Supplemental data

Methods

Immunohistochemistry

Staining was performed manually or using the Ventana BenchMark Ultra automated IHC/ISH slide staining system (Ventana Medical Systems Inc., Tucson, USA) (suppl. table 1a). The cytoplasmic expression of TFF3 and MUCL3 was evaluated using the immunoreactivity scoring system (IRS) based on staining intensity (0: negative, 1: mild, 2: moderate, 3: intense) and the percentage of stained cells (0: no positive cells, 1: <10% positive cells, 2: 10-50% positive cells, 3: 51-80% positive cells, 4: >80% positive cells). The final score (0-12) was found by multiplying the positive cells proportion score (0-4) and the staining intensity score (0-3). The mean value of Ki67 proliferation rate of five randomly selected high power (40x) fields (HPFs) was calculated using the percentage of stained cells.

DNA/ RNA Isolation from FFPE samples

For genomic DNA or total RNA Isolation, 5-8 8-µm-thick tissue sections were prepared, and lesions were dissected manually or by laser-capture microdissection (LMD), depending on their size to ensure adequate cellularity (>80%) for subsequent molecular analysis. For LMD, cresyl violet staining was done before using the Palm Microbeam System (Carl Zeiss, Oberkochen, Germany) according to the manufacturer's instructions. In some samples containing larger lesions, manual microdissection was used, as previously described.[1] The obtained cell clusters were isolated using the QIAamp DNA micro Kit or the GeneRead FFPE DNA Kit for DNA and the RNeasy FFPE Kit for RNA (all from Qiagen, Hilden, Germany) following the manufacturer's instructions. The genomic DNA quality control was performed by quantitative PCR using the Power SYBR™ Green PCR Master Mix on a StepOnePlus™ Real-Time PCR System. Quantification was performed with a self-designed primer assay (HML-2 for: 5' AAACGCCAATCCTGAGTGTC-3'; HML-2 rev: 5' CATAGCTCCTCCGATTCCAT-3'). These primers are complementary to long terminal repeats (LTRs) of the HML 2 human endogenous retroviruses and have a length of about 115 bp.

Liffers S-T, *et al. Gut* 2022;0:1–13. doi: 10.1136/gutjnl-2021-326550

Targeted NGS

A PDAC-Panel with two primer pools was created by the Ion AmpliSeqTM Designer (v5.6, ThermoFisher Scientific, Dreieich, Germany). The panel consists of 217 amplicons of 21 genes covering hot-spot mutational sites of 18 and the whole coding sequence of 3 (*ARID1A*, *TP53* and *RNF43*) additional genes relevant for PDAC (suppl. table 2). Barcoded libraries from gDNA (up to 10 ng per pool) were prepared using the Ion AmpliSeq Library kit 2.0 with Ion Xpress™ Barcode adapters. The Ion library TaqMan™ Quantitation Kit was used for quantification of the libraries. The libraries were pooled and amplified in an emulsion PCR reaction using the Ion 520™ & Ion 530™ Kit-OT2. The resulting Ion Sphere particles (ISPs) were loaded on a 520™ or 530™ Chip and sequenced on the Ion S5™ system (all reagents from ThermoFisher).

The results of the next generation sequencing from the Ion S5™ system were aligned to the human reference genome (GRCh37/hg19) using the S5 Ion Torrent Server VM (ThermoFisher). The Ion Reporter software (Version 5.12.0.0) was used for variant calling and annotations of the DNA panel sequencing. The parameters for variant calling were set equal for all samples. Following thresholds were defined: 3% allele frequency with a minimum coverage of 500 and a Phred Score of ≥ 30. Detected variants were validated using the Integrative Genomics Viewer (IGV), ClinVar database from National Institutes of Health (NIH) and University of California Santa Cruz (UCSC) Genome Browser. Variants not present in the above mentioned databases were classified according to the American College of Medical Genetics and Genomics (ACMG) guidelines using the ACMG database (varsome.com; v7.3.7).[2]

Fusion transcript analysis

50 ng of isolated RNA were used for cDNA synthesis by QuantiTect Reverse Transcription Kit (Qiagen) and were subsequently subjected to library preparation using the Oncomine Comprehensive Assay Plus RNA (ThermoFisher) targeting over 1,300 isoforms of 49 tumor driver genes including approximately 200 known *BRAF* fusion transcripts. NGS was performed (as described above) and data analysis was done using the Oncomine Comprehensive Plus w2.1 - Fusion workflow implemented within the IonReporter Software package (V5.18; ThermoFisher).

Isolation of epithelial cells from the main pancreatic duct and from peripheral (branch) ducts

Liffers S-T, *et al. Gut* 2022;0:1–13. doi: 10.1136/gutjnl-2021-326550

Specimens were obtained fresh from the operating theater and immediately subjected to gross examination. The main pancreatic duct was probed, and the specimen dissected by a pathologist along the probe. The main duct was then carefully dissected with a scissor and then fixed in 10% buffered formalin and embedded in paraffin. Peripheral tissue blocks were prepared, and branch-ducts were isolated by LMD, as described above. DNA extraction was performed as described above.

Generation of β-cells

β-cell populations from FFPE tissue were generated from 50-µm-thick sections. Tissue sections were dewaxed with xylol and rehydrated in descending ethanol concentrations. Antigen retrieval was done at 80°C for one hour in a pressure cooker before tissue was digested with 1% (w/v) collagenase Ia (Sigma, Steinheim, Germany) and 1%(w/v) dispase (Gibco, Grand Island, USA) for 45 min at 37°C to obtain single cells. The cell suspension was subsequently filtered (30 µm mesh) and the cells were collected by centrifugation. Single cells were stained Insulin (Abcam, Cambridge, UK; 1:200). The stained cells were sorted with a BD FACS AriaTM III System. DNA was isolated from the sorted cells as described above.

Transcriptome analysis

After total RNA isolation, the samples were shipped to Macrogen (Seoul, Korea) for sequencing. Libraries from total RNA were prepared using the Illumina TruSeq™ Stranded mRNA Library Prep kit and sequenced with 2 x 100 bp on the Illumina NovaSeq 6000 (Illumina Inc, San Diego, USA). The raw data processing of the transcriptome data was performed by Macrogen. Briefly, adapter and low-quality base trimming was carried out with Trimmomatic (v0.38).[3] Trimmed reads were mapped against the GRCh38/hg38 human reference genome using the Bowtie2 (v2.3.4.1) aligner.[4] Afterwards, the aligned reads were assembled with Cufflinks (v2.2.1).[5] After assembly the abundance of gene was calculated in read counts per gene. Before differential gene expression analysis lowly expressed genes were filtered from the data set. Therefore, genes which showed a lower read count as 0.5 transcripts per million reads and were missing in more than one sample per group were excluded from further analysis. The filtered raw count matrix was normalized and batch-corrected using the DESeq2 package (v3.14).[6] Finally, differentially expressed genes were calculated pairwise and defined as followed $\log_2$ fold change of < -1 and > 1, respectively, and the significance level of

the adjusted p-value was set to < 0.05. PCA, heatmap and expression plots were calculated based on the variance stabilizing transformation output of DESeq2

Pathway analysis

Gene set enrichment analysis

For methylation data, enrichment of KEGG terms was estimated for all differentially methylated probes (DMP) in a pair-wise manner. DMPs were defined as displaying a beta value change of 0.4 and an adjusted p-value < 0.05. Gene set enrichment was calculated with the gometh function of the missMethyl package (v.1.26.1).[7]

The single sample Gene Set Enrichment Analysis (ssGSEA) was performed only for RNA seq derived data. Briefly, the normalized enrichment scores (NES) were calculated on the variance stabilized transformation data with the GSVA package (v.1.40.1).[8] Differentially activated gene sets were calculated between the different precursor lesions as described by Larsen *et al.* with a p-value of < 0.05.[9]

VIPER analysis

The activation of transcription factors was calculated with the VIPER algorithm (v1.26.0).[9] For the analysis, the paad regulon was taken from the arcane.networks package (1.18.0). Activated transcription factors were defined as displaying a p-value < 0.005 and a NES score of >3 or >-3.

