Pathology - Research and Practice 254 (2024) 155129

Contents lists available at ScienceDirect

Pathology - Research and Practice

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

Comparative transcriptomic analyses reveal activation of the epithelial-mesenchymal transition program in non-metastasizing low grade pseudomyxoma peritonei

Elise Pretzsch [a,b], Jens Neumann [b,c,*], Hanno Nieß [a], Charlotte M. Pretzsch [d], F.O. Hofmann [a,b], Thomas Kirchner [b,c], Frederick Klauschen [b,c], Jens Werner [a,b], Martin Angele [a,1], Jörg Kumbrink [b,c,1]

[a] Department of General, Visceral, and Transplant Surgery, Ludwig-Maximilians-University Munich, Munich, Germany
[b] German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg, partner site Munich, Germany
[c] Institute of Pathology, Faculty of Medicine, Ludwig-Maximilians-Universität München, Munich, Germany
[d] Department of Forensic and Neurodevelopmental Sciences, Institute of Psychiatry, Psychology and Neuroscience, King's College London, De Crespigny Park, London SE5 8AF, UK

ARTICLE INFO

Keywords:
LAMN
EMT
CRC
Metastasis

ABSTRACT

Epithelial-mesenchymal transition (EMT), angiogenesis, cell adhesion and extracellular matrix (ECM) interaction are essential for colorectal cancer (CRC) metastasis. Low grade mucinous neoplasia of the appendix (LAMN) and its advanced state low grade pseudomyxoma peritonei (lgPMP) show local aggressiveness with very limited metastatic potential as opposed to CRC. To better understand the underlying processes that foster or impede metastatic spread, we compared LAMN, lgPMP, and CRC with respect to their molecular profile with subsequent pathway analysis.

LAMN, lgPMP and (mucinous) CRC cases were subjected to transcriptomic analysis utilizing Poly(A) RNA sequencing. Successfully sequenced cases (LAMN n = 10, 77%, lgPMP n = 13, 100% and CRC n = 8, 100%) were investigated using bioinformatic and statistical tests (differential expression analysis, hierarchical clustering, principal component analysis and gene set enrichment analysis).

We identified a gene signature of 28 genes distinguishing LAMN, lgPMP and CRC neoplasias. Ontology analyses revealed that multiple pathways including EMT, ECM interaction and angiogenesis are differentially regulated. Fifty-three significantly differentially regulated gene sets were identified between lgPMP and CRC followed by CRC vs. LAMN (n = 21) and lgPMP vs. LAMN (n = 16). Unexpectedly, a substantial enrichment of the EMT gene set was observed in lgPMP vs. LAMN (FDR=0.011) and CRC (FDR=0.004). Typical EMT markers were significantly upregulated (Vimentin, TWIST1, N-Cadherin) or downregulated (E-Cadherin) in lgPMP. However, MMP1 and MMP3 levels, associated with EMT, ECM and metastasis, were considerably higher in CRC.

We show that the different tumor biological behaviour and metastatic spread pattern of midgut malignancies is reflected in a different gene expression profile. We revealed a strong activation of the EMT program in non-metastasizing lgPMP vs. CRC. Hence, although EMT is considered a key step in hematogenous spread, successful EMT does not necessarily lead to hematogenous dissemination. This emphasizes the need for further pathway analyses and forms the basis for mechanistic and therapy-targeting research.

1. Introduction

Metastatic spread is the major cause of death in colorectal cancer patients [1]. Despite rapid advances in local and systemic therapies, 5-year relative survival rates for patients with metastatic colorectal cancer (CRC) only range between 14–17% [2]. The process of tumor cells leaving their primary site and forming new colonies in distant organs is summed up under the umbrella term "invasion-metastasis

* Correspondence to: Institute of Pathology, Medical Faculty, Ludwig-Maximilians-University München (LMU), Thalkirchner Straße 36, 80337 Munich, Germany.
E-mail address: jens.neumann@med.uni-muenchen.de (J. Neumann).
[1] Both last authors contributed equally to this work

https://doi.org/10.1016/j.prp.2024.155129
Received 14 November 2023; Received in revised form 8 January 2024; Accepted 10 January 2024
Available online 14 January 2024
0344-0338/© 2024 The Authors. Published by Elsevier GmbH. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

cascade" [3,4]. The underlying tumor biological mechanisms of this cascade are only partly understood yet. However, epithelial-mesenchymal transition (EMT), local invasion, systemic circulation, tumor cell adhesion and interaction with extracellular matrix (ECM) components, and angiogenesis are considered essential in this highly dynamic and complex process and pivotal steps for hematogenous spread [5,6]. Different tumor entities from colorectal origin that show different tumor biological behaviour could serve as models to further identify and better characterize the respective processes underlying metastatic spread. In this respect, extended gene expression analysis has the potential to pinpoint markers that are critical for metastatic progression, identify biomarkers that allow precise risk stratification, establish druggable targets to advance current treatment options and overall improve clinical decision making in a precision medicine approach.

The low grade mucinous neoplasia of the appendix (LAMN) and its advanced state the low grade pseudomyxoma peritonei (lgPMP) are rare malignancies with very limited potential of metastasis. LAMN is characterized by local aggressiveness exhibiting an expansile growth with pushing borders whereby it can traverse the wall of the appendix. Whereas LAMN does not invade, the produced mucin can push on the appendix with the risk to spill out resulting in lgPMP. In this respect, LAMN often maintains indolent histological features, whereas the lgPMP disseminates aggressively throughout the abdominal cavity. Mortality in lgPMP is ascribed to the excessive accumulation of extracellular mucin rather than metastatic spread that seldom occurs [7–9].

In this respect, we believe that LAMN and lgPMP could serve as excellent models in comparison to CRC to investigate mechanisms that foster or impede metastatic progression. Transcriptomic analyses using Poly(A) RNA sequencing followed by differential expression analysis (DEseq2), hierarchical clustering, principal component analysis (PCA) and gene set enrichment analysis were performed to molecularly characterize the tumor biological behaviour of the non-metastasizing LAMN and lgPMP as opposed to CRC.

2. Methods

2.1. Study collective and tissue selection

Patients undergoing surgical resection for LAMN, lgPMP or non-metastasized mucinous colorectal cancer of the midgut at the Department of General, Visceral and Transplant Surgery, LMU Munich between 2008 and 2019 were registered in a prospectively maintained database. We used this database and identified a cohort of 51 patients. The corresponding clinicopathological data were obtained from the databases of the Department of General, Visceral, and Transplant Surgery and the Institute of Pathology. Because LAMN and lgPMP are usually microsatellite stable [10], only right-sided CRC patients with microsatellite stability (MSS) status and KRAS mutation were included to generate a study cohort as homogenous as possible, leaving a study collective of 44 patients. In a next step, formalin fixed and paraffin embedded (FFPE) samples were obtained from the archives of the Institute for Pathology, LMU Munich. As not all cases had tissue samples available, the final study collective available for NGS incorporated 34 patients. The study was conducted according to the recommendations of the local ethics committee of the Medical Faculty of the Ludwig-Maximilians-Universität München (ethics vote 20–0984).

2.2. Histopathological samples

Histopathological diagnosis and classification was reviewed for every available tumor specimen by two experienced pathologists (JN, TK) of the Institute of Pathology of the University of Munich (Germany). Histopathological grade was confirmed by an experienced pathologist (JN).

