factors regulating the development and differentiation of the esophagus, stomach, and intestine in murine embryogenesis [10-15]. In the murine gastrointestinal (GI) tract, loss of *Barx1* prevents stomach epithelial differentiation and induces intestine-specific genes instead [11]. The reported mechanism includes Barx1-mediated inhibition of Wnt signaling through induction Wnt antagonists such as secreted frizzled proteins (sFRP1, sFRP2). Moreover, *Barx1*$^{-/-}$ mice display impairment of stratified squamous epithelium formation in the esophagus [10]. *FOXP1* regulates the development of the muscle compartment in the mouse GI tract, including the esophagus, and its loss leads to motility dysfunction [12-14]. *FOXF1* is downstream of the Hedgehog signaling pathway, which is essential for foregut separation. Mice heterozygous for a *Foxf1* null allele have major structural abnormalities, including esophageal atresia and mispatterning of the esophageal epithelium [15, 16]. *Foxp1* is ubiquitously expressed in the esophagus, fundus, duodenum, ileum, and colon of adult mice, throughout all layers of the murine GI tract including the myenteric plexus [17]. *Foxf1* expression is widespread in embryonic murine GI tract, although little is reported regarding its expression pattern in adult tissues. Barx1 expression is limited to the embryonal stomach, where is expressed in fibroblasts exerting their paracrine effects on the epithelial compartment, and its expression is gradually reduced until almost undetectable in adult mice. This begs the question as to why single nucleotide variants in this gene would be associated to the carcinogenic cascade in the esophagus. However, cancer itself may be seen as a reversal of adult tissue to an embryonic state, with embryonic genes being re-

expressed [18]. As such, a role for *BARX1*, *FOXF1*, *FOXP1* transcription factors in carcinogenic pathways in humans has been proposed [12, 19-21] where depending on the tissue and context, they may function as tumor suppressor or oncogene. However, while a functional role for *BARX1*, *FOXF1*, *FOXP1* in the etiology of BE and EAC thus may be speculated, little is known regarding the expression and function of these factors in the adult human GI tract. Thus, the aim of our study is to investigate the expression of *BARX1*, *FOXF1*, *FOXP1* in the healthy human GI tract, BE and EAC.

MATERIAL AND METHODS

Collection of human material

All human tissues used in this study were obtained at the Erasmus University Medical Center, department of Gastroenterology & Hepatology. The use of these samples was approved by the Erasmus MC medical ethical committee (MEC-2015-208, MEC-2015-209, MEC-2015-199, MEC-2010-093; tissues were handled according to the FEDERA code of conduct and informed consent was obtained where necessary [22]. Biopsy specimens to investigate *BARX1*, *FOXF1* and *FOXP1* expression were obtained by double balloon enteroscopy. Nine biopsy specimens were obtained from each patient (n=13) at different locations along the GI-tract. Sequentially these locations were: 1: esophagus, 2: stomach, 3: duodenum, 4: jejunum, 5: proximal ileum, 6: distal ileum, 7: ascending colon, 8: descending colon and 9: sigmoid/rectum. Included patients had unexplained symptoms, mostly anemia, while patients with inflammatory bowel disease were excluded. All biopsies for RNA isolation were stored at -80˚C.

BE and normal adjacent squamous mucosal biopsy specimens were obtained from the same patients. Squamous esophageal biopsies originated 5 cm above the squamocolumnar junction (SCJ). Barrett's (BE) biopsies were obtained caudal of the SCJ and cranial of the gastric folds (all patients were on PPI therapy). Forceps biopsy specimens of EACs were obtained and included when pathological examination of simultaneously taken forceps biopsies around the study specimens were positive for EAC.

Animal studies

For the *Hoxa13* mRNA expression analysis throughout the murine gastrointestinal (GI) tract, four *C57BL/6J* wildtype mice aged between three to five months were used. The GI-tract was divided into 1: esophagus; 2: stomach; 3: duodenum; 4: jejunum; 5: proximal ileum; 6: distal ileum; 7: cecum; 8: proximal colon; 9: distal colon, of which sections were opened and rinsed in PBS followed by storage in RNAlater at -80˚C. These murine experiments were approved by the Ethical Committee for Animal Experiments of the Erasmus MC.

Surgical rat model of BE

Chronic exposure of esophagus to mixed gastroduodenal contents was caused by creation of an esophago-gastroduodenal anastomosis between the gastro-esophageal junction and the duodenum on its anterior mesenteric border as described before [23, 24]. Briefly, 8-week-old male Wistar rats fasted for 24 h prior to the surgery performed under general isoflurane anesthesia. After midline laparotomy, two longitudinal incisions of ~7 mm in length were made: one extending along the lower part of anterior esophagus wall including the gastroesophageal junction area, and one 4 cm distal from the Treitz ligament on the anterior mesenteric border of the duodenum. Immediately after surgery, the rats were infused subcutaneously with 10–15 mL of isotonic sodium chloride, and then were placed in individual cages, and kept under standard housing conditions (room temperature 22°C, relative humidity 50 ± 5%, and 12:12-h light-dark cycle), fasted, and fed only tap water for next 48 h [23].

The study was approved by the Local Animal Care and Use Ethical Committee held by Jagiellonian University Medical College in Cracow and was run in compliance with the European Union regulations, ARRIVE guidelines and with implications for replacement, refinement or reduction (the 3Rs) principles, regarding handling of experimental animals.

RNA isolation and qPCR

RNA was isolated using the NucleoSpin RNA isolation kit (Macherey Nagel, Düren, Germany). Biopsies and animal tissues were homogenized by the TissueRuptor obtained from Qiagen Inc. RNA concentrations were measured using a Nanodrop spectrophotometer and kept at -80 °C.

cDNA was made from 1µg RNA using Primescript RT Master Mix according to manufacturer's instructions (Takara, Otsu, Japan), for 15min at 37 °C and 5 sec at 95 °C, and stored at -20 °C. qPCR was performed for 40 cycles in the iQ5 Real-Time PCR detection system that was obtained from BioRad Laboratories (Veenendaal, The Netherlands). For each reaction 10 µL cDNA template, 12.5 µL SYBR GreenER purchased from Invitrogen (Carlsbad, CA), and 2.5 µL 10 pM/µl primer were used. Reactions were performed in duplicate. Primers used are shown in **Supplementary Table 1** and were ordered at Sigma-Aldrich (Darmstadt, Germany). Specificity and intron-spanning amplification of *BARX1, FOXP1, FOXF1* primers were verified by PCR on human genomic DNA (cell-line derived), RNA and derived cDNA. To further confirm the specificity of the *BARX1* primers, qPCR-product was been sequenced at MACROGEN Europe B.V. A blast analysis of sequencing product confirmed the amplicon to be a *BARX1* mRNA transcript. qPCR data were analyzed with Microsoft Excel using the ΔΔCt method with *RP2* as a reference gene on human materials. Reference genes used for qPCRs on mouse materials was *Leng8,* on rat materials *HPRT1*. Differences in expression were analyzed with a two-sided Student's t-test using Prism 8.01, obtained from GraphPad Software (San Diego, USA).

Cell culture

All cells were cultured with penicillin (100u/mL) and streptomycin (100u/mL) and were regularly STR-verified and checked for mycoplasma by handing in samples prepared according to instructions at GATC Biotech (Konstanz, Germany). **HET1A**, a primary immortalized human squamous esophageal cell line, was a gift of J.W.P.M. van Baal (University Utrecht, The Netherlands). Primary human esophageal epithelial cells transformed with hTERT (**EPC2-hTERT**) was a gift of K.K. Krishnadath [25]. Cells were cultured with Keratinocyte SFM medium, supplemented with bovine pituitary extract at 50 µg/ml and EGF at 1 ng/ml (Thermo-Fisher Scientific). Culture medium for **SK-GT-4,** (gift from Prof W.N.M. Dinjens, Erasmus MC, Netherlands [26]) consisted of RPMI-1640, 2mM Glutamine and 10% Fetal Bovine Serum (FBS), culture medium for **FLO-1** (gift from Prof W.N.M. Dinjens, Erasmus MC, Netherlands [26]) was DMEM with 2mM Glutamine and 10% Fetal Bovine Serum (FBS).

Acid and bile exposure

For assessment of mRNA expression, EPC2-hTERT cells were treated for 30 minutes with cell culture medium containing bile acids, HCl or a combination of these. The bile acid mixture was used in concentrations of 0, 100, and 200 µmole/L and consisted of 25% deoxycholic acid, 45% glycocholic acid and 30% taurochenodeoxycholic acid. For HCl exposure, medium was adjusted to a pH of 7.0, 4.5 or 4.0. Cells were subsequently washed using PBS and given standard medium. All cell culture experiments were performed at least three times. After 24 hours, the cells were harvested and mRNA was isolated. Methods were derived from Bus *et al.* [27].

siRNA-mediated gene knock-down

Smartpool ON-TARGETplus siRNAs targeting *BARX1* and nontargeting siRNA control #2 were introduced into cells using DharmaFECT. Successful knockdown was confirmed by qPCR.

Statistical analysis

For comparison of two groups paired Wilcoxon signed rank test or Mann-Whitney U test were used. For comparison more 3 and more groups ANOVA with Holm-Šídák's multiple comparisons test was used. The type of statistical test was based on the result of the normality test (either a Komogorov-Smirnov test, the D'Agostino, Pearson omnibus normality test or the Shapiro-Wilk normality test). Tests were performed in Graphpad Prism 8; GraphPad Software Inc., USA). P-values <0.05 were considered to be statistically significant.

RESULTS

Expression of *BARX1*, *FOXP1* and *FOXF1* in the healthy human GI tract

We considered that expression pattern of *BARX1*, *FOXP1* and *FOXF1* along the human GI tract might be important as we and others showed that dysregulated location-specific gene expression may be involved in the etiology of BE *[28-30]*. To investigate expression of these transcriptional factors in the healthy human GI tract, biopsies from different regions along the GI tract were collected (esophagus, stomach, duodenum, ileum, and colon) and expression was analyzed by qPCR. *BARX1* expression highly varied per tissue and per individual (N=13) (**Figure 1A, B**). Of the tested specimens, 91% (10/11) of jejunum samples showed detectable *BARX1* levels, versus 22% (2/9) of esophagus, 18% (2/11) of stomach, 54% (7/13) of duodenum, 70% (7/10) of proximal ileum, 44% (4/9) of distal ileum, 50% (6/12) of ascending colon, 54% (7/13) of descending colon and 45% (5/11) of sigmoid/rectum samples (**Figure 1B**). Unexpectedly, a significantly higher level of *BARX1* was seen in jejunum as compared to esophagus, stomach, duodenum, ileum, and colon (**Figure 1A**), which is in contrast to previously reported data on adult mouse physiology [31]. Thus, we also investigated expression of *Barx1* in the mouse GI tract and verified that expression of *Barx1* in the adult mouse is limited to the esophagus and stomach (**Figure 1C**), while it is not detected in the intestine. Thus, there are clear interspecies differences in the expression of *BARX1*.

FOXP1 and *FOXF1* were abundantly expressed and detectable in 100% of tested specimens (N=7 patients, 9 different locations per patient). A gradual increase of *FOXP1* was noted from the stomach to sigmoid/rectum, with the exception of the proximal ileum (R = 0.87, **Supplementary figure 1**). Levels of *FOXP1* were significantly higher in the sigmoid/rectum compared to the esophagus, stomach, duodenum and proximal ileum (**Figure 1D**). There was significantly less *FOXF1* in the proximal ileum compared to the esophagus, duodenum, stomach, duodenum, distal ileum, descending colon, ascending colon and sigmoid/rectum (**Figure 1E**). *FOXF1* and *FOXP1* expression correlated in all locations except the distal ileum and ascending colon (**Supplementary table 2, Supplementary figure 3**). While for most locations a correlation of *BARX1* with *FOXP1* or *FOXF1* could not reliably be established due to low expression levels, for jejunum (where *BARX1* was detected in all samples), no correlation with either *FOXP1* or *FOXF1* was found. Taken together, *BARX1*, *FOXP1* and *FOXF1* show region-specific gastrointestinal expression, with limited *BARX1* but ample *FOXP1* and *FOXF1* expression seen at the normal esophagus.

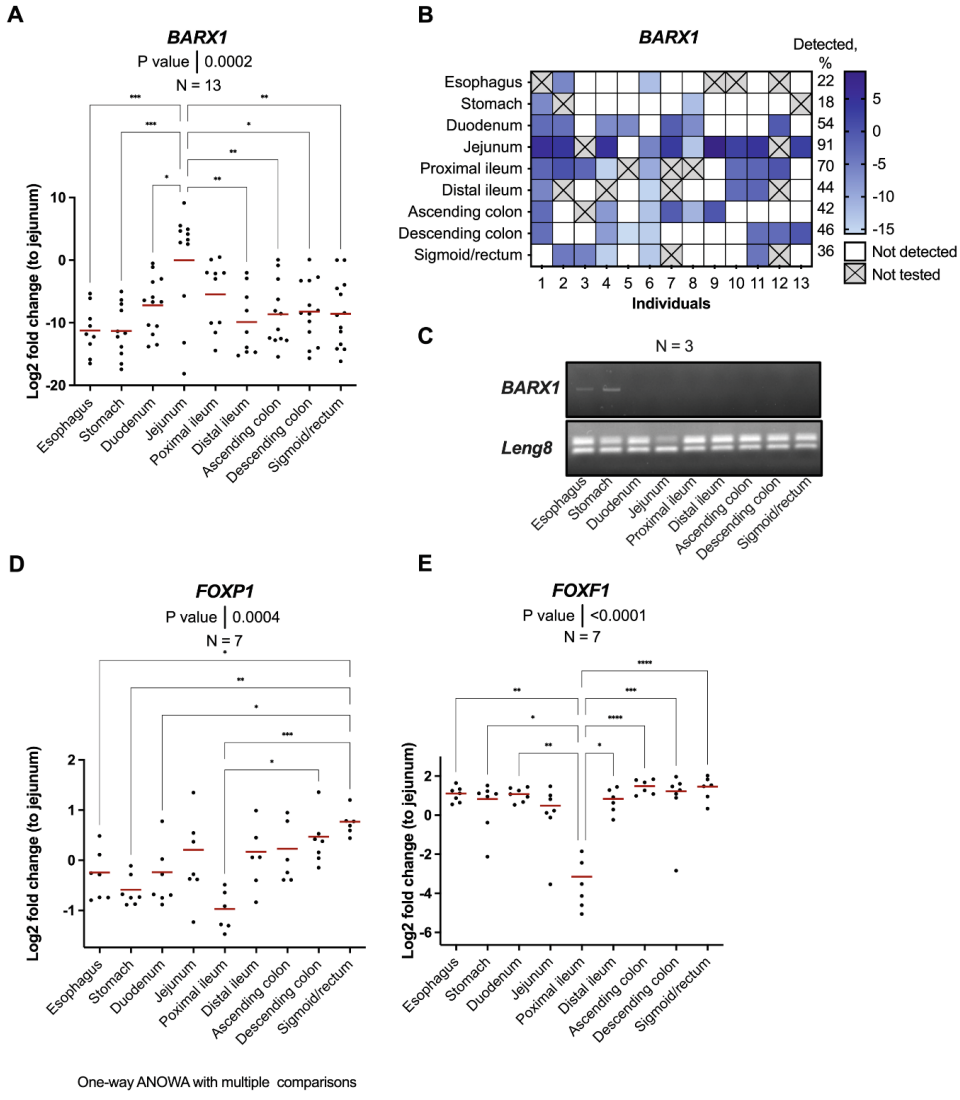


Figure 1. Expression of *BARX1*, *FOXP1* and *FOXF1* in the adult GI tract. A) *BARX1* expression in different segments of the adult human GI tract as determined by qPCR. B) *BARX1* expression highly varies per tissue and per individual. C*) BARX1* expression is limited to the esophagus and stomach of adult mice (representative example of three independent mice shown). D) *FOXP1* expression in different segments of the adult human GI tract. E) *FOXF1* expression in different segments of the adult human GI tract.

Expression of *BARX1*, *FOXP1* and *FOXF1* in Barrett's esophagus and esophageal adenocarcinoma

Having established that *BARX1*, *FOXP1* and *FOXF1* are to variable extent present in the normal esophagus, we next investigated expression of these factors in BE. *BARX1* was overexpressed in BE

(FC9.9) when compared to squamous esophageal tissue obtained from the same patient (P=0.036, N=23) (**Figure 2A**). In contrast, in a rat model of Barrett's esophagus, *BARX1* mRNA levels were decreased in columnar BE tissues compared to squamous esophagus (**Supplementary figure 2**) [23], further confirming interspecies variation, and in line with our data that the rat *in vivo* model shares about 45% of transcriptional expression pattern changes with human clinical BE [23]. For human tissues, no differences were found for *FOXP1* (P=0.57, N= 10) or *FOXF1* (P=0.23, N=10) between BE and squamous tissue (**Figure 1C, E**).

