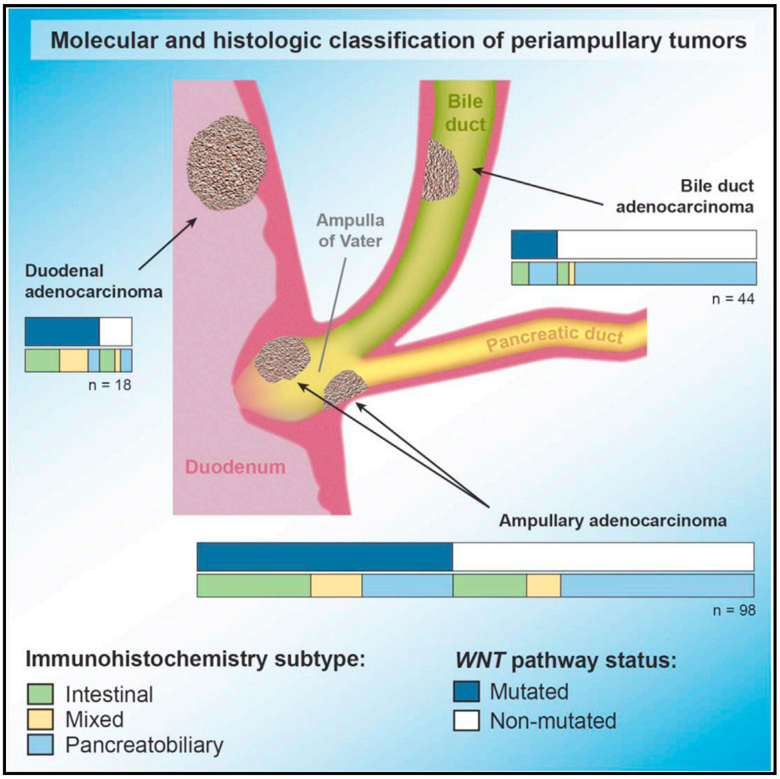
Cell Reports

Article

Ampullary Cancers Harbor *ELF3* Tumor Suppressor Gene Mutations and Exhibit Frequent *WNT* Dysregulation

Graphical Abstract



Highlights

- Three periampullary tumor types share a common molecular blueprint
- Frequent *WNT* pathway disruption can be a potential therapeutic target
- *ELF3* is a frequently mutated tumor suppressor gene of periampullary tumors

Authors

Marie-Claude Gingras, Kyle R. Covington, David K. Chang, ..., Andrew V. Biankin, David A. Wheeler, Richard A. Gibbs

Correspondence

mgingras@bcm.edu (M.-C.G.),
wheeler@bcm.edu (D.A.W.)

In Brief

Gingras et al. compare the genomic profiles of ampullary, distal bile duct, and duodenal adenocarcinomas. They find disruption of the *WNT* pathway, which could be used as a tumor subclassification for targeted therapy, a high frequency of *ELF3* inactivating mutations, and microsatellite instability.

Gingras et al., 2016, Cell Reports *14*, 907–919
February 2, 2016
http://dx.doi.org/10.1016/j.celrep.2015.12.005

Ampullary Cancers Harbor *ELF3* Tumor Suppressor Gene Mutations and Exhibit Frequent *WNT* Dysregulation

Marie-Claude Gingras,[1,2,3,*] Kyle R. Covington,[1] David K. Chang,[4,5,6,7] Lawrence A. Donehower,[1,8] Anthony J. Gill,[6,9,10] Michael M. Ittmann,[3,11,12] Chad J. Creighton,[1,3] Amber L. Johns,[6] Eve Shinbrot,[1] Ninad Dewal,[1] William E. Fisher,[2,3,13] Australian Pancreatic Cancer Genome Initiative, Christian Pilarsky,[14] Robert Grützmann,[15] Michael J. Overman,[16] Nigel B. Jamieson,[4,5,17] George Van Buren II,[2,3,13] Jennifer Drummond,[1] Kimberly Walker,[1] Oliver A. Hampton,[1] Liu Xi,[1] Donna M. Muzny,[1] Harsha Doddapaneni,[1] Sandra L. Lee,[1] Michelle Bellair,[1] Jianhong Hu,[1] Yi Han,[1] Huyen H. Dinh,[1] Mike Dahdouli,[1] Jaswinder S. Samra,[10,18] Peter Bailey,[4] Nicola Waddell,[19,20] John V. Pearson,[19,20] Ivon Harliwong,[19] Huamin Wang,[21] Daniela Aust,[22] Karin A. Oien,[4,23] Ralph H. Hruban,[24] Sally E. Hodges,[2,13] Amy McElhany,[2,13] Charupong Saengboonmee,[1,25] Fraser R. Duthie,[4,23] Sean M. Grimmond,[4,19] Andrew V. Biankin,[4,5,6,7] David A. Wheeler,[1,3,*] and Richard A. Gibbs[1]