Statistical analysis

Statistical analysis was performed using the GraphPad Prism 8 software (GraphPad Software Inc., San Diego, USA) or R v. 3.6.0 (R Core Team 2018). Statistical significance in immunohistochemistry was determined by Kruskal-Wallis test with Dunn's multiple comparison test. Results are presented as means ± standard error of the mean (SEM). P values less than 0.05 were considered statistically significant (* p< 0.05; ** p < 0.01; *** p<0.001).

References

1 Schlitter AM, Born D, Bettstetter M, *et al.* Intraductal papillary neoplasms of the bile duct: Stepwise progression to carcinoma involves common molecular pathways. *Mod Pathol* 2014;27:73–86.

2 Richards S, Aziz N, Bale S, *et al.* Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* 2015;17:405–24.

3 Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30:2114-2120.

4 Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012;9:357-9.

5 Li H, Handsaker B, Wysoker A, *et al.* The sequence alignment/map format and SAMtools. *Bioinformatics* 2009;25:2078-9.

6 Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014;15:550.

7 Phipson B, Maksimovic J, Oshlack A. missMethyl: an R package for analyzing data from Illumina's HumanMethylation450 platform. *Bioinformatics* 2016;32:286-8.

8 Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics* 2013;14:7.

9 Larsen BM, Kannan M, Langer LF, *et al.* A pan-cancer organoid platform for precision medicine. *Cell Rep* 2021;36:109429.

10 Alvarez MJ, Shen Y, Giorgi FM, *et al.* Functional characterization of somatic mutations in cancer using network-based inference of protein activity. *Nat Genet* 2016;48:838-47.

Supplementary tables

Supplementary table 1a: Antibodies and protocols for immunohistochemistry.

Antibody	Type	Dilution	Antigen Demasking	Source
Anti-MUC1	Mo Mono	1:100	CC1	Biocare
Anti-MUC2	Mo Mono	1:100	CC1	Dako
Anti-MUC5AC	Mo Mono	1:1000	CC1	Chemicon
Anti-CDX2	Mo Mono	1:40	CC1	BioGenex
Anti-MIB1	Mo Mono	1:100	CC1	Dako
Anti-TFF3	Rb Mono	1:2000	EDTA buffer pH 9	Abcam
Anti-MUCL3	Rb Poly	1:500	Citrate buffer pH 6	LSBio

*Rb: rabbit. Mo: mouse Mono: monoclonal. Poly: polyclonal. CC1: Cell Conditioning 1 (Ventana Medical System, Tucson, AZ, USA).

Supplementary table 1b: Tissue collective used for Ki67, TFF3 and anti-MUCl3 staining.

Type of lesion	**Number of lesions**
PanIN	**31**
Low grade	26
High grade	5
Gastric IPMN	**28**
Low grade	20
High grade	8
Intestinal IPMN	**20**
Low grade	9
High grade	11
PDAC	**24**

Liffers S-T, *et al. Gut* 2022;0:1–13. doi: 10.1136/gutjnl-2021-326550

1 **Supplementary table 2: Genes and amplicons in targeted NGS.**

Gene Symbol	Chr	Ion AmpliSeq Fwd Primer (5'-3')	Ion AmpliSeq Rev Primer (5'-3')	Amplicon ID
ALK	chr2	TCTCTCGGAGGAAGGACTTGAG	GCCCAGACTCAGCTCAGTTAAT	CHP2_ALK_1
ALK	chr2	ACAGGGTACCAGGAGATGATGTAAG	GGAAGAGTGGCCAAGATTGGA	CHP2_ALK_2
APC	chr5	GAGAGAACGCGGAATTGGTCTA	GTATGAATGGCTGACACTTCTTCCA	CHP2_APC_1
APC	chr5	AGCACTGATGATAAACACCTCAAGTT	ATCTTCTTGACACAAAGACTGGCT	CHP2_APC_2
APC	chr5	TTCATTATCATCTTTGTCATCAGCTGAA	TTTGGTTCTAGGGTGCTGTGAC	CHP2_APC_3
APC	chr5	GCAGACTGCAGGGTTCTAGTT	GTGAACTGACAGAAGTACATCTGCT	CHP2_APC_4
APC	chr5	AGCCCCAGTGATCTTCCAGATA	CCCTCTGAACTGCAGCATTTACT	CHP2_APC_5
APC	chr5	AGAGGGTCCAGGTTCTTCCA	TCATTTTCCTGAACTGGAGGCATT	CHP2_APC_6
APC	chr5	ATGAAACAGAATCAGAGCAGCCTAAA	CGTGATGACTTTGTTGGCATGG	CHP2_APC_7
ARID1A	chr1	CAAAATGAACAACAAGGCAGATGGG	TCAGAGACTATCTAGTCCGGTGTC	ARID1A_10.112972
ARID1A	chr1	CAGCTAAACTTACTGGACTTGAGAATTTTT	GAGTCAAGACAAAAATCACTACCTTGG	ARID1A_10.135473
ARID1A	chr1	CATGATGGGAACTGGACCTCCTTA	TTAGCTGTGATGTGACTCTTGAAGAAAT	ARID1A_10.143283
ARID1A	chr1	CCCCCAGCCTACGGCTTC	CCCCGCGTAGGGCTCCA	ARID1A_1.1.15178
ARID1A	chr1	CCCTAGGCCCGCCCTGA	GGCTCCGGCCGTAGGGT	ARID1A_1.1.16654
ARID1A	chr1	CAGTCAAGAGACTTCTGAGACCCTTA	CAGATAACGGTCCACCCACATC	ARID1A_11.181180
ARID1A	chr1	CCGCTGGGAAAGGAGCTG	GCCTAGGGCCCGCGTTC	ARID1A_1.1.20289
ARID1A	chr1	CTATCGCCTCTATGTGTCTGTGAAG	GTACCACATGAAGCCAGTGAGTAC	ARID1A_11.248116
ARID1A	chr1	ACAACTCCTACTACCCCAACC	CTGCTGAGCGAAGGACGA	ARID1A_1.1.2481
ARID1A	chr1	CTTCCAGAAATCCAGTTCTTCTACTACA	ATAGAGGTCCAGAGGTTTCCTACC	ARID1A_11.279375
ARID1A	chr1	CTCAGCAGCGCTTCGGG	GGGCCCGCCACTGTAGT	ARID1A_1.1.36612
ARID1A	chr1	CTCGGAGCTGAAGAAAGCCG	GCTCTCGGCCCCGTCCT	ARID1A_1.1.38056
ARID1A	chr1	GAGCCCGTCTGCCGTCG	GGAGTTGTACTGGTGGTTGGG	ARID1A_1.1.42139
ARID1A	chr1	GGCCCCAGCAGAACTCTCAC	AGCCCGGAGTGCCACCTC	ARID1A_1.1.52554
ARID1A	chr1	GGCTGCCGGCTCCAAGC	GCTGGGCGACGTGAGCA	ARID1A_1.1.54514
ARID1A	chr1	GGGATCATGGCCGCGCA	CCGGCGGCTGCCTTCAT	ARID1A_1.1.54590
ARID1A	chr1	TTATCTGGCCTTCACTGAGGAGAA	CTCACCTGAGTCAATCCACCAAT	ARID1A_11.550938
ARID1A	chr1	AGCCGGACCTGAAGAACTCG	GGCCGCGGCTGAGTGAG	ARID1A_1.1.6484
ARID1A	chr1	CTCGCCCGGACCCCTCAG	GCCAGACAATGGCAGCTCC	ARID1A_1.2.19161
ARID1A	chr1	GGGCTACCAGGGCTACCC	GGGCTCATGGGCGCGTG	ARID1A_1.2.26067
ARID1A	chr1	GATATACCTCGACTCCTTTGGTTTGG	AGGGTCTTCTCCCCGTTCAAT	ARID1A_12.293039
ARID1A	chr1	GCCAGCTCCTTGAAAAAGCAGTATATC	GACCCCATCCTTACCAGGAGAG	ARID1A_12.311881
ARID1A	chr1	AGACATCTTTGCAGCTGCTGATT	CACAGATCCTTGGCATATCCTGTTG	ARID1A_12.73402
ARID1A	chr1	CCGGCGGACATGGCCTC	CCTCCCCACTCAGCTGTGTA	ARID1A_1.2.9363
ARID1A	chr1	CTCAACTTGTATCTCTGTCCACAGC	CTGCTCTTGGCCTTACCTCATG	ARID1A_13.224100
ARID1A	chr1	CTCCTGCGTGTCCTTTGTTATATTGG	TGGAGTCATGGAATTCCGCTT	ARID1A_13.228066
ARID1A	chr1	GAGGAGACTTAAAGCCACCAACTC	CAAGGAGTTCCCATGCACTTATCT	ARID1A_13.262576
ARID1A	chr1	GCCTTGTAGATCCTCTGCTAAGAAG	GCCCCTGCATAGATCCTGATCC	ARID1A_13.286741
ARID1A	chr1	CTTTAATGATGGAAGTGACTCCACATTC	CAAGTTCAAATAGCAATCAGATCAGTCA	ARID1A_14.234479
ARID1A	chr1	TGACTCCAAACCCTGGGTATCA	CATTTCACTGGCCCTGTCTTTACG	ARID1A_14.440936
ARID1A	chr1	GACCACGACAGCACTATCCCTA	TCATGTTTCCCTCAGGCCCTATT	ARID1A_15.209989
ARID1A	chr1	TCACCGCTTGCCTTTCTACG	TCACTCTGTCATAAGGACCTCCA	ARID1A_15.321878
ARID1A	chr1	CCAATTTTGTTTAGGACGGAGCCT	CACCGAGACCAGGCTTTACTC	ARID1A_15.99688