2.3. RNA extraction from FFPE samples

Sections from formalin fixed paraffin embedded (FFPE) tissue samples were prepared followed by hematoxylin-eosin staining of one slide. Neoplastic/tumor cell content (30–90%, average 62.1%) was estimated by an experienced pathologist. Neoplastic/tumor tissues were microdissected from subsequent unstained sections and used for RNA preparation. Total RNA was extracted from four to 12 sections of FFPE tissue sections using the RNeasy FFPE Kit (Qiagen, Hilden, Germany) following the manufacturer's instructions. RNA yield and purity were quantified with the NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Rockland, USA).

2.4. RNA sequencing, data processing, and analyses

Poly(A) RNA sequencing was performed using an adapted protocol for FFPE samples of the previously described protocol also containing primer sequences [11,12]. Briefly, 50 ng of fragmented total RNA (5 ng/µl) were reverse transcribed with Maxima H Minus reverse transcriptase (Thermofisher, Waltham, ᵱMA, USA) using oligo(dT) primers including barcode and unique molecular identifier (UMI) sequences (E3V6NEXT primers) as well as the E5V6NEXT universal adapter primer. cDNAs of a maximum of 24 samples were pooled and purified with DNA Clean & Concentrator-5 columns (Zymo Research, Irvine, CA, USA), treated with Exonuclease I (New England Biolabs, Ipswich, MA, USA) and afterwards PCR amplified (15 cycles) with KAPA Hifi Hot Start polymerase (Roche, Basel, Switzerland) and SINGV6 primer. AMPure beads (Beckman Coulter, Brea, CA, USA) were used for subsequent purification of the PCR products. The purified DNA was quantified utilizing the Qubit dsDNA HS Assay Kit (Thermofisher, Waltham, ᵱMA, USA). Subsequently, the Nextera XT DNA library preparation kit (Illumina, San Diego, CA, USA) was used for rapid library preparation according to the manufacturer's instructions, with the exception that the i5 primer was replaced by the P5NEXTPT5 primer. The quality of the library was checked by Qubit quantification and with an Aligent 2100 bioanalyzer using Agilent High sensitivity DNA chips (Agilent, Santa Clara, CA, USA). The library was denatured and diluted to 20 pM. Sequencing was performed on an Illumina NextSeq 500 using the NextSeq 500/550 high output v2 kit (Illumina, San Diego, CA, USA) for dual-indexed sequencing according to the manufacturer's instructions.

Sequencing raw data were processed as follows. Demultiplexing and conversion: bcl2fastq2. The subsequent steps were performed using the modules available at the Galaxy platform [13]. Quality control: FastQC/MultiQC. FastQ to BAM (Alignment): RNAStar. Counting/Annotation: HTSeq-count/hg19 GTF/GFF File. Normalization of expression data, quantification of expression changes and identification of differentially expressed genes (DEGs) were conducted with DESeq2 [14]. Benjamini and Hochberg correction was used for calculation of adjusted p values/false discovery rates (FDR) [15]. The following sequencing quality metrics were applied: 1) FastQC mean Phred score > 30; 2) reads mapped $\geq$ 1.5 M; 3) reads aligned $\geq$ 1 M. In addition, principal component analysis (PCA) visual outliers were excluded. Detailed quality metrics are presented for each sample in Supplementary Fig. 1 and Supplementary Table 1.

Unsupervised hierarchical clustering and PCA were performed with the ClustVis web tool [16]. Default settings were used except for heatmap clustering distance for rows and columns, which was set to Euclidean. VENN diagram was created with a web tool (default settings) available at https://bioinformatics.psb.ugent.be/webtools/Venn/. Volcano plots were visualized with the volcano plot module available at the Galaxy platform [13]. Gene Set Enrichment Analyses (GSEA) were performed using the Hallmark, C2 KEGG and C6 oncogenic signature data sets with GSEA v4.3.2 (1000 permutations; Broad Institute, Cambridge, MA, USA) [17]. Single gene expression statistics (Mann-Whitney test) in Fig. 5 and Supplementary Fig. S3 were conducted from normalized expression data from DESeq2 with Graphpad Prism (v.8.2.1,

GraphPad Software, Inc., San Diego, CA, USA). All p values < 0.05 (two-sided) were regarded significant.

3. Results

3.1. Study design

This study aims to identify differentially regulated pathways or expressed genes (DEG) that might help to understand the distinct aggressiveness of LAMN, lgPMP and right-sided CRC. Therefore, tissue samples from 51 patients were inspected for sufficient tissue for transcriptome analyses. Sufficient tissue was available from 34 patients. RNA extracted from these FFPE samples (LAMN n = 13, lgPMP n = 13, CRC n = 8) was subjected to RNA sequencing followed by transcriptomic analyses. Successful sequencing was achieved in 91% of the samples (LAMN n = 10 (77%), lgPMP n = 13 (100%), CRC n = 8 (100%). Patient characteristics (including TNM stage, grading etc.) and a cohort diagram/study design are shown in Table 1 and Fig. 1, respectively.

3.2. Identification of DEGs/ a DEG signature distinguishing LAMN, lgPMP and CRC

To classify LAMN (group 1), lgPMP (group 2) and CRC (group 3) based on their transcriptomic profile, a comparative expression analysis was performed with DESeq2 followed by PCA and unsupervised hierarchical clustering [14,16]. The by far highest number of DEG ($\log_2$fold change >0.5; $p_{adjusted}$ <0.05) was observed when comparing group 2 vs 3 (1376 of 15254 successfully investigated genes; 9%), followed by 2 vs 1 (471/15255; 3.1%) and 1 vs 3 (331/14877; 2.2%) (Fig. 2A). Only two genes were differentially expressed in all group comparisons (Fig. 2B), whereas the largest number of DEG (n = 995) was exclusive to 2 vs 3. The TOP20 DEG between all groups are presented in Table 2. Importantly, a perfect separation of the three groups was achieved by the TOP10 genes of each comparison (28 genes after removal of duplicates) in PCA and hierarchical clustering (Fig. 2C and D), suggesting that this 28 gene signature could be utilized to distinguish LAMN, lgPMP and right-sided CRC.

3.3. Identification of pathways differentially regulated in LAMN, lgPMP and CRC

The large number of identified DEG suggests a strong influence on the activity of signalling networks and cellular programs especially when comparing lgPMP and CRC. To gain insight into signalling pathways and cellular systems that are distinctly regulated, GSEA (Hallmark, C2 KEGG and C6 oncogenic gene sets; FDR <0.3) was conducted (Fig. 3). Group 2 vs 1 revealed in total 11 sets (seven oncogenic, avg. FDR 0.168; three ECM/adhesion, avg FDR 0.056; one angiogenesis, FDR 0.102) associated with group 2 and one with group 1 (oncogenic, FDR 0.209) (Fig. 3A). Fourteen gene sets were correlated with group 3 (13 oncogenic, avg. FDR 0.087; one ECM/adhesion, FDR 0.296) and two with group 1 (one ECM/adhesion, FDR 0.236; one angiogenesis, FDR 0.076) in the group 3 vs 1 comparison. The group 2 vs 3 analysis identified 22 gene sets associated with group 2 (11 oncogenic, avg. FDR 0.146; eight ECM/adhesion, avg. FDR 0.099; three angiogenesis, avg. FDR 0.160) and 12 sets (all oncogenic, avg. FDR 0.090) enriched in group 3. Detailed results for significantly enriched gene sets in the hallmark and C2 KEGG sets as well as selected sets in the C6 oncogenic collection are displayed in Fig. 3 B-D. The results clearly show that multiple cellular programs associated with ECM/adhesion (e.g. KEGG ECM receptor interaction, FDR 0.003), angiogenesis (Hallmark angiogenesis, FDR 0.16), MAPK signalling (KEGG MAPK Signaling Pathway, FDR 0.22) and EMT (Hallmark EMT, FDR 0.004, Hallmark TGF Beta Signaling, FDR 0.014) are upregulated in lgPMP whereas DNA repair programs (KEGG Mismatch Repair, FDR 0.055; KEGG Base Excision Repair, FDR 0.108) are downregulated compared with CRC.