[1]Department of Molecular and Human Genetics, Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX 77030, USA
[2]Michael DeBakey Department of Surgery, Baylor College of Medicine, Houston, TX 77030, USA
[3]Dan L. Duncan Cancer Center, Baylor College of Medicine, Houston, TX 77030, USA
[4]Wolfson Wohl Cancer Research Centre, Institute for Cancer Sciences, University of Glasgow, Garscube Estate, Bearsden, Glasgow G61 1BD, UK
[5]West of Scotland Pancreatic Unit, Glasgow Royal Infirmary, Glasgow G31 2ER, UK
[6]The Kinghorn Cancer Centre and the Cancer Research Program Garvan Institute of Medical Research, Darlinghurst, Sydney, NSW 2010, Australia
[7]South Western Sydney Clinical School, Faculty of Medicine, University of New South Wales, Liverpool, NSW 2170, Australia
[8]Department of Molecular Virology and Microbiology, Baylor College of Medicine, Houston, TX 77030, USA
[9]Department of Anatomical Pathology, Royal North Shore Hospital, St Leonards, Sydney, NSW 2065, Australia
[10]Sydney Medical School, University of Sydney, Sydney, NSW 2006, Australia
[11]Department of Pathology and Immunology, Baylor College of Medicine, Houston, TX 77030, USA
[12]Michael E. DeBakey Department of Veterans Affairs Medical Center, Houston, TX 77030, USA
[13]The Elkins Pancreas Center at Baylor College of Medicine, Houston, TX 77030, USA
[14]Department of Surgery, TU Dresden, 01307 Dresden, Germany
[15]Department of Surgery, Universitätsklinikum Erlangen, 91054 Erlangen, Germany
[16]Department of Gastrointestinal Medical Oncology, University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA
[17]Academic Unit of Surgery, Institute of Cancer Sciences, Glasgow Royal Infirmary, Level 2, New Lister Building, University of Glasgow, Glasgow G31 2ER, UK
[18]Department of Surgery, Royal North Shore Hospital, St Leonards, Sydney, NSW 2065, Australia
[19]Queensland Centre for Medical Genomics, Institute for Molecular Bioscience, The University of Queensland, St Lucia, Brisbane, QLD 4072, Australia
[20]QIMR Berghofer Medical Research Institute, Herston, Brisbane, QLD 4006, Australia
[21]Department of Pathology, University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA
[22]Department of Pathology, TU Dresden, 01307 Dresden, Germany
[23]Department of Pathology, Southern General Hospital, Greater Glasgow and Clyde NHS, Glasgow G51 4TF, UK
[24]Department of Pathology, The Sol Goldman Pancreatic Cancer Research Center, the Johns Hopkins University School of Medicine, Baltimore, MD 21231, USA
[25]Department of Biochemistry and Liver Fluke and Cholangiocarcinoma Research Center, Faculty of Medicine, Khon Kaen University, Khon Kaen 40002, Thailand
*Correspondence: mgingras@bcm.edu (M.-C.G.), wheeler@bcm.edu (D.A.W.)
http://dx.doi.org/10.1016/j.celrep.2015.12.005

SUMMARY

The ampulla of Vater is a complex cellular environment from which adenocarcinomas arise to form a group of histopathologically heterogenous tumors. To evaluate the molecular features of these tumors, 98 ampullary adenocarcinomas were evaluated and compared to 44 distal bile duct and 18 duodenal adenocarcinomas. Genomic analyses revealed mutations in the *WNT* signaling pathway among half of the patients and in all three adenocarcinomas irrespective of their origin and histological morphology. These tumors were characterized by a high frequency of inactivating mutations of *ELF3,* a high rate of microsatellite instability, and common focal deletions and amplifications, suggesting common attributes in the molecular pathogenesis are at play in these tumors. The high frequency of *WNT* pathway activating mutation, coupled with small-molecule inhibitors of β-catenin in clinical trials, suggests future

treatment decisions for these patients may be guided by genomic analysis.

