

Primordial germ cells as a potential shared cell of origin for mucinous cystic neoplasms of the pancreas and mucinous ovarian tumors

Kevin M Elias^{1,2,3}, Petros Tsantoulis^{4,5}, Jean-Christophe Tille⁶, Allison Vitonis⁷, Leona A Doyle^{3,8}, Jason L Hornick^{3,8}, Gurkan Kaya^{4,9}, Laurent Barnes⁹, Daniel W Cramer^{3,7}, Giacomo Puppa⁶, Sarah Stuckelberger¹⁰ , Jagmohan Hooda¹⁰, Pierre-Yves Dietrich^{4,5}, Michael Goggins¹¹ , Candace L Kerr¹², Michael Birrer¹³, Michelle S Hirsch^{3,8}, Ronny Drapkin¹⁰  and Sana Intidhar Labidi-Galy^{4,5*} 

¹ Division of Gynecologic Oncology, Department of Obstetrics and Gynecology and Reproductive Biology, Brigham and Women's Hospital, Boston, MA, USA

² Division of Gynecologic Oncology, Dana-Farber Cancer Institute, Boston, MA, USA

³ Harvard Medical School, Boston, MA, USA

⁴ Department of internal medicine specialties, Faculty of Medicine, Université de Genève, Geneva, Switzerland

⁵ Department of Oncology, Hôpitaux Universitaires de Genève, Geneva, Switzerland

⁶ Division of Pathology, Hôpitaux Universitaires de Genève, Geneva, Switzerland

⁷ Department of Obstetrics and Gynecology, Epidemiology Center, Brigham and Women's Hospital, Boston, MA, USA

⁸ Department of Pathology, Brigham and Women's Hospital, Boston, MA, USA

⁹ Division of Dermatology, Hôpitaux Universitaires de Genève, Geneva, Switzerland

¹⁰ Penn Ovarian Cancer Research Center, Department of Obstetrics and Gynecology, University of Pennsylvania, Philadelphia, PA, USA

¹¹ Department of Pathology, Johns Hopkins Hospital, Baltimore, MD, USA

¹² Department of Biochemistry and Molecular Biology, University of Maryland, Baltimore, MD, USA

¹³ Division of Hematology–Oncology, University of Alabama at Birmingham Comprehensive Cancer Center, Birmingham, AL, USA

*Correspondence to: S Labidi-Galy, Department of Oncology, Hôpitaux Universitaires de Genève, Rue Gabrielle Perret-Gentil 4, 1204, Geneva, Switzerland. E-mail: intidhar.labidi-galy@hcuge.ch

Abstract

Mucinous ovarian tumors (MOTs) morphologically and epidemiologically resemble mucinous cystic neoplasms (MCNs) of the pancreas, sharing a similar stroma and both occurring disproportionately among young females. Additionally, MOTs and MCNs share similar clinical characteristics and immunohistochemical phenotypes. Exome sequencing has revealed frequent recurrent mutations in *KRAS* and *RNF43* in both MOTs and MCNs. The cell of origin for these tumors remains unclear, but MOTs sometimes arise in the context of mature cystic teratomas and other primordial germ cell (PGC) tumors. We undertook the present study to investigate whether non-teratoma-associated MOTs and MCNs share a common cell of origin. Comparisons of the gene expression profiles of MOTs [including both the mucinous borderline ovarian tumors (MBOTs) and invasive mucinous ovarian carcinomas (MOCs)], high-grade serous ovarian carcinomas, ovarian surface epithelium, Fallopian tube epithelium, normal pancreatic tissue, pancreatic duct adenocarcinomas, MCNs, and single-cell RNA-sequencing of PGCs revealed that both MOTs and MCNs are more closely related to PGCs than to either eutopic epithelial tumors or normal epithelia. We hypothesize that MCNs may arise from PGCs that stopped in the dorsal pancreas during their descent to the gonads during early human embryogenesis, while MOTs arise from PGCs in the ovary. Together, these data suggest a common pathway for the development of MCNs and MOTs, and suggest that these tumors may be more properly classified as germ cell tumor variants.

Copyright © 2018 Pathological Society of Great Britain and Ireland. Published by John Wiley & Sons, Ltd.

Keywords: mucinous ovarian tumor; mucinous cystic neoplasm; primordial germ cells; pathogenesis; RHOB

Received 20 December 2017; Revised 13 August 2018; Accepted 25 August 2018

No conflicts of interest were declared.

Introduction

Epithelial ovarian cancers (EOCs) are the leading cause of gynecologic cancer death in the developed world [1]. EOC are histologically classified into four major subtypes: serous, clear cell, endometrioid, and mucinous. Of these, the mucinous type has been the least

studied, probably because of its less frequent incidence, comprising about 3% or less of EOCs [2]. Mucinous tumors are morphologically distinct from all other epithelial ovarian cancers. They tend to be borderline or low grade, have an indolent course and a favorable prognosis, and occur in young female smokers. Some mucinous ovarian tumors (MOTs) have been shown to

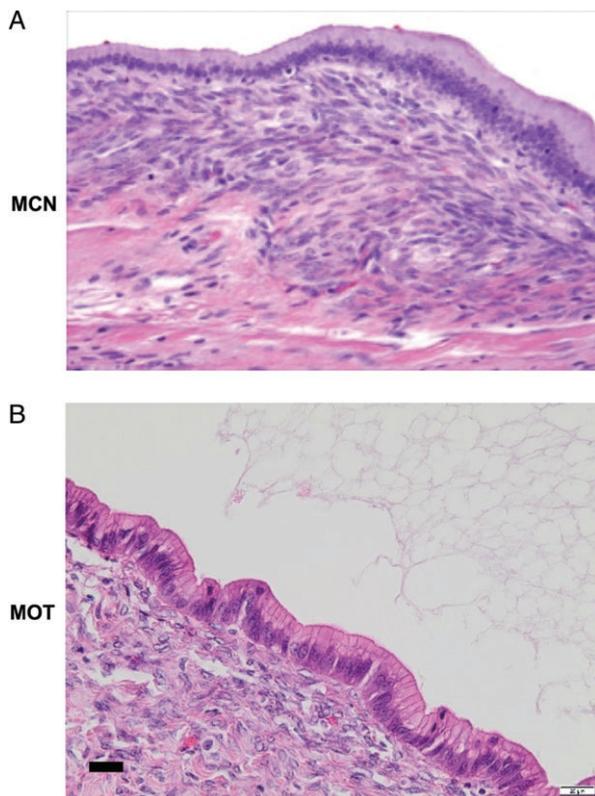


Figure 1. Hematoxylin and eosin sections of a mucinous cystic neoplasm (MCN) of the pancreas (A) and a mucinous ovarian tumor (MOT) (B). Note the similar morphology of the overlying, mucin-filled epithelium and dense underlying stroma. Scale bar = 20 μ m.

arise from germ cell tumors [3]. This has raised the hypothesis that these tumors might arise from a different cell of origin than other EOCs.

Morphologically, MOTs closely resemble a rare pancreatic tumor, mucinous cystic neoplasm (MCN) (Figure 1). MCN is distinct among pancreatic lesions by the presence of a unique ovarian-like stroma [4]. MCN, like MOT, is typically a low-grade neoplasm that occurs mainly in young women. Like MOTs, MCNs tend to be indolent, usually localized to the body and/or tail of the pancreas [5,6]. Whole-exome sequencing studies have revealed molecular similarities between MOT and MCN, with mutations at similar frequencies in *RNF43* and *KRAS* [7,8]. However, if MOT and MCN are somehow related, why do MCNs arise mainly in women (sex ratio 1:10), in the pancreas (a non-gynecologic organ), and why do they have such specific anatomic localization in the body/tail of the pancreas?