Table 1
Study cohort characteristics.

Case number	Tumor entity	Group	Tumor cells (%)	TNM stage	UICC	Grading
1	CRC (cecum)	3	80%	pT4b, pN1a, pM1c (PER), L0, V0, Pn0	IVc	G1
2	CRC (ascending)	3	60%	pT4a, pN2a, L1, V0, Pn0	IIIc	G3
4	lgPMP	2	30%	pT4b, pN0, L0, V0, Pn0	IIc	G1
5	LAMN	1	70%	pT4a, pN0, pM1b (PER, OTH), L0, V0, Pn0	IVa	G1
9	lgPMP	2	70%	pT4b, pN0, L0, V0, Pn0	IIc	G1
10	lgPMP	2	60%	pT3, pN0, L0, V0, Pn0, M1b		G1
11	LAMN	1	50%	pTis, pN0, L0, V0, Pn0	I	G1
12	LAMN	1	50%	pT1, pNX, L0, V0, Pn0	I	G1
13	LAMN	1	70%	pT2, pNx, L0, V0, Pn0	I	G1
16	lgPMP	2	30%	n.a.	n.a.	n.a.
17	lgPMP	2	40%	pT4b, pNX, L0, V0, Pn0	IIc	G1
18	LAMN	1	90%	n.a.		
19	LAMN	1	60%	pT4a, pN0, L0, V0, Pn0, pM1a	IVa	G1
20	lgPMP	2	80%	n.a.	n.a.	n.a.
21	LAMN	1	60%	pT2, N0, L0, V0, Pn0	I	G1
22	LAMN	1	60%	pT4a, pN0, L0, V0, Pn0	IIb	G1
25	LAMN	1	70%	pTis, pNX, L0, V0, Pn0, M1b	n.a.	n.a.
26	lgPMP	2	90%	pT4b, pNX, L0, V0, Pn0	IIc	G1
27	LAMN	1	60%	pT4a, pN0, L0, V0, Pn0	IIb	G1
28	lgPMP	2	40%	pTX, pN0, L0, V0, Pn0, M1b		G1
29	CRC (ascending)	3	70%	pT3, pN0, L0, V0, Pn0	IIa	G3
30	CRC (ascending)	3	80%	pT2, pN0, V0, L0, Pn0	I	G3
31	LAMN	1	50%	pTis, pN0, L0, V0, Pn0	0	G1
34	lgPMP	2	30%	n.a.	n.a.	n.a.

(continued on next page)

Table 1 (*continued*)

Case number	Tumor entity	Group	Tumor cells (%)	TNM stage	UICC	Grading
35	LAMN	1	80%	pT4a, pNX, L0, V0, Pn0	IIb	G1
36	LAMN	1	80%	T4a, pN0, L0, V0, Pn0	IIb	G1
39	lgPMP	2	40%	pT3, pN0, L0, V0, Pn0	n.a.	G1
40	CRC (ascending)	3	70%	pT3, pN0, L0, V0, Pn0	II	G3
41	CRC (ascending)	3	50%	pT3, pN0, L0, V0, Pn0	IIa	G3
42	CRC (ascending)	3	70%	pT4a, pN2, L1, V0, Pn0	IV	G3
44	lgPMP	2	70%	n.a.	n.a.	n.a.
46	CRC (ascending)	3	70%	pT3, pN1a, L0, V0, Pn0	IIa	G3
47	lgPMP	2	70%	n.a.	n.a.	n.a.
48	lgPMP	2	60%	n.a.	n.a.	n.a.

n.a., not available

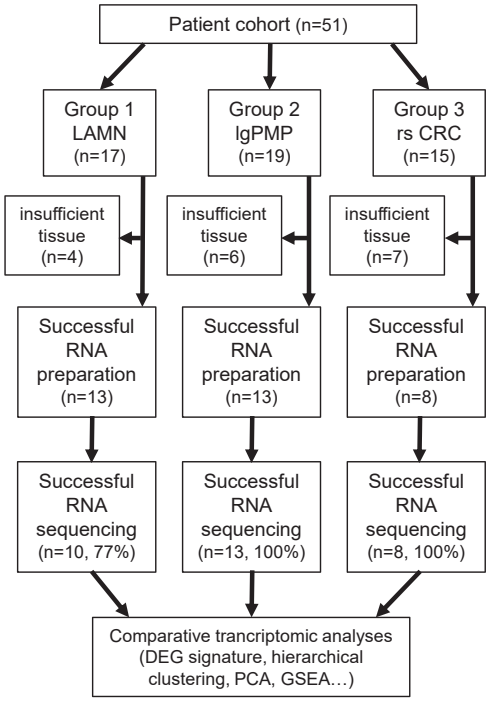


Fig. 1. Cohort diagram. DEG, differentially expressed gene. GSEA, gene set enrichment analysis. LAMN, low grade mucinous neoplasia of the appendix. lgPMP, low grade pseudomyxoma peritonei. PCA, principal component analysis. Rs CRC, right-sided colorectal cancer.

3.4. The EMT program is strongly upregulated in lgPMP compared with LAMN and CRC

Unexpectedly, the most significant activation within the investigated GSEA sets was observed for the Hallmark EMT set in lgPMP patients (Fig. 4A and B). Compared with group 1, a normalized enrichment score (NES) of 1.82 (FDR 0.012) and with group 3 of 1.92 (FDR 0.004) were found. Therefore, we further focussed on the role of EMT especially in lgPMP patients. The Hallmark EMT gene set contains 200 genes of which 197 were successfully analysed in group 1 vs 2 and 2 vs 3 as well as 196 in 1 vs 3 (Fig. 2C). Again, the highest number of DEG was observed in group 2 vs 3 (86 DEG, 43.7%), whereas only 44 (22.3%) and 14 (7.1%) were found in 1 vs 2 and 1 vs 3, respectively. The DEG within the EMT gene set identified in comparison of group 2 vs 3 are depicted in Fig. 4D. Moreover, the detailed results of the TOP30 DEG of each comparison are shown in Supplementary Table S2. PCA using the 86 DEG identified in group 2 vs 3 resulted in a very good separation of all three groups (Fig. 4E). Similar results were obtained with PCA and hierarchical clustering (91% correctly assigned) when utilizing the Top30 genes of each comparison (26 DEG after duplicate removal) (Supplementary Fig. 2). Single gene analyses of typical EMT markers further supported the strong EMT activation in lgPMP (Fig. 5). A significantly lower expression of the epithelial marker E-Cadherin (*CDH1*; group 2 vs 1: 56% lower; 2 vs 3: 61% lower) and corresponding higher levels of the mesenchymal markers Vimentin (*VIM*; group 2 vs 1: 108% higher; 2 vs 3: 186% higher), N-Cadherin (*CDH2*; group 2 vs 1: 139% higher; 2 vs 3: 579% higher) and TWIST1 (group 2 vs 1: 610% higher; 2 vs 3: 74% higher) were measured in group 2. To obtain a better view of the EMT process in midgut cancers, appendix carcinoma (primary, high grade, n = 6, group 4) and appendix carcinoma (high grade) displaying peritoneal carcinomatosis (n = 8, group 5) samples were additionally analysed (Supplementary Fig. 3). Confirming the strong activation of the EMT program in lgPMP (group 2) patients, the Hallmark EMT set was significantly enriched in group 2 compared with group 4 (FDR 0.016) and group 5 (FDR 0.081). Moreover, higher levels of the EMT markers VIM, CDH2 and TWIST1 were found in group 2 vs 4 (VIM: 108% higher; CDH2: 252% higher; TWIST1: 165% higher) and 2 vs 5 (VIM: 60% higher; CDH2: 161% higher; TWIST1: 69% higher). E-Cadherin levels were lower in group 2 vs 4 (40% lower) and similarly lower in group 2 and 5 (group 2: 46% lower; group 5: 28% lower) compared with group 1.