INTRODUCTION

Though the pancreas, bile duct, and intestinal duodenum share common embryologic origins in the ventral endoderm, the adenocarcinomas arising in this region presumably originate from different epithelial cellular constituents present at the site (Zaret and Grompe, 2008). These tumors have been described in many different ways: intra-ampullary, periampullary, intra-ampullary papillary-tubular neoplasm, ampullary/ductal, periampullary-duodenal, and ampullary/not otherwise specified. The tumors clearly separated from the ampulla of Vater and localized in the bile duct, duodenum, or pancreatic duct have been identified as distal cholangiocarcinomas or distal bile duct (CAC), duodenal (DUOAC), or pancreatic ductal (PDAC) adenocarcinomas.

As recommended in the AJCC seventh edition 2009 staging system (Edge et al., 2009), the current subtype classification of ampullary adenocarcinoma (AMPAC) is based on the anatomical location from which the tumor is thought to arise (Edge et al., 2009), sometimes supplemented by histopathology and expression of differential markers (Adsay et al., 2012; Chang et al., 2013; Ehehalt et al., 2011; Morini et al., 2013). This classification is subjective and prone to inter-observer variability and can significantly impact treatment selection and therapeutic development (Amptoulach et al., 2011; Heinrich and Clavien, 2010; Romiti et al., 2012; Westgaard et al., 2013). Current treatment approaches do not distinguish patients based on subtypes, yet tumors may arise from at least the three epithelia that converge at that site, and some may arise from the ampulla itself, where little is known of the specialized epithelium that may be present. Malignancies that arise from different cellular origins often have vastly differing sensitivities to therapeutics. Post hoc analyses of clinical trials using histopathological criteria have not discerned such a difference and likely represent the inaccuracy of such a classifier. However, as most therapeutic development is focused on agents that target specific molecular mechanisms, a molecular characterization that would allow selection of patients for specific therapies would facilitate therapeutic development with the aim of improving outcomes and alleviate the impact of an inaccurate subjective classification.

For this study, we have assembled a large cohort of AMPAC with nearby DUOAC and CAC for comparison. We show that tumors from the duodenum, ampulla of Vater, and distal bile duct exhibit a common spectrum of features irrespective of their morphology, marker expression, and cellular origin. Here, we use the term ‘‘periampullary tumors’’ in this study to refer to the three tumor types of AMPAC, DUOAC, and CAC collectively, as defined by the AJCC seventh edition 2009 staging system (Edge et al., 2009), excluding cases that clearly arise from the pancreas (pancreatic adenocarcinoma [PDAC]).

RESULTS

In order to develop a molecular taxonomy for periampullary cancers and define subtypes with clinical relevance, we performed exome sequencing and copy-number analysis of 160 cancers arising in the periampullary region, 62 of these clearly arising from either the bile duct (n = 44) or the duodenum (n = 18) and 98 for which the epithelium of origin could not be clearly defined morphologically (AMPAC). Mutations were validated by deep and ultra-deep sequencing on a limited target region consisting of 71 recurrently mutated genes. RNA sequencing (RNA-seq) was performed on 30 patients: a 28-patient subset of the 98 ampullary tumors and a two-patient subset of the 18 duodenal tumors.

Clinical Characteristics and Subtyping

The clinical characteristics of our patient cohort are described in Table S1A. In this study, the anatomical primary site of origin of all tumors was defined using the AJCC seventh edition 2009 staging system (Edge et al., 2009). In addition, the tumors were also classified independently by cellular morphology and immunohistochemistry (IHC) staining (see Experimental Procedures) into intestinal, pancreatobiliary, or mixed subtypes (Table S1B). Since treatment may be determined based on subtypes defined by the combination of morphology and IHC even if these measures are somewhat subjective, it was an important objective of our study to assess the reliability and meaning of these subtypes. Subtyping according to IHC, the AMPAC tumors were 51% pancreatobiliary and 34% intestinal, with the remainder mixed. CAC was dominated by the pancreatobiliary subtype, 86% as expected; however, 11% of CAC exhibited an intestinal phenotype. In DUOAC, the intestinal subtype was 44%, with 22% pancreatobiliary and the remainder mixed.