Next, we contrasted EAC tissues to normal squamous esophagus obtained from BE patients. In EAC, *BARX1* (P<0.001, N=18 for EAC, N=19 for SQ) and *FOXF1* (P=0.035, N=10) were significantly upregulated, while no difference was observed for *FOXP1* (P=0.96, N=10) (**Figure 1B, D, F**). Again, a correlation between *FOXF1* and *FOXP1* levels was seen within the same BE tissues, this correlation was not observed for BARX1 (**Supplementary Table 3**). However, in AEC tissues, expression of *FOXF1* and *BARX1* did correlate. Thus, upregulated expression of *BARX1* in pre-malignant lesions and both *BARX1* and *FOXF1* in cancer of the esophagus is seen.

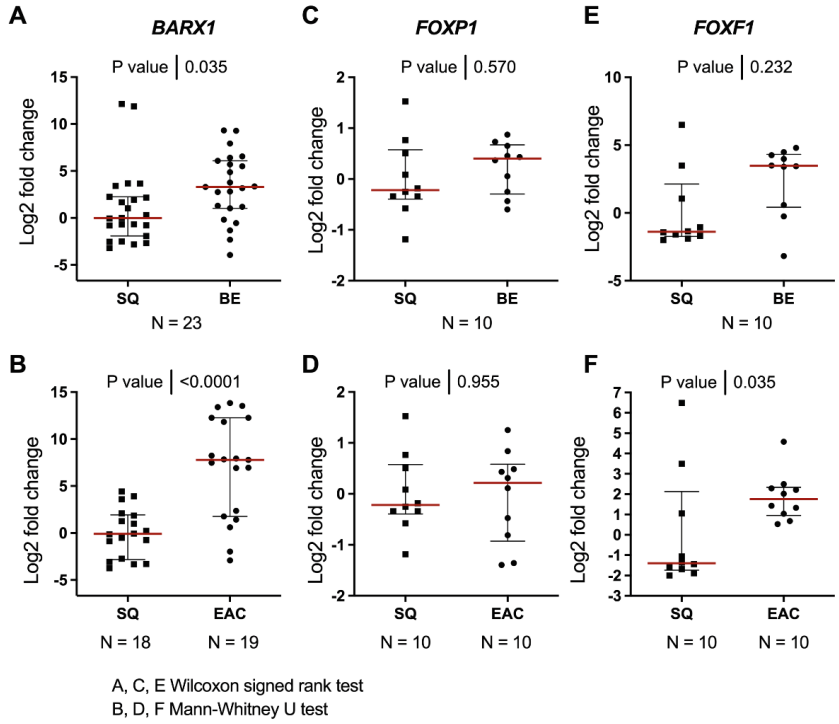


Figure 2. *BARX1* is upregulated in (pre)-malignant lesions of the esophagus, while upregulation of *FOXF1* is present in esophageal adenocarcinoma. mRNA expression of *BARX1*, *FOXP1* and *FOXF1* in Barrett's esophagus (BE) segments compared to matched normal esophageal squamous epithelium (A, C, E), as compared to matched normal squamous epithelium and EAC (B, D, F).

WNT signaling and *BARX1*

In mouse embryogenesis, *BARX1*-induced expression of *SFRP1* and *sFRP2* inhibits the Wnt/β-catenin signaling which drives intestinal differentiation [11]. As BE is characterized by intestinal metaplastic features including a columnar epithelium, we investigated whether upregulation of *BARX1* would be paralleled by expression of molecules of the Wnt/β-catenin signaling pathway in these samples, including *sFRP1*, *sFRP2* and *AXIN2* as downstream targets of this pathway. In BE, *AXIN* tended to be increased but this did not reach significant difference (P=0.09, N=10). *sFRP1* (P=0.005, N=10) was decreased in BE samples, while no changes in *sFRP2* were observed (P=0.50, N=10) (**Figure 3A, C, E**). There was no significant difference in EAC for *AXIN2* (P=0.10, N=18 and N=19) and *sFRP1* (P=0.42, N=10) while *sFRP2* showed a trend towards upregulation in EAC as compared to normal squamous epithelium (P=0.063, N=10) (**Figure 3B, D, F**). Correlation analysis did not show a direct correlation between BARX1 expression and either *AXIN1*, *SFRP1* or *sFRP2* as measured in the same sample. Thus, *BARX1* does not seem to modulate molecules of Wnt/β-catenin signaling pathway in these samples.

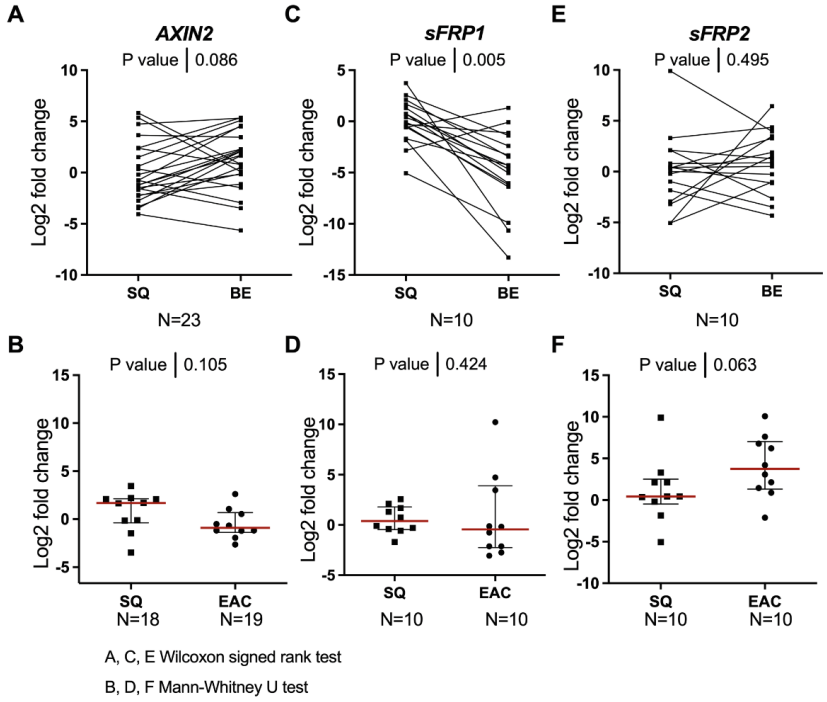


Figure 3. There is no significant difference in expression of *AXIN2*, *sFRP1*, *sFRP2* in BE (A, C, E) or EAC (B, D, F). mRNA expression of *AXIN2*, *sFRP1* and *sFRP2* in Barrett's esophagus (BE) segments compared to matched normal esophageal squamous epithelium (A, C, E), as compared to matched normal squamous epithelium and in EAC compared to squamous epithelium of BE patients (B, D, F).

To further confirm this point we reduced *BARX1* expression by siRNA knock-down in EAC cell lines. Of three tested EAC cells lines (JH-esopAd1, SK-GT-4, FLO-1), SK-GT-4 and FLO-1 expressed *BARX1* at the highest levels (ΔCT of 3 compared to housekeeping gene) and were chosen for the experiment. *BARX1* knockdown in EAC cells did not affect the expression of *AXIN2*, *sFRP1*, and *APC* [expression of *sFRP2* was not detected in these lines] **(Figure 4).** Hence, in the human esophagus, *BARX1* expression is not associated with modulation of Wnt signaling.

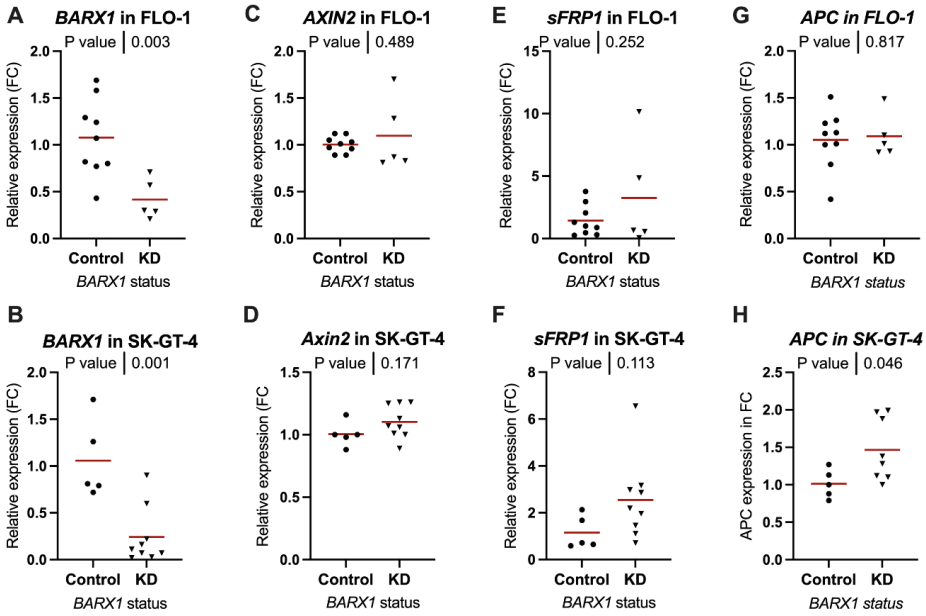


Figure 4. SiRNA-mediated *BARX1* knockdown did not change expression of *AXIN2*, *sFRP1* and *APC* in esophageal adenocarcinoma EAC cells (A, C, E, G FLO-1 and B, D, F, H -SK-GT-4) compared to nontargeting siRNA control.

Expression of *BARX1*, *FOXP1* and *FOXF1* upon bile/acid exposure

To glean insight into the role of *BARX1*, *FOXP1* and *FOXF1* in formation of esophageal (pre-)malignant lesions, we next explored how exposure to either bile, acid, or their combination will affect the expression of these genes in an *in vitro* model of GERD. For this, we investigated two primary immortalized squamous esophageal cell lines (EPC2-hTERT and Het1a). Similar to primary squamous esophagus biopsies, untreated EPC2-hTERT cells had a low expression of *BARX1* (ΔCT of 7 compared to housekeeping gene). In contrast, Het1a had a high expression of *BARX1 (*ΔCT of -6 compared to housekeeping gene). For this reason, EPC2-hTERT was used for further experiments. Bile/acid treatment induced expression of *BARX1* in these cells, while no difference was observed for *FOXF1*, *FOXP1* or Wnt signaling molecules (*AXIN2, sFRP1, sFRP2*) (**Figure 5**).

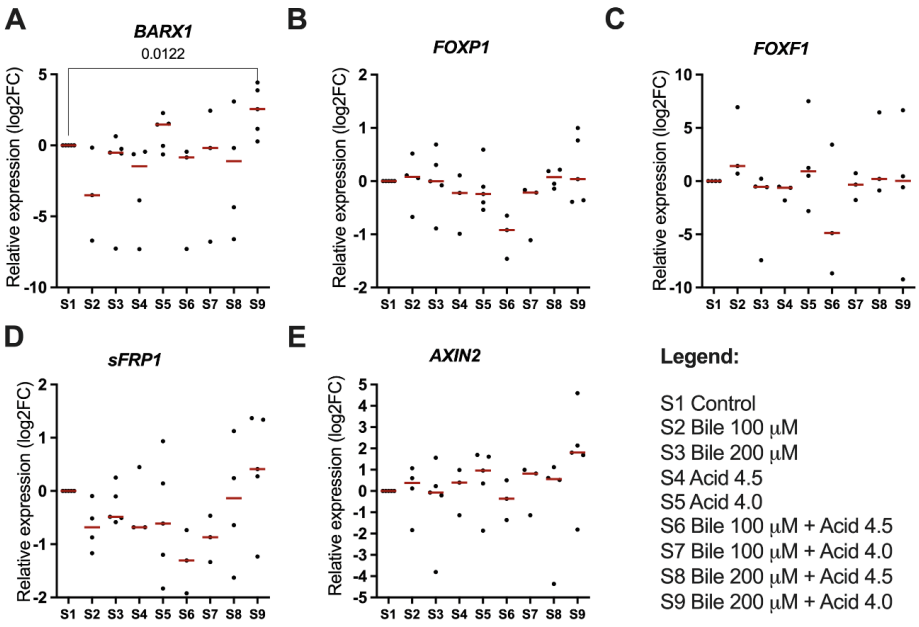


Figure 5. Combined of bile/acid treatment increased expression of *BARX1* mRNA in immortalized EPC2-hTERT cells, but not *FOXP1*, *FOXF1*, *sFRP1*, *AXIN2*. sFRP2 was not detected. EPC2-hTERT cells were treated for 30 minutes with cell culture medium containing bile acids or HCl or their combination. The bile acid mixture consisted of 25% deoxycholic acid, 45% glycocholic acid and 30% taurochenodeoxycholic acid. For HCl exposure, medium adjusted to a pH of 7.0, 4.5 or 4.0. Cells were harvested 24h after treatment.

DISCUSSION

To the best of our knowledge, we are the first to characterize the expression of *BARX1*, *FOXF1* and *FOXP1* along the different regions of the adult healthy human GI tract. *BARX1* expression highly varied per tissue and per individual and, surprisingly, was the highest in the jejunum. In the murine GI tract, were *BARX1* was studied before, *BARX1* expression in the adult GI tract was restricted to the stomach and esophagus. Thus, there is a clear interspecies difference in expression of *BARX1* between the human and murine GI tract. *FOXF1* and *FOXP1* were present throughout the GI tract, and their expression pattern correlated, with in particular the proximal ileum showing a low expression of these two genes. We further demonstrate that *BARX1* is overexpressed in BE and EAC, while *FOFX1* is overexpressed only in EAC.

Genetic factors can contribute to development of both BE and to EAC. According to Ek et al [32], 25% of EAC cases and 35% of BE cases arise as a composite effect of many common mutations with small individual relative risk and substantial polygenic overlap between these two diseases [33]. GWAS studies identified SNPs associated with BE and EAC [33], but most await functional studies to confirm a driver role in these diseases. One of the susceptibility loci for BE and EAC is the SNP rs11789015 located in the intron of the *BARX1* gene, a member of the Bar subclass of homeobox transcription factors [8, 9]. The minor G allele of rs11789015 at 9q22 was significantly associated with a decreased

risk of BE and EAC. Recent data show that this polymorphisms of the *BARX1* also associated with GERD [34]. In addition, *BARX1* was proposed as genetic contributor to development of human congenital Hiatal hernia, a condition characterized by a protrusion of the stomach into the thoracic cavity through a widening of the right crus of the diaphragm which may predispose to GERD and BE [35]. Our data showing a gradual increase in *BARX1* expression from squamous esophagus of BE patients to BE and to EAC combined with bile/acid-induced expression of *BARX1,* indicate that rs11789015 may contribute to BE and EAC by regulating the function of *BARX1* potentially via its role in GERD. Indeed**,** evidence of the variant rs11789015 influencing *BARX1* expression by regulation of promotor functions exists. First, Yan et al showed that rs11789015 is located upstream of the short transcript of *BARX1* where histone marks denote likely promoter activity [19]. Second, based on the publicly available GTEx database, these authors showed rs11789015 to be a cis-eQTL for *BARX1*, with the protective G-allele associated with decreased level of *BARX1* mRNA both in esophageal mucosa and gastroesophageal junction tissues [19]. Interesting, this locus at 9q22 also confers risk to another histological subtype of esophageal cancer the esophageal squamous cell carcinoma (ESCC) [19]. Although these constitute cancers with different etiology, some other genes are also known to play role in both EAC and ESCC, including *HOXA13* and *TP53* [28, 36, 37]. The study by Yan et al [19] is also the only known study investigating the mechanistic and functional role of rs11789015 in esophageal carcinogenesis, showing enhanced protein expression of BARX1 in ESCC samples [19].

Little is known about transcriptional regulatory function *of BARX1*. Our data indicate that unlike the developing murine foregut, the main function of *BARX1* in EAC cells is not to regulate Wnt/β-catenin signaling [10, 31]. In the developing stomach of mice*, Barx1* induces expression of 14 transcripts in mesenchymal cells including Pitx1, Igfbp4 and secreted sFRP1 and sFRP2, genes that reduce local Wnt activity [31]. These signals paracrinely directed stomach epithelial differentiation from endoderm [31]. In contrast to our study, Boccellato at al. [38] showed that conditioned medium from gastric stromal cells reduces the level of epithelial *LGR5* mRNA, reduces Wnt reporter signal and induces partial differentiation of stomach epithelium in ALI cultures. Simultaneously, they observed strong expression of *sFRP1*, *DKK1* and *DKK3*, as well as *BARX1* in gastric stromal cells suggesting that these soluble inhibitors from gastric stromal cells caused the observed effects [38]. KEGG enrichment analysis of human ESCC tumors present in the the Cancer Genome Atlas (TCGA) based on genes co-expressed with *BARX1* showed multiple significant immune-related pathways, including IL-17 signaling pathway, RIG-I-like receptor signaling pathway and Toll-like receptor signaling pathway [19]. Although direct targets of *BARX1* transcriptional factor are not known, functionally, *BARX1* induces ESCC cell proliferation, migration, and invasion [19]. Interestingly, loss of Barx1 promotes hepatocellular carcinoma metastasis through up-regulating MGAT5 and MMP9 expression [39], indicating that for other types of cancer *BARX1* may be protective [39].