Interestingly, the expression of the EMT-, ECM- and metastasis-associated genes MMP1 and MMP3 [6] is substantially higher in group 3 (CRC) and group 5 (peritoneal carcinomatosis) than in groups 1, 2, and 4.

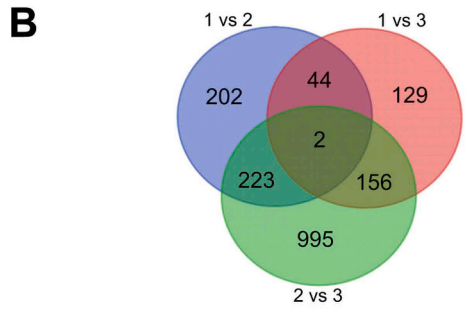
Taken together, these results suggest that the EMT program is unexpectedly active in lgPMP. However, MMPs involved in metastatic processes only present at very low levels in lgPMP in part explaining the low metastatic potential and point to a partial EMT in lgPMP.

4. Discussion

To gain a more comprehensive understanding of key steps in the metastatic process of midgut malignancies, RNA sequencing was performed to compare CRC cases to LAMN and lgPMP cases as models for locally aggressive yet seldom metastasizing neoplasias. Bioinformatic analysis revealed a gene signature of 28 genes that could perfectly separate the three groups, namely distinguishing between LAMN, lgPMP and CRC neoplasias. This demonstrates that the tumor biological behaviour and metastatic spread pattern of midgut cancers is reflected in different molecular profiles. Further ontology analyses revealed that cellular programs that are considered key steps for metastatic progression were differentially regulated between the groups as follows: upregulation of cellular programs associated with EMT, ECM/adhesion, angiogenesis and MAPK signalling in lgPMP (compared with CRC); downregulation of DNA repair programs in lgPMP (compared with CRC). Of interest, lgPMP showed a significant enrichment of the EMT gene set (in comparison to LAMN, CRC and appendix carcinoma as well as peritoneal carcinomatosis) reflected in an upregulation of EMT markers that promote EMT (N-Cadherin, Vimentin, TWIST1) and downregulation of markers that sustain the epithelial state thereby inhibiting EMT (E-Cadherin). This strong activation of the EMT program in a locally aggressive yet non metastasizing neoplasia not only provides new insight into the tumor biological behaviour of lgPMP but also indicates that successful EMT, that is considered a key step in

A

test	DEG	genes investigated	% DEG
group 2 vs 1	471	15255	3.09
group 1 vs 3	331	14877	2.22
group 2 vs 3	1376	15254	9.02



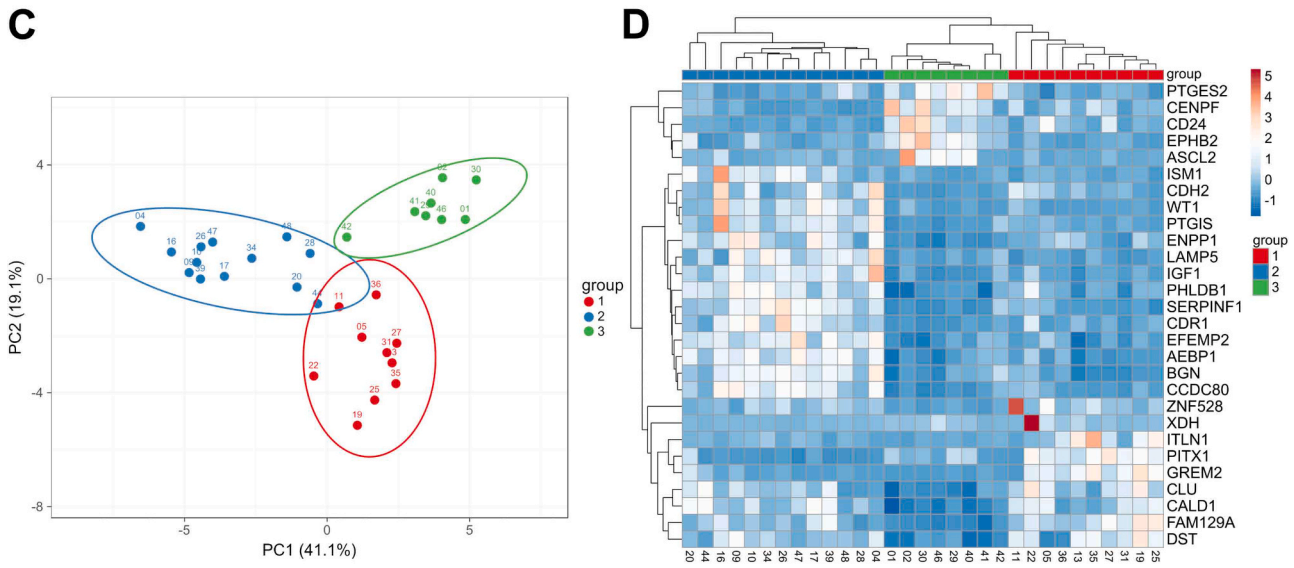


Fig. 2. Identification of DEGs/ a DEG signature distinguishing LAMN (group 1), lgPMP (group 2) and CRC (group 3). A. Number of identified DEG and percentage of all succesfully investigated genes in each group comparison. **B.** VENN diagram of DEG found in each group comparison. **C** and **D**. PCA (**C**) and unsupervised hierarchical clustering (**D**) using the TOP10 FDR DEG found in each gropup comparison. DEG, differentially expressed gene. FC, fold change. FDR, false discovery rate. PCA, principal component analysis.

Table 2
TOP20 differentially expressed genes in all group comparisons.