By histological morphology, a smaller proportion of each tumor type was classified as pancreatobiliary (AMPAC, 37%; CAC, 77%; and DUOAC, 6%). The two methods of classification yielded concordant subtypes only 62% of the time for AMPAC tumors, 77% of the time in CAC, and 53% of the time in DUOAC. Although the two methods often disagreed, all three tumor types included in their numbers concordant cases of all three subtypes. Thus, tumors originating in each organ site in the periampullary region may be classified as any of the three subtypes, though this classification system is rarely applied to DUOAC or CAC tumors. These tumors were analyzed by genomic methods to further characterize their molecular properties.

Mutation Analysis

Exomes were sequenced to an average of 120-fold coverage resulting in 28,795 mutations across 152 patients. Eight additional patients were sequenced with targeted custom sequencing and were included in the study (see Supplemental Experimental Procedures, ‘‘Sequencing design and Mutation analysis,’’ and Tables S1C–S1E). Microsatellite instable phenotypes were observed in 12 patients representing each organ cohort (Figure 1), accounting for 18,572 of the whole-exome sequencing (WES) discovery set. Using a method we developed based on the enzyme slippage of the homopolymer region (E.S., unpublished data), we identified two other patients among the targeted sequencing set (Figure S1A).

Excluding microsatellite instable (MSI) tumors and correcting for tumor purity, the median mutation rate did not vary

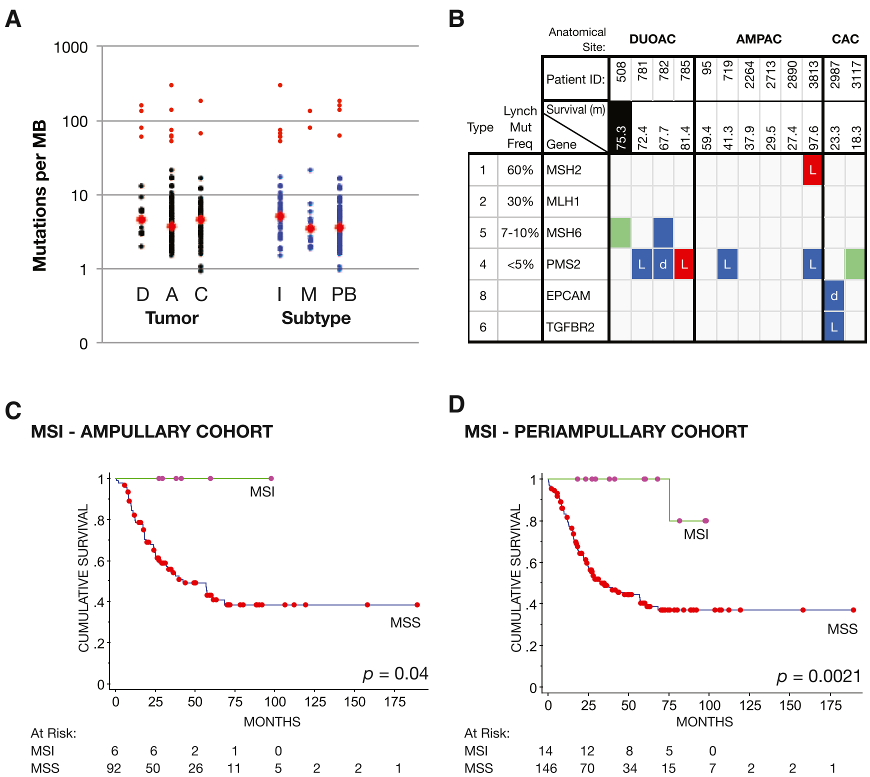


Figure 1. Mutation Frequencies and MSI Characteristics

(A) Mutation frequencies for all patients by anatomical site (D, DUOAC; A, AMPAC; C, CAC) and subtype (I, intestinal; M, mixed; PB, pancreatobiliary). Black dots, microsatellite stable (MSS); red dots, microsatellite instable (MSI).