Deregulation of BARX1 may contribute to metaplastic processes. SiRNA-induced *Barx1* deficiency had a dramatic effect for overlying gastric endoderm with activation of intestinal markers and reduced expression of stomach-specific genes [31]. Similarly, forced expression of *Barx1* in intestinal epithelial-mesenchymal cocultures reduced expression of intestine-specific transcripts and enhanced expression

of the stomach-specific gene Muc1 [31]. In the murine GI tract, defective stomach development with formation of intestine-like phenotype is seen in *Barx1* null mice. Authors [31] suggested that intestinal differentiation represents a default state for gut endoderm and that active signals, in the form of Wnt inhibition, are needed to specify the stomach epithelium, which is in line with evolutional processes considering that the intestine is a primitive structure, whereas both lungs and stomach are recent adaptations [31]. They speculated that this is why intestinal metaplasia occurs commonly with foregut epithelial injury, whereas the reverse condition, gastric metaplasia of midgut or hindgut derivatives, is rare [31]. Indeed, in another study, when *Barx1* was ectopically expressed in intestinal mesenchyme at levels similar to those present in the native fetal stomach it produced severe defects in development of the alimentary canal and other abdominal organs in mice [40]. However unlike its potent role in the stomach, ectopic *Barx1* expression in the intestine was insufficient to impose stomach differentiation of intestinal endoderm [40]. Epithelium of proximal *Barx1*-/- foregut showed anomalies and extensive mixing of cell types such as a squamous mucosa and a cuboidal epithelium [41], reminiscent of BE. However, in our study, BE showed overexpression of *BARX1,* suggesting that at least *BARX1* does not counteract the intestine-like BE phenotype. Clearly, human and mouse physiology differ in this respect, but functional studies are needed to investigate whether *BARX1* supports this phenotype in the human esophagus and what cell types express *BARX1* in the adult human esophagus.

Not only does the role of *BARX1* in human esophageal carcinogenesis require further elucidation, for *FOXP1* and *FOXF1*, the roles are also unclear. While we did not observe a correlation between *BARX1* and *FOXF1* in human tissues, murine *Barx1* mRNA expression parallels expression of *FoxF1* and Vimentin [31]. There was significantly less *FOXF1* and *FOXP1 i*n the proximal ileum compared to other parts tested. *FOXP1* expression gradually increased from proximal to distal GI tract and was the highest in the sigmoid/rectum. *FOXP1* is under regulation of *HOX* genes and with the exception of the proximal ileum followed a collinear pattern, similar to *HOX* genes involved in BE pathogenesis [30, 42]. Nevertheless, *FOXP1* was not found to be overexpressed in BE or EAC in our study, and rs2687201 did not pose an eQTL for expression of *FOXP1* in the human GI tract (analysis of GTEx data, not shown). *FOXF1*, however, was overexpressed in EAC in this study. In contrast to reported before [43] we did not confirm its significant overexpression in BE, nor was its expression induced by bile/acid. Others showed that *FOXF1* mRNA and protein levels were upregulated in biopsy specimens from patients with GERD, and BE [43] with both stroma and epithelium being positive for FOXF1. Functionally, *FOXF1* induced columnar phenotype and epithelial-to-mesenchymal transition in esophageal squamous cells [43]. Differences between our studies are pH and bile salts ratios in *in vitro* GERD model.

This was the first step towards understanding of the role of *BARX1, FOXF1* and *FOXP1* in the adult human GI tract. Future studies it should be determine which cell types express *BARX1* under physiological and pathological conditions, what molecular targets of this transcriptional factor are and what its functional role is in esophageal lesions. *In toto*, *BARX1* is overexpressed in BE and EAC, but does not act via Wnt signaling in adult human esophageal cells.

ACKNOWLEDGMENT

Investigations within rat model of BE were supported by National Science Centre Poland (UMO-2016/23/D/NZ4/01913) and by the National Centre for Research and Development, Poland (LIDER/9/0055/L-8/16/NCBR/2017).

REFERENCES

1. Siegel, R.L., K.D. Miller, and A. Jemal, *Cancer statistics, 2020.* CA Cancer J Clin, 2020. **70**(1): p. 7-30.
2. Grotenhuis, B.A., et al., *Delay in diagnostic workup and treatment of esophageal cancer.* Journal of Gastrointestinal Surgery, 2010. **14**(3): p. 476-483.
3. Thrift, A.P. and D.C. Whiteman, *The incidence of esophageal adenocarcinoma continues to rise: analysis of period and birth cohort effects on recent trends.* Ann Oncol, 2012. **23**(12): p. 3155-62.
4. Jenkins, G.J.S., et al., *The bile acid deoxycholic acid has a non-linear dose response for DNA damage and possibly NF-κB activation in oesophageal cells, with a mechanism of action involving ROS.* Mutagenesis, 2008. **23**(5): p. 399-405.
5. Arnold, M., et al., *Global incidence of oesophageal cancer by histological subtype in 2012.* Gut, 2015. **64**(3): p. 381.
6. Schoofs, N., R. Bisschops, and H. Prenen, *Progression of barrett's esophagus toward esophageal adenocarcinoma: An overview.* Annals of Gastroenterology, 2017. **30**(1): p. 1-6.
7. Krishnamoorthi, R., et al., *Rates and predictors of progression to esophageal carcinoma in a large population-based Barrett's esophagus cohort.* Gastrointestinal Endoscopy, 2016. **84**(1): p. 40-46.e7.
8. Levine, D.M., et al., *A genome-wide association study identifies new susceptibility loci for esophageal adenocarcinoma and Barrett's esophagus.* Nature Genetics, 2013. **45**(12): p. 1487-1493.
9. Becker, J., et al., *Supportive evidence for FOXP1, BARX1, and FOXF1 as genetic risk loci for the development of esophageal adenocarcinoma.* Cancer Med, 2015. **4**(11): p. 1700-4.
10. Woo, J., et al., *Barx1-mediated inhibition of Wnt signaling in the mouse thoracic foregut controls tracheo-esophageal septation and epithelial differentiation.* PLoS One, 2011. **6**(7): p. e22493.
11. Kim, B.-M., et al., *The Stomach Mesenchymal Transcription Factor Barx1 Specifies Gastric Epithelial Identity through Inhibition of Transient Wnt Signaling.* Developmental Cell, 2005. **8**(4): p. 611-622.
12. Koon, H.B., et al., *FOXP1: a potential therapeutic target in cancer.* Expert Opinion on Therapeutic Targets, 2007. **11**(7): p. 955-965.
13. Shu, W., et al., *Foxp2 and Foxp1 cooperatively regulate lung and esophagus development.* Development, 2007. **134**(10): p. 1991.
14. Fröhlich, H., et al., *Gastrointestinal dysfunction in autism displayed by altered motility and achalasia in Foxp1 (+/-) mice.* Proc Natl Acad Sci U S A, 2019. **116**(44): p. 22237-22245.
15. Mahlapuu, M., S. Enerbäck, and P. Carlsson, *Haploinsufficiency of the forkhead gene Foxf1, a target for sonic hedgehog signaling, causes lung and foregut malformations.* Development, 2001. **128**(12): p. 2397-406.
16. Ustiyan, V., et al., *FOXF1 transcription factor promotes lung morphogenesis by inducing cellular proliferation in fetal lung mesenchyme.* Dev Biol, 2018. **443**(1): p. 50-63.
17. Fröhlich, H., et al., *Gastrointestinal dysfunction in autism displayed by altered motility and achalasia in.* Proc Natl Acad Sci U S A, 2019. **116**(44): p. 22237-22245.
18. Monk, M. and C. Holding, *Human embryonic genes re-expressed in cancer cells.* Oncogene, 2001. **20**(56): p. 8085-91.
19. Yan, C., et al., *An esophageal adenocarcinoma susceptibility locus at 9q22 also confers risk to esophageal squamous cell carcinoma by regulating the function of BARX1.* Cancer Lett, 2018. **421**: p. 103-111.
20. Banham, A.H., et al., *The FOXP1 winged helix transcription factor is a novel candidate tumor suppressor gene on chromosome 3p.* Cancer Res, 2001. **61**(24): p. 8820-9.
21. Zhang, J., et al., *Genetic variants of FOXP1 and FOXF1 are associated with the susceptibility of oesophageal adenocarcinoma in Chinese population.* J Genet, 2018. **97**(1): p. 213-218.
22. Federa, *Human Tissue and Medical Research: Code of conduct for responsible use (2011)*. 2011.
23. Korbut, E., et al., *Molecular Profile of Barrett's Esophagus and Gastroesophageal Reflux Disease in the Development of Translational Physiological and Pharmacological Studies.* Int J Mol Sci, 2020. **21**(17).
24. Nishijima, K., et al., *Impact of the biliary diversion procedure on carcinogenesis in Barrett's esophagus surgically induced by duodenoesophageal reflux in rats.* Ann Surg, 2004. **240**(1): p. 57-67.
25. Harada, H., et al., *Telomerase induces immortalization of human esophageal keratinocytes without p16INK4a inactivation.* Molecular cancer research : MCR, 2003. **1**(10): p. 729-738.
26. Boonstra, J.J., et al., *Verification and unmasking of widely used human esophageal adenocarcinoma cell lines.* J Natl Cancer Inst, 2010. **102**(4): p. 271-4.
27. Bus, P., et al., *Upregulation of miRNA-143, -145, -192, and -194 in esophageal epithelial cells upon acidic bile salt stimulation.* Diseases of the esophagus : official journal of the International Society for Diseases of the Esophagus / I S D E, 2013.
28. Janmaat, V.T., et al., *HOXA13 in etiology and oncogenic potential of Barrett's esophagus.* Nature Communications, 2021. **12**(1): p. 3354.
29. Barros, R., et al., *Dynamics of SOX2 and CDX2 Expression in Barrett's Mucosa.* Dis Markers, 2016. **2016**: p. 1532791.
30. di Pietro, M., et al., *Evidence for a functional role of epigenetically regulated midcluster HOXB genes in the development of Barrett esophagus.* Proc Natl Acad Sci U S A, 2012. **109**(23): p. 9077-82.
31. Kim, B.M., et al., *The stomach mesenchymal transcription factor Barx1 specifies gastric epithelial identity through inhibition of transient Wnt signaling.* Dev Cell, 2005. **8**(4): p. 611-22.
32. Ek, W.E., et al., *Germline genetic contributions to risk for esophageal adenocarcinoma, Barrett's esophagus, and gastroesophageal reflux.* J Natl Cancer Inst, 2013. **105**(22): p. 1711-8.
33. Contino, G., et al., *The Evolving Genomic Landscape of Barrett's Esophagus and Esophageal Adenocarcinoma.* Gastroenterology, 2017. **153**(3): p. 657-673.e1.
34. Argyrou, A., et al., *Polymorphisms of the BARX1 and ADAMTS17 Locus Genes in Individuals With Gastroesophageal Reflux Disease.* J Neurogastroenterol Motil, 2019. **25**(3): p. 436-441.
35. Chang, C.A., et al., *Congenital hiatal hernia segregating with a duplication in 9q22.31q22.32 in two families*. Am J Med Genet A, 2020. **182**(12): p. 3040-3047.
36. Nesteruk, K., et al., *Achalasia and associated esophageal cancer risk: What lessons can we learn from the molecular analysis of Barrett's–associated adenocarcinoma?* Biochimica et Biophysica Acta (BBA) - Reviews on Cancer, 2019. **1872**(2): p. 188291.
37. Nesteruk, K., et al., *Forced expression of HOXA13 confers oncogenic hallmarks to esophageal keratinocytes.* Biochim Biophys Acta Mol Basis Dis, 2020. **1866**(8): p. 165776.
38. Boccellato, F., et al., *Polarised epithelial monolayers of the gastric mucosa reveal insights into mucosal homeostasis and defence against infection.* Gut, 2019. **68**(3): p. 400-413.

39. Wang, G., et al., *Loss of Barx1 promotes hepatocellular carcinoma metastasis through up-regulating MGAT5 and MMP9 expression and indicates poor prognosis.* Oncotarget, 2017. **8**(42): p. 71867-71880.
40. Jayewickreme, C.D. and R.A. Shivdasani, *Control of stomach smooth muscle development and intestinal rotation by transcription factor BARX1.* Dev Biol, 2015. **405**(1): p. 21-32.
41. Kim, B.M., et al., *Independent functions and mechanisms for homeobox gene Barx1 in patterning mouse stomach and spleen.* Development, 2007. **134**(20): p. 3603-13.
42. Rousso, D.L., et al., *Coordinated actions of the forkhead protein Foxp1 and Hox proteins in the columnar organization of spinal motor neurons.* Neuron, 2008. **59**(2): p. 226-40.
43. De, A., et al., *Forkhead box F1 induces columnar phenotype and epithelial-to-mesenchymal transition in esophageal squamous cells to initiate Barrett's like metaplasia.* Lab Invest, 2021. **101**(6): p. 745-759.

SUPPLEMENTARY DATA

Supplementary tables

Supplementary Table 1: forward and reverse sequences of primers that were used.

	Forward primer sequence (5' to 3'):	**Reverse primer sequence (5' to 3'):**
BARX1 human	AACGCTTCGAGAAGCAGAAG	CTCGCTCGTTGGAATTGAGT
BARX1 mouse	CCAAGAAAGGACGCCGGAGTC	CTGACACCTGGGATTGGCTTC
BARX1 rat	CACCGTATTCACTGAGCTGC	CGTCTTCACCTGTAACTGGCT
FOXP1	CGAATGTTTGCTTACTTCCGACGC	ACTTCATCCACTGTCCATACTGCC
FOXF1	CGTATCTGCACCAGAACAGC	GACAAACTCCTTTCGGTCACA
Leng8 mouse	GGTTGTCTTGAAGCTGCCTT	GACCTTGGGGTGTAGGGAAT
RP2 human	AAGCTGAGGATGCTCAAAG	CCCATTAAACTCCAAGGCAA
HPRT1 rat	AGGCCAGACTTTGTTGGATT	GCTTTTCCACTTTCGCTGAT
AXIN2 human	TATCCAGTGATGCGCTGACG	TTACTGCCCACACGATAAGG
SFRP1 human	GGCTTCTTCTTCTTGGGGAC	ATCTCTGTGCCAGCGAGTTT
SFRP2 human	TCTTGCTCTTGGTCTCCAGG	CGACATAATGGAAACGCTTTG
APC human	AAAATGTCCCTCCGTTCTTATGG	CTGAAGTTGAGCGTAATACCAGT

Supplementary table 2. *FOXF1* and *FOXP1* mRNA expression correlates in different segments along human GI tract.