ARID1A	chr1	CTAATCCTGTGTTTCTTTGCCTCCT	TTTTCAAGGCGAACCTGCATG	ARID1A_16.147847
ARID1A	chr1	GGATGTATTCTCCTAGCCGCTAC	TTGGGTGGAGAACTGATTGCCATA	ARID1A_16.243588
ARID1A	chr1	AGCGTGCCATACAGCACT	GGCAGTGGCAGGATAGGCA	ARID1A_18.122838
ARID1A	chr1	AACCGCACCTCTCCTAGC	TCCCGCCGAATCATGGG	ARID1A_18.17117
ARID1A	chr1	CAGATGAAATGCTGCACACAGATC	GATACCTGAGGAATGTGATTCTGCAT	ARID1A_18.249269
ARID1A	chr1	CAGGTATCCAGCCCTGCTC	TGCTATGTGCGAGGCAGGT	ARID1A_18.260793
ARID1A	chr1	CCACTGCCACAGCTGCTAC	GCTGAGCAACCTCAGCTGAT	ARID1A_18.303487
ARID1A	chr1	AAGGCTCGTGGCCTTCCC	GTGCGGTTCTCCATTGGC	ARID1A_18.33212
ARID1A	chr1	CTGTGTCCACCAAGCATCTGG	GGCACGCTGTACATCTCC	ARID1A_18.457891
ARID1A	chr1	GCAAAACATGCCACCACAAATGATG	TGTTCGGTTCACGCCATGATAG	ARID1A_18.536845
ARID1A	chr1	GCCTTCCCCTCAGCAAGATGTATA	GGTCTCGGCCAAACTGGAATG	ARID1A_18.584475
ARID1A	chr1	ACATAGCACCTGCCCCTGT	GGGCAGATTAGGCAACCGAATG	ARID1A_18.63843
ARID1A	chr1	TGCTCAGCAAGGCACCATG	CGAGCCTTCGTGGTTGG	ARID1A_18.820768
ARID1A	chr1	TTGTCTCTGCCTTAGAATTACAAGCG	GCTGGGCAGCTTGTTGCT	ARID1A_18.880618
ARID1A	chr1	AGACGACATGGAGGTTTATTTCAGG	CCCCAGGCACTGATACTCA	ARID1A_19.54023
ARID1A	chr1	ATCTTCAGAGTAGCTTCACTGATGGG	GTTGATGGTATCTAATGCCCATGTG	ARID1A_19.79292
ARID1A	chr1	CAACATCCTGCTGTATGATGACAAC	GGCATGGAAGATATCTACAAGAGAGAAA	ARID1A_19.96133
ARID1A	chr1	ATCCTGGGAGGTTCAGCAA	GAAGCTGGCTTGTCCTTGC	ARID1A_20.1.247050
ARID1A	chr1	CCAGCTTCAGAGAATAGTGAGGA	GCAGCAGGCCACTGTCAAA	ARID1A_20.1.351511
ARID1A	chr1	CCCTCGGAAGCATGTGACAAC	CTTGATGGCCTCTGAACTCTTAGC	ARID1A_20.1.374770
ARID1A	chr1	CCTGCTGCACTGGCGGAT	GGCCCCTCCTGGTCTGTTG	ARID1A_20.1.397870
ARID1A	chr1	CGAAGCCTGTCATTTGTGCCA	GTTAGTGGTGCCTGCTTCCG	ARID1A_20.1.400147
ARID1A	chr1	CTGTTCTTAGGCCACTTTTCTCC	CCCAGGATCCAGTAGCGTT	ARID1A_20.1.464959
ARID1A	chr1	GAGGAAGTAGTTGAAAATGATGAGGAGA	TCCACCACAAATGGATCATTCTTCTGTA	ARID1A_20.1.565326
ARID1A	chr1	GCAGCAAGTTTCCATTTGGCATTAG	AGGCTTCGAATGGTATTGGACAC	ARID1A_20.1.612574
ARID1A	chr1	GCCTGATTGAGATCTTTGGCATTTTAAA	ACTTCCTCTTCTTCTTCCTCTTCTAGTTTA	ARID1A_20.1.637352
ARID1A	chr1	GGGCCCCCACCTGATGGA	GTTCCGGTGGCTCTGTGC	ARID1A_20.1.733579
ARID1A	chr1	GTGGTGGACTGCTCAGATAAGCT	AGCTCTGTCTTGCTCTCGAAGT	ARID1A_20.1.787955
ARID1A	chr1	ACCGGAACATCAAGATCCTAGAG	CCCTGGGTGTTTGGACATC	ARID1A_20.1.90003
ARID1A	chr1	ATGGTGCGCTTCCTCAGT	CGGCAAGGCTGTCCTCTAG	ARID1A_20.2.152846
ARID1A	chr1	CAGTGCAGAAGGGCAGTATCG	CCGCATCATGTCCACACTAGTTG	ARID1A_20.2.222602
ARID1A	chr1	CCACTAACTTATGAAAAGGAGGAGGAA	CCCGAGATGTTGGCGAGTGTA	ARID1A_20.2.247152
ARID1A	chr1	CCTTGCCGCCACACAGTT	CAACAGCCGTGATTCGTACA	ARID1A_20.2.313338
ARID1A	chr1	CTTGGAGATGCTCCGGGAA	AGGGCAAACTGCCCAGTGTA	ARID1A_20.2.409280
ARID1A	chr1	CTTCAGCTGAAGCCCAGGAC	CGCACCATAGTGCTATACAACTTCT	ARID1A_20.2.432258
ARID1A	chr1	AAACTCAGCATCCAGGACAACAAT	GCCAGGTTGGCCAGCAGTA	ARID1A_20.2.5295
ARID1A	chr1	ACCCAGGGCTGCTGCTCAT	TCTCCAAGCAGTCCCACCA	ARID1A_20.2.54967
ARID1A	chr1	GGCTGTTGGACATCTCGGT	GTTTTGCATAAATAAAGGGCAACAGTC	ARID1A_20.2.603909
ARID1A	chr1	GGGCAGTTGGACCTATCTCCATAC	CTGAGTTTGCTGAGGGTTTCCAA	ARID1A_20.2.613243
ARID1A	chr1	TGGACGAGAACCACTCAGAGTTTAC	GCTGTCATGACTGGCCAATCAAAA	ARID1A_20.2.760156
ARID1A	chr1	CATGGGCGGCCTCTCTTATAC	TAGTAGCACTCTGTAATTAACTGAGCCA	ARID1A_2.198073
ARID1A	chr1	CTAACCCATACTCGCAGCAACA	TCACAATCACCATCTACCTGCTG	ARID1A_2.263887
ARID1A	chr1	GCCATCCAGTCCAATGGATCAG	CCTGCATGGTCATCGGGTAC	ARID1A_2.310812
ARID1A	chr1	AAACCTGTGTACTTGGGTTATATATTCAGT	CCATATGGCTGAGGTCTCATCTTG	ARID1A_2.6808
ARID1A	chr1	AGTCCCAGCAAACTGCCTATTC	ACCCAGAGTTTAATTGGTCTTTAAGTG	ARID1A_3.115759