group 1 vs 2				group 1 vs 3				group 2 vs 3			
gene	log2 FC	*p* value	FDR	gene	log2 FC	*p* value	FDR	gene	log2 FC	*p* value	FDR
PITX1	2.33	1.23×10^{-12}	1.88×10^{-8}	GREM2	3.20	4.95×10^{-14}	7.36×10^{-10}	CCDC80	1.95	2.46×10^{-14}	3.75×10^{-10}
BGN	-1.35	5.06×10^{-11}	3.86×10^{-7}	FAM129A	1.64	4.16×10^{-12}	3.1×10^{-8}	ENPP1	2.00	2.48×10^{-12}	1.89×10^{-8}
AEBP1	-1.08	6.04×10^{-10}	2.82×10^{-6}	EPHB2	-1.94	2.03×10^{-9}	1.01×10^{-5}	SERPINF1	1.56	5.05×10^{-12}	2.57×10^{-8}
XDH	2.80	7.39×10^{-10}	2.82×10^{-6}	ZNF528	2.18	2.50×10^{-8}	9.28×10^{-5}	CDR1	2.16	1.11×10^{-11}	4.22×10^{-8}
EFEMP2	-1.00	3.15×10^{-9}	9.25×10^{-6}	ASCL2	-2.41	4.93×10^{-8}	0.000105	PTGIS	1.93	2.65×10^{-11}	8.1×10^{-8}
LAMP5	-2.14	3.64×10^{-9}	9.25×10^{-6}	CLU	1.35	4.32×10^{-8}	0.000105	CD24	-2.06	4.49×10^{-11}	1.14×10^{-7}
SERPINF1	-1.25	4.48×10^{-9}	9.76×10^{-6}	ITLN1	2.37	4.09×10^{-8}	0.000105	IGF1	1.88	9.7×10^{-11}	2.11×10^{-7}
ISM1	-1.99	7.71×10^{-9}	1.47×10^{-5}	PTGES2	-1.55	6.33×10^{-8}	0.000118	PHLDB1	1.22	1.51×10^{-10}	2.88×10^{-7}
WT1	-2.10	1.22×10^{-8}	2.07×10^{-5}	DST	0.88	1.52×10^{-7}	0.000251	CDH2	2.29	1.99×10^{-10}	3.38×10^{-7}
CCDC80	-1.34	2.82×10^{-8}	3.7×10^{-5}	CALD1	0.95	2.32×10^{-7}	0.000308	CENPF	-1.91	2.36×10^{-10}	3.61×10^{-7}
HSD11B2	2.06	2.91×10^{-8}	3.7×10^{-5}	DLG2	2.35	2.81×10^{-7}	0.000308	GPX2	-1.60	3.22×10^{-10}	3.97×10^{-7}
WT1-AS	-2.17	2.48×10^{-8}	3.7×10^{-5}	DMBT1	-2.17	2.57×10^{-7}	0.000308	VIM	1.43	3.09×10^{-10}	3.97×10^{-7}
TWIST1	-1.93	5.31×10^{-8}	6.23×10^{-5}	NEXN	1.34	2.10×10^{-7}	0.000308	WT1	2.40	3.39×10^{-10}	3.97×10^{-7}
ANKRD18A	1.96	6.41×10^{-8}	6.27×10^{-5}	SORBS1	1.71	3.09×10^{-7}	0.000308	KIF14	-2.58	4.71×10^{-10}	5.13×10^{-7}
GREM2	2.13	6.58×10^{-8}	6.27×10^{-5}	SYNPO2	1.68	3.11×10^{-7}	0.000308	AEBP1	1.14	5.65×10^{-10}	5.75×10^{-7}
SPARC	-1.26	6.25×10^{-8}	6.27×10^{-5}	MRVI1	1.4	3.41×10^{-7}	0.000318	ID4	1.75	1.33×10^{-9}	1.27×10^{-6}
LGALS4	1.18	8.15×10^{-8}	7.31×10^{-5}	PTGER3	1.7	4.7×10^{-7}	0.000411	EFEMP2	1.07	2.04×10^{-9}	1.64×10^{-6}
IGF1	-1.47	9.08×10^{-8}	7.7×10^{-5}	FHL1	1.76	5.19×10^{-7}	0.000429	HOOK1	-2.02	1.88×10^{-9}	1.64×10^{-6}
MRC2	-1.18	1.15×10^{-7}	8.47×10^{-5}	GNG7	2.28	5.71×10^{-7}	0.000446	MAP7	-1.54	2.03×10^{-9}	1.64×10^{-6}
PTGIS	-1.46	1.17×10^{-7}	8.47×10^{-5}	LPP	1.06	5.99×10^{-7}	0.000446	DPYSL3	1.34	2.66×10^{-9}	1.84×10^{-6}

FDR, false discovery rate. FC, fold change.

hematogenous metastatic spread, does not necessarily need to lead to hematogenous dissemination. Also, it may indicate, that EMT alone, which is not only relevant for metastasis but also for local invasion does not suffice to initiate invasion, a necessary step that precedes metastatic spread. This might also be reflected in the typical kind of growth of LAMN, where the basal membrane is not infiltrated but rather mucus is

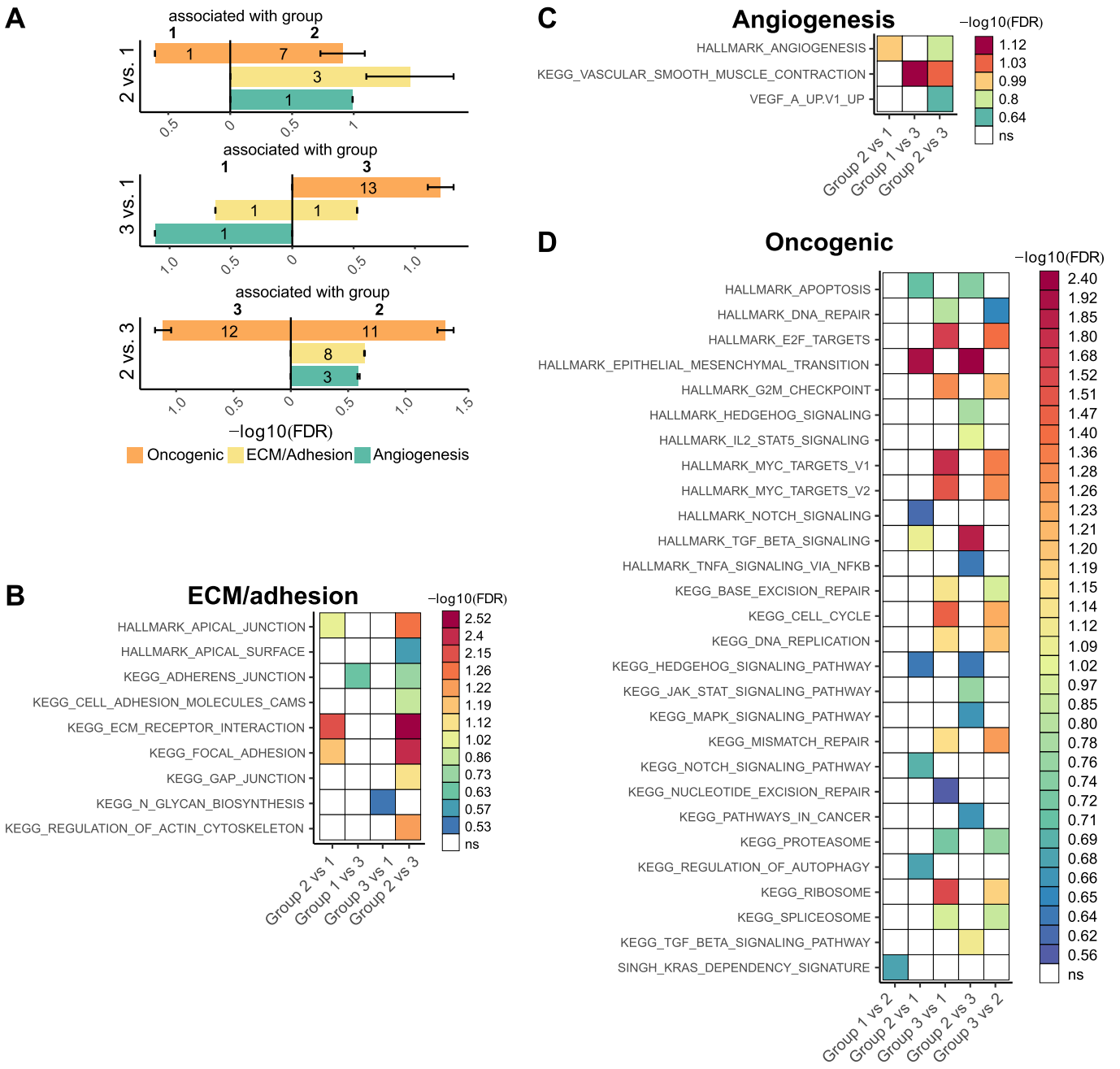


Fig. 3. Identification of pathways differentially regulated in LAMN (group 1), lgPMP (group 2) and CRC (group 3). GSEA results from analyses with Hallmark, C2 KEGG and C6 oncogenic signature data sets. **(A)** Total number of associated gene sets (FDR <0.3) in the indicated groups are presented within the bars. Average FDR (-$\log_{10}$) and SD of all significantly enriched gene sets are shown for each group comparison. **B to D.** Significantly enriched gene sets (FDR <0.3) associated with ECM/adhesion (**B**), angiogenesis (**C**) or oncogenic progression (**D**) in indicated group comparisons are depicted. FDR are presented as -$\log_{10}$. Gene sets are enriched in the first mentioned group (e.g. 1 vs 2 = enrichment in 1). ECM, extracellular matrix. FDR, false discovery rate. GSEA, gene set enrichment analysis. ns, not significant.

pushed through, referred to as pushing borders, as well as the down-regulation of MMPs that commonly promote invasion.