Pearson correlation		
Esophagus		
R	*FOXF1* dCT	*BARX1* dCT
FOXP1 dCT	0.99	N/A
FOXF1 dCT		N/A
P-value	*FOXF1* dCT	*BARX1* dCT
FOXP1 dCT	**4.90E-05**	N/A
FOXF1 dCT		N/A
Stomach		
R	*FOXF1* dCT	*BARX1* dCT
FOXP1 dCT	1.00	N/A
FOXF1 dCT		N/A
P-value	*FOXF1* dCT	*BARX1* dCT
FOXP1 dCT	**2.92E-04**	N/A
FOXF1 dCT		N/A
Duodenum		
R	*FOXF1* dCT	*BARX1* dCT
FOXP1 dCT	1.00	N/A
FOXF1 dCT		N/A
P-value	*FOXF1* dCT	*BARX1* dCT
FOXP1 dCT	**1.87E-08**	N/A
FOXF1 dCT		N/A
Jejunum		

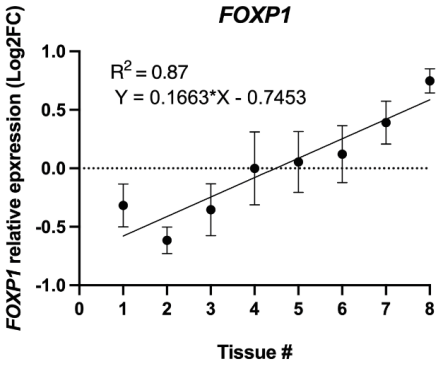
R	*FOXF1* dCT	*BARX1* dCT
FOXP1 dCT	0.97	-0.43
FOXF1 dCT		-0.20
P-value	FOXF1dCT	BARX1dCT
FOXP1 dCT	**1.35E-03**	0.40
FOXF1 dCT		0.71
Proximal ileum		
R	*FOXF1* dCT	*BARX1* dCT
FOXP1 dCT	0.99	N/A
FOXF1 dCT		N/A
P-value	FOXF1dCT	BARX1dCT
FOXP1 dCT	**1.24E-04**	N/A
FOXF1 dCT		N/A
Distal ileum		
R	*FOXF1* dCT	*BARX1* dCT
FOXP1 dCT	0.84	N/A
FOXF1 dCT		N/A
P-value	*FOXF1* dCT	*BARX1* dCT
FOXP1 dCT	0.08	N/A
FOXF1 dCT		N/A
Ascending colon		
R	*FOXF1* dCT	*BARX1* dCT
FOXP1 dCT	0.65	N/A
FOXF1 dCT		N/A
P-value	FOXF1dCT	BARX1dCT
FOXP1 dCT	0.23	N/A
FOXF1 dCT		N/A
Descending colon		
R	*FOXF1* dCT	*BARX1* dCT
FOXP1 dCT	0.96	N/A
FOXF1 dCT		N/A
P-value	FOXF1dCT	BARX1dCT
FOXP1 dCT	**5.72E-04**	N/A
FOXF1 dCT		N/A
Sigmoid/rectum		
R	*FOXF1* dCT	*BARX1* dCT
FOXP1 dCT	0.99	N/A
FOXF1 dCT		N/A
P-value	*FOXF1* dCT	*BARX1* dCT
FOXP1 dCT	**2.17E-04**	N/A
FOXF1 dCT		N/A

Supplementary table 3. *FOXF1* and *FOXP1* mRNA expression correlates in BE and EAC tissues.

Pearson correlation		
Squamous esophagus of BE patients		
R	*FOXF1* dCT	*BARX1* dCT
FOXP1 dCT	0.17	0.58
FOXF1 dCT		
P-value	*FOXF1* dCT	*BARX1* dCT
FOXP1 dCT	0.17	0.08
FOXF1 dCT		0.55
BE		
R	*FOXF1* dCT	*BARX1* dCT
FOXP1 dCT	0.74	0.06
FOXF1 dCT		0.23
P-value	*FOXF1* dCT	*BARX1* dCT
FOXP1 dCT	**0.01**	0.87
FOXF1 dCT		0.52
EAC		
R	*FOXF1* dCT	*BARX1* dCT
FOXP1 dCT	0.77	0.50
FOXF1 dCT		0.66
P-value	*FOXF1* dCT	*BARX1* dCT
FOXP1 dCT	**0.01**	0.15
FOXF1 dCT		**0.04**

Supplementary figures

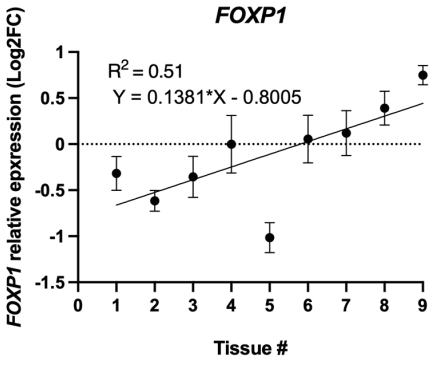
A



Esophagus	1
Stomach	2
Duodenum	3
Jejunum	4
Distal ileum	5
Ascending colon	6
Descending colon	7
Sigmoid/rectum	8

Simple linear regression, Mean with SEM

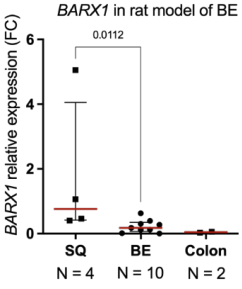
B



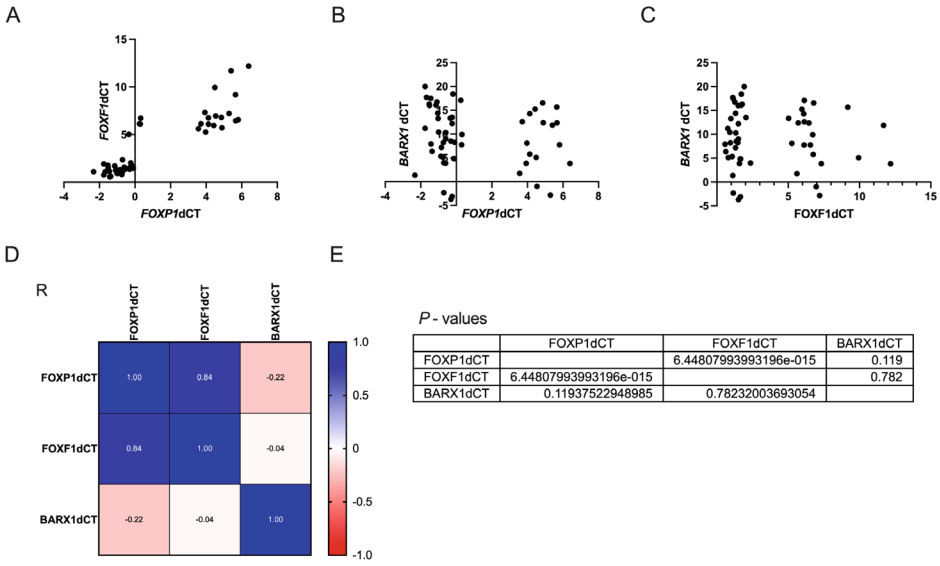
Esophagus	1
Stomach	2
Duodenum	3
Jejunum	4
Poximal ileum	5
Distal ileum	6
Ascending colon	7
Descending colon	8
Sigmoid/rectum	9

Simple linear regression, Mean with SEM

Supplementary figure 1. A gradual increase of *FOXP1* mRNA was noted from the stomach to sigmoid/rectum. A) Without proximal ileum, which is exception from the trend (*P*=0.0007). B) With proximal ileum (*P*=0.016), N=7.



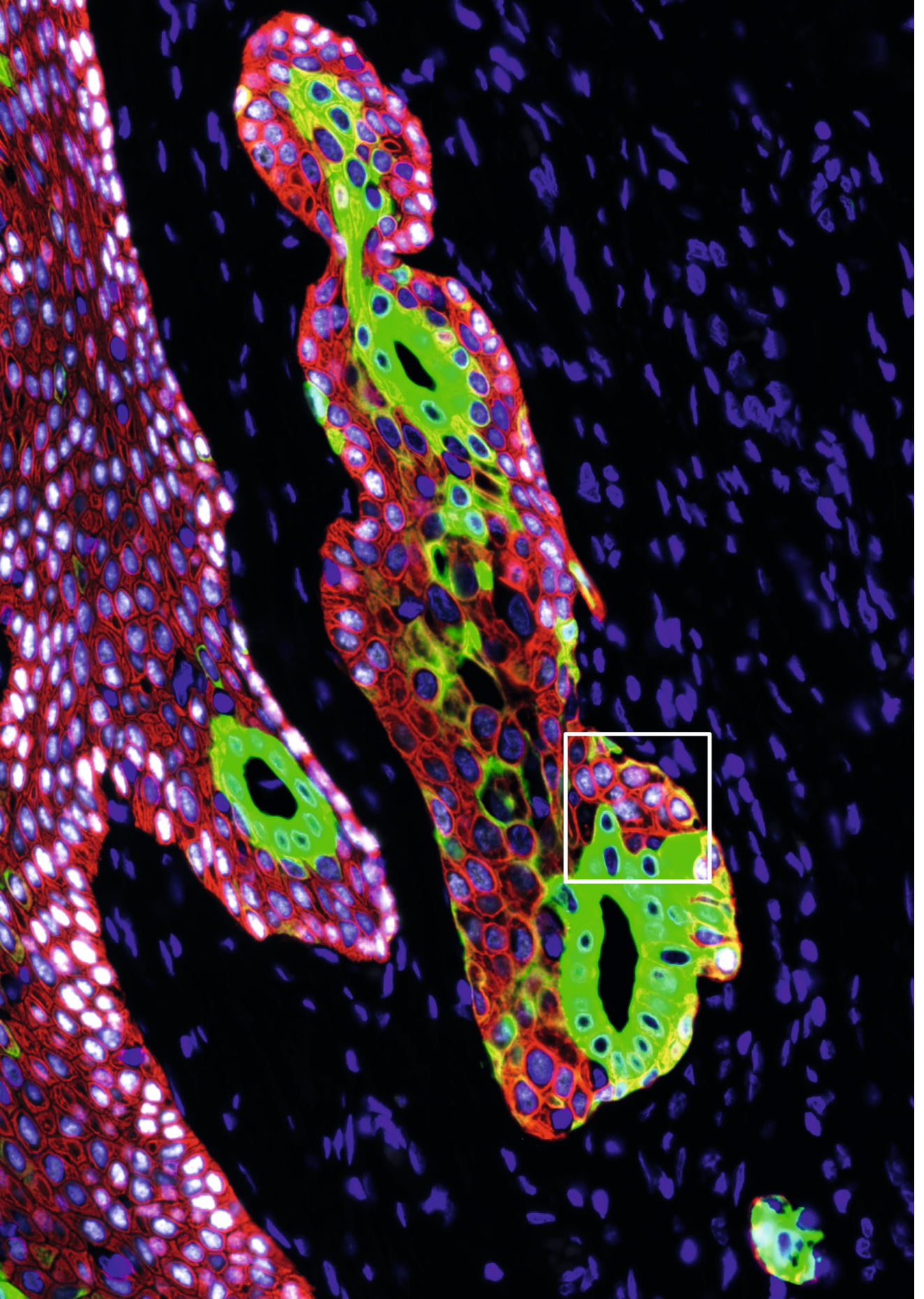
Supplementary figure 2. *BARX1* is downregulated in rat model of BE (Mann-Whitney U test).



P - values

	FOXP1dCT	FOXF1dCT	BARX1dCT
FOXP1dCT		6.44807993993196e-015	0.119
FOXF1dCT	6.44807993993196e-015		0.782
BARX1dCT	0.11937522948985	0.78232003693054	

Supplementary figure 3. *FOXF1* and *FOXP1* mRNA expression correlates in GI tract (Spearman correlation). Spearman correlation of A) *FOXF1* with *FOXP1* mRNA B) *BARX1* and *FOXP1* mRNA C) *BARX1* and *FOXF1* mRNA. D) Spearman R E) *P*-values



Chapter 9

Summary discussion and future perspectives

The current thesis combines basic and translational studies on gastrointestinal carcinogenesis. The first part (**Chapters 2-5**) focused on the detection and treatment of pancreatic cancer. The second part (**Chapters 6-8**) investigated gene-regulators of gastrointestinal embryogenesis in esophageal cancer pathogenesis.

PANCREATIC LESIONS

Early detection of pancreatic cancer

Pancreatic cancer has a dismal prognosis. Imaging-based surveillance of high-risk individuals fails to detect tumor masses on time, creating a need for biomarkers [1]. We aimed to explore whether molecules in pancreatic juice can support a diagnosis of pancreatic cancer. The first step was to determine how best to collect the most molecule-rich (secretin-stimulated) pancreatic juice from the duodenal lumen during endoscopic ultrasound (EUS). In **Chapter 2,** comparing two collection methods (through the endoscope suction channel or through-the-scope catheter) and three timeframes, we made a recommendation based on pancreatic juice yield/purity and levels of candidate biomarkers (cfDNA, pancreatic exosomes, ex-miR, and cytokines [2-7]). The optimal protocol was determined to be collection of juice through the endoscopic suction channel for up to 8 min. This process resulted in the highest total yield of pancreatic juice, DNA, and concentration of pancreas specific PLA2G1B. A possible reason is a wider diameter of the endoscopic channel suiting better for viscous pancreatic juice aspiration compared to a catheter.

Regarding the collection time, juice harvested during 4-8 min, compared to 0-4 min, contained more IgG suggesting contamination with peripheral blood. However, these blood impurities did not affect the concentration of determined cytokines, miRNA, or mutated KRAS ratio. Pancreatic juice collected from the duodenum might also contain bile and duodenal content. A technique to avoid contamination is direct cannulation of the pancreatic duct with a catheter during endoscopic retrograde cholangiopancreatography. However, this technique poses a risk of pancreatitis in up to 25% of cases [8] making it unsuitable for life-long surveillance. In contrast, such risk has not been described for collection from the duodenal lumen after secretin stimulation to our best knowledge. Nevertheless, EUS is an invasive procedure; thus, it is not recommended for screening in the general population, rather for high-risk individuals under surveillance for pancreatic cancer. Suenaga *et al*.[9] showed that prolonged collection increased the chances to detect mutated KRAS in pancreatic juice. However, in our study, prolonging collection to 15-min led to a drop in yield of pancreatic juice per min, and reduced levels of ex-miR-155, IL-10, and IFNγ, suggesting that the bulk of these biomarkers are released with a more concentrated juice at earlier timepoints. Moreover, 8-min collections will fit better in routine clinical practice with multiple EUS procedures daily. Thus, the optimal timeframe might depend on the target being investigated. We advise collection for up to 8-min when developing a biomarker panel with a variety of molecules. A panel may be essential eventually as pancreatic adenocarcinoma is a highly heterogeneous disease [10].

Our studies showed that organoid culture from pancreatic juice was feasible. The highest yield of organoids was with an alternative collection technique when suction is performed by a through-the-scope catheter positioned close to the ampulla without occluding it. When collecting during 4-8 min, the yield of organoids reached 66.7%. Further analysis and characterization of organoids derived from pancreatic juice are needed to show that organoids are of pancreatic origin (taking into account possible contaminant from duodenum or bile). While genetic markers commonly expressed in pancreatic tissue were present in PJ juice organoids, these markers are not specific to pancreas and better genetic panels discriminating pancreatic from duodenal and bile-derived epithelial cells are eagerly awaited. The next step may be to investigate whether it is possible to grow pancreatic cancer organoids from the pancreatic juice. There are indications for the success of such studies. The cellular component of pancreatic juice has been intensively investigated for the detection of pancreatic cancer [11-13]. Pancreatic juice cytology had high AUC and sensitivity values (AUC 0.84, sensitivity 54%, specificity 91%) according to a recent meta-analysis (N of paper =193) [14]. Thus, pancreatic juice contains cancer cells that may grow into organoids. Additionally, the creation of organoids was feasible for another body fluid: bile [15]. Bile-derived organoids recapitulated an inflammatory immune profile of patients with primary sclerosing cholangitis. Organoids presenting a molecular profile resembling the original tissue grew from a variety of healthy adult tissues (intestine, prostate, liver, and pancreas) and tumor masses [16]. PDAC-derived organoids showed a specific tumor phenotype and could progress into locally invasive and metastatic carcinomas upon orthotopic transplantation [17]. Modulation of organoid culture media composition allows functional selection for oncogenic mutations in patient-derived organoid lines [18]. Thus, using a selection medium might discriminate cancer from normal pancreatic juice-derived organoids. In addition, an effective and accurate selection method will level the presence of contaminants. Pancreatic juice organoids may be used to diagnose pancreatic cancer, predict response to treatment, or as a research model. However, to our best knowledge, only one study (conference abstract) reported the establishment of pancreatic juice-derived organoids [19]. Pancreatic juice was collected preoperatively by endoscopic retrograde cholangiopancreatography in individuals with intraductal papillary mucinous neoplasms [19]. Organoids developed into subcutaneous tumors after transplantation into nude mice. In contrast to our study with a success rate of 66.7%, they had a success rate of 78.2%, likely resulting from both the technique of collection (cannulation of ampulla) and a higher amount of juice used (15 mL vs 500-1000μL in our study).