In human embryos, the precursors of gametes, termed primordial germ cells (PGCs), are initially located in extragonadal regions (yolk sac) at the third week of development and migrate caudally [9]. In the 5-week-old embryo, PGCs reach the dorsal mesentery, which becomes the body, tail, and isthmus of the pancreas. PGCs then continue to move laterally around both sides of the coelomic angle, pass beyond the primitive mesonephros bodies, and eventually enter the gonadal ridges at the ninth week [9]. This raises the possibility

that MOT and MCN could derive from a common embryologic precursor, PGCs that stopped in the body or tail of the pancreas during their migration to the gonads.

Here, we consider the possibility of a biological relationship among ovarian and pancreatic mucinous tumors based on the gene expression patterns of tumor and normal tissue samples. We demonstrate that the closest normal cell type to either of these tumors is in fact the PGC, rather than either pancreatic or gynecologic epithelia. We suggest that MOT and MCN may be related tumors and more properly be classified as unusual germ cell tumor variants.

Materials and methods

Ethics approval

Review of patient medical records and use of archival specimens were approved by the Brigham and Women's Hospital Institutional Review Board Protocols 2013P000553 and 2016P002742.

Isolation and gene expression profile of human PGCs

Data were pooled from two published gene expression datasets [10,11]. We used gene expression profiles from 8- to 11-week human PGCs (one male and one female) [10] (E-MTAB-6851). As described by the authors, PGCs had been isolated using magnetic cell sorting technology (MACS) and indirect labeling of cells with magnetically tagged goat anti-mouse IgM antibodies toward a mouse anti-SSEA1 antibody [Miltenyi Biotec, Bergisch Gladbach, Germany, clone REA321; diluted (1:5)] [10]. To analyze gene expression profiles, the Affymetrix[®] Human U133 Plus 2.0 GeneChip (Affymetrix, Santa Clara, CA, USA) was used. In addition, single-cell RNA-sequencing data were obtained from five samples of human female PGCs from week 4 to week 11 (GSE63818) (see supplementary material, Table S1) [11]. For that study, human gonads from 7- to 11-week embryos were dissected in Dulbecco's phosphate buffered saline (DPBS) (plus 1% fetal bovine serum). The gonads were washed in DPBS twice before digestion in 250 μ l of Accutase Cell Detachment Solution (Millipore #SCR005; Merck KGaA, Darmstadt, Germany) for 5 min at 37 $^{\circ}$ C. For the isolation of pure human PGCs from 7- to 11-week embryos, 100 μ l of FcR Blocking Reagent and 100 μ l of CD117 MicroBeads (#130-091-332; Miltenyi Biotec) were added to the 300 μ l of gonad cell suspension and mixed well by gently pipetting. After magnetic enrichment, the fraction containing PGCs (CD117-positive cells) was further sorted by BD FACSAria (BD Biosciences, Franklin Lakes, NJ, USA), and CD117-positive cells were collected for downstream analysis. The purity of PGCs was assayed by immunostaining for OCT4 and single-cell RT-qPCR for human *OCT4* transcripts.

For 4-week embryos, the aorta–gonads–mesonephros regions were dissected and a single-cell suspension was obtained by digestion with Accutase. Then the cell suspension was inspected under the microscope carefully and the large cells (less than 0.1%) were manually picked out with a mouth pipette. For RNA sequencing, a DNA library prep kit for Illumina® (New England Biolabs Inc, Ipswich, MA, USA) was used to prepare the sequencing library following the manufacturer's protocol. Libraries were pooled and sequenced on Illumina® HiSeq2500 (Illumina Inc, San Diego, CA, USA) sequencers using 100-bp paired-end sequencing, as previously described [11].

Gene expression profile of pancreatic and ovarian samples

Data were pooled from several published gene expression datasets [12–16] (see supplementary material, Table S1). We compiled gene expression data from 16 samples of normal pancreatic tissue (GSE16515) [16], 36 pancreatic ductal adenocarcinomas (PDACs) (GSE16515) [16], the epithelial component of seven microdissected MCNs of the pancreas (E-MTAB-6853) [13] (two distinct samples from MCN4 were analyzed), the epithelial component of eight microdissected mucinous borderline ovarian tumors (MBOTs) and nine microdissected invasive mucinous ovarian carcinomas (MOCs) (E-MTAB-6844) [12], 24 microdissected samples of Fallopian tube epithelium (FT) (GSE10971) [14], the epithelial component of 13 microdissected high-grade serous ovarian carcinomas (HGSOCs) (GSE10971) [14], and six microdissected samples of ovarian surface epithelium (OSE) (GSE40595) [15]. All array expression data were generated with Affymetrix® microarrays (Affymetrix, Santa Clara, CA, USA). Our raw data are available at ArrayExpress (<https://www.ebi.ac.uk/arrayexpress/>): E-MTAB-6844, E-MTAB-685, and E-MTAB-6853.

Data processing

For all array expression datasets, we recovered the raw data in the form of .CEL files, which were imported and processed in the R language and environment for statistical computing [17]. Raw data from each dataset was treated with the Robust Multi-array Average (RMA) algorithm [18], from the 'Affy' BioConductor package [19] using default settings. For genes mapping to multiple probes, the mean expression was used as a gene-level summary. Batch effects were attenuated by using a previously published list of endogenous control genes [20] to correct for unwanted variation [21].

The single-cell RNAseq data (GSE63818) were obtained as RPKM estimates and then filtered to exclude genes with extremely low expression (< 0.3 RPKM) or extremely low variance (bottom 1%). For every sample, the gene expression estimates were aggregated by calculating the mean expression in the isolated single cells. The RNAseq expression estimates were mean-centered

against the microarray datasets to obtain a similar distribution of values. The mean-variance plot for all common genes from both platforms was used as a quality control and confirms that the results have comparable distributions for most of the expression range.

Class comparison

Differentially expressed genes between groups were identified using the *limma* package with default options, with gene expression ordered by absolute log fold-change (FC) using a false discovery rate (FDR) of 0.05 or less [22]. These comparisons identified genes whose expression was significantly altered between PGCs, MCN specimens, PDACs, and normal pancreatic tissue. Similarly, we identified genes whose expression was different between PGCs, MBOT, HGSOC, OSE, and FT specimens. Dendrograms were created using hierarchical clustering (hclust) and a Euclidean distance metric on genes expressed by all samples ($N = 9626$). The most discriminative and statistically significant genes were ordered by log FC and summarized in heatmaps. During the revision of the manuscript, the MOC samples were added and a second distinct analysis was done in order to identify genes whose expression was different between PGCs, MBOT, MOC, HGSOC, OSE and FT specimens. A new gene list was also generated ($N = 9626$).

Gene set enrichment analysis

Common biologic pathways for the 1000 most similarly expressed genes (as measured by absolute log FC) between PGCs and MCN or PGCs and MBOT were examined by gene set enrichment analysis using DAVID v.6.8 [23]. The top ten gene ontology terms most over-represented by adjusted P value after a Bonferroni correction are shown.