ARID1A	chr1	CAGCAAAGTCCTCACCCTCAG	GGGATGGCTGCTGGGAGTAT	ARID1A_3.189854
ARID1A	chr1	CAGCCTCCACATCAGCAGTC	AGCCTGCTGGGAGAGCGT	ARID1A_3.203145
ARID1A	chr1	CAGGCTCAGTCTCCTTACCA	GCAGGAGGCAGGGATATCTT	ARID1A_3.210697
ARID1A	chr1	TGCTTTCTATACTCATCATCAGTGCAT	CTTTGCTGGTTGTAATATGGAGTCTG	ARID1A_3.663071
ARID1A	chr1	TTTTCCTTTCCTACAGATTCCTCCTT	CTGCTGCTGATACGAAGGTTG	ARID1A_3.701387
ARID1A	chr1	AGCAGCAGCCACAGTCTCAA	TGAGCCTGTGGCTGTGAGTA	ARID1A_3.77267
ARID1A	chr1	CCATCACAGCTTTTGTTTTTCTTGTTGTAG	ACCTTTCAGAAGGTGCAGAAATACT	ARID1A_4.173246
ARID1A	chr1	CTGGCCTTCACATAATACTTTTCGC	GATGCCTGAGACCCAAATGAATC	ARID1A_4.215850
ARID1A	chr1	GAGGGCAAGAAGATATGAACCTGAG	AGGTCAAAATTAGCTAAACTTCCAACCA	ARID1A_4.255932
ARID1A	chr1	GAGTCCTGGAGTGAGCACATC	CGAGAGTGGTCCTGAGCGA	ARID1A_5.220317
ARID1A	chr1	AGAATCTTTCTGCCTAATATTACTAATCCATG	AGGAGACTGAGCTGGATTACTCT	ARID1A_5.34011
ARID1A	chr1	AGTCTCCTTTCTCTCCTCATACCT	CAGTCACCTTTCCCTCTCCCTAA	ARID1A_5.70816
ARID1A	chr1	GATATGCTTATGTTGTTCTTTGTCTGGA	CACTCAAATGTCTGCCCTAGCTC	ARID1A_6.227771
ARID1A	chr1	AGCCATTTCTAGCTCTGAATTAACTTCC	TCGATCTTGGGCAATGCTTGAT	ARID1A_6.45476
ARID1A	chr1	CAGCCTTATCTCCGCGTCAG	ACTGTTTTCTCCTCTCACCCGTAT	ARID1A_7.149596
ARID1A	chr1	CAGGATAAGGATGGAGAGCATTTGTTC	TGTGTGTATCTGTCCTCCGGAA	ARID1A_7.152412
ARID1A	chr1	CATGGCCAATATGCCACCTCA	ATAATACATTTTCTTGCACTGACACCCT	ARID1A_8.210031
ARID1A	chr1	CCAATGCCAACTACCCCAGTG	GGCCATGTTAGGGCCATAAGG	ARID1A_8.229785
ARID1A	chr1	GTTGCTAGTGAGTGACTAACCAAGTC	GGCTGTCCATGCATTTGACCTC	ARID1A_8.517555
ARID1A	chr1	AGGATGAGTCACGCCTCCATG	GGCCTTACCTGTTTTGGATAGAGTTG	ARID1A_8.91537
ARID1A	chr1	AGCACTATTTGGCTCCAGTTCAAATC	GGTTGATCATGCCAGCCATACTATTAA	ARID1A_9.65769
BRAF	chr7	CATACTTACCATGCCACTTTCCCTT	TTTCTTTTTCTGTTTGGCTTGACTTGA	CHP2_BRAF_1
BRAF	chr7	CCACAAAATGGATCCAGACAACTGT	GCTTGCTCTGATAGGAAAATGAGATCTA	CHP2_BRAF_2
CDKN2A	chr9	CACCAGCGTGTCCAGGAA	CCCTGGCTCTGACCATTCTGT	CHP2_CDKN2A_1
CDKN2A	chr9	CATCTATGCGGGCATGGTTACT	CGCTGGTGGTGCTG	CHP2_CDKN2A_2
CTNNB1	chr3	ACTGTTTCGTATTTATAGCTGATTTGATGGA	CCTCTTCCTCAGGATTGCCTTT	CHP2_CTNNB1_1
EGFR	chr7	CCTCATTGCCCTCAACACAGT	TCAGTCCGGTTTTATTTGCATCATAGTT	CHP2_EGFR_1
EGFR	chr7	CACCACGTACCAGATGGATGT	CCCAAAGACTCTCCAAGATGGGATA	CHP2_EGFR_2
EGFR	chr7	AGACATGCATGAACATTTTTCTCCAC	TCCAGACCAGGGTGTTGTTTTC	CHP2_EGFR_3
EGFR	chr7	TGTGGAGCCTCTTACACCCA	GTGCCAGGGACCTTACCTTATAC	CHP2_EGFR_4
EGFR	chr7	ACGTCTTCCTTCTCTCTCTGTCA	CTGAGGTTCAGAGCCATGGA	CHP2_EGFR_5
EGFR	chr7	CATGCGAAGCCACACTGAC	ACATAGTCCAGGAGGCA	CHP2_EGFR_6
EGFR	chr7	GACTATGTCCGGGAACACAAAGA	CCCCATGGCAAACTCTTGCTA	CHP2_EGFR_7
EGFR	chr7	CGCAGCATGTCAAGATCACAGAT	GCATGTGTTAAACAATACAGCTAGTG	CHP2_EGFR_8
FBXW7	chr4	TGACAATGTTTAAAGGTGGTAGCTGTT	ACTCATTGATAGTTGTGAACCAACACA	CHP2_FBXW7_1
FBXW7	chr4	CCTGTGACTGCTGACCAAACTTTTA	CACATCTTTCTTATAGGTGCTGAAAGG	CHP2_FBXW7_2
FBXW7	chr4	CCCAACCATGACAAGATTTTCCC	GGTCATCACAAATGAGAGACAACATCA	CHP2_FBXW7_3
FBXW7	chr4	ACTAACAACCCTCCTGCCATCATA	TCTGCAGAGTTGTTAGCGGTT	CHP2_FBXW7_4
FBXW7	chr4	GTAGAATCTGCATTCCCAGAGACAA	TCTCTTGATACATCAATCCGTGTTTGG	CHP2_FBXW7_5
FGFR2	chr10	CATCACTGTAAACCTTGCAGACAAAC	TGGTCTCTCATTCTCCCATCCC	CHP2_FGFR2_1
FGFR2	chr10	CATCCTCTCTCAACTCCAACAGG	AGTGGATCAAGCACGTGGAAAA	CHP2_FGFR2_2
FGFR2	chr10	GCTTCTTGGTCGTGTTCTTCATT	CTCCTCCTGTGATCTGCAATCT	CHP2_FGFR2_3
FGFR2	chr10	TGGAAGCCCAGCCATTTCTAAA	GATGATGAAGATGATTGGGAAACACAAG	CHP2_FGFR2_4
GNAS	chr20	TTGGTGAGATCCATTGACCTCAATTT	TGAATGTCAAGAAACCATGATCTCTGTT	CHP2_GNAS_1
GNAS	chr20	CCTCTGGAATAACCAGCTGTCC	TGATCCCTAACAACACAGAAGCAA	CHP2_GNAS_2