Pancreatic juice is receiving attention as a resource for biomarkers for detection of early pancreatic cancer. Pancreatic ductal adenocarcinoma (PDAC) can evolve from acinar or ductal pancreatic cells [20]. Both belong to the exocrine pancreas producing and transporting pancreatic juice into the duodenum [20]. Thus, pancreatic juice will directly be in contact with newly developed cancer cells when these are not yet expected to be connected with a distinct blood supply. It has been shown to be possible to detect cancer cells and cancer-derived molecules in pancreatic juice [3, 11-13]. However, the search for cancer-specific molecules with high diagnostic performance is ongoing. While the general composition of pancreatic juice in physiology and main pancreatic enzymes are known for several decades [21, 22], a new stage in the characterization of pancreatic juice in health and diseases occurred due to the recent development of technologies. The pancreatic juice proteome was revised in pancreatic

adenocarcinoma [23-28], pre-cancer [29], benign pancreatic diseases [24], and with no apparent pancreatic pathology [30]. Pancreatic juice is protein rich and proteomics approaches have revealed hundreds of proteins in human pancreatic juice, including pancreatic enzymes and other pancreas-associated proteins. Pancreatic juice in PDAC had distinct proteins from noncancerous controls such as kallikrein 1, IGFBP2, lithostathine 1, pancreatic secretory granule membrane major glycoprotein, tumor-associated trypsin inhibitor, pancreatitis-associated protein 1, pancreatic ribonuclease, and T-cell receptor beta chain [23]. ^{1}H nuclear magnetic resonance allowed the comparison of the metabolome of pancreatic juice in physiology and pathological conditions [31] showing that PDAC can be discriminated from benign pancreatic diseases based on metabolic profile. Specifically, PDAC samples were overtly glycolytic, with accumulation of lactate. Methylated DNA markers in pancreatic juice were also strongly associated with pancreatic cancer [32]. Digital next-generation sequencing detected low-abundance cancer mutations in pancreatic juice [3], with PDAC and IPMN being more likely to have mutant DNA in pancreatic juice than controls [3]. These examples demonstrate that PJ serves as a basis for biomarker research and that novel approached are identifying ever more discriminating features between PDAC cases and controls.

Our study also aimed to contribute to the current understanding of pancreatic juice composition. Most cells release double-layer phospholipid membrane vesicles [33]. Initially, these extracellular vesicles (EVs) were considered to be cellular debris. However, currently EVs are attracting significant research interest due to the discovery of their specific functions [33]. EVs carry biologically active molecules between cells therefore mediate in intercellular communication. They are highly heterogeneous in size, cargo, membrane composition, biogenesis, and specific biological function [33]. The field of EV biology is rapidly expanding [33]. EVs and their molecular content have potential as biomarkers for the detection of pancreatic cancer [34]. In **Chapter 3,** we characterized the composition of pancreatic juice in terms of EV size and concentration. Nanoparticle tracking analysis allows detecting nanoparticles (from 30 to 1,000 nm) scattering the light from a laser beam [35]. A digital camera records Brownian motion of nanoparticles in a liquid suspension. After this, software automatically analyzes the particles individually and calculates their hydrodynamic diameters using the Stokes-Einstein equation. This technique showed that pancreatic juice contains fewer extracellular vesicles than serum, but that they are larger in diameter. As EV size indicates vesicular subtypes [36], this suggests that this body fluid has a different proportional composition of different types of extracellular vesicles. The concentration of extracellular vesicles did not differ between controls and cancer patients in either pancreatic juice or serum. Importantly, however, there was an enhanced proportion of large vesicles (>350) in pancreatic cancer compared to non-malignant controls. This difference was seen in pancreatic juice but not in serum, suggesting that the nature of pancreatic juice EVs not only differs from serum EVs, but that patient-derives PJ EVs are of even a different nature. The next step might be investigating the nature of large vesicles produced by cancer cells and to what extent is their function different from EVs from normal cells? Specific molecular markers of large vesicles might be useful for the diagnosis of pancreatic cancer.

EVs transport signaling molecules such as proteins, miRNA, mRNA, tRNA-derived small RNAs, Y RNAs, circRNA, and lncRNAs [37]. In several human malignancies [38-40] including pancreatic cancer [41, 42], when a tumor arises, progresses, or resists chemotherapy, it parallels with altered expression of certain free miRNA and miRNA in EVs (EV-miRNA). In **Chapter 4,** we isolated microRNA from pancreatic juice-derived and serum-derived extracellular vesicles and found that while only EV-miR-210 was overexpressed in serum from pancreatic cancer cases, EV-miR-21, EV-miR-25, EV-miR-210, EV-miR-16 were overexpressed in pancreatic juice, suggesting that pancreatic juice might be a better source for biomarkers. A model with EV-miR-21, EV-miR-25, EV-miR-16, and serum EV-miR-210, CA19-9 could distinguish pancreatic cancer from controls with a sensitivity of a specificity of 84.2% and a sensitivity of 81.5%. At present, it remains unclear whether these EV-miRNAs are also overexpressed at early-stage pancreatic cancer which is not yet diagnosed with imaging techniques. Long term collection of pancreatic juice from surveillance cohorts may in future allow us to determine whether EV-miRNAs are able to discriminate patients who go on to develop PDAC from those who do not. Moreover, sequencing technologies might help to determine EV-miRNA in pancreatic juice with better diagnostic performance. We observed a difference between cancer and controls in the expression of miRNA-16, which was previously used as control miRNA [43]. The current study had a balanced number of controls and cases per batch, and we used the same volume of pancreatic juice for each sample; thus, we do not expect that possible variations affected results. However, an internal control might be needed for developing a reliable tool suitable for routine clinical practice.

Treatment of pancreatic cancer - novel molecular targets

Treatment of PDAC depends on the stage of cancer at diagnosis (resectable, borderline, locally advanced, and metastatic disease), and consists of surgery, chemotherapy, chemoradiotherapy, and supportive care [44]. Combination of surgery with chemotherapy is remaining the only potentially curative treatment but available for low percentage of eligible patients. The main of treatment for advanced PDAC is Systemic chemotherapy such as [FOLFIRINOX (5-fluorouracil, folinic acid [leucovorin], irinotecan, and oxaliplatin) and gemcitabine plus nab-paclitaxel] [45]. For locally advanced disease and metastatic disease, other treatment options are being investigated such as personalized medicine, innovative targets, immunotherapy, therapeutic vaccines, adoptive T-cell transfer, or stemness inhibitors. However, any of them have resulted yet in the significant improvement of survival [44]. Radiation, chemotherapy, immunotherapy, and the use of targeted drugs just mildly increase survival rate and reduce cancer-related symptoms [46]. Thus, new approaches and findings in diagnosis and cure are very valuable.

Newly synthesized molecule Metavet has been recently described as promising anti-cancer agent for PDAC [47]. However, this dual inhibitor of GSK3b and HDAC Metavert upregulated β-catenin protein level in pancreatic cancer cells. In **Chapter 5,** we demonstrate that an increase in β-catenin expression levels in PDAC cells during dual GSK3b and HDAC inhibition does not contribute to killing of these cells in response to these inhibitors. Knock-down of β-catenin, as expected, reduced cell viability in itself.

Upregulated β-catenin might affect other cancerous processes induced by dual inhibition (for example, by Metavert [47]).

We used two inhibitors of GSK-3b: CHIR99021 and TWS119. TWS119 is a less potent inhibitor with an IC50 towards the purified enzyme of 30 nM compared to 6.7 nM for CHIR99021. Expectedly, with the same dosage of 5 μM, TWS119 showed activation of β-catenin signaling as compared to CHIR99021, as determined by luciferase activity (in two cell lines) and expression of the well-established β-catenin target *AXIN2* (in all three cell lines) and synergism with HDAC inhibitor Vorinostat. Of note, TWS119 showed enhanced killing efficacy as compared to CHIR99021 while displaying lower β-catenin pathway activation, again pointing towards a β-catenin-independent killing mechanism. The limitation is that inhibitor CHIR99021 has almost the same IC50 for inhibiting GSK-3b and -3a (6.7 nM and 10 nM, respectively); thus, using a specific inhibitor of GSK3b would have been preferable. Nevertheless, CHIR99021 is considered to be the most potent and selective GSK3b inhibitor available to date. To our best knowledge, only one inhibitor with superior kinome-wide selectivity has been described [48], but this compound is not commercially available. TWS119 is quoted to be a non-reversible inhibitor of GSK-3b with an IC50 of 60 nM, based on studies of Juan Manuel Domínguez *et al.* [49], but its activity towards GSK3a has not been tested as far as we are aware. However, selectivity of Metavert is also not shown – while it inhibits GSK3b, the binding-site on GSK3b could not be established, and its activity towards GSK3a was not tested to our best knowledge.

In contrast to inhibition of GSK3 which targets the β-catenin destruction complex, the mechanism of β-catenin upregulation by an inhibitor of HDAC is less clear. Nevertheless, an effect of HDAC inhibitors on Wnt/β-catenin signaling was reported before. HDAC inhibitors increased β-catenin signaling in various colorectal cancer cell lines [50], breast cancer cell lines [51], and HDAC regulated Wnt signaling in ureteric bud epithelium [52]. One of the possible mechanisms described by Bordonaro et al. [50] was that HDAC inhibitors with different chemical structures increase the level of active (Ser-37 and Thr-41 dephosphorylated) β-catenin by increasing protein phosphatase activity. Another possible explanation of the effect of HDAC inhibition on Wnt/β-catenin signaling described is that HDAC also has non-histone targets [53] and can modulate activity of LEF/TCF. These transcriptional factors transduce β-catenin signaling to the nucleus [54]. In the absence of β-catenin, Lef-1 is repressed by HDAC1 activity. When the ratio of β-catenin to HDAC increases, HDAC is dissociated from LEF-1, and a signal can be transduced. Billin et al. [54] showed that HDAC inhibitors stimulated β-catenin signaling in a LEF-reporter assay in HEK 293 cells. In our study, we, at the same time, increase β-catenin level by GSK3 inhibition and inhibit HDAC, which can explain the synergistic effect seen in qPCR but not in a reporter assay. It might also explain minor differences in reporter assay and qPCR. Hence, increase in Wnt/β-catenin signaling in our study during HDAC inhibition is not completely clear but more likely can be explained by described above mechanisms.

In summary, we demonstrate in part 1 of this thesis that detection of pancreatic cancer with biomarkers in pancreatic juice may feasible in the future. A better understanding of the molecular mechanisms underlying tumor formation and treatment may serve as a starting point for the development of better treatments for this disease.

ESOPHAGEAL LESIONS

Gene expression deregulation in gastrointestinal (pre)malignant lesions including Barrett's esophagus

Barrett's esophagus (BE), a precursor of esophageal adenocarcinoma, is most often diagnosed in patients suffering from gastric esophageal reflux disease. BE is characterized by a metaplastic epithelium resembling that of the distal gut. Although BE and colon epithelia are not identical (for example, in terms of heterogeneity), histological features of BE greatly overlaps with colonic epithelium, including the presence of enterocytes, enteroendocrine cells, and goblet cells [55]. The two biggest conceptual challenges faced in the field are the molecular basis of Barrett's esophagus and its cell of origin. We addressed these questions in **Chapter 6**. Homeobox (*HOX*) genes are important mediators of homeotic transformations [56]. We demonstrate that BE epithelia show molecular overlap with colonic epithelia, exhibiting a similar *HOX* gene expression pattern. Moreover, *HOX* coding is established in the intestine at the stem cell level in both the normal epithelium, where *HOX* coding is already present in the location-specific stem cells along the gastrointestinal tract, as well as BE epithelium, where colon-specific pattern is also present in the BE stem cell. Thus, this data shows that inherent location identity drivers are a feature of the early gastrointestinal stem cells and that the BE stem cells resemble colonic stem cells in their *HOX* expression pattern. This evidence is of importance because existing theories regarding cells of origin of BE (except the transdifferentiation theory) suggest that BE arises from stem-like progenitor cells that can give rise to all different cell types of BE [57].

The distal *HOX* paralogue *HOXA13* was most prominently expressed in Barrett's esophagus, its stem cells, gastric intestinal metaplasia, and intestinal heterotopias. *HOXA13* conferred the phenotypical and premalignant characteristics of Barrett's esophagus as shown by an *in vivo* organotypic culture system, which appears mediated via *HOXA13*-dependent downregulation of the epidermal differentiation complex on chromosome 1. These data on *HOXA13* provide a molecular explanation of the pathogenesis of Barrett's esophagus but do not yet explain the cell of origin for this condition. Employing a mouse model that contains a reporter coupled to the *Hoxa13* promotor, we identified single *HOXA13*-positive cells in an otherwise *HOXA13* depleted environment in the physiological gastroesophageal junction and distally thereof. This was parallel to the human GI tract expressing *HOXA13* in individual cells of gastroesophageal junction and esophageal submucosal glands. The number of these *HOXA13*+ cells increased in Barrett's esophagus and intestinal metaplasia and was accompanied by a genetic profile suggesting a submucosal origin and a proliferative advantage. The fact that some *HOXA13*-positive cells were found in human esophagus, while not in mice, may represent species differences. Technical differences may also play a part: cutting sections from tissue reduces the number of cells assessed simultaneously while sorting cells for RNAseq enriches for cells able to withstand flow-sort shear stress without undergoing apoptosis – *HOXA13*-positive cells may be among this population. We conclude that rare, single *HOXA13*$^{+}$ cells are indeed present in normal physiology, including at the GEJ, and the number of these cells is increased in BE.

HOXA13 expression conferred a competitive advantage in cell line models and is expected to mediate repopulation of the esophagus following reflux disease. Therefore, we propose that Barrett's esophagus and associated esophageal adenocarcinoma result from an expansion of this novel gastroesophageal *HOXA13*-expressing compartment following epithelial injury. Indeed, *HOXA13*-overexpressing cells were less sensitive to bile/acid treatment suggesting that *HOXA13* cells are activated upon tissue damage. Data from the literature suggests that *HOXA13* also confers protection against other damaging factors, as, for instance, *HOXA13* protected ESCC cells from chemotherapy [58]. BE can be considered as a wound healing process initiated by esophageal injury from GERD [57]. Data from the literature support idea that *HOXA13* is involved in processes related to wound healing and regeneration after tissue damage of zebrafish bony fin ray [59], Xenopus limb and tail [60], spermatogenic cells of mice and cell lines in response to UV irradiation [61]. Furthermore, *HOXA13* promoted wound healing of *in vitro* gastric cancer cell cultures [62]. Thus, our new data is in line with other studies suggesting a role for *HOXA13* resistance to stressors and wound healing, which might be further investigated with lineage tracing experiments.

We focused on *HOXA13* is because it was the highest overexpressed paralogue group 13 member identified in Barrett's tissue compared to squamous tissue in absolute measurements. According to the collinear theory, a paralogue group 13 member would confer more distal characteristics than a more anterior paralogue group member. An additional reason to focus on cluster A, with particular emphasis on *HOXA13*, is that it was less well studied than the *HOXB* cluster. However, the role of other *HOX* genes should not be discarded. A11, B6,9,13, *HOXA11*, and *HOXC10* had a similar expression change in Barrett's vs squamous epithelium and air-liquid interface cultures. This suggests *HOX* genes are likely to act in tandem and that *HOXB 6, 9, 13,* and probably also *HOXA11* and *HOXC10* would be interesting candidates for study as well. Having seen the major consequences of cellular morphology and proliferation by manipulating only one *HOX* gene (*HOXA13*), one may speculate whether manipulation of multiple *HOX* genes simultaneously would have an even stronger effect on the described processes. Additive and synergistic effects of *HOX* genes were described before [63-65]. Analysis of publicly available single-cell RNA seq data (#EGAS00001003144, recently described by Owen *et al* [66] demonstrates that *HOXA13* is expressed in non-squamous (*TFF3*+) ESMG cells of healthy esophagus and overlaps with other *HOX* genes within the same cells. This overlap is imperfect and mainly limited to the caudal *HOX* genes *HOXA11* and *HOXB13* and *HOXB6*. However, in our dataset, modulation of *HOXA13* in two cell lines did not significantly affect any of the other *HOX* genes (as indicated by RNAseq data), demonstrating that in this instance, only *HOXA13* effects were studied in these cell models.

The data presented in both **Chapter 6** and **7** of this thesis, as well as literature suggest a role of *HOXA13* in esophageal carcinogenesis. Using qPCR, we found in **Chapter 6** that *HOXA13* mRNA levels are increased in EAC indicative of these cells' capacity to proliferate and outcompete other cells. Also, *HOXA13* provided a proliferative advantage in cell line models and activated cancer-related Notch signaling. *HOXA13* has previously been shown to be involved in ESCC [67] and other cancer types [62, 68-70]. In **Chapter 7** we further investigated the mechanistic role of *HOXA13* in ESCC and examined

whether aberrant *HOXA13* might drive carcinogenesis in esophageal keratinocytes. We overexpressed *HOXA13* in immortalized human esophageal cell line EPC2, performed gene expression profiling, phosphoprotein profiling, and investigated the effect of overexpression on cell functions. We found that *HOXA13* expression provides esophageal keratinocytes with oncogenic characteristics. This transcription factor provided a proliferation advantage to keratinocytes, reduced sensitivity to chemical agents, regulated MHC class I expression and differentiation status, and promoted cellular migration. Our data indicate a crucial role of *HOXA13* at the early stages of esophageal carcinogenesis. High *HOXA13* expression was associated with poor clinic-pathological characteristics of ESCC and high TNM stage [71]. To our best knowledge, there is no clinical data in precursor lesions and early ESCC supporting a role of *HOXA13*. Thus, clinical studies in squamous dysplasia could complement our *in vitro* data and further contribute to understanding of *HOXA13* in ESCC.