Laser capture microdissection and RNA extraction for RT-qPCR

New samples were selected for validation of gene expression profiles by RT-qPCR. Sections were cut at 4 μm thickness from each sample and stained with hematoxylin and eosin (H&E) for review to ensure proper tissue orientation and histology (JCT and GP). Fresh-frozen samples of HGSOC ($n = 5$), PDAC ($n = 5$), and normal pancreatic tissue ($n = 5$) were macrodissected. The epithelium component from frozen samples of MCNs ($n = 4$) and MBOTs ($n = 4$) were laser capture-microdissected using a Leica LMD7000 instrument (Leica Microsystems, Wetzlar, Germany) (see supplementary material, Figure S1). Shortly before microdissection, 10 μm sections were cut, adhered onto frame slides, immediately fixed in ethanol 75% for 2 min, stained with hematoxylin, washed in water, dehydrated in graded alcohols, then xylene, and microdissected. RNA was also isolated from immortalized OSE cell lines ($n = 2$) and immortalized FT cell lines ($n = 4$) [24]. Total RNA was extracted from

tissues or cell lines using a QIAGEN® RNeasy kit (#74104; QIAGEN, Valencia, CA, USA) following the manufacturer's protocol. Measurement of total RNA concentration was performed with a Qubit fluorimeter (Thermo Fisher Scientific, Waltham, MA, USA) and quality assessed with an Agilent Bioanalyzer (Agilent Technologies, Lexington, MA, USA).

RT-qPCR

To confirm expression of identified genes, cDNA was synthesized from 150 ng of total RNA using a mix of random hexamers – oligo d(T) primers and Primer-Script reverse transcriptase enzyme (Takara Bio, Inc, Kusatsu, Japan), following the supplier's instructions. SYBR Green assays were designed using the program Primer Express v 2.0 (Applied Biosystems, Waltham, MA, USA) with default parameters. Amplicon sequences were aligned against the human genome by BLAST to ensure that they were specific for the gene being tested. Oligonucleotides were obtained from Thermo Fisher Scientific (Waltham, MA, USA). The efficiency of each design was tested with serial dilutions of cDNA. PCR reactions (10 µl volume) contained diluted cDNA, 2× Power Up SYBR Green Master Mix (Thermo Fisher Scientific), and 300 nM of forward and reverse primers. PCR primers are listed in the supplementary material, Table S2. PCR reactions were performed on an SDS 7900 HT instrument (Applied Biosystems) with the following parameters: 50 °C for 2 min, 95 °C for 10 min, and 45 cycles of 95 °C for 15 s and 60 °C for 1 min. Each reaction was performed in three replicates on a 384-well plate. Raw C_T values obtained with SDS 2.2 (Applied Biosystems) were imported in Excel (Microsoft, Redmond, WA, USA). Normalization factor and fold-changes were calculated using the GeNorm method [25]. Target gene C_T values were normalized to β -tubulin (*TUBB*) and β -actin (*ACTB*) transcripts.

Chart review of MCN and MOT

Twenty-three cases of MCNs were identified by querying the Brigham and Women's Hospital (BWH) pathology database. All cases were confirmed by a staff GI pathologist (JLH or LAD) to be MCN and not an intraductal pancreatic mucinous neoplasm (IPMN), mucinous adenocarcinoma, or other form of pancreatic neoplasm. Medical charts for the cases were reviewed to extract clinical data after approval by the BWH Institutional Review Board (Protocol 2013P000553). MOT cases were identified from the New England Case Control (NECC) study as previously described [26]. In brief, cases were enrolled from July 1984 to September 1987 (NECC2), May 1992 to March 1997 (NECC3), August 1998 to April 2003 (NECC4), and October 2003 to November 2008 (NECC5). The four phases enrolled 2475 cases including 2274 with epithelial ovarian cancers, of which 287 were mucinous. Controls for NECC3 were identified by random-digit dialing supplemented

Table 1. Patient clinical characteristics

	Mucinous ovarian tumors N = 287N (%)	Mucinous cystic neoplasms N = 23N (%)	P value
Sex			
Female	287 (100)	23 (100)	1.0
Male	0 (0)	0 (0)	
Age, years			
< 44	128 (44.6)	6 (26)	0.12
44–53	69 (24.0)	6 (26)	
54–62	46 (16.0)	8 (35)	
> 62	44 (15.3)	3 (13)	
Race			
White	273 (95.1)	18 (78.2)	0.008
Non-white	14 (4.9)	5 (21.7)	
Ever-smoker			
No	123 (42.9)	12 (63.1)	0.10
Yes	164 (57.1)	7 (36.8)	
Unknown		4	
Stage*			
I	213 (88.8)	23 (100)	
II–IV	27 (11.3)	0 (0)	0.14

*Missing for 47 cases.

with residents' lists for older controls. About 10% of households contacted had an eligible control and of these, 421 (72%) agreed to participate.

Statistical methods

Unless specified otherwise, all statistical tests were performed in R v.3.4.3 using *P* value correction for multiple comparisons to account for the false discovery rate (FDR) [17]. Clinical characteristics of patients with MCN and MOT were compared using a *z*-test for population proportions. Median fold-changes of gene expression from RT-qPCR data were compared using a pairwise Student's *t*-test.

Results

Clinical presentation for patients with mucinous ovarian tumors or mucinous cystic neoplasms

We performed a chart review to investigate clinical similarities between patients with MOTs or MCNs. From BWH, we identified 23 cases of MCNs. Tumors occurred exclusively in women (Table 1). Compared with the cases in the NECC study, women with MCNs were similar in terms of age, smoking history, and stage at diagnosis. The only notable difference appeared to be in racial distribution, as a larger proportion of MCNs than MOTs were diagnosed among non-white women.

Similar gene expression profiles for MCNs of the pancreas and PGCs

Unsupervised hierarchical clustering was performed for gene expression profiles from seven MCNs and 36 PDACs, 16 normal pancreatic tissues, and seven human