IDH1	chr2	CCAACATGACTTACTTGATCCCCAT	ATCACCAAATGGCACCATACGA	CHP2_IDH1_1
IDH2	chr15	ACCCTGGCCTACCTGGTC	AGTTCAAGCTGAAGAAGATGTGGAA	CHP2_IDH2_1
KRAS	chr12	CAAAGAATGGTCCTGCACCAGTAATAT	AGGCCTGCTGAAAATGACTGAATATAA	CHP2_KRAS_1
KRAS	chr12	TCCTCATGTACTGGTCCCTCATT	GTAAAAGGTGCACTGTAATAATCCAGACT	CHP2_KRAS_2
KRAS	chr12	CAGATCTGTATTTATTTCAGTGTTACTTACCT	GACTCTGAAGATGTACCTATGGTCCTA	CHP2_KRAS_3
NRAS	chr1	CCTCACCTCTATGGTGGGATCATAT	GTTCTTGCTGGTGTGAAATGACTG	CHP2_NRAS_1
NRAS	chr1	TTCGCCTGTCCTCATGTATTGG	CACCCCCAGGATTCTTACAGAAAA	CHP2_NRAS_2
NRAS	chr1	GCACAAATGCTGAAAGCTGTACC	CAAGTGTGATTTGCCAACAAGGA	CHP2_NRAS_3
PIK3CA	chr3	CCATAAAGCATGAACTATTTAAAGAAGCAAGA	GGTTGAAAAAGCCGAAGGTCAC	CHP2_PIK3CA_1
PIK3CA	chr3	TGGAATGCCAGAACTACAATCTTTTGAT	AAGATCCAATCCATTTTTGTTGTC	CHP2_PIK3CA_10
PIK3CA	chr3	TGGATCTTCCACACAATTAAACAGCAT	TGCTGTTCATGGATTGTGCAATTC	CHP2_PIK3CA_11
PIK3CA	chr3	CCCTTTTTAAAAGTAATTGAACCAGTAGGC	TTTAAGATTACGAAGGTATTGGTTTAGACAGAA	CHP2_PIK3CA_2
PIK3CA	chr3	GACGCATTTCCACAGCTACAC	AGCATCAGCATTTGACTTTACCTTATCA	CHP2_PIK3CA_3
PIK3CA	chr3	CATAGGTGGAATGAATGGCTGAATTATG	TCAATCAGCGGTATAATCAGGAGTTTTT	CHP2_PIK3CA_4
PIK3CA	chr3	TCCCATTATTATAGAGATGATTGTTGAATTTTCCT	CAAACAAGTTTATATTTCCCCATGCCA	CHP2_PIK3CA_5
PIK3CA	chr3	GCTTTGAATCTTTGGCCAGTACCT	CATAAGAGAGAAGGTTTGACTGCCATA	CHP2_PIK3CA_6
PIK3CA	chr3	CAGAGTAACAGACTAGCTAGAGACAATGA	GCACTTACCTGTGACTCCATAGAAA	CHP2_PIK3CA_7
PIK3CA	chr3	CACGATTCTTTTAGATCTGAGATGCACA	CCTTTTGTGTTTCATCCTTCTTCTCCTG	CHP2_PIK3CA_8
PIK3CA	chr3	GATGCAGCCATTGACCTGTTTAC	AGAAAACCATTACTTGTCCATCGTCT	CHP2_PIK3CA_9
PTEN	chr10	GCCATCTCTCTCCTCCTTTTTCTT	GCCGCAGAAATGGATACAGGTC	CHP2_PTEN_1
PTEN	chr10	TGTTAATGGTGGCTTTTTGTTTGTTTGT	TCTACCTCACTCTAACAAGCAGATAACT	CHP2_PTEN_2
PTEN	chr10	CCATAACCCACCACAGCTAGAA	TGCCCCGATGTAATAAATATGCACAT	CHP2_PTEN_3
PTEN	chr10	GGCTACGACCCAGTTACCATAG	TGCCACTGGTCTATAATCCAGATGAT	CHP2_PTEN_4
PTEN	chr10	TGAGATCAAGATTGCAGATACAGAATCC	ACCTTTAGCTGGCAGACCAC	CHP2_PTEN_5
PTEN	chr10	AGGTGAAGATATATTCCTCCAATTCAGGAC	TTGGATATTTCTCCCAATGAAAGTAAAGTAC	CHP2_PTEN_6
PTEN	chr10	CACTTTTGGGTAAATACATTCTTCATACCAGGA	TATACTGCAAATGCTATCGA	CHP2_PTEN_7
PTEN	chr10	GCAGTATAGAGCGTGCAGATAATGA	CATCACATACATACAAGTCAACAACCC	CHP2_PTEN_8
RNF43	chr17	CCAAACACATCTGGAGCACACT	GCCTGACCCTCAATGACCTCTT	RNF43_1.100174
RNF43	chr17	CCGCTTTTTGTAGTGGTGGT	TGACTTTGACCCCCTAGTGTACT	RNF43_2.1.213215
RNF43	chr17	ACAACCACACTGGCTGTGAA	GCACCCAGCTTGCCAGATT	RNF43_2.1.22754
RNF43	chr17	CGGTGTCAGAACTCCATTCAGAAG	GACAAGAGGCTGCTACCAGAAA	RNF43_2.1.264668
RNF43	chr17	CTCTCCCTACCACACCCACTT	GTGGTTGTGCCTGACTCCTC	RNF43_2.1.277654
RNF43	chr17	CTGGGTGCACAGTTGCATC	CCCTGGCCCAGTTGACG	RNF43_2.1.308947
RNF43	chr17	GAAACCTGGGTTTCCCCTGT	GGGTCCATGGCAGCAGTTC	RNF43_2.1.342797
RNF43	chr17	AAAGTCACTGCTTAGGGAGCT	AGAAAGCTATTGCACAGAACGC	RNF43_2.1.4249
RNF43	chr17	GGGACCAAGGATATGCCACACT	TGCAAAAATCCAGCCTCTCTGC	RNF43_2.1.479773
RNF43	chr17	GGGCACTGTGGGTTAGAGAG	AAAAGCGGTTCCAGTGGCA	RNF43_2.1.483268
RNF43	chr17	GTGACTTGCTGATCAGGAGAAGGT	GTTTCCAGCCATGTCCACTACC	RNF43_2.1.554471
RNF43	chr17	TTTTTGCAAGTTGAACAGACTGCT	CAAGTCACCAGATCCAACTCAGC	RNF43_2.1.716452
RNF43	chr17	CTCCAGATCCACTGCTGTCA	TTCCCCAGAGCTGCACATC	RNF43_2.2.168519
RNF43	chr17	GTAGGCTGATGTCCGTGCAG	GCTTGCCCAGTGCCCCTA	RNF43_2.2.335644
RNF43	chr17	GTGCTGTGAGGTGGATTGGAG	CCCACGACCTGGTCCCTT	RNF43_2.2.348341
RNF43	chr17	GTGATGCCGAGGGCCCAT	CAGGTCGAAGACTCCACCTC	RNF43_2.2.359855
RNF43	chr17	AGGTGGTAGTGGGCATGGC	TGTCTTTCTGAATGCATTCTCTGTAGG	RNF43_2.2.62203
RNF43	chr17	CCTCCTACCTGTGATGTTGAACATG	CCTGATTCCTGGCAATTCCTATGG	RNF43_3.144168