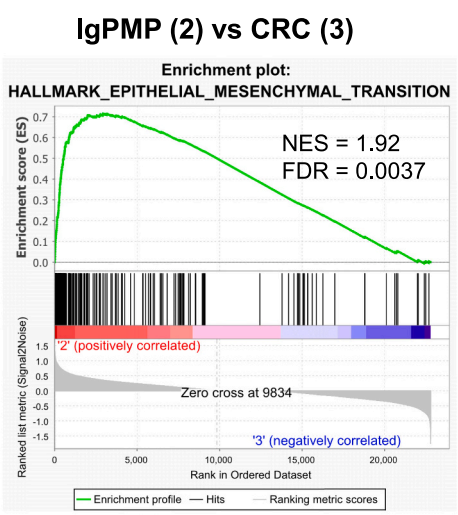
Due to their rarity, only few studies have investigated the molecular profile of LAMN and lgPMP. To our knowledge, this is the first study subjecting LAMN and lgPMP cases to transcriptomic analysis by RNA sequencing and subsequent comparison of gene expression results to CRC. Previous studies indicate that some molecular characteristics in the pathogenesis are common to both lgPMP and CRC (KRAS mutation, MSS status) which is why we only selected KRAS mutated, MSS positive mucinous CRC cases as a comparator arm to ensure a homogenous study collective [8–10]. However, despite these shared characteristics and being both of colorectal origin, lgPMP and CRC show significantly different metastatic behaviour. In this respect, cellular programs such as EMT, tumor cell adhesion and interaction with extracellular matrix (ECM) components, and angiogenesis which are described as hallmarks of metastatic spread were activated in the non-metastasizing lgPMP [3].

EMT enables epithelial tumor cells to lose their cell-cell adherence and acquire mesenchymal properties that are a prerequisite for migration, intravasation and invasion into distant tissue sites. The transition from an epithelial to a mesenchymal state is reflected by a down-regulation of the epithelial marker E-Cadherin [18] that stabilizes cell-cell contacts and maintains epithelial cell polarity and subsequent upregulation of mesenchymal markers such as N-Cadherin, Vimentin [19,20] and TWIST1 [21,22]. This phenotype that is characteristic for

A

HALLMARK_EMT	NES	NOM *p*	FDR
group 2 vs 1	1.83	0.005	0.011
group 1 vs 3	1.16	0.33	0.48
group 2 vs 3	1.92	0.0006	0.0037

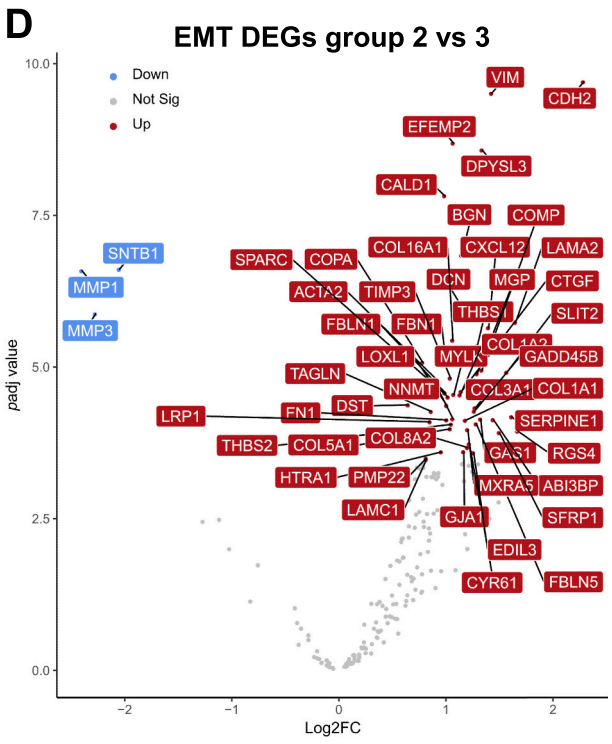
B



C

Test group	DEG	HALLMARK_EMT (investigated genes)	% DEG
1 vs 2	44	197	22.34%
1 vs 3	14	196	7.14%
2 vs 3	86	197	43.65%

D



E

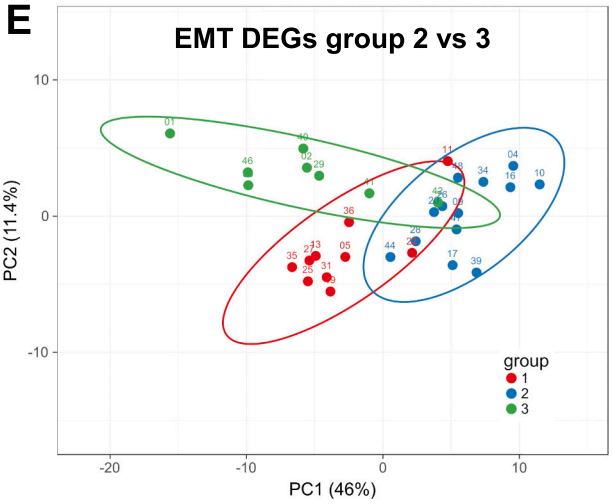


Fig. 4. The EMT program is strongly upregulated in lgPMP (group 2) compared with LAMN (group 1) and CRC (group 3). A and B. GSEA results of the Hallmark EMT gene set comparing the indicated groups. **C.** Number of identified DEG and percentage of all succesfully investigated genes within the Hallmark EMT gene set in each group comparison. **D.** Volcano plot visualizing the DEG in group 2 vs. 3. Labeled are genes with *p*adj< 0.01 and log2FC> 0.5 **E.** PCA using the 86 DEG within the Hallmark EMT gene set identified in the comparison group 2 vs 3. DEG, differentially expressed gene. FC, fold change. FDR, false discovery rate. NES, normalized enrichment score. Nom *p*, nominal *p* value. PCA, principal component analysis.