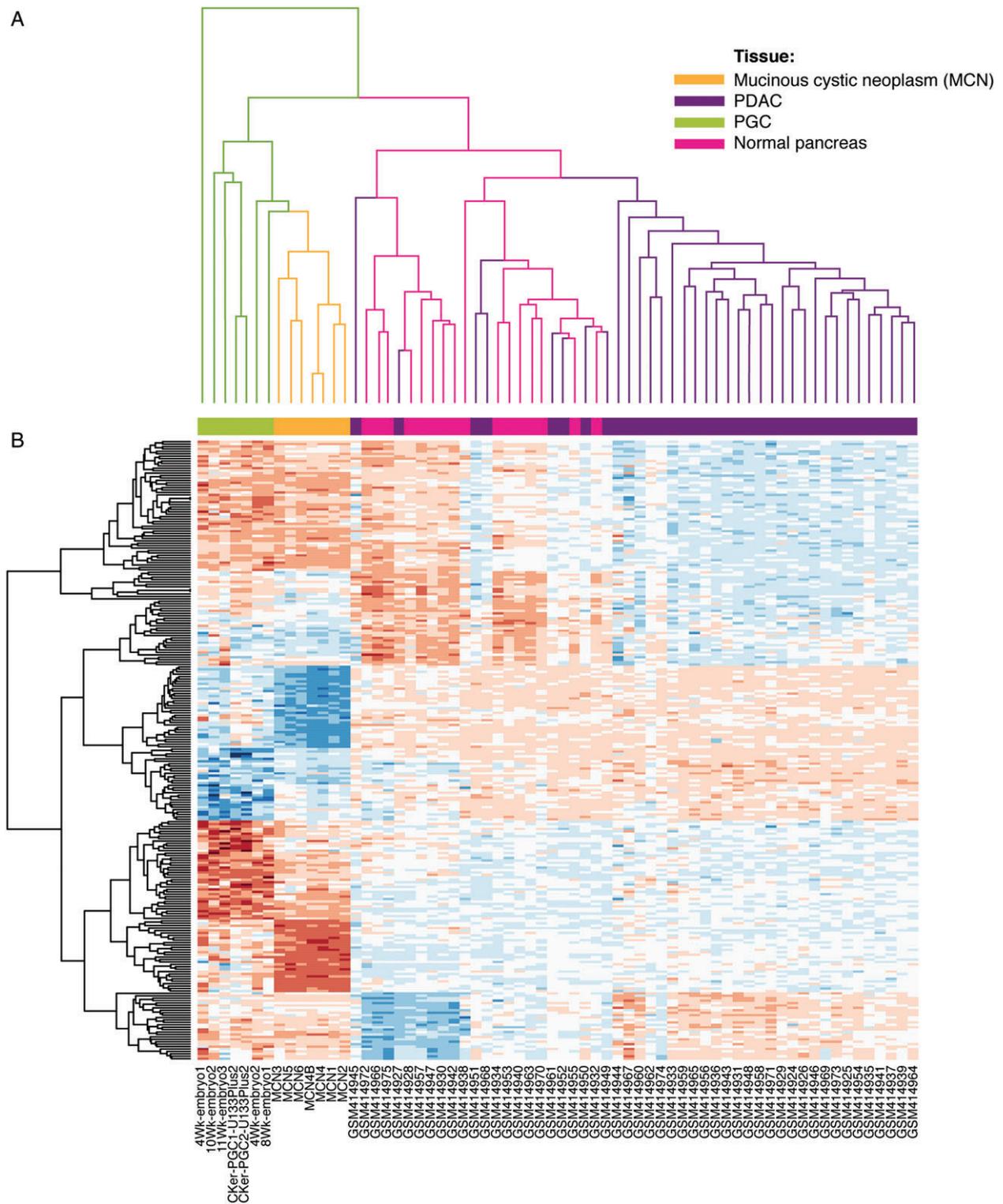


Figure 2. Unsupervised hierarchical clustering of PGCs and pancreatic and MCN samples. (A) Expression of 9626 genes in MCNs, PGCs, PDACs, and normal pancreatic tissue. Dendrogram of the 66 experimental samples. Hierarchical clustering illustrates that MCN specimens are closely associated with PGCs (left branch), whereas pancreatic ductal adenocarcinomas (PDACs) group together with normal pancreatic tissue (right branch). (B) Heatmap of the differentially expressed genes in MCNs, PGCs, PDACs, and normal pancreatic tissue. Blue = high expression; brown = low expression.

PGC samples. The dendrogram in Figure 2A shows the unsupervised hierarchical clustering of all 66 samples based on the expression of 9626 genes (the intersection of available genes over all datasets). The samples separated into two main branches. The left branch

contains PGCs and MCNs. The right branch includes PDACs and normal pancreatic tissues. This dendrogram suggests that PGCs are more closely related to MCNs and PDACs are more closely associated with normal pancreatic tissue. Importantly, a small number

of PDAC samples clustered together with normal pancreatic samples, but all MCN samples were clearly distinct. The heatmap displays the most differentially expressed, statistically significant genes per sample type (Figure 2B). It clearly demonstrates important similarities between PGC and MCN samples, compared with normal pancreas and PDAC. Importantly, the single-cell PGC RNA sequencing dataset and array expression PGC dataset are highly concordant, emphasizing the consistency of these gene expression profiles. Together, these data suggest that global gene expression in MCNs more closely resembles that of PGCs rather than normal pancreatic tissue or other pancreatic tumors.

We used *limma* and a cut-off at a FDR of 0.05 to identify common differentially expressed genes between the MCN and PGCs samples on the one hand and the PDAC and normal pancreatic tissue samples on the other hand. The list of 1000 top differentially expressed genes are listed in the supplementary material, Dataset S1. Gene set enrichment analysis of the shared gene sets between MCNs and PGCs showed overrepresentation of genes related to phosphoproteins (see supplementary material, Table S3).

Similar gene expression profiles for MOTs and PGCs

Unsupervised hierarchical clustering was performed from microarrays of eight MBOTs, 24 FT and six OSE samples, 13 HGSOCs, and seven PGCs. The dendrogram in the supplementary material, Figure S2A represents the unsupervised hierarchical clustering of all 58 samples. MBOT samples were aligned more closely with PGCs than OSE, whereas HGSOCs grouped with either OSE or FT samples. The heatmap summarizes differentially expressed genes in each sub-group of samples (see supplementary material, Figure S2B). It indicates that MBOT and PGC samples have important similarities and are more distantly related to OSE, HGSOCs, and FT, respectively. MBOTs are closely related to MOCs and a continuum appears to be present from borderline to carcinoma, which is different from other epithelial EOCs [27]. Thus, we questioned whether MOCs resemble PGCs. Unsupervised hierarchical clustering was performed from microarrays of eight MBOTs, nine MOCs, 24 FT and six OSE samples, 13 HGSOCs, and seven PGCs. MBOT and MOC samples clustered together and were aligned more closely with PGCs than OSE (Figure 3A,B). These data suggest that global gene expression in MOTs more closely resembles that of PGCs than Müllerian epithelia. The normal FT cells clustered closely with HGSOCs, adding further evidence to the theory that FT cells give rise to most HGSOCs [24,28,29]. Genes similarly expressed among MBOTs, MOCs, and PGCs versus HGSOCs, OSE, and FT were selected using *limma* at a FDR of 0.05. A complete list of the 1000 top genes is shown in the supplementary material, Dataset S2. Gene set enrichment analysis again showed a high percentage of phosphoproteins (see supplementary material, Table S4).

Shared gene expression among PGCs, MCNs, and MOTs

By comparing the lists of genes generated in the supplementary material, Datasets S1 and S2, we observed that many genes had similar expression patterns shared among the MOTs, PGCs, and MCN samples. Genes differentially expressed between MOTs (MBOTs and MOCs), MCNs, and PGCs, and the pancreatic and adnexal eutopic normal and tumor samples were selected and ordered by log FC (see supplementary material, Figure S3). A complete list of the 411 most differentially expressed genes among all the mucinous tumor types and PGCs versus other tissue types is shown in the supplementary material, Dataset S3. Three genes highly expressed and common among MOTs (MBOTs and MOCs), MCNs, and PGCs were *CPM*, *RHOB*, and *ASPN*. To validate the microarray results, these three genes were selected for RT-qPCR analysis. Expression levels for the three genes were determined on independent samples, not included in the microarray analyses. In agreement with our microarray data, MCNs were found to express higher levels of *ASPN*, *RHOB*, and *CPM* compared with normal pancreatic tissue or PDAC, although only *RHOB* compared with PDAC reached statistical significance ($p = 7.2 \times 10^{-3}$) (Figure 4A). Consistent with our profiling data, MBOT samples expressed much higher levels of *ASPN* compared with OSE ($p = 10^{-5}$), FT ($p = 4.3 \times 10^{-5}$) or HGSOC ($p = 3.14 \times 10^{-3}$). *RHOB* was also strongly expressed by MBOTs and this was statistically significant compared with OSE ($p = 3.8 \times 10^{-4}$), FT ($p = 1.3 \times 10^{-4}$) or HGSOC ($p = 0.01$). *CPM* was not tissue-specific (Figure 4B). Since our RT-qPCR data were obtained on specimens distinct from those of the microarray data, our observations suggest that *RHOB* and *ASPN* could be specific new markers for these mucinous neoplasms.