RNF43	chr17	AAGCCACATTCTAGACCTGTCTG	CTCTTTTTCTCCAGGAGCTACGG	RNF43_3.9926
RNF43	chr17	TCCTCCAGACAGATGGCACA	CCCAATCTGAGCCCCATTCCT	RNF43_4.332752
RNF43	chr17	TTCAATCTCCCCAGTCTGGTCAT	AGCTGGCCACCAGGAGGTA	RNF43_4.381754
RNF43	chr17	TACTCCTTCCTTCTCCCTAACCAC	ATGATGTGTGGATCCTAATGACAGT	RNF43_5.252108
RNF43	chr17	AAGCCAGGATGATCACAAAGATGG	CTCAAGGGAACCTCCAGTTAGCTAT	RNF43_5.9080
RNF43	chr17	CCCTGAGAGCTTTATCTTCCTCCATC	GACCTCAGCCCAACCTCTACT	RNF43_6.126819
RNF43	chr17	GTCTGGAGGTCTAGTGTGCT	TGGGCACTTTCCCCCTGTA	RNF43_6.251138
RNF43	chr17	ATCAGCTTCTCAGCGTCATTACC	CTGGATGGAGGAAGATAAAGCTCTCA	RNF43_6.62210
RNF43	chr17	CACAGGACAAAGTAGGGCTAAGTG	AAGCTGATGGAGTTTGTGTACAAGAA	RNF43_6.83735
RNF43	chr17	CCAGCTTGACGATGCTGATGAAT	TGGTACCTCCCTAGAAAAATGGAGAG	RNF43_8.145848
RNF43	chr17	GTGTAGGGCGAAGTGTGAGTC	CCTAACCCAAGTCTGTCTCTCTCTG	RNF43_8.334246
RNF43	chr17	ATTTCCACTTCTCTCAGACCAGTCAT	CCTGTCACTGGCTAGCAAGGTA	RNF43_8.99948
RNF43	chr17	GCAAACACACCTTCCAAAGTGAGATT	TGGACGCACAGGACTGGTAC	RNF43_9.251835
RNF43	chr17	ACAAAAGAAGAAAGACATATTTCAAACAGATG	TTATCAGAGTGATCCCCTTGAAAATGG	RNF43_9.43347
RNF43	chr17	AGCTTTCTGTTCTGCTGATCTTTCA	GTATGTATGGTTGAAGTGCATTGCTG	RNF43_9.83487
SMAD4	chr18	CTCATGTGATCTATGCCCGTCT	AGTCTACTTACCAATTCCAGGTGATACA	CHP2_SMAD4_1
SMAD4	chr18	TGCTACTTCTGAATTGAAATGGTTCA	GATTACCTACCATTACTCTGCAGTGTT	CHP2_SMAD4_2
SMAD4	chr18	ATGGTGAAGGATGAATATGTGCATGA	GCTGGTAGCATTAGACTCAGATGG	CHP2_SMAD4_3
SMAD4	chr18	GTGAAGGACTGTTGCAGATAGCAT	AAGGCCCACATGGGTTAATTTG	CHP2_SMAD4_4
SMAD4	chr18	TTTCTTTAGGGCCTGTTCACAATGA	CTGAGAAGTGACCCCATAATTCCATT	CHP2_SMAD4_5
SMAD4	chr18	GCTCCTGAGTATTGGTGTTCCAT	CCTGTGGACATTGGAGAGTTGA	CHP2_SMAD4_6
SMAD4	chr18	TGTAATTTCTTTTTTCTTCCTAAGGTTGCACATAG	ACTTGGGTAGATCTTATGAACAGCAT	CHP2_SMAD4_7
SMAD4	chr18	AGGTCTTTGATTTGCGTCAGTGT	GCTGGAGCTATTCCACCTACTG	CHP2_SMAD4_8
SMAD4	chr18	GCTGCTGGAATTGGTGTTGATG	AGTACTTCGTCTAGGAGCTGGAG	CHP2_SMAD4_9
STK11	chr19	GAGCTGATGTCGGTGGGTAT	CTCCGAGTCCAGCACCTC	CHP2_STK11_1
STK11	chr19	CTCCCAGGCAGCTGCAA	CCGGTGGTGAGCAGCAG	CHP2_STK11_2
STK11	chr19	CCGGTGGCACCCTCAAA	CTGGTCCGGCAGGTGTC	CHP2_STK11_3
STK11	chr19	AACATCACCACGGGTCTGTAC	GATGAGGCTCCCACCTTTCAG	CHP2_STK11_4
STK11	chr19	GAAGAAACATCCTCCGGCTGAA	ACCGTGAAGTCCTGAGTGTAGA	CHP2_STK11_5
TP53	chr17	TCCACTCACAGTTTCCATAGGTCT	GTTGGAAGTGTCTCATGCTGGAT	CHP2_TP53_1
TP53	chr17	GGCTGTCCCAGAATGCAAGAA	GATGAAGCTCCCAGAATGCCA	CHP2_TP53_2
TP53	chr17	TGCACAGGGCAGGTCTTG	CCGTCTTCCAGTTGCTTTATCTGT	CHP2_TP53_3
TP53	chr17	ACCAGCCCTGTCGTCTCT	GTGCAGCTGTGGGTTGATTC	CHP2_TP53_4
TP53	chr17	CCAGTTGCAAACCAGACCTCA	AGGCCTCTGATTCCTCACTGAT	CHP2_TP53_5
TP53	chr17	GGCTCCTGACCTGGAGTCTT	CTCATCTTGGGCCTGTGTTATCTC	CHP2_TP53_6
TP53	chr17	CGCTTCTTGTCCTGCTTGCT	TTCTCTTTTCCTATCCTGAGTAGTGGT	CHP2_TP53_7
TP53	chr17	GGAAGGGGCTGAGGTCACT	CCCCTCCTCTGTTGCTGC	CHP2_TP53_8
VHL	chr3	CTCCCAGGTCATCTTCTGCAAT	GTACCTCGGTAGCTGTGGATG	CHP2_VHL_1
VHL	chr3	GTGGCTCTTTAACAACCTTTGCT	GTCAGTACCTGGCAGTGTGATA	CHP2_VHL_2
VHL	chr3	GGCAAAGCCTCTTGTTCGTTC	TGACGATGTCCAGTCTCCTGTAAT	CHP2_VHL_3

1

1 **Supplementary table 3: Mutation profile of precursor lesions detected by targeted NGS.**