The place of HOXA13 in cell signaling is not completely clear. Functionally, HOXA13 is a sequence-specific, AT-rich binding transcription factor, the overexpression of which changes expression of 2995 genes in our study. However, it remains unclear which of these genes are directly regulated by HOXA13 transcriptional activity. As far as we are aware, only one study performed Chip-DSL experiments to show genes targeted by *HOXA13* on an esophageal cell line [72]. The authors analyzed protein expression in EC-109 cells after *HOXA13*-knockdown**.** Then they performed CHIP-DSL experiments with MAS software and identified the signaling pathways targeted by HOXA13. The most affected were the MAPK signaling pathway, Wnt signaling pathway, TGF-b signaling pathway, cell cycle, gap junction, focal adhesion, tight junctions, mTOR signaling, and p53 signaling pathway, apoptosis. The biological functions affected by HOXA13 in this study are cancer, cellular growth and proliferation, cell death, cellular movement, hematology system development, nervous system development, embryo development, DNA replication, recombination and repair, tumor morphology. They did not provide information regarding the direction of the changes or functional experiments. Raw data for comparison with ours are also not available, however, the overall picture emerging appears to be similar to ours: HOXA13 affects cytoskeletal and proliferation signaling, cancerous processes, and migration. However, ChiP experiments rely on precipitation of the transcription factor with robust antibodies. Antibodies were not validated in this study. Thus, it is complicated to draw conclusions based on these antibody-based assays. Similarly, one of our study's limitations that we could not demonstrate changes in HOXA13 protein because western blot and immunohistochemistry for HOXA13 were not possible. Antibody selectivity is very rarely established to be exclusive for a single gene product due to functionally important regions of amino acid homology, therefore frequent epitope sharing by antigens. Thus, investigation of *HOXA13* remains impossible unless tagged (e.g., with Flag-tag, or GFP), which in turn may affect protein function. Given the large-scale gene expression modulation effects of HOXA13, it is not surprising that phosphoprotein profiling in our study was also affected upon HOXA13 overexpression. We were interested to investigate the final effect and direction of kinomic changes regardless of the mechanism by which these changes were reached for each particular signaling molecule. Thus, only phosphorylated forms of these proteins were investigated. However, we cannot exclude the possibility that the observed changes in phospho-levels of kinases upon *HOXA13*

overexpression are due to changes in their respective protein levels. Nevertheless, expression of corresponding genes on mRNA level did not change except for a slight downregulation of RhoK2.

Another an interesting question remains whether endogenous inducer of *HOXA13*, FGF2, might induce activation of *HOXA13* related pathways in EPC-hTERT cells. FGF2 correlates with poor survival of patients with ESCC [73] and activated Erk signaling in ESCC. It has also been suggested that HOXA13 itself regulates *FGF2*-expression, suggesting a positive feedback loop. However, there were not any changes in *FGF2* expression upon overexpression of *HOXA13* in our cell model suggesting that this pathway is perhaps not relevant in our setting.

In Chapter 8 we extended our search for driving genetic factors for esophageal malignancy. We found that *BARX1* is expressed in the adult human gastrointestinal tract and overexpressed in BE and EAC. Bile/acid treatment stimulates expression of BARX1 in cell models. It is remains unclear 1) whether epithelial cells or fibroblast are expressing *BAXR1*, 2) what the functional role is of *BARX1* in these conditions, 3) whether BARX1 regulates the Wnt/β-catenin pathway in adult human esophagus similarly to mouse embryogenesis. The fact that this gene comes up in GWAS analysis implies that it is relevant to these diseases. But alternatively, the SNP may also be an eQTL for other genes which mediate these effects.

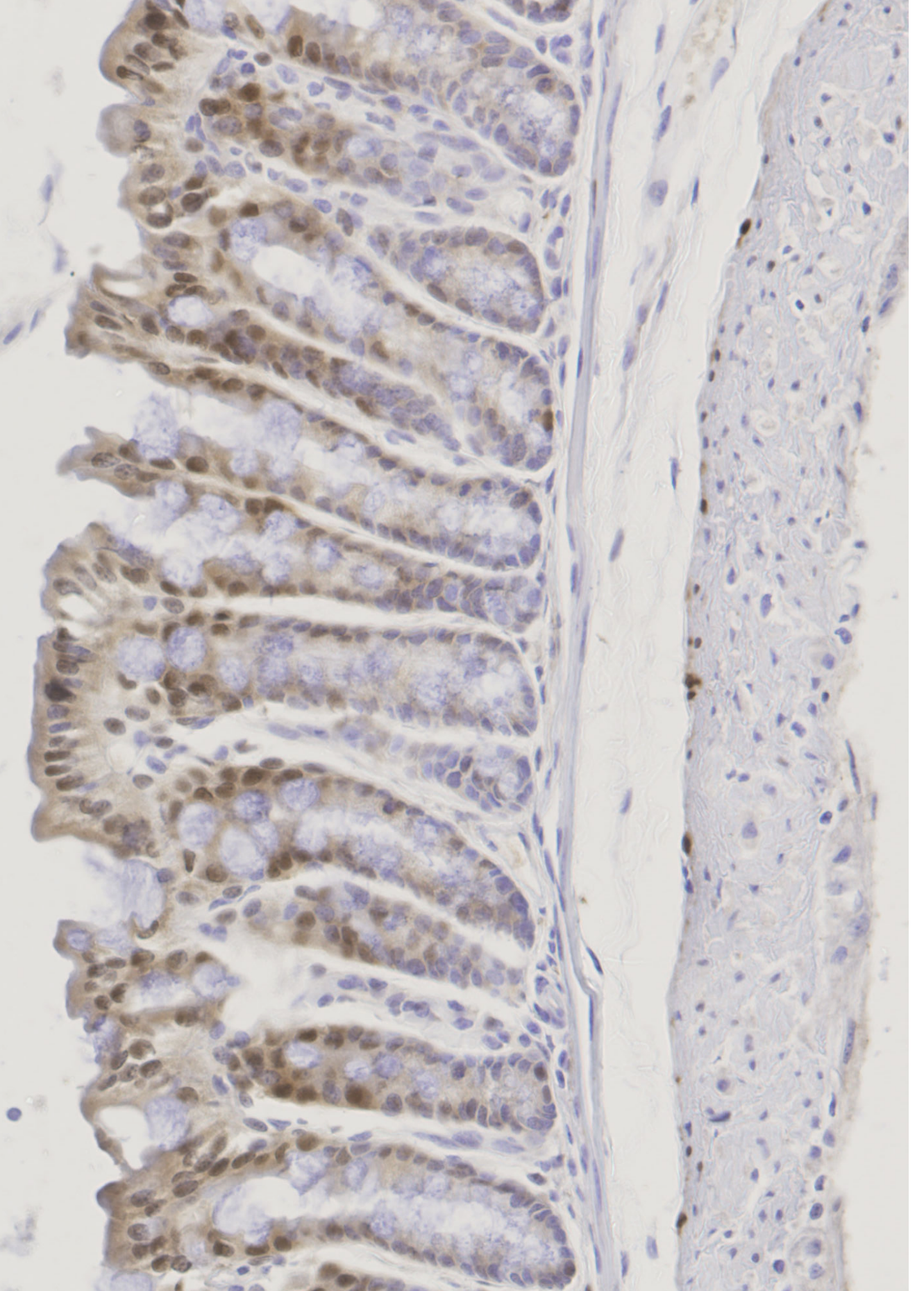
Our data highlights the importance of human clinical data and accurate models for BE research. Conservative *HOX* genes showed similar expression pattern in both human and mouse gastrointestinal tract. This allowed us to employ C57BL/6J-Hoxa13-GFP heterozygous mutant mouse model for *in vivo* study. But *BARX1* expression in human gut was distinct from both mouse and rat. Moreover, the surgical rat model of BE showed opposite *BARX1* expression patterns to the human setting. This is not unexpected as rodents have anatomical differences from the human upper gastrointestinal tract [74]. Rodent's forestomach is squamous, they lack esophageal submucosal glands and do not have regurgitation reflex [75] and do not naturally experience BE or GERD [74]. Thus, we should carefully consider a model and interpret animal data. In this context, development of *in vitro* models is important, especially cells growing in 3D culture. Cell lines are lacking complexity and undergo selection for continuous 2D growth [76]. 3D and air-liquid interface culture can help overcome some of these limitations. BAR-T cells differentiated similarly in both air-liquid interface culture and rat trachea *in vivo* reconstitution model. The stimulating effect of *HOXA13* on the growth of keratinocytes was seen only upon 3D culture of EPC2-hTERT cells but not in 2D MTT assay. Thus, 2D growth can mask important differences. Another critical point is heterogeneity between cell lines. Squamous esophageal cell line EPC2-hTERT cells had low or undetectable expression of *BARX1* similarly to biopsies. However, analogical Het1-A cells were highly positive for *BARX1*. Indeed, there is 57% similarity between human biopsies and the EPC2 cell model in gene expression and only 26% similarity for the Het-1A model [77]. Combining both expression profiles resulted in 73% overlap with the human biopsy data. Organoid culture or differentiated embryonic stem culture might fill the gap and complement data from cell line models or animal models.

In summary, this part of the thesis shows that genes that determine positional identity during embryogenesis also play an important role in adult physiology, and that disruption in these genes can contribute to gastrointestinal pathologies.

REFERENCES

1. Pereira, S.P., et al., *Early detection of pancreatic cancer.* Lancet Gastroenterol Hepatol, 2020. **5**(7): p. 698-710.
2. Choi, M.H., et al., *Mutation analysis by deep sequencing of pancreatic juice from patients with pancreatic ductal adenocarcinoma.* BMC Cancer, 2019. **19**(1): p. 11.
3. Yu, J., et al., *Digital next-generation sequencing identifies low-abundance mutations in pancreatic juice samples collected from the duodenum of patients with pancreatic cancer and intraductal papillary mucinous neoplasms.* Gut, 2017. **66**(9): p. 1677-1687.
4. Suenaga, M., et al., *Pancreatic Juice Mutation Concentrations Can Help Predict the Grade of Dysplasia in Patients Undergoing Pancreatic Surveillance.* Clin Cancer Res, 2018. **24**(12): p. 2963-2974.
5. Wang, J., et al., *Circulating microRNAs in Pancreatic Juice as Candidate Biomarkers of Pancreatic Cancer.* J Cancer, 2014. **5**(8): p. 696-705.
6. Sadakari, Y., et al., *MicroRNA expression analyses in preoperative pancreatic juice samples of pancreatic ductal adenocarcinoma.* Jop, 2010. **11**(6): p. 587-92.
7. Nakamura, S., et al., *Pancreatic Juice Exosomal MicroRNAs as Biomarkers for Detection of Pancreatic Ductal Adenocarcinoma.* Annals of Surgical Oncology, 2019.
8. Yamakawa, K., et al., *Evaluation of efficacy of pancreatic juice cytology for risk classification according to international consensus guidelines in patients with intraductal papillary mucinous neoplasm; a retrospective study.* Pancreatology, 2019. **19**(3): p. 424-428.
9. Suenaga, M., et al., *The Effect of Pancreatic Juice Collection Time on the Detection of KRAS Mutations.* Pancreas, 2018. **47**(1): p. 35-39.
10. Jenkinson, C., et al., *Biomarkers for early diagnosis of pancreatic cancer.* Expert Rev Gastroenterol Hepatol, 2015. **9**(3): p. 305-15.
11. Kawada, N., et al., *Pancreatic juice cytology as sensitive test for detecting pancreatic malignancy in intraductal papillary mucinous neoplasm of the pancreas without mural nodule.* Pancreatology, 2016. **16**(5): p. 853-858.
12. Koshita, S., et al., *Pancreatic juice cytology with immunohistochemistry to detect malignancy and histologic subtypes in patients with branch duct type intraductal papillary mucinous neoplasms of the pancreas.* Gastrointest Endosc, 2017. **85**(5): p. 1036-1046.
13. Nakaizumi, A., et al., *Cytologic examination of pure pancreatic juice in the diagnosis of pancreatic carcinoma. The endoscopic retrograde intraductal catheter aspiration cytologic technique.* Cancer, 1992. **70**(11): p. 2610-4.
14. Tanaka, M., et al., *Cytologic Analysis of Pancreatic Juice Increases Specificity of Detection of Malignant IPMN - A Systematic Review.* Clin Gastroenterol Hepatol, 2019.
15. Soroka, C.J., et al., *Bile-Derived Organoids From Patients With Primary Sclerosing Cholangitis Recapitulate Their Inflammatory Immune Profile.* Hepatology, 2019. **70**(3): p. 871-882.
16. Hohwieler, M., et al., *Pancreatic Progenitors and Organoids as a Prerequisite to Model Pancreatic Diseases and Cancer.* Stem Cells International, 2019. **2019**: p. 9301382.
17. Boj, Sylvia F., et al., *Organoid Models of Human and Mouse Ductal Pancreatic Cancer.* Cell, 2015. **160**(1): p. 324-338.
18. Driehuis, E., et al., *Pancreatic cancer organoids recapitulate disease and allow personalized drug screening.* Proceedings of the National Academy of Sciences, 2019. **116**(52): p. 26580.
19. Kita, E., et al., *Tu2013 - Human Organoid Models of Intraductal Papillary Mucinous Neoplasm (IPMN) Derived from Pancreatic Juice.* Gastroenterology, 2018. **154**(6): p. S-1369.
20. Ferreira, R.M.M., et al., *Duct- and Acinar-Derived Pancreatic Ductal Adenocarcinomas Show Distinct Tumor Progression and Marker Expression.* Cell Reports, 2017. **21**(4): p. 966-978.
21. Göke, B., et al., *Resolution of human exocrine pancreatic juice proteins by reversed-phase high performance liquid chromatography (HPLC).* Pancreas, 1990. **5**(3): p. 261-6.
22. Beaudoin, A.R., P. St-Jean, and G. Grondin, *Pancreatic juice composition: new views about the cellular mechanisms that control the concentration of digestive and nondigestive proteins.* Dig Dis, 1989. **7**(4): p. 210-20.
23. Chen, R., et al., *Quantitative proteomic profiling of pancreatic cancer juice.* Proteomics, 2006. **6**(13): p. 3871-9.
24. Chen, R., et al., *Comparison of pancreas juice proteins from cancer versus pancreatitis using quantitative proteomic analysis.* Pancreas, 2007. **34**(1): p. 70-9.
25. Grønborg, M., et al., *Comprehensive proteomic analysis of human pancreatic juice.* J Proteome Res, 2004. **3**(5): p. 1042-55.
26. Lv, S., et al., *Transthyretin, identified by proteomics, is overabundant in pancreatic juice from pancreatic carcinoma and originates from pancreatic islets.* Diagn Cytopathol, 2011. **39**(12): p. 875-81.
27. Makawita, S., et al., *Integrated proteomic profiling of cell line conditioned media and pancreatic juice for the identification of pancreatic cancer biomarkers.* Mol Cell Proteomics, 2011. **10**(10): p. M111 008599.
28. Tian, M., et al., *Proteomic analysis identifies MMP-9, DJ-1 and A1BG as overexpressed proteins in pancreatic juice from pancreatic ductal adenocarcinoma patients.* BMC Cancer, 2008. **8**: p. 241.
29. Chen, R., et al., *Elevated level of anterior gradient-2 in pancreatic juice from patients with pre-malignant pancreatic neoplasia.* Mol Cancer, 2010. **9**: p. 149.
30. Doyle, C.J., et al., *The proteome of normal pancreatic juice.* Pancreas, 2012. **41**(2): p. 186-94.
31. Cortese, N., et al., *Metabolome of Pancreatic Juice Delineates Distinct Clinical Profiles of Pancreatic Cancer and Reveals a Link between Glucose Metabolism and PD-1(+) Cells.* Cancer Immunol Res, 2020. **8**(4): p. 493-505.
32. Majumder, S., et al., *Methylated DNA in Pancreatic Juice Distinguishes Patients With Pancreatic Cancer From Controls.* Clin Gastroenterol Hepatol, 2020. **18**(3): p. 676-683 e3.
33. Margolis, L. and Y. Sadovsky, *The biology of extracellular vesicles: The known unknowns.* PLoS Biol, 2019. **17**(7): p. e3000363.
34. Yee, N.S., et al., *Extracellular Vesicles as Potential Biomarkers for Early Detection and Diagnosis of Pancreatic Cancer.* Biomedicines, 2020. **8**(12).
35. Filipe, V., A. Hawe, and W. Jiskoot, *Critical evaluation of Nanoparticle Tracking Analysis (NTA) by NanoSight for the measurement of nanoparticles and protein aggregates.* Pharm Res, 2010. **27**(5): p. 796-810.
36. Maas, S.L.N., X.O. Breakefield, and A.M. Weaver, *Extracellular Vesicles: Unique Intercellular Delivery Vehicles.* Trends in Cell Biology, 2017. **27**(3): p. 172-188.
37. Veziroglu, E.M. and G.I. Mias, *Characterizing Extracellular Vesicles and Their Diverse RNA Contents.* Frontiers in genetics, 2020. **11**: p. 700-700.
38. Calin, G.A., et al., *Familial Cancer Associated with a Polymorphism in ARLTS1.* New England Journal of Medicine, 2005. **352**(16): p. 1667-1676.