Discussion

MOTs and MCNs are both rare, indolent mucinous neoplasms with a propensity to develop in young women who smoke. They share common epidemiologic, clinical, morphologic, and genomic features. The histological presence of a unique ovarian-type stroma is mandatory to diagnose MCNs and distinguishes MCNs from other pancreatic neoplasms [30]. MCNs nearly always occur in the body or tail of the pancreas, and unlike PDAC, do not communicate with the pancreatic ducts. Likewise, MOTs are unique among ovarian tumors in that they usually arise within large parenchymal cysts, not along the Fallopian tube–ovarian interface. The tumors share expression of CK7, and unlike gastrointestinal tumors, have only variable expression of CK20 and do not express MUC2 [31]. The stroma of MOTs and MCNs also show similar expression patterns for sex hormone receptors [32]. Here, we consider the possibility that MOTs and MCNs could derive from a

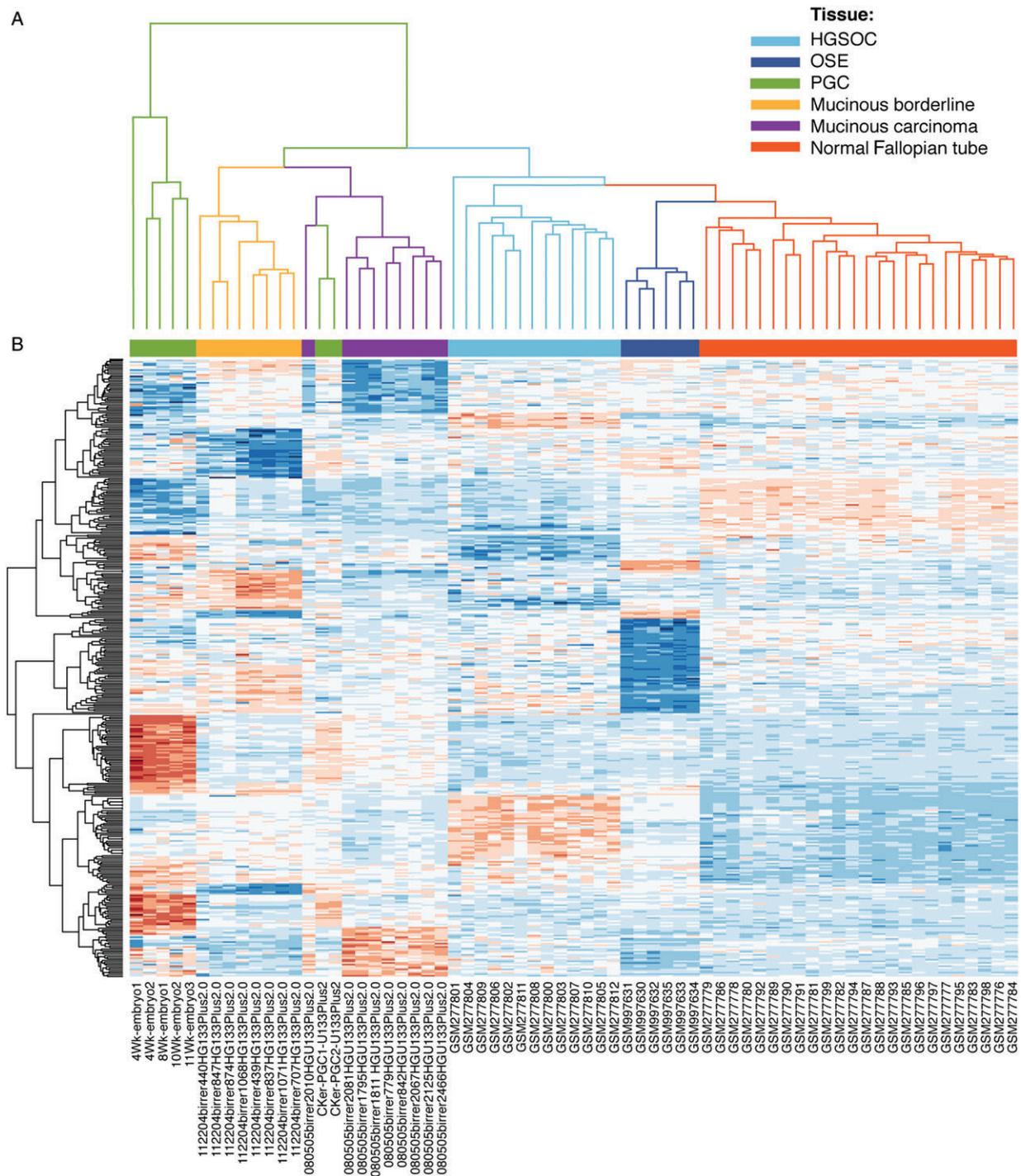


Figure 3. Unsupervised hierarchical clustering of PGCs, and ovarian, Fallopian tube, and MOT samples. (A) Expression of 9626 genes in MOCs, MBOTs, PGCs, HGSOCs, OSE, and FT. Dendrogram of the 67 experimental samples based on hierarchical clustering. MOC and MBOT specimens cluster together and are closely associated with PGCs, whereas high-grade serous ovarian carcinomas (HGSOCs) group with normal Fallopian tube (FT) or ovarian surface epithelium (OSE). (B) Heatmap of the differentially expressed genes in MOCs, MBOTs, PGCs, HGSOCs, OSE, and FT. Blue = high expression; brown = low expression.

common embryologic precursor, PGCs. Approximately 100 PGCs start the journey but by the time they arrive at the gonads, they number about 1700 because they proliferate *en route* [33]. MCNs could arise from some PGCs that stop in the body or tail of the pancreas during their migration to the gonads. Thus, embryological remnants of PGCs that stopped in the body/tail of the pancreas would give rise to MCNs, whereas MOTs

would develop from PGCs that reached the ovaries (Figure 5A,B).