Grade	Gene	Sample	Variant	VAF [%]	Variant Effect	Transcript
low-grade PanIN	*ARID1A*	**113**	Gln802fs	4.83	frameshift/insertion	NM_006015.5
	CDKN2A	74	Arg58Ter	3.48	nonsense	NM_001195132.1
	GNAS	127	Arg201His	20.17	missense	NM_000516.5
	KRAS	52	Gly12Val	15.38	missense	NM_033360.3
		55	Gly12Val	9.66	missense	NM_033360.3
		56	Gly12Asp	6.13	missense	NM_033360.3
		69	Gln61His	4.08	missense	NM_033360.3
		74	Gly12Asp	12.05	missense	NM_033360.3
		111	Gly12Val	11.19	missense	NM_033360.3
		125	Gly12Val	4.29	missense	NM_033360.3
		127	Gly12Asp	18.24	missense	NM_033360.3
		128	Gly12Asp	3.42	missense	NM_033360.3
		129	Gly12Asp	3.7	missense	NM_033360.3
		113	Gly12Arg	4.18	missense	NM_033360.3
		113	Gly12Val	8.75	missense	NM_033360.3
		43	Gly12Val	12.88	missense	NM_033360.3
	PIK3CA	68	Arg349Ter	3.74	nonsense	NM_006218.3
		127	Phe83fs	4.52	frameshift/deletion	NM_006218.3
	PTEN	128	Asn323fs	22.37	frameshift/deletion	NM_000314.6
high-grade PanIN	*ARID1A*	73	Trp2091Ter	7.46	nonsense	NM_006015.4
	GNAS	96	Arg201His	3.42	missense	NM_000516.5
	KRAS	73	Gly12Asp	8.29	missense	NM_033360.3
		80	Gly12Val	16.59	missense	NM_033360.3
		101	Gly12Asp	13.7	missense	NM_033360.3
		104	Gly12Asp	5.15	missense	NM_033360.3
		114	Gly12Val	18.3	missense	NM_033360.3

Liffers S-T, *et al. Gut* 2022;0:1–13. doi: 10.1136/gutjnl-2021-326550

	TP53	**80**	Arg213Ter	34.5	nonsense	NM_000546.5
		104	Arg196Ter	4.04	nonsense	NM_000546.5
low-grade IPMN gastric	*ARID1A*	21	Asp1850fs	7.45	frameshift/insertion	NM_006015.5
	GNAS	**2**	Arg201His	33.98	missense	NM_000516.5
		7	Arg201Cys	11.34	missense	NM_000516.5
		35	Arg201His	41.52	missense	NM_000516.5
		62	Arg201Cys	15.56	missense	NM_000516.5
		86	Arg201His	30.46	missense	NM_000516.5
		89	Arg201His	5.56	missense	NM_000516.5
		94	Arg201His	21.97	missense	NM_000516.5
		97	Arg201Cys	33.81	missense	NM_000516.5
		99	Arg201His	25.08	missense	NM_000516.5
		110	Arg201Cys	22.37	missense	NM_000516.5
		112	Arg201His	24.32	missense	NM_000516.5
		54	Arg201Cys	23.16	missense	NM_000516.5
	KRAS	21	Gly12Val	18.86	missense	NM_033360.3
		2	Gly12Val	35.1	missense	NM_033360.3
		7	Gly12Val	12.14	missense	NM_033360.3
		35	Gly12Asp	38.57	missense	NM_033360.3
		39	Gly12Val	35	missense	NM_033360.3
		61	Gly12Val	18.03	missense	NM_033360.3
		62	Gly12Val	18.5	missense	NM_033360.3
		64	Gly12Val	9.01	missense	NM_033360.3
		86	Gly12Val	27.81	missense	NM_033360.3
		89	Gln61His	36.49	missense	NM_033360.3
		94	Gly12Asp	21.16	missense	NM_033360.3
		97	Gly12Asp	33.68	missense	NM_033360.3

		99	Gly12Val	26.83	missense	NM_033360.3
		110	Gly12Val	24.76	missense	NM_033360.3
		57	Gly12Arg	6.96	missense	NM_033360.3
	STK11	**61**	Tyr60Ter	61.89	nonsense	NM_000455.4
high-grade IPMN gastric	*ARID1A*	130	Gln2115Ter	30.84	nonsense	NM_006015.5
	GNAS	51	Arg201Cys	31.03	missense	NM_000516.5
	KRAS	51	Gly12Val	33.12	missense	NM_033360.3
		65	Gly12Asp	16.51	missense	NM_033360.3
		66	Gly12Asp	33.64	missense	NM_033360.3
		130	Gly12Asp	22.65	missense	NM_033360.3
		131	Gly12Asp	23.98	missense	NM_033360.3
	TP53	65	Arg248Trp	17.99	missense	NM_000546.5
low-grade IPMN intestinal	*GNAS*	**92**	Arg201Cys	80.59	missense	NM_000516.5
	KRAS	**92**	Gly12Arg	37.95	missense	NM_033360.3
high-grade IPMN intestinal	*GNAS*	**33**	Arg201Cys	45.42	missense	NM_000516.5
		88	Arg201Cys	36.81	missense	NM_000516.5
		90	Arg201His	35.51	missense	NM_000516.5
		17	Gln227Lys	42.36	missense	NM_000516.5
	KRAS	**88**	Gly12Ser	33.81	missense	NM_033360.3
	TP53	**88**	Met237Thr	49.16	missense	NM_000546.5

PanIN/IPMN in cases without associated/concomitant PDAC are indicated in bold.

Supplementary table 4. Distribution of CNV-positive and negative precursor lesions according to the degree of dysplasia.

		CNV pos	*CNV neg*
PanIN	**Low-grade**	16 (57%)	12 (43%)
	High-grade	6 (75%)	2 (25%)
Gastric IPMN	**Low-grade**	22 (76%)	7 (24%)
	High-grade	6 (75%)	3 (25%)
Intestinal IPMN	**Low-grade**	8 (100%)	0
	High-grade	13 (100%)	0

Percentages refers to the total number of cases in each group

Liffers S-T, *et al. Gut* 2022;0:1–13. doi: 10.1136/gutjnl-2021-326550

Supplementary table 5: Overview of log2 copy number ratios sorted in ascending order. Values of genomic alterations were detected by low-coverage sequencing (n=28) and DNA methylation data (n=67), respectively.

Genomic location	affected samples (n)	PanIN (n=36)	gIPMN (n=38)				iIPMN (n=21)				
deleted regions											
chr01:010875000-013052998	3						-0.26	-0.25	-0.24		
chr01:015375000-016825000	3						-0.26	-0.25	-0.20		
chr06:074175000-074375000	5		-0.29	-0.22	-0.21		-0.44	-0.24			
chr06:133664400-143100000	5		-0.37	-0.28			-0.50	-0.45	-0.28		
chr06:143620678-151100000	7	-0.21	-0.37	-0.34	-0.29		-0.50	-0.45	-0.28		
chr09:005958053-023802212	5	-0.23	-0.51	-0.36	-0.36	-0.24					
chr10:071075000-120925000	6		-0.45	-0.22			-0.36	-0.30	-0.28	-0.22	
chr10:120925000-125869472	5		-0.30				-0.41	-0.30	-0.28	-0.22	
chr11:057325000-058807232	4						-0.44	-0.40	-0.27	-0.22	
chr11:058807232-069089801	5						-0.44	-0.40	-0.28	-0.27	-0.22
chr11:096437584-114325000	5						-0.51	-0.44	-0.40	-0.28	-0.22
chr11:114325000-134898258	4						-0.51	-0.44	-0.28	-0.22	
chr17:006225000-009675000	4		-0.50	-0.32	-0.21		-0.23				
chr17:009675000-012500000	5		-0.53	-0.31	-0.21		-0.50	-0.22			
chr17:015792977-021566608	6		-0.58	-0.23	-0.21		-0.49	-0.32	-0.25		
amplified regions											
chr01:035225000-037325000	3						0.25	0.33	0.61		