39. He, L., et al., *A microRNA polycistron as a potential human oncogene.* Nature, 2005. **435**(7043): p. 828-833.
40. Lu, J., et al., *MicroRNA expression profiles classify human cancers.* Nature, 2005. **435**(7043): p. 834-838.
41. Rawat, M., et al., *MicroRNA in Pancreatic Cancer: From Biology to Therapeutic Potential.* Genes, 2019. **10**(10): p. 752.
42. Daoud, A.Z., et al., *MicroRNAs in Pancreatic Cancer: biomarkers, prognostic, and therapeutic modulators.* BMC Cancer, 2019. **19**(1): p. 1130.
43. Nakamura, S., et al., *Pancreatic Juice Exosomal MicroRNAs as Biomarkers for Detection of Pancreatic Ductal Adenocarcinoma.* Ann Surg Oncol, 2019. **26**(7): p. 2104-2111.
44. Hosein, A.N., R.A. Brekken, and A. Maitra, *Pancreatic cancer stroma: an update on therapeutic targeting strategies.* Nat Rev Gastroenterol Hepatol, 2020. **17**(8): p. 487-505.
45. Mizrahi, J.D., et al., *Pancreatic cancer.* The Lancet, 2020. **395**(10242): p. 2008-2020.
46. Adamska, A., A. Domenichini, and M. Falasca, *Pancreatic Ductal Adenocarcinoma: Current and Evolving Therapies.* International journal of molecular sciences, 2017. **18**(7): p. 1338.
47. Edderkaoui, M., et al., *An Inhibitor of GSK3B and HDACs Kills Pancreatic Cancer Cells and Slows Pancreatic Tumor Growth and Metastasis in Mice.* Gastroenterology, 2018. **155**(6): p. 1985-1998 e5.
48. An, W.F., et al., *Discovery of Potent and Highly Selective Inhibitors of GSK3b.* 2010.
49. Domínguez, J.M., et al., *Evidence for irreversible inhibition of glycogen synthase kinase-3β by tideglusib.* The Journal of biological chemistry, 2012. **287**(2): p. 893-904.
50. Bordonaro, M., D.L. Lazarova, and A.C. Sartorelli, *The activation of beta-catenin by Wnt signaling mediates the effects of histone deacetylase inhibitors.* Experimental cell research, 2007. **313**(8): p. 1652-1666.
51. Debeb, B.G., et al., *Histone deacetylase inhibitors stimulate dedifferentiation of human breast cancer cells through WNT/β-catenin signaling.* Stem cells (Dayton, Ohio), 2012. **30**(11): p. 2366-2377.
52. Chen, S., et al., *Histone deacetylase 1 and 2 regulate Wnt and p53 pathways in the ureteric bud epithelium.* Development, 2015. **142**(6): p. 1180.
53. Lin, H.Y., et al., *Targeting histone deacetylase in cancer therapy.* Medicinal Research Reviews, 2006. **26**(4): p. 397-413.
54. Billin, A.N., H. Thirlwell, and D.E. Ayer, *Beta-catenin-histone deacetylase interactions regulate the transition of LEF1 from a transcriptional repressor to an activator.* Molecular and cellular biology, 2000. **20**(18): p. 6882-6890.
55. Nakagawa, H., K. Whelan, and J.P. Lynch, *Mechanisms of Barrett's oesophagus: intestinal differentiation, stem cells, and tissue models.* Best practice & research Clinical gastroenterology, 2015. **29**(1): p. 3-16.
56. di Pietro, M., et al., *Evidence for a functional role of epigenetically regulated midcluster HOXB genes in the development of Barrett esophagus.* Proceedings of the National Academy of Sciences of the United States of America, 2012. **109**(23): p. 9077-9082.
57. Que, J., et al., *Pathogenesis and Cells of Origin of Barrett's Esophagus.* Gastroenterology, 2019. **157**(2): p. 349-364.e1.
58. Shi, Q., et al., *Downregulation of HOXA13 sensitizes human esophageal squamous cell carcinoma to chemotherapy.* Thorac Cancer, 2018. **9**(7): p. 836-846.
59. Geraudie, J. and V. Borday Birraux, *Posterior hoxa genes expression during zebrafish bony fin ray development and regeneration suggests their involvement in scleroblast differentiation.* Dev Genes Evol, 2003. **213**(4): p. 182-6.
60. Christen, B., et al., *Regeneration-specific expression pattern of three posterior Hox genes.* Dev Dyn, 2003. **226**(2): p. 349-55.
61. Liang, M. and K. Hu, *Involvement of lncRNA-HOTTIP in the Repair of Ultraviolet Light-Induced DNA Damage in Spermatogenic Cells.* Molecules and Cells, 2019. **42**(11): p. 794-803.
62. Qin, Z., et al., *Elevated HOXA13 expression promotes the proliferation and metastasis of gastric cancer partly via activating Erk1/2.* OncoTargets and therapy, 2019. **12**: p. 1803-1813.
63. Di-Poï, N., et al., *Additive and global functions of HoxA cluster genes in mesoderm derivatives.* Developmental Biology, 2010. **341**(2): p. 488-498.
64. Condie, B.G. and M.R. Capecchi, *Mice with targeted disruptions in the paralogous genes hoxa-3 and hoxd-3 reveal synergistic interactions.* Nature, 1994. **370**(6487): p. 304-307.
65. Yamamoto, S., et al., *Hoxa13 regulates expression of common Hox target genes involved in cartilage development to coordinate the expansion of the autopodal anlage.* Dev Growth Differ, 2019. **61**(3): p. 228-251.
66. Owen, R.P., et al., *Single cell RNA-seq reveals profound transcriptional similarity between Barrett's oesophagus and oesophageal submucosal glands.* Nature Communications, 2018. **9**(1): p. 4261.
67. Gu, Z.D., et al., *HOXA13 promotes cancer cell growth and predicts poor survival of patients with esophageal squamous cell carcinoma.* Cancer Res, 2009. **69**(12): p. 4969-73.
68. Deng, Y., et al., *The expression of HOXA13 in lung adenocarcinoma and its clinical significance: A study based on The Cancer Genome Atlas, Oncomine and reverse transcription-quantitative polymerase chain reaction.* Oncol Lett, 2018. **15**(6): p. 8556-8572.
69. Quagliata, L., et al., *High expression of HOXA13 correlates with poorly differentiated hepatocellular carcinomas and modulates sorafenib response in in vitro models.* Lab Invest, 2018. **98**(1): p. 95-105.
70. Dong, Y., et al., *HOXA13 is associated with unfavorable survival and acts as a novel oncogene in prostate carcinoma.* Future Oncol, 2017. **13**(17): p. 1505-1516.
71. Lin, C., et al., *Transcriptional and posttranscriptional regulation of HOXA13 by lncRNA HOTTIP facilitates tumorigenesis and metastasis in esophageal squamous carcinoma cells.* Oncogene, 2017. **36**: p. 5392.
72. Shen, L.Y. and K.N. Chen, *Exploration of target genes of HOXA13 in esophageal squamous cell carcinoma cell line.* Cancer Lett, 2011. **312**(1): p. 18-23.
73. Barclay, C., et al., *Basic fibroblast growth factor (FGF-2) overexpression is a risk factor for esophageal cancer recurrence and reduced survival, which is ameliorated by coexpression of the FGF-2 antisense gene.* Clin Cancer Res, 2005. **11**(21): p. 7683-91.
74. Kapoor, H., et al., *Animal Models of Barrett's Esophagus and Esophageal Adenocarcinoma-Past, Present, and Future.* Clin Transl Sci, 2015. **8**(6): p. 841-7.
75. Horn, C.C., et al., *Why can't rodents vomit? A comparative behavioral, anatomical, and physiological study.* PLoS One, 2013. **8**(4): p. e60537.
76. Bus, P., P.D. Siersema, and J.W.P.M. van Baal, *Cell culture models for studying the development of Barrett's esophagus: a systematic review.* Cellular oncology (Dordrecht), 2012. **35**(3): p. 149-161.
77. Korbut, E., et al., *Molecular Profile of Barrett's Esophagus and Gastroesophageal Reflux Disease in the Development of Translational Physiological and Pharmacological Studies.* Int J Mol Sci, 2020. **21**(17).



Chapter 10

Nederlandse samenvatting

Van alle tumoren (niet-melanoma huidkanker niet meegerekend) wordt wereldwijd zo'n 28% uitgemaakt door kanker van het maag-darm kanaal. De meest voorkomende maag-darm tumor is dikkedarmkanker, gevolgd door maagkanker, leverkanker, slokdarmkanker en alvleesklierkanker, die respectievelijk op plaats 3, 4, 5, 6, en 12 van meest voorkomende tumoren staan [1]. Van deze tumoren zijn slokdarmkanker en alvleesklier kanker het gevaarlijkst, met een 5-jaar overlevingskans van slechts 19% en 8.2%. Een betere vroegtijdige opsporing alsmede een betere behandeling van deze tumoren zou mogelijk de overlevingskans van patiënten met deze tumoren kunnen verhogen.
In dit proefschrift heb ik basale en translationele aspecten van gastro-intestinale carcinogenese bekeken. In het eerste deel van dit proefschrift heb ik mij gericht op de detectie en behandeling van alvleesklierkanker. In het tweede deel van het proefschrift heb ik de rol van enkele genen die betrokken zijn bij de embryonale aanleg van het maag-darm kanaal bestudeerd in de context van slokdarmkanker.

DEEL 1

Een belangrijke reden voor de lage overlevingskans bij alvleesklierkanker is dat de tumor vaak pas in een laat stadium wordt ontdekt. Huidige beeldvormende technieken lijken niet toereikend om de tumor te detecteren in stadia die nog goed te behandelen zijn met chirurgie en/of chemo/stralings therapie. Biomarkers om alvleesklierkanker in een vroeger stadium op te sporen zouden mogelijk een uitkomst kunnen bieden. Een mogelijke bron voor dergelijke biomarkers is het alvleeskliersap, aangezien dit in nauw contact staat met de cellen waaruit de tumor ontstaat. In **Hoofdstuk 2** hebben we gekeken wat de beste manier is om dit alvleeskliersap te verzamelen, op zodanige manier dat verschillende biomarkers erin gedetecteerd zouden kunnen worden. Hierbij werden twee collectie methoden met elkaar vergeleken (endoscopiekanaal vs katheter), verschillende tijdsduur (0-4 minuten, 4-8 minuten, 8-15 minuten) en werden DNA, micro-RNA, eiwitten en cellulaire markers bestudeerd. We toonden aan dat verzamelen van alvleesklier door de Endoscopiekanaal voor 4-8 minuten resulteerde in de hoogste opbrengst aan alvleeskliersap, waarbij het sap ook de beste kwaliteit leek te hebben voor detectie van pancreas-specifieke biomarkers.
In **Hoofdstuk 3** hebben we de proef op de som genomen en gekeken of biomarkers in alvleeskliersap van patiënten met alvleesklierkanker meer aanwezig waren dan in controle monsters. Hierbij hebben we in eerste instantie gekeken naar de aanwezigheid van extracellulaire blaasjes, die door tumor cellen worden vrijgegeven. We toonden aan dat in het alvleesvliersap van kanker patiënten meer extracellulaire blaasjes van een groter formaat aanwezig zijn. Dit leek specifiek voor alvleesvliersap te zijn, want in serum van alvleesklierkanker patiënten werden deze grotere blaasjes niet aangetroffen. Dit duidt er op dat alvleeskliersap wellicht een betere bron is voor biomarker detectie dan serum, en dat het mogelijk is om kanker patiënten van niet-kanker patiënten te onderscheiden op basis van de grootte van de extracellulaire blaasjes aanwezig in dit sap. In **Hoofdstuk 4** gingen we een stap verder en bekeken we de inhoud van de extracellulaire blaasjes, waarbij vijf verschillende micro-RNA moleculen werden bestudeerd. Ook hier leken 4 van de 5 gemeten micro-RNA moleculen afkomstig uit extracellulaire blaasjes verhoogd aanwezig te zijn in alvleeskliersap van kanker patiënten in vergelijking tot niet-kanker patiënten. Slechts 1 van deze biomarkers was ook verhoogd aanwezig in serum van kanker patiënten, waarmee opnieuw alvleeskliersap een betere afspiegeling lijkt te geven van de

aanwezigheid van kankercellen. Toekomstig onderzoek zal moeten uitwijzen of het combineren van verschillende alvleeskliersap en serum biomarkers in staat is om vroegtijdige alvleesklierkanker opsporing te faciliteren.

Naast het opsporen van alvleesklierkanker is ook een betere bestrijding van deze ziekte noodzakelijk. Een recent onderzoek toonde aan dat gelijktijdige remming van de kinase GSK3b en het DNA methylatie enzym HDAC door het nieuwe middel Metavert resulteert in celdood van pancreaskankercellen. Remming van GSK3c resulteert in een opregulatie van β-Catenin signalering, waarvan de rol in alvleesklierkanker nog niet duidelijk is. In **Hoofdstuk 5** laten wij zien dat deze β-Catenin signalering niet bijdraagt aan de anti-tumorcel werking van Metavert. Het exacte mechanisme van de werking van Metavert blijft dus onduidelijk, maar het is veelbelovend dat nieuwe middelen op de markt verschijnen die mogelijk een rol kunnen spelen in de strijd tegen alvleesklierkanker.

Samenvattend demonstreren we in deel 1 van dit proefschrift dat opsporing van alvleesklierkanker middels biomarkers in alvleeskliersap mogelijk in de toekomst uitkomst kan bieden. Een beter begrip van de moleculaire mechanismen die tumorvorming en behandeling onderliggen wellicht een aangrijpingspunt kunnen vormen voor de ontwikkeling van betere behandelingen van dit ziektebeeld.

DEEL 2

In het tweede deel van dit proefschrift verlegde ik mijn aandacht naar de moleculaire mechanismen die bijdragen aan de ontwikkeling van slokdarmkanker. Twee verschillende types slokdarmkanker kunnen worden onderscheiden; slokdarmadenocarcinoom en plaveiselcelcarcinoom. Een belangrijke risicofactor voor het ontwikkelen van slokdarmkanker, met name adenocarcinoom, is Barrett's slokdarm, een conditie waarbij het epitheel van de slokdarm een ander aspect aanneemt. Plaatselijk vertoont het slokdarm epitheel karakteristieken van darmepitheel, inclusief het vormen van een kubisch epitheel (in plaats van plaveisel epitheel) en de aanwezigheid van bijvoorbeeld slijmvormende cellen en Paneth cellen. De aanleg van verschillende weefsels in het humane lichaam worden tijdens de embryogenese gereguleerd door verschillende genen die positionele identiteit aan een cel kunnen meegeven. Ook in het volwassen lichaam lijken deze genen nog een rol te spelen bij de aanmaak van nieuwe cellen van verschillende weefseltypes. In **hoofdstuk 6** laten we zien dat zogenoemde HOX, waarvan er meerdere clusters zijn een specifiek expressie patroon vertonen over het gehele maag-darm kanaal. Opvallend hierbij is dat HOX genen (met name HOXA13) die normaliter alleen in de dikke darm tot uiting komen, niet alleen in Barrett's slokdarm aanwezig is, maar ook in andere ziektebeelden waarbij het epitheel het aspect van lager gelegen tractus delen heeft aangenomen. Dit suggereert dat een verstoorde positionele identiteit van cellen bijdraagt aan het ontstaan van onder andere Barret's slokdarm. We laten voor het eerst zien dat individuele HOXA13-positieve cellen verscholen liggen in het slokdarm en maagepitheel, en dat HOXA13 expressie ervoor zorgt dat cellen beter in staat zijn om te gaan met stressoren zoals gal en zuur. Het is dus mogelijk dat deze, normaal wellicht onschuldige, HOXA13-positieve cellen in het geval van schade de omliggende HOXA13-negatieve epitheel kunnen overgroeien en een darmepitheel laten ontstaan op plekken waar deze niet zou moeten voorkomen. Het veranderde aspect van dit epitheel op een plek waar het niet had moeten zitten draagt wellicht bij aan een gevoeligheid tot ontwikkeling tot adenocarcinoma in Barrett's slokdarm patiënten.