the metastatic progression of neoplasias was observed in lgPMP. Our finding is in line with a previous study that described the respective phenotype in PMP using immunohistochemistry [23]. In contrast, another study reported opposite results with an upregulation of E-Cadherin, however, the sample size was small and included not only lgPMP but also PMP from adenocarcinoma [24]. Although the question of why lgPMP exhibit an EMT phenotype but do not present with hematogenous spread remains to be elucidated, our data suggest that in lgPMP cellular programs that impede metastatic spread may be upregulated and mechanisms that facilitate metastatic dissemination may be downregulated. In this respect, cellular programs regulating epithelial integrity and tumor cell adhesion to epithelium and ECM (hallmark apical junction, hallmark adherens junction, KEGG focal adhesion, KEGG ECM receptor interaction etc.) were all activated, whereas MMPs (MMP1,3) that are essential for invasion [25,26] were downregulated [3,27] compared to CRC and peritoneal carcinomatosis cases. As CRC cases were all non-metastatic, which means the EMT program may not have been activated yet, we performed the same analyses comparing lgPMP to peritoneal carcinomatosis cases, tumors that are clearly metastatic. Of interest, we found very equal results namely a strong activation of the EMT program in lgPMP. This points towards lgPMP exhibiting an EMT state, whereas at the same time, mechanisms that sustain epithelial cell polarity and hinder detachment and dissemination are also active. This was reflected in an upregulation of tight junctions, adherens junctions and desmosomes, potentially indicating a state of "partial EMT" [28]. In this respect, the expression of genes related to EMT could be necessary to sustain the peritoneal growth of PMP. This might be morphologically reflected by PMP tumor cells that accumulate in large amounts of extracellular mucin within in the abdominal cavity but do not have the ability to spread outside the peritoneum. The upregulation of adhesion

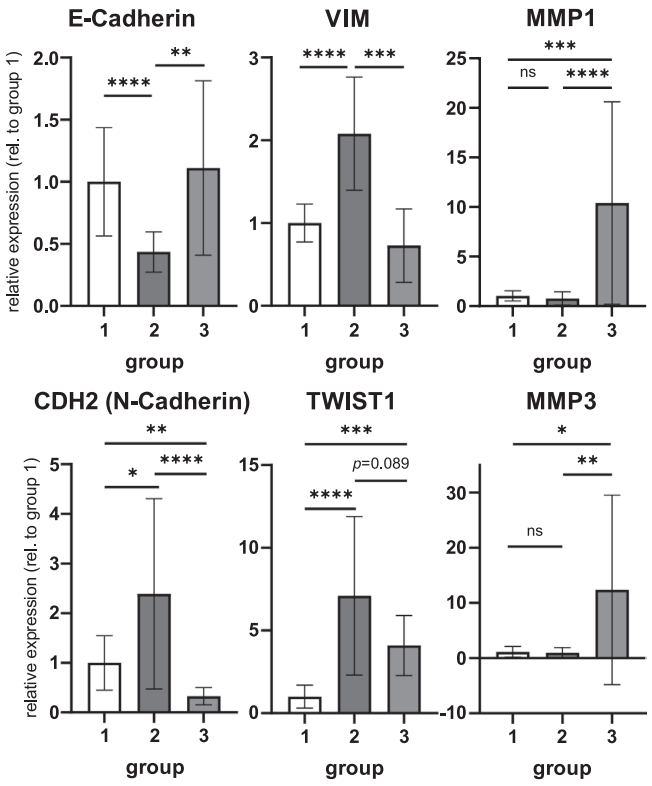


Fig. 5. The EMT program is enriched in lgPMP (group 2) – single gene analysis of EMT markers. Relative expression of the indicated EMT associated genes normalized to group 1 is depicted. *p* values were calculated with Mann-Whitney test. * , $p < 0.05$. * *, $p < 0.01$. * ** , $p < 0.001$, * ** *, $p < 0.0001$. ns, not significant.

molecules was confirmed by another study that suggested that this might contribute to the noninvasive phenotype [24]. Another explanation for the observed gene expression pattern might also be that the activation of the EMT program in lgPMP is not the cause for the dissemination into the peritoneum but its consequence due to the change of microenvironment from appendix to peritoneal cavity. Potentially, as opposed to acellular mucin in perforated LAMNs, lgPMP cells have the ability to survive in the peritoneum only through an upregulation of the EMT pathway. In this respect, the loss of contact of tumor cells to ECM stroma components in the peritoneal cavity may lead to a shift in gene expression patterns by which tumor cells react with an upregulation of the EMT program in order to survive and interact with the new environment [29].

Angiogenesis associated programs were also upregulated in lgPMP. In line with this, vessels can sometimes be observed throughout the mucinous tissues that provide the nutritional environment required for tumor growth [24]. However, potentially due to the overall "partial EMT" status, PMP cells may not be able to enter the blood stream and disseminate which needs to be further explored.

Limitations of this study include the small sample size, which is owed to the rarity of LAMN and lgPMP. We admit that only limited conclusions can be drawn from a cohort of 34 patients. Nevertheless, we believe that this work contributes to the very limited number of studies to date addressing molecular changes in the extremely rare diseases LAMN and lgPMP. Due to the high technical effort, low chance of successful immunohistochemistry in PMP samples, and also lack of leftover tissue of all samples we unfortunately had to decide to abstain from immunohistochemistry. Nevertheless, in order to generate the best quality data, we decided to use hematoxylin-eosin staining to mark tumor regions of interest followed by precise microdissection of the marked tumor tissues. This way, we aimed to ensure a very high tumor cell content in actual tissue used for transcriptomic analyses and to minimize the chance of accidently using neighboring stromal tissue. Further, due to the retrospective approach of our molecular analyses, validation using independent datasets would be desirable, however, to our knowledge, these are currently not available. We continue to maintain a prospective database, however, due to the rarity of the disease, data is only slowly growing.

In summary, our gene expression data show that different tumor biological behaviour and metastatic spread patterns in midgut malignancies are reflected in different molecular profiles. Successful EMT, even though a key process in metastatic spread does not necessarily need to lead to metastatic dissemination. Further studies are warranted to identify underlying mechanisms that impede metastatic progression after EMT is activated. Rather than focusing on single parameters, characterization of cellular programs and pathways might bring immense benefit with respect to the development of targeted therapies that could prevent metastatic spread in CRC in the future.

5. Conclusion

Th different tumor biological behaviour and metastatic spread pattern of midgut malignancies is reflected in different gene expression profiles. There is a strong activation of the EMT program in non-metastasizing lgPMP vs. CRC. In this respect, although EMT is considered a key step in hematogenous spread, successful EMT does not necessarily lead to hematogenous dissemination. This emphasizes the need for further pathway analyses and forms the basis for mechanistic and therapy-targeting research.

Funding

The authors received no specific funding for this work.

CRediT authorship contribution statement

Pretzsch Elise: Conceptualization, Formal analysis, Methodology, Writing – original draft, Writing – review & editing. **Kumbrink Jörg:** Conceptualization, Data curation, Formal analysis, Investigation, Software, Supervision, Visualization, Writing – original draft. **Angele Martin:** Conceptualization, Funding acquisition, Resources, Supervision, Writing – review & editing. **Werner Jens:** Funding acquisition, Writing – review & editing. **Klauschen Frederick:** Conceptualization, Funding acquisition, Writing – review & editing. **Kirchner Thomas:** Conceptualization, Funding acquisition, Writing – review & editing. **Hofmann F.O.:** Writing – review & editing. **Pretzsch Charlotte M.:** Formal analysis, Supervision, Writing – review & editing. **Nieß Hanno:** Writing – review & editing. **Neumann Jens:** Formal analysis, Funding acquisition, Supervision, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability statement

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Acknowledgements

We thank Sabine Sagebiel-Kohler for organisatory help and for establishing and performing the FFPE RNA sequencing library preparation protocol.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.prp.2024.155129.

References

[1] J. Ferlay, H.R. Shin, F. Bray, D. Forman, C. Mathers, D.M. Parkin, Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008, Int. J. Cancer vol. 127 (12) (2010) 2893–2917, https://doi.org/10.1002/ijc.25516.

[2] "Colorectal Cancer Early Detection, Diagnosis, and Staging," *American Cancer Society*, 2022.