chr03:176225000-188875000	6	0.26	0.21	0.23	0.28	0.38	0.44				
chr05:028950000-044925000	5				0.24	0.25	0.31	0.74	0.79		
chr06:024125000-033575000	5		0.25	0.36	0.31	0.35	0.56				
chr06:033575000-042725000	4		0.42		0.32	0.34	0.56				
chr07:000282484-007150000	5				0.22	0.30	0.32	0.36	0.41		
chr07:054725000-055775000	5				0.21	0.26	0.27	0.33	0.36		
chr07:061967157-074715724	4				0.27	0.32	0.36	0.37			
chr07:112425436-130154523	5				0.21	0.33	0.36	0.38	0.57		
chr07:139404377-142048195	5				0.22	0.32	0.36	0.33	0.58		
chr07:143397897-154270634	5				0.24	0.33	0.34	0.38	0.59		
chr08:086726451-089550000	3				0.21	0.31	0.36				
chr08:127450000-129175000	7	0.42			0.21	0.27	0.29	0.36	0.56	0.66	
chr09:001992685-035698318	3				0.24	0.32	0.61				
chr09:070835468-092343416	4				0.22	0.26	0.32	0.60			
chr09:096718222-097575000	4				0.21	0.24	0.32	0.60			
chr09:097775000-114750000	4				0.22	0.26	0.32	0.61			
chr09:124994207-133073060	3				0.22	0.32	0.61				
chr12:006475000-007169938	8	0.21			0.24	0.33	0.56	0.71	0.88	0.89	1.70
chr12:024993545-028938805	4	0.21			0.21	0.38	1.32				
chr14:020700000-022050000	3				0.21	0.28	0.33				
chr14:022800000-050175000	3				0.22	0.25	0.35				
chr14:097258910-107289540	3				0.25	0.22	0.36				
chr17:061125000-062410760	3				0.32	0.32	0.74				
chr17:062775000-063525000	3				0.32	0.34	0.49				
chr17:068117898-077546461	3				0.21	0.32	0.38				

chr20:008050000-016400000	7	0.24	0.22	0.23	0.24	0.26	0.33	0.84
chr20:016625000-021300000	7	0.25	0.22	0.23	0.25	0.28	0.33	0.82
chr20:030025000-034897085	7	0.26	0.22	0.22	0.23	0.31	0.33	0.81
chr20:036958189-042991501	6	0.26	0.22	0.23	0.35	0.33	0.79	
chr20:052650000-061091437	7	0.23	0.22	0.22	0.23	0.33	0.35	0.80
chr21:032825000-034475000	3		0.21	0.26	0.30			

1

Supplementary table 6a: Differentially methylated probes in low- and high-grade pancreatic cancer precursors.

	Low-grade samples	High-grade samples	DMPs (hypermethylated in high-grade)	DMP associated genes (hypermethylated in high-grade)
iPMN	**8**	**12**	**0**	**0**
gIPMN	**24**	**8**	**0**	**0**
PanIN	**20**	**7**	**86 (62)**	**59 (45)**

iIPMN: intestinal IPMN; gIPMN: gastric IPMN

Supplementary table 6b: Genes associated with significantly hyper- or hypomethylated CpG in PanIN high-grade

Genes associated with at least one significantly hypermethylated CpG	Genes associated with at least one significantly hypomethylated
TAC1	GLRX
AKAP13	BCL11B
POLR1D	ITFG3
GLYATL3	IFT140
HOXA5	CTNNA3
ZIC2	RUNX1
SIM2	SPARCL1
ARID4B	PHLDB1
MON2	NINJ2
CNKSR3	METTL9
SP8	SLC51A
ADD2	EMID2
HOXB1	CACNA1A
ST3GAL6	C19orf35
FBN2	
ZIK1	
LRP1B	
NTRK3	
GLI3	
NTM	
RASGRF1	
FAM46C	

NXPH1	
LBX2	
LOC101929710	
ADRA1A	
GRIK3	
PARP8	
KIAA1026	
SLC6A15	
IRX1	
GRIA4	
TLE4	
DGKI	
PACSIN2	
DOK5	
ZIC4	
MYLK	
DAPK1	
C5orf66-AS1	
AMER3	
CPEB1	
ST6GAL2	
INA	
AP2A2	

1

Supplementary Figure Legends

Supplementary figure 1: Overview of lesions and methods

55 PanINs, 46 gastric IPMNs (gIPMN) and 21 intestinal IPMNs (iIPMN) were subjected to 4 main analyses: targeted next generation sequencing (n=52), low-coverage whole-genome sequencing (n=28), genome-wide DNA methylation analysis (n=79) and transcriptome analysis (n=34). Each circle of the diagram represents one of the mentioned methods and includes the number of samples used for related analyses. The samples that could not be placed into the diagram were shown at the right bottom corner of the figure. Pancreatobiliary and mixed-type IPMNs were excluded from further analyses due to small sample size.

Supplementary figure 2: Allele frequency of *KRAS* and *GNAS* mutations in different precursor lesions.

(A) VAF of *KRAS* mutations; (B) VAF of *GNAS* mutations; (C) scatterplot of the VAF of the *KRAS* mutations (G12) against the VAF of *GNAS* mutations (R201) detected in low- and high-grade gastric IPMN. The grey lines represent the 95% confidence interval of the Pearson's correlation coefficient r (n=9). (Kruskal-Wallis-test * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$).

Supplementary figure 3: Morphology of lesions with different genetic status according to targeted NGS

Representative HE images of low- and high-grade PanINs, gastric IPMNs and intestinal IPMNs with variable mutation profiles detected by targeted NGS are shown. No specific morphology was identified related to the mutation status between the samples in the same diagnostic group. In particular, gastric lesions with GNAS mutations (case 127, 96, 99 and 112) did not show relevant intestinal differentiation; only in case 112, focal (<5% of the cells) expression of MUC2 and CDX2 was observed (not shown). Scale bars represent 200 µm. Detailed mutation profile of the samples is provided in Suppl. Table 3.

Supplementary figure 4: Quality control of DNA methylation data derived from normal pancreas cell preparations.

(A) Multi-dimensional scaling of the 1000 most variable probes. (B) Hierarchical clustering of probes for known acinar and ductal marker genes.

Supplementary figure 5: Proliferation activity, expression of TFF3 and of MUCL3 protein in PanIN, IPMN and PDAC.

Representative images (A) and related graphs (B-D) of IHC staining performed in whole tissue sections in 31 PanIN, 28 gastric IPMN, 20 intestinal IPMN and 24 PDAC. Intestinal IPMNs and PDACs showed higher proliferation rates with Ki67 staining than PanIN and gastric IPMN. TFF3 was strongly expressed in intestinal IPMN. Gastric IPMN revealed higher expression of MUCL3 compared to PanIN. Scale bars represent 100 µm. (IRS: immunreactivity score) (*p<0.05).

Supplementary figure 6: Hierarchical clustering of DNA methylation data based on published marker genes for distinct normal pancreas cell populations. The mean methylation beta-value for all gene associated probes is displayed, respectively.

Supplementary figure 7: Differentially activated gene sets. Displayed pathways were detected by pairwise comparison between the indicated lesions. Odds ratios below 0.8 indicate the activation in the first listed lesion whereas 1.1 is associated with the second group. The analyzed gene sets based on the KEGG pathway (A) and hallmarks (B) from the MSigDB.

Liffers S-T, *et al. Gut* 2022;0:1–13. doi: 10.1136/gutjnl-2021-326550