Recent studies suggest that teratoma-associated MOTs are of germ cell origin [3,34]. Using unsupervised clustering of gene expression profiles and RNA sequencing of different ovarian and pancreatic tissues and tumors, we have shown for the first time that gene expression in non-teratoma-associated pancreatic and

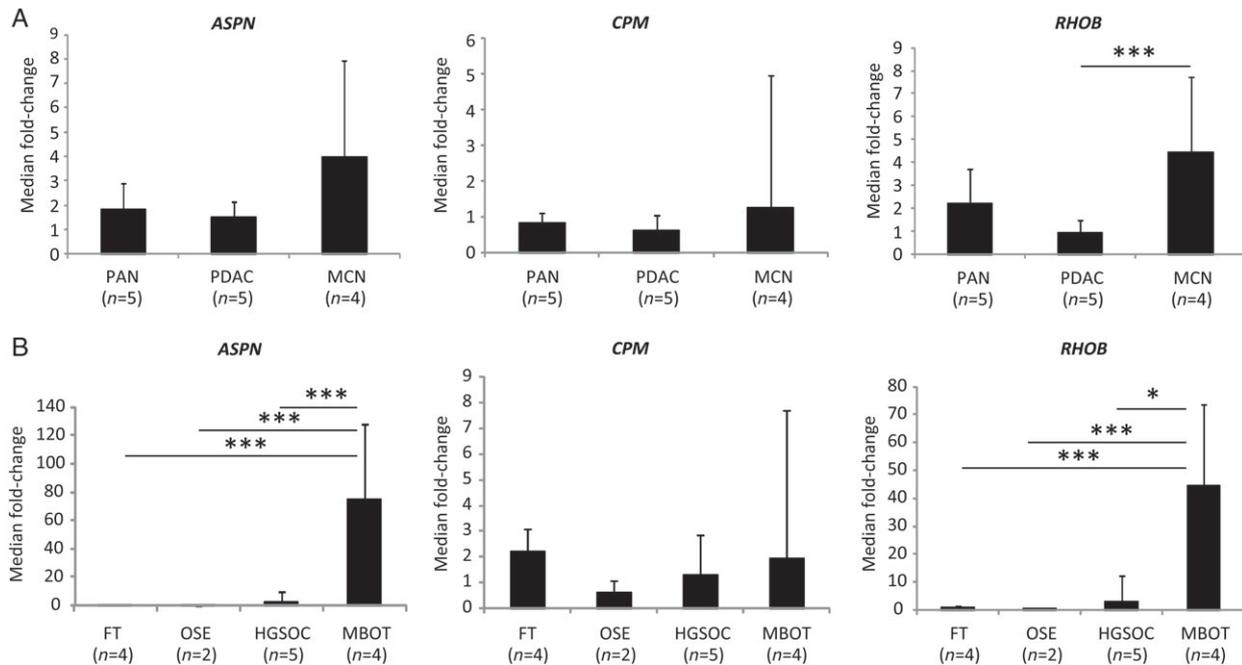


Figure 4. Expression of *ASPN*, *RHOB*, and *CPM* in ovarian and pancreatic samples. To confirm the increased expression of *ASPN*, *RHOB*, and *CPM* in MBOTs and MCNs, qPCR was performed using cDNA generated from new samples of ovarian and pancreatic tumors and normal tissues. (A) RT-qPCR confirmed an increase of *RHOB* in MCN compared with PDAC. (B) RT-qPCR confirmed the differential expression of *ASPN* and *RHOB* uniquely expressed in MBOT. FT, immortalized Fallopian tube cell lines; OSE, immortalized ovarian surface epithelium cell lines; HGSOc, high-grade serous ovarian carcinomas; PDACs, pancreatic ductal adenocarcinomas; PAN, normal pancreas; MBOTs, mucinous borderline ovarian tumors; MCNs, mucinous cystic neoplasms of the pancreas. * $p = 0.01$; *** $p \leq 0.001$.

ovarian mucinous tumors also resembles PGCs. Validation on independent samples and by RT-qPCR of the microarray data for selected genes further strengthens our microarray analysis.

We acknowledge several limitations of our study. First, while we offer several observations to advance the theory that MOTs and MCNs share a common cell of origin, we cannot exclude convergent evolution of pancreatic and ovarian cells to a common PGC-like phenotype. Second, we have used bulk pancreas as a control. Microdissected pancreatic ductal epithelium would be more ideal, but this is technically challenging as pancreatic enzymes degrade the quality of RNA. Third, we did not investigate the expression of specific PGC markers. This important question needs to be addressed in future work.

This hypothesis can explain many characteristics of MCNs and MOTs: their rarity; their development outside the normal epithelial interfaces; and their clinical, histological, and molecular similarities [30]. Although not profiled in this study, MCNs of the liver and kidney have similar clinical and pathological characteristics of pancreatic MCNs (almost exclusively in young women and the tumor having two components: mucinous epithelium associated with ovarian-like stroma). These could similarly arise from PGCs stopping in the right part of the abdomen (specifically the left lobe of the liver) or retroperitoneum [35–37]. Likewise, another rare tumor that arises mainly in young women and is located on the PGC migration trajectory is the mixed epithelial stromal tumor (MEST). These

tumors develop in the kidney and rarely contain mucinous epithelium, and the epithelial elements are always PAX8-positive; however, like MCNs, MESTs are always associated with ovarian-like stroma and occur almost exclusively in middle-aged females [38]. Nevertheless, several questions remained unanswered: why do MCNs develop only in women since PGCs have the same migration process in male and female embryos? And why are these epithelial tumors mucinous rather than other histological subtypes? It is possible that exposure to female hormones, smoking, and other factors yet to be identified play an important role in the pathogenesis of MCNs.

Our microarray and RT-qPCR data on two independent cohorts suggest that *RHOB* and *ASPN* are common among MOTs, MCNs, and PGCs. Deregulations of these genes have been observed in several types of tumors. *RHOB* codes for Ras homolog B (RhoB) protein, a Rho family GTPase that is itself a subset of the Ras superfamily. RhoB plays an important role in cell migration, membrane trafficking, cell proliferation, and DNA repair. RhoB alteration seems crucial for the response of Ras-transformed cells to farnesyltransferase inhibitors [39]. Because *KRAS* is frequently mutated in MOTs [7] and MCNs [8], it would be interesting to investigate whether *KRAS* mutations correlate with *RHOB* expression.

ASPN codes for asporin, a small leucine-rich proteoglycan (SLRP). In the tumor microenvironment, asporin is mainly secreted by cancer-associated fibroblasts [40,41]. Its expression in prostate cancer samples