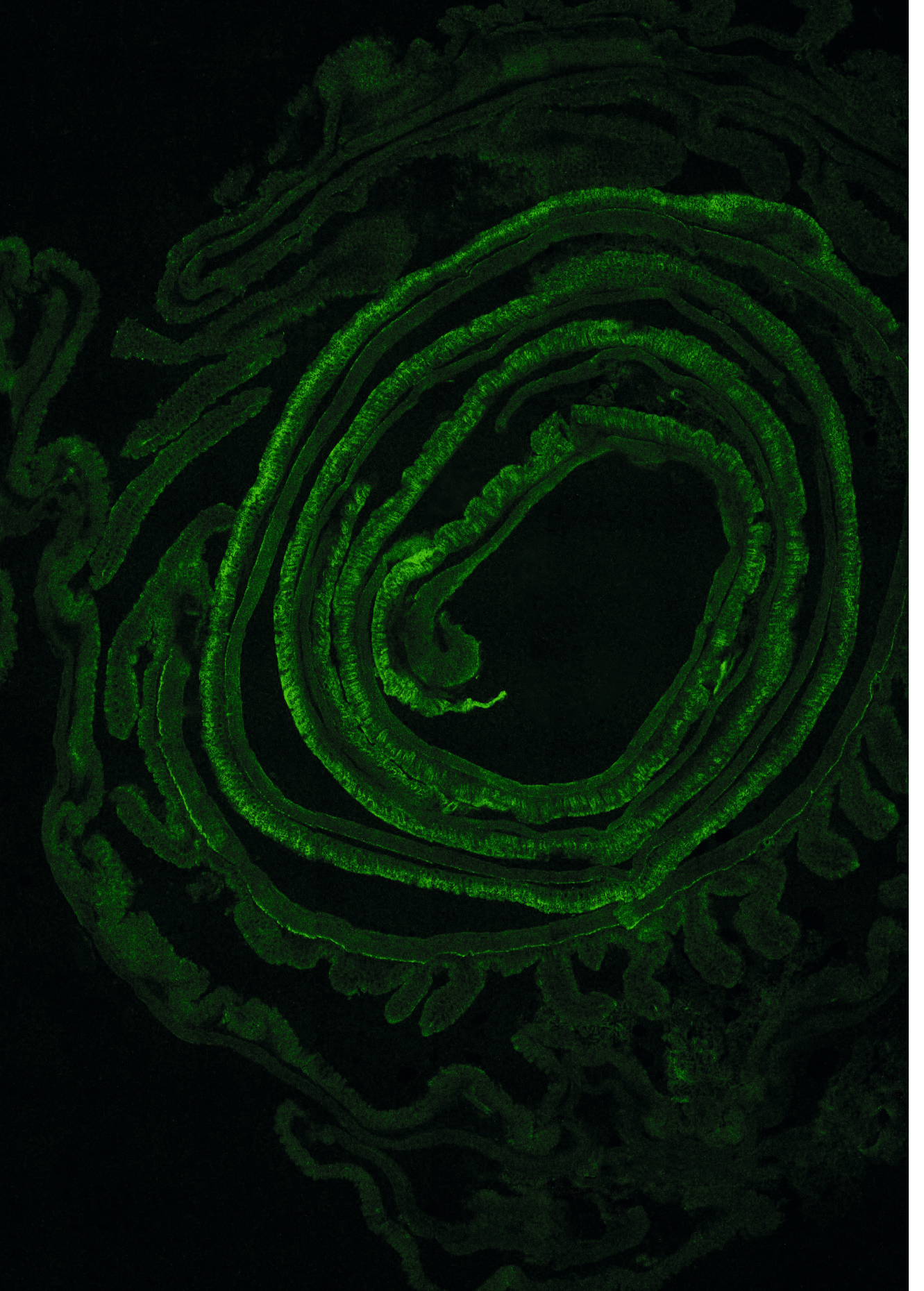
Hoe HOXA13 precies bijdraagt aan ontwikkeling van kanker, en of dit ook een rol speelt voor plaveiselcelcarcinoma, onderzochten we in **Hoofdstuk 7**. Hier introduceerden we HOXA13 in een geimmortaliseerde slokdarmcellijn, en lieten zien dat expressie van HOXA13 resulteert in een verhoogde celdeling, verhoogde celmigratie, minder differentiatie en verminderde gevoeligheid tot therapeutische middelen. Al deze kankereigenschappen zouden kunnen bijdragen aan de ontwikkeling van tumor, en suggereren dat overexpressie van HOXA13 inderdaad kan bijdragen aan de slokdarmkanker.

Naast HOX genen, zijn ook andere genen betrokken bij de aanleg van het maagdarm kanaal tijdens de embryogenese. Het BARX1 gen speelt een belangrijke rol bij de aanleg van de maag, in verhouding tot de darm. Muisstudies hebben aangetoond dat tijdens de embryogenese, expressie van BARX in fibroblasten bij de maag er voor zorgt dat WNT signalering geremd wordt, waardoor het aanwezige epitheel zich ontwikkelt tot maag epitheel, terwijl in darm is BARX1 juist niet aanwezig is en WNT kan bijdragen aan ontwikkeling van darmepitheel. In volwassen muizen lijkt BARX1 geen rol meer te spelen, daar de expressie van dit gen niet meer aanwezig is. In **hoofdstuk 8** laten we echter zien dat hoewel BARX inderdaad niet gevonden wordt in volwassen muis weefsels, het wel aanwezig is in humane darmbiopten, waarbij met name het jejunum hogere expressie van dit gen laten zien. Daarnaast werd expressie van dit gen ook gevonden in Barrett's slokdarm, wat suggereert dat dit gen een belangrijkere rol speelt in volwassen humane setting dan eerder aangenomen, en dat muizen niet altijd een goed modelsysteem vormen om de humane fysiologie in na te bootsen en te bestuderen. De precieze rol van BARX1 in humane fysiologie en ziektebeelden vraagt verdere aandacht.

Samenvattend laat dit onderdeel van het proefschrift zien dat genen die positionele identiteit gedurende de embryogenese bepalen ook een belangrijke rol spelen in de volwassen fysiologie, en dat verstoring in deze genen kan bijdragen aan ziektebeelden van het maag-darm kanaal.

REFERENCE

1. Bray, F., et al., *Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries.* CA Cancer J Clin, 2018. **68**(6): p. 394-424.



Chapter 11

Acknowledgements

PhD Portfolio

Curriculum vitae

Publications

ACKNOWLEDGEMENTS

I would like to thank everyone who helped and supported me during PhD and have made my research and life in the Netherlands a wonderful time!

My dear co-promotor Dr. Gwenny Fuhler and Promotor Prof. Dr. Maikel Peppelenbosch, I am very proud to be in your team. You both were so kind and constantly supportive to me last 3.5 years that I can't express my gratitude. Gwenny, I couldn't have asked for a more brilliant daily supervisor. I am so lucky to have you as a guiding force and inspiration during my PhD. I appreciate the level of freedom I had while working on the projects, your critical thinking, practical approach, and great ideas on how to make everything working the most effectively. Thank you for our Friday meetings, for being always patient, optimistic, for seeing positive side in each experimental result and inspiring me with your sparkling enthusiasm and insightful comments. I highly admire your writing skills and thank you for your help with writing this thesis. Maikel, I have met many wonderful people in the Netherlands and was amazed by this beautiful country and I thank you for this chance. I sincerely appreciate your broad knowledge and impressive scientific creativity.

Dr. Ron Smits, I appreciate so much that you were always there to help with your good advice, technical expertise and high scientific standards. Thank you for teaching me how to do siRNA KO and for your supervision during the β-catenin project and all your valuable input.

Dear Leonie Moree, I am very thankful for all your help with documents and for your kindness.

My excellent collaborators Dr. Iris Levink and Dr. Vincent Janmaat, I believe we were a good team with both of you. I learned a lot from you and enjoyed very much working together on esophageal and pancreatic projects. I wish you both bright future.

Thank you, Shalini Arupillai and Esther de Vries, for your efforts with BARX1 and miRNA projects. It was a pleasure for me to work with you both during your internships.

I would like to express my gratitude to the people who shaped the most my professional growth before starting a PhD. Thank you to Valentyna Lasarchuk, my school biology teacher, Ganna Tolstanova, supervisor of my Bachelor's and Master's thesis, and Petro Borysko, my supervisor at Bienta/Enamine Ltd., for everything that they taught me and for mentoring me during many years. Without all those knowledge it would not be possible for me to complete this thesis.

I am very grateful to all co-authors and collaborators for their contribution to this research and all MDL-lab members!

Thank you for Xiaopei, Elisa, Suk Yee, and Binting for all the good time we spent together inside and outside the laboratory. Suk Yee, you are a very supportive and patient person connecting people from different cultures and I consider this as a superpower. Binting, thank you for taking care of my cell lines and organoids when I needed this and for the analysis of single cell RNA seq. This was a significant contribution! Xiaopei, thank you for your help with SPSS, BARX1 project and all other things!! I wish

you all the best with your PhD. Elisa, thank you for all the wonderful Italian experience! I will return to Italy one day and hope to see you and your baby Chiara!

Thank you for the good advices and fun to my colleagues Sunrui, Ksenia, Floris, Lisanne, Alessandra, Sharida, Michiel, Zhouhong, Celio, and Rebecca.

Thank you for other members of the department. Jan, Henk, Auke, Amy, Natasha, Kelly, Petra, Luc, Rachid, Stefano, Rachid, Robbie, Gülce, Lucia, Amy, Jan, Abdullah, Ruby, Jorke, Gilles, Monique. It was a pleasure to meet you. Best wishes to all of you.

I also thank my family end true friends who stayed with me through all the distances! Дякую мамі, татові, Аліку. Inna Danilova, my 311.3 (Іра велика, Іра маленька, Богданка, Вероніка). Ilona and Natalia, I am proud to be your friend through all these years and that we managed to be in three different countries but so close during this time. Thank for being always there for me no matter what.

Very special thanks to my boyfriend and partner in crime Sergii. These years were so exciting because of you! Thanks for your huge support! I am thrilled about building our future together in London and cannot wait to join you there!

PHD PORTFOLIO

Name PhD candidate: K. Nesteruk **Erasmus MC Department: Gastroenterology and Hepatology** **Research School: Molmed**	**PhD period: Nov 2017- Apr 2021** **Promotor: Prof. dr. M.P. Peppelenbosch** **Co-promotor: dr. G.M. Fuhler**	
1. PhD training	**Year**	**ECTS**
General courses		
The course on Biomedical English Writing course	2018	2
Research Integrity	2019	0.3
Specific courses		
Introduction to fluorescent microscopy	2019	0.39
A microscopic image analysis: from theory to practice	2018	0.8
The Photoshop and Illustrator CC 2018 Workshop for PhD-students and other researchers	2018	1
Seminars, workshops, and conferences		
Weekly MDL seminar program in experimental gastroenterology and hepatology (attending)	2017-20	6.75
Weekly research group education (attending)	2017-20	6.75
Weekly organoid culture meeting (attending)	2018-19	3.375
UEG week (attending)	2020	0.9
Presentations		
Weekly MDL seminar program in experimental gastroenterology and hepatology (presenting)	2017-20 (3x)	1.71
Weekly research group education (presenting)	2017-20	6.86
Weekly organoid culture meeting (presenting)	2018-20 (3x)	1.143
UEG week (presenting)	2020	0.29
Reviewing for scientific journals		
Reviewing for scientific journal (review)	2018	0.2857
Reviewing for scientific journal (research manuscript)	2020	0.2857
Teaching		
Supervise 4 students during 3 weeks for Bachelor project “Visualising multiple (immunological) biomarkers in pancreatic ductal adenocarcinomas and software for overlay of sequential IHC images: a proof-of-principle”	2019	2.57
Supervise a student during 10 months for HBO graduation thesis “The role of Barx-1 in Barrett’s esophagus”	2019-20	11.4
Supervise a student during 7 months for HBO graduation thesis “MicroRNAs of extracellular vehicles from pancreatic juice as biomarkers for detection of pancreatic cancer”	2020	7.98

CURRICULUM VITAE

Kateryna Nesteruk was born in Ivano-Frankivsk, Ukraine, on November 30th, 1992. She and her brother Oleksii were raised by their beloved parents Ruslana and Volodymyr Nesteruk, and grandmother Pavlina Petsak. In 2010, she finished Rivne Oblast Lyceum. In 2016, she graduated from Taras Shevchenko National University of Kyiv with Master degree in biochemistry. Her graduation research "Mechanisms of blood vessels permeability in intestine of rats with hemiparkinsonism," was performed in the Department of Biochemistry of the Taras Shevchenko National University of Kyiv under the supervision of Prof. Dr. Ganna Tolstanova. After one year of working in the high-throughput screening laboratory of Enamine ltd. in Kyiv, she started a PhD project on gastrointestinal carcinogenesis at Gastroenterology and Hepatology, Erasmus Medical Center, Rotterdam, The Netherlands. This work was done under the supervision of Prof. Dr. Maikel Peppelenbosch and Dr. Gwenny Fuhler.

PUBLICATIONS

Introduction

1. Achalasia and associated esophageal cancer risk: What lessons can we learn from the molecular analysis of Barrett's-associated adenocarcinoma?

Nesteruk K, Spaander MCW, Leeuwenburgh I, Peppelenbosch MP, Fuhler GM.
Biochim Biophys Acta Rev Cancer. 2019 Dec;1872(2):188291. IF=7.4

Pancreatic lesions

2. Optimization of Pancreatic Juice Collection: A First Step Toward Biomarker Discovery and Early Detection of Pancreatic Cancer.
Levink IJM*, **Nesteruk K***, Visser DI, Sieuwerts AM, Fernandes CJC, Jansen MPHM, van Driel LMJW, Poley JW, Peppelenbosch MP, Cahen DL, Fuhler GM, Bruno MJ.
Am J Gastroenterol. 2020 Dec;115(12):2103-2108. IF=10.2

3. Upregulated β-catenin signaling does not affect survival of pancreatic cancer cells during dual inhibition of GSK3B and HDAC.

Nesteruk K, Smits R, Bruno M, Peppelenbosch MP, Fuhler GM.
Pancreatology. 2020 Apr;20(3):558-561. IF=3.6

4. Size and concentration of extracellular vesicles in pancreatic juice from pancreatic ductal adenocarcinoma patients.
Nesteruk K, Levink IJM, Dits NFJ, Cahen DL, Peppelenbosch MP, Bruno MJ, Fuhler GM.
(submitted)

5. Extracellular vesicles derived microRNA in pancreatic juice as biomarkers for detection of pancreatic ductal adenocarcinoma
Nesteruk K*, Levink* IJM, de Fries E, Visser IJ, Peppelenbosch MP, Cahen DL, Bruno MJ, Fuhler GM
(manuscript in preparation)

Esophageal lesions

6. *HOXA13* in etiology and oncogenic potential of Barrett's esophagus.
Janmaat VT., **Nesteruk K**, Spaander MCW, Verhaar AP, Yu B, Silva RA, Phillips WA, Magierowski M, van de Winkel A, Stadler HS, Sandoval-Guzmán T, van der Laan LJW, Kuipers EJ, Ron Smits, Bruno MJ, Fuhler GM#, Clemons NJ, & Peppelenbosch MP#.
Nature Communications. 2021 Jun 7; 7;12(1):3354. IF=12.1

7. Forced expression of HOXA13 confers oncogenic hallmarks to esophageal keratinocytes.
Nesteruk K, Janmaat VT, Liu H, Ten Hagen TLM, Peppelenbosch MP, Fuhler GM.
Biochim Biophys Acta Mol Basis Dis. 2020 Aug 1;1866(8):165776. IF=4.4

8. BARX1 is overexpressed in (pre-)malignant esophageal lesions, but does not act via Wnt signaling in adult human esophageal cells.
Nesteruk K, Guo X, Janmaat VT, Arupillai S, Magierowski M, Smits R, Peppelenbosch MP, Fuhler GM (*manuscript in preparation*)

Other

9. Straightforward hit identification approach in fragment-based discovery of bromodomain-containing protein 4 (BRD4) inhibitors.
Borysko P, Moroz YS, Vasylchenko OV, Hurmach VV, Starodubtseva A, Stefanishena N, **Nesteruk K**, Zozulya S, Kondratov IS, Grygorenko OO.
Bioorg Med Chem 2018 Jul 26 (12): 3399-3405. IF=3.1