[3] E. Pretzsch, et al., Mechanisms of metastasis in colorectal cancer and metastatic organotropism: hematogenous versus peritoneal spread, " *J. Oncol.*, Vol. 2019 (2019), https://doi.org/10.1155/2019/7407190.

[4] K.A. Paschos, A.W. Majeed, N.C. Bird, "Natural history of hepatic metastases from colorectal cancer - Pathobiological pathways with clinical significance,", World J. Gastroenterol. vol. 20 (14) (2014) 3719–3737, https://doi.org/10.3748/wjg.v20.i14.3719.

[5] S.M. Weis, D.A. Cheresh, Tumor angiogenesis: molecular pathways and therapeutic targets, Nat. Med. vol. 17 (11) (2011) 1359–1370, https://doi.org/10.1038/nm.2537.

[6] K. George S, P. Theofilos, E. Susan E, K. Richard, R. Robert H, D. Eleftherios P, Cancer-associated fibroblasts drive the progression of metastasis through both paracrine and mechanical pressure on cancer tissue, Mol. Cancer Res vol. 10 (11) (2015) 1403–1418, https://doi.org/10.1158/1541-7786.MCR-12-0307.Cancer-Associated.

[7] W.L. Shaib, et al., Appendiceal mucinous neoplasms: diagnosis and management, 137–137, Oncologist vol. 23 (1) (2018), https://doi.org/10.1634/theoncologist.2017-0081erratum.

[8] E.M. Gleeson, et al., Appendix-derived pseudomyxoma peritonei (PMP): molecular profiling toward treatment of a rare malignancy, Am. J. Clin. Oncol. Cancer Clin. Trials vol. 41 (8) (2018) 777–783, https://doi.org/10.1097/COC.0000000000000376.

[9] P. Nummela, et al., Genomic profile of pseudomyxoma peritonei analyzed using next-generation sequencing and immunohistochemistry, Int. J. Cancer vol. 136 (5) (2015) E282–E289, https://doi.org/10.1002/ijc.29245.

[10] J. Misdraji, L.J. Burgart, G.Y. Lauwers, Defective mismatch repair in the pathogenesis of low-grade appendiceal mucinous neoplasms and adenocarcinomas, Mod. Pathol. vol. 17 (12) (2004) 1447–1454, https://doi.org/10.1038/modpathol.3800212.

[11] M. Soumillon, D. Cacchiarelli, S. Semrau, A. van Oudenaarden, T.S. Mikkelsen, Characterization of directed differentiation by high-throughput single-cell RNA-Seq, bioRxiv (2014) 003236.

[12] J. Kumbrink, P. Li, A. Pók-udvari, F. Klauschen, T. Kirchner, A. Jung, P130cas is correlated with ereg expression and a prognostic factor depending on colorectal cancer stage and localization reducing folfiri efficacy, Int. J. Mol. Sci. vol. 22 (22) (2021), https://doi.org/10.3390/ijms222212364.

[13] V. Jalili, et al., The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2020 update, Nucleic Acids Res vol. 48 (W1) (2021) W395–W402, https://doi.org/10.1093/NAR/GKAA434.

[14] M.I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2, Genome Biol. vol. 15 (12) (2014) 1–21, https://doi.org/10.1186/s13059-014-0550-8.

[15] Y. Benjamini, Y. Hochberg, "Controlling the false discovery rate: a practical and powerful approach to multiple testing,", J. R. Stat. Soc. vol. 57 (June) (1995) 289–300, https://doi.org/10.2307/2346101.

[16] T. Metsalu, J. Vilo, ClustVis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap, Nucleic Acids Res vol. 43 (W1) (2015) W566–W570, https://doi.org/10.1093/nar/gkv468.

[17] A. Subramanian, et al., Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles, Proc. Natl. Acad. Sci. vol. 102 (43) (2005) 15545–15550, https://doi.org/10.1073/pnas.0506580102.

[18] F. Van Roy, Beyond E-cadherin: roles of other cadherin superfamily members in cancer, Nat. Rev. Cancer vol. 14 (2) (2014) 121–134, https://doi.org/10.1038/nrc3647.

[19] Y. Toiyama, et al., Increased expression of slug and vimentin as novel predictive biomarkers for lymph node metastasis and poor prognosis in colorectal cancer, Carcinogenesis vol. 34 (11) (2013) 2548–2557, https://doi.org/10.1093/carcin/bgt282.

[20] C. Velez-Delvalle, M. Marsch-Moreno, F. Castro-Muñozledo, I.J. Galván-Mendoza, W. Kuri-Harcuch, Epithelial cell migration requires the interaction between the vimentin and keratin intermediate filaments, Sci. Rep. vol. 6 (November 2015) (2016) 1–10, https://doi.org/10.1038/srep24389.

[21] T. Okada, Y. Suehiro, K. Ueno, S. Mitomori, S. Kaneko, M. Nishioka, TWIST1 hypermethylation is observed frequently in colorectal tumors and its overexpression is associated with unfavorable outcomes in patients with colorectal cancer, Genes, Chromosom. Cancer vol. 462 (February) (2010) 452–462, https://doi.org/10.1002/gcc.

[22] M.L. Hui Cao, Enping Xu, Hong Liu, Ledong Wan, Epithelial-mesenchymal transition in colorectal cancer metastasis: a system review, Pathol. - Research Pract. vol. 211 (8) (2015) 557–569, https://doi.org/10.1016/j.prp.2015.05.010.

[23] R. Bibi, et al., A specific cadherin phenotype may characterise the disseminating yet non-metastatic behaviour of pseudomyxoma peritonei, Br. J. Cancer vol. 95 (9) (2006) 1258–1264, https://doi.org/10.1038/sj.bjc.6603398.

[24] K. Flatmark, B. Davidson, A. Kristian, H.T. Stavnes, M. Førsund, W. Reed, Exploring the peritoneal surface malignancy phenotype-a pilot immunohistochemical study of human pseudomyxoma peritonei and derived animal models, Hum. Pathol. vol. 41 (8) (2010) 1109–1119, https://doi.org/10.1016/j.humpath.2009.12.013.

[25] X. Wang, Y. Liu, Y. Ding, G. Feng, CAMSAP2 promotes colorectal cancer cell migration and invasion through activation of JNK/c-Jun/MMP-1 signaling pathway, Sci. Rep. vol. 12 (1) (2022) 1–14, https://doi.org/10.1038/s41598-022-21345-7.

[26] Y. Wen, et al., Histone deacetylase (HDAC) 11 inhibits matrix metalloproteinase (MMP) 3 expression to suppress colorectal cancer metastasis, J. Cancer vol. 13 (6) (2022) 1923–1932, https://doi.org/10.7150/jca.66914.

[27] J. Mikuła-Pietrasik, P. Uruski, A. Tykarski, K. Książek, The peritoneal 'soil' for a cancerous 'seed': a comprehensive review of the pathogenesis of intraperitoneal cancer metastases, Cell. Mol. Life Sci. vol. 75 (3) (2018) 509–525, https://doi.org/10.1007/s00018-017-2663-1.

[28] S. Brabletz, H. Schuhwerk, T. Brabletz, M.P. Stemmler, Dynamic EMT: a multi-tool for tumor progression, EMBO J. vol. 40 (18) (2021) 1–22, https://doi.org/10.15252/embj.2021108647.

[29] H. Jung, L. Fattet, J. Yang, Molecular pathways: linking tumor microenvironment to epithelial–mesenchymal transition in metastasis, Clin. Cancer Res. vol. 21 (5) (2015) 962–968, https://doi.org/10.1158/1078-0432.CCR-13-3173.Molecular.