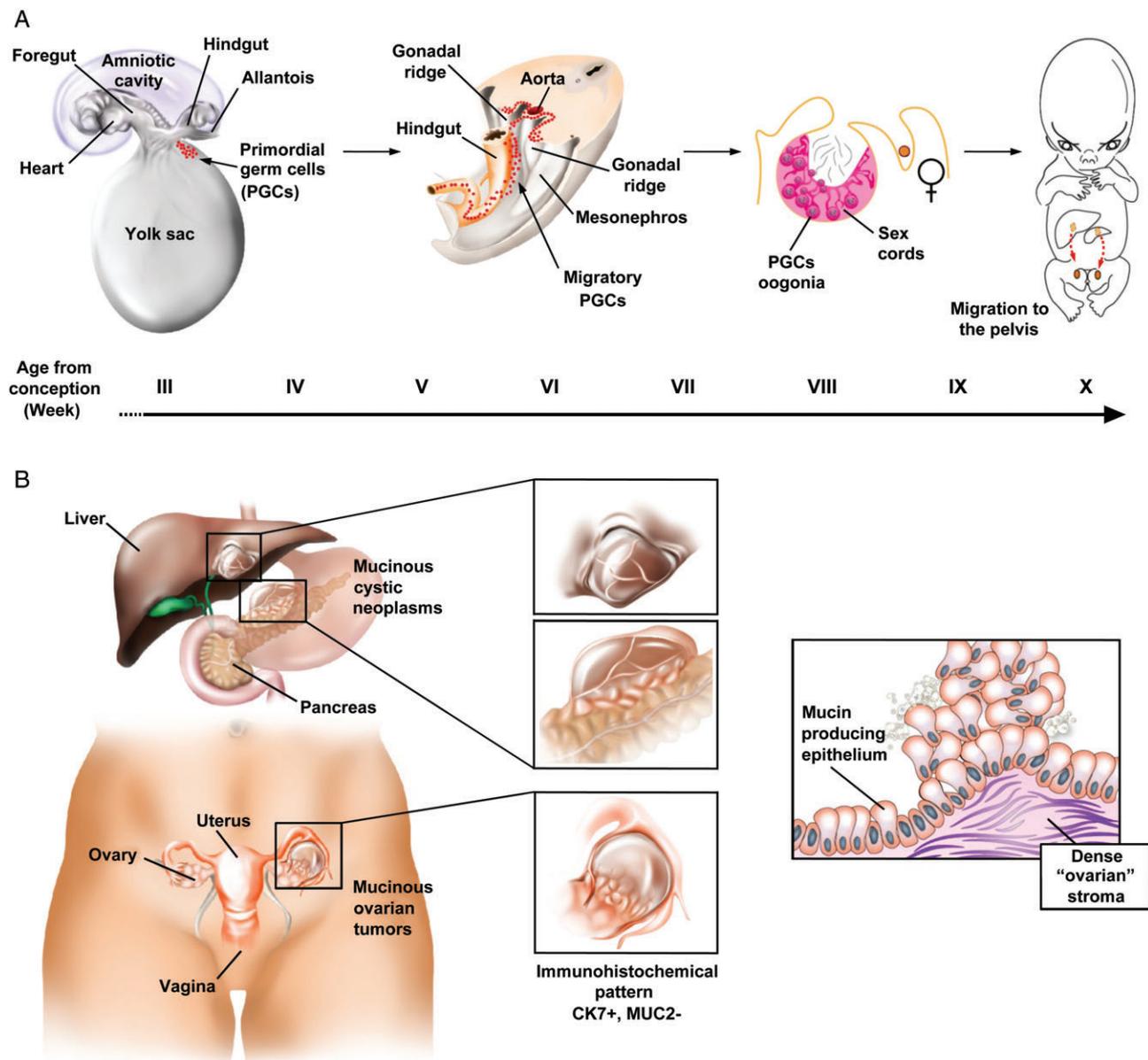


Figure 5. Proposed model for the origin of mucinous ovarian tumors and mucinous cystic neoplasms of the pancreas from primordial germ cells. (A) Migration of primordial germ cells (PGCs) in the human embryo starts from the dorsal wall of the yolk sac near the developing allantois (III weeks). At VII weeks, PGCs migrate into the midgut and hindgut, passing through the dorsal mesentery into the gonadal ridges (VIII weeks). At IX weeks, PGCs colonize gonadal ridges [9]. (B) Mucinous cystic neoplasms (MCNs) of the pancreas would arise from left embryological remnant of migrating PGCs that stopped in the body/tail of the pancreas. MCNs of the liver would arise from right embryological remnant of migrating PGCs that stopped in the left lobe of the liver. In the ovaries, MOTs would develop from PGCs that did not develop oogonia. These three mucinous tumors each occur only in women, have CK7⁺ MUC2⁻ immunohistochemical staining, and are surrounded by ovarian-like stroma.

correlates with disease progression [42]. Certainly, these data need further validation in larger cohorts and at the protein level to elucidate this relationship further.

In conclusion, we present molecular data that may provide a better understanding of the pathogenesis of mucinous ovarian and pancreatic cystic tumors. Our data support the hypothesis that MOTs resemble MCNs of the pancreas at the macroscopic, microscopic, and molecular levels, and share a possible common cell of origin in PGCs. Knowledge of the cell of origin may accelerate translational and clinical research for these rare diseases.

Acknowledgments

We thank Mrs Laurence Zulianello for iconographic support. We thank the Genomic platform at the Faculty of Medicine of Geneva for technical support. KME and RD are supported by the Honorable Tina Brozman Foundation. KME is supported by the Robert and Deborah First Family Fund, Saltonstall Research Fund, the Minnesota Ovarian Cancer Alliance, and the Reproductive Scientist Development Program from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) award

K12-HD000849. RD is supported by the Dr Miriam and Sheldon G Adelson Medical Research Foundation, the Department of Defense, the Claneil Foundation, and the Run & Walk 4 Family & Friends with Cancer Foundation. JH is supported by the Ann and Sol Schreiber Mentored Investigator Award from the Ovarian Cancer Research Fund Alliance (OCRFA) and the Foundation for Women's Wellness. SS is supported by the Ann and Sol Schreiber Mentored Investigator Award from the OCRFA. SILG is supported by the Research Fund of the Department of Internal Medicine of the Hôpitaux Universitaires de Genève and the Faculty of Medicine of Geneva; this fund receives an unrestricted grant from AstraZeneca Switzerland.

Author contributions statement

KME, RD, and SILG conceived and designed the study. KME, PT, RD, and SILG developed the methodology. KME, PT, JCT, AV, LAD, JLH, DWC, SS, JH, MG, CLK, MB, MSH, RD, and SILG acquired data. KME, PT, AV, LAD, JLH, DWC, MSH, RD, and SILG analyzed and interpreted data. KME, PT, JCT, MG, CLK, MB, MSH, RD, and SILG were responsible for writing, review, and/or revision of the manuscript. KME, PT, JCT, AV, LAD, JLH, GK, LB, DWC, GP, SS, JH, PYD, MG, CLK, MB, MSH, RD, and SILG gave administrative, technical, or material support. RD and SILG supervised the study.

References

- Siegel RL, Miller KD, Jemal A. Cancer statistics, 2015. *CA Cancer J Clin* 2015; **65**: 5–29.
- Seidman JD, Kurman RJ, Ronnett BM. Primary and metastatic mucinous adenocarcinomas in the ovaries: incidence in routine practice with a new approach to improve intraoperative diagnosis. *Am J Surg Pathol* 2003; **27**: 985–993.
- Kerr SE, Flotte AB, McFalls MJ, et al. Matching maternal isodisomy in mucinous carcinomas and associated ovarian teratomas provides evidence of germ cell derivation for some mucinous ovarian tumors. *Am J Surg Pathol* 2013; **37**: 1229–1235.
- Testini M, Gurrado A, Lissidini G, et al. Management of mucinous cystic neoplasms of the pancreas. *World J Gastroenterol* 2010; **16**: 5682–5692.
- Crippa S, Salvia R, Warshaw AL, et al. Mucinous cystic neoplasm of the pancreas is not an aggressive entity: lessons from 163 resected patients. *Ann Surg* 2008; **247**: 571–579.
- Yamao K, Yanagisawa A, Takahashi K, et al. Clinicopathological features and prognosis of mucinous cystic neoplasm with ovarian-type stroma: a multi-institutional study of the Japan Pancreas Society. *Pancreas* 2011; **40**: 67–71.
- Ryland GL, Hunter SM, Doyle MA, et al. *RNF43* is a tumour suppressor gene mutated in mucinous tumours of the ovary. *J Pathol* 2013; **229**: 469–476.
- Wu J, Jiao Y, Dal Molin M, et al. Whole-exome sequencing of neoplastic cysts of the pancreas reveals recurrent mutations in components of ubiquitin-dependent pathways. *Proc Natl Acad Sci U S A* 2011; **108**: 21188–21193.
- De Felici M. Origin, migration and proliferation of human primordial germ cells. In *Oogenesis*, Coticchio G, Albertini DF, De Santis L (eds). Springer: London, 2013; 19–37.
- Pashai N, Hao H, All A, et al. Genome-wide profiling of pluripotent cells reveals a unique molecular signature of human embryonic germ cells. *PLoS One* 2012; **7**: e39088.
- Guo F, Yan L, Guo H, et al. The transcriptome and DNA methylation landscapes of human primordial germ cells. *Cell* 2015; **161**: 1437–1452.
- Wamunyokoli FW, Bonome T, Lee JY, et al. Expression profiling of mucinous tumors of the ovary identifies genes of clinicopathologic importance. *Clin Cancer Res* 2006; **12**: 690–700.
- Fukushima N, Sato N, Prasad N, et al. Characterization of gene expression in mucinous cystic neoplasms of the pancreas using oligonucleotide microarrays. *Oncogene* 2004; **23**: 9042–9051.
- Tone AA, Begley H, Sharma M, et al. Gene expression profiles of luteal phase fallopian tube epithelium from *BRCA* mutation carriers resemble high-grade serous carcinoma. *Clin Cancer Res* 2008; **14**: 4067–4078.
- Yeung TL, Leung CS, Wong KK, et al. TGF-beta modulates ovarian cancer invasion by upregulating CAF-derived versican in the tumor microenvironment. *Cancer Res* 2013; **73**: 5016–5028.
- Pei H, Li L, Fridley BL, et al. FKBP51 affects cancer cell response to chemotherapy by negatively regulating Akt. *Cancer Cell* 2009; **16**: 259–266.
- R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing: Vienna, 2013. Available from: <https://www.r-project.org>.
- Irizarry RA, Hobbs B, Collin F, et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 2003; **4**: 249–264.
- Gautier L, Cope L, Bolstad BM, et al. Affy – analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* 2004; **20**: 307–315.
- Eisenberg E, Levanon EY. Human housekeeping genes are compact. *Trends Genet* 2003; **19**: 362–365.
- Gagnon-Bartsch JA, Speed TP. Using control genes to correct for unwanted variation in microarray data. *Biostatistics* 2012; **13**: 539–552.
- Ritchie ME, Phipson B, Wu D, et al. *limma* powers differential expression analyses for RNA-seq and microarray studies. *Nucleic Acids Res* 2015; **43**: e47.
- Huang da W, Sherman BT, Stephens R, et al. DAVID gene ID conversion tool. *Bioinformatics* 2008; **2**: 428–430.
- Karst AM, Levanon K, Drapkin R. Modeling high-grade serous ovarian carcinogenesis from the fallopian tube. *Proc Natl Acad Sci U S A* 2011; **108**: 7547–7552.
- Vandesompele J, De Preter K, Pattyn F, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 2002; **3**: RESEARCH0034.
- Elias KM, Labidi-Galy SI, Vitonis AF, et al. Prior appendectomy does not protect against subsequent development of malignant or borderline mucinous ovarian neoplasms. *Gynecol Oncol* 2014; **132**: 328–333.
- Brown J, Frumovitz M. Mucinous tumors of the ovary: current thoughts on diagnosis and management. *Curr Oncol Rep* 2014; **16**: 389.
- Perets R, Wyant GA, Muto KW, et al. Transformation of the fallopian tube secretory epithelium leads to high-grade serous ovarian cancer in *Brca*/*Tp53*/*Pten* models. *Cancer Cell* 2013; **24**: 751–765.
- Labidi-Galy SI, Papp E, Hallberg D, et al. High grade serous ovarian carcinomas originate in the fallopian tube. *Nat Commun* 2017; **8**: 1093.
- Matthaei H, Schlick RD, Hruban RH, et al. Cystic precursors to invasive pancreatic cancer. *Nat Rev Gastroenterol Hepatol* 2011; **8**: 141–150.

31. Chu PG, Chung L, Weiss LM, *et al.* Determining the site of origin of mucinous adenocarcinoma: an immunohistochemical study of 175 cases. *Am J Surg Pathol* 2011; **35**: 1830–1836.
32. Suzuki Y, Sugiyama M, Abe N, *et al.* Immunohistochemical similarities between pancreatic mucinous cystic tumor and ovarian mucinous cystic tumor. *Pancreas* 2008; **36**: e40–e46.
33. Jones RE and Lopez KH. Sexual differentiation and development. In *Human Reproductive Biology (4th edn)*. Elsevier: London, 2014; Ch 5.
34. Wang Y, Shwartz LE, Anderson D, *et al.* Molecular analysis of ovarian mucinous carcinoma reveals different cell of origins. *Oncotarget* 2015; **6**: 22949–22958.
35. Bosman F, Carneiro F, Hruban R, *et al.* *WHO Classification of Tumours of the Digestive System (4th edn)*. IARC: Lyon, 2010.
36. Matsubara T, Sato Y, Sasaki M, *et al.* Immunohistochemical characteristics and malignant progression of hepatic cystic neoplasms in comparison with pancreatic counterparts. *Hum Pathol* 2012; **43**: 2177–2186.
37. Adsay NV, Eble JN, Srigley JR, *et al.* Mixed epithelial and stromal tumor of the kidney. *Am J Surg Pathol* 2000; **24**: 958–970.
38. Calio A, Eble JN, Grignon DJ, *et al.* Mixed epithelial and stromal tumor of the kidney: a clinicopathologic study of 53 cases. *Am J Surg Pathol* 2016; **40**: 1538–1549.
39. Prendergast GC. Actin' up: RhoB in cancer and apoptosis. *Nat Rev Cancer* 2001; **1**: 162–168.
40. Maris P, Blomme A, Palacios AP, *et al.* Asporin is a fibroblast-derived TGF-beta1 inhibitor and a tumor suppressor associated with good prognosis in breast cancer. *PLoS Med* 2015; **12**: e1001871.
41. Satoyoshi R, Kuriyama S, Aiba N, *et al.* Asporin activates coordinated invasion of scirrhous gastric cancer and cancer-associated fibroblasts. *Oncogene* 2015; **34**: 650–660.
42. Rochette A, Boufaied N, Scarlata E, *et al.* Asporin is a stromally expressed marker associated with prostate cancer progression. *Br J Cancer* 2017; **116**: 775–784.

SUPPLEMENTARY MATERIAL ONLINE

Supplementary figure legends

Figure S1. Laser-capture microdissection of the epithelium of mucinous tumors

Figure S2. Unsupervised hierarchical clustering of PGCs, HGSOCS, Fallopian tube epithelium, ovarian surface epithelium, and MBOT samples

Figure S3. Heatmap of differentially expressed genes among MCNs, MBOTs, MOCs, and PGCs versus adnexal and pancreatic samples

Table S1. Published datasets used in array expression analysis

Table S2. Primers for RT-qPCR

Table S3. Gene set enrichment analysis for genes highly expressed by PGCs and mucinous cystic neoplasms of the pancreas

Table S4. Gene set enrichment analysis for genes highly expressed by PGCs and mucinous ovarian tumors

Dataset S1. Top 1000 genes highly expressed by PGCs and mucinous cystic neoplasms of the pancreas

Dataset S2. Top 1000 genes highly expressed by PGCs and mucinous ovarian tumors

Dataset S3. Top 411 genes highly expressed by PGCs, mucinous ovarian tumors, and mucinous cystic neoplasms of the pancreas