(B) Germline mutations in MMR gene associated with Lynch syndrome were detected in 66% of the MSI samples. Survival (m) is in months; black tile, patient died of disease; white tiles, patient alive; Lynch Mutation Freq, frequency each gene is observed in Lynch syndrome patients; blue tiles, missense mutations; green, frameshift mutations; red, nonsense mutations; "L" = known Lynch syndrome mutation; "d" deleterious mutation by PolyPhen2.

(C) Kaplan-Meier plot for survival based on MSI status in AMPAC (log rank p = 0.04, n = 96).

(D) Kaplan-Meier plot for survival based on MSI status in all periampullary tumors (p = 0.0028, n = 160).

See also Figures S1A–S1C and Tables S1A–S1E.

significantly across the AMPAC, CAC, and DUOAC (3.8, 4.6, and 4.7 per Mb, respectively) but was clearly distinct from the MSI mutation rate (68, 127, and 108 per Mb, respectively) (Figures 1A, S1B, and S1C). Two-thirds of the hypermutated WES samples had germline mutation in genes associated with Lynch syndrome. Interestingly, *PMS2*, a gene that accounts for less than 5% of Lynch syndrome patients overall (Thompson et al., 2004) (OMIM #600259), was mutated in one half of our MSI patients (Figure 1B). Although MSI was more common in DUOAC than CAC patients, every morphologic category harbored at least one *PMS2* germline mutation in this study. Leaving aside germline contribution, the overall frequency of MSI in AMPAC was 3%. MSI appeared to confer a survival advantage in AMPAC, as it does in other gastrointestinal (GI) cancers, as all six AMPAC patients were alive ranging from 2 to 8 years after diagnosis (p = 0.04 with a lack of negative event) (Figure 1C). Taking all three anatomical sites into consideration, MSI have better survival, $p < 0.0021$ (Figure 1D).

Non-negative matrix factorization was used to evaluate the mutation signatures associated with periampullary tumors. We identified five prominent signatures, out of 21 observed (Figures 2A, S2A, and S2B; Table S2). The most common signature was C > T at CpG islands (#6). Indeed, this signature is most common across all tumor types. A few CAC and AMPAC tumors had a strong T > G, > C signature (#7) associated with the digestive track tumors and consistent with DNA damage and exposure to arsenic (Martinez et al., 2013). A C > G signature (#4) characteristic of DNA damage by *APOBEC* enzymes was also present in a few patients (Roberts et al., 2013).

We observed signature #1 at greater than 20% of the total signature in 9.6% of our entire tumor set (6% AMPAC and 21% CAC). Signature #1 is characterized by AC, AT > AN and is enriched in non-transcribed regions of the genome in samples from several cancer types (PDAC, medulloblastoma, breast tumor, AML, and CLL). However, signature #1 was also observed in the coding region of 18 out of 486 hepatocellular carcinoma (4%) and 31 out of 450 colorectal carcinoma (CRC) (7%) (Lawrence et al., 2013; Totoki et al., 2014; K.R.C., unpublished data). Whereas none of the known signatures have yet been associated with a difference in outcome, signature #1 was associated with poor outcomes in our study set (multivariate Cox proportional hazards p = 0.02) (Figure 2B).

The analysis of the periampullary tumors, excluding MSI patients, revealed 19 genes mutated significantly above background using MutSig-CV (Lawrence et al., 2013) (Figure 3A; Table S3A). Considering the ratio of inactivating to missense mutations, an additional three genes were brought in to the significantly mutated gene list (Table S3B) including *PBRM1*, *RECQL4*, and *KDM6A*. Gene expression data confirmed that the variants harboring missense mutation in the driver genes were expressed between 85% and 88% of the time (Table S3C).

Most interestingly, *ELF3* a transcriptional regulator of *TGFBR2* was mutated in 10.6% of the periampullary tumors with predominantly inactivating frameshift or nonsense mutations (Figure 3B). This mutation frequency is three times higher than in any other cancer (Table S3D) (Cerami et al., 2012; Gao et al., 2013; Lawrence et al., 2014) (http://www.cBioPortal.org). In agreement with our finding, *ELF3* mutations were found in 9.5% of extrahepathic CAC in a recent study of 74 samples with four inactivating mutations out of seven (Nakamura et al., 2015). *ELF3* mutation occurred 71% of the time with *WNT* pathway mutations in all three periampullary groups (Figure S3A). (chi-square test, p = 0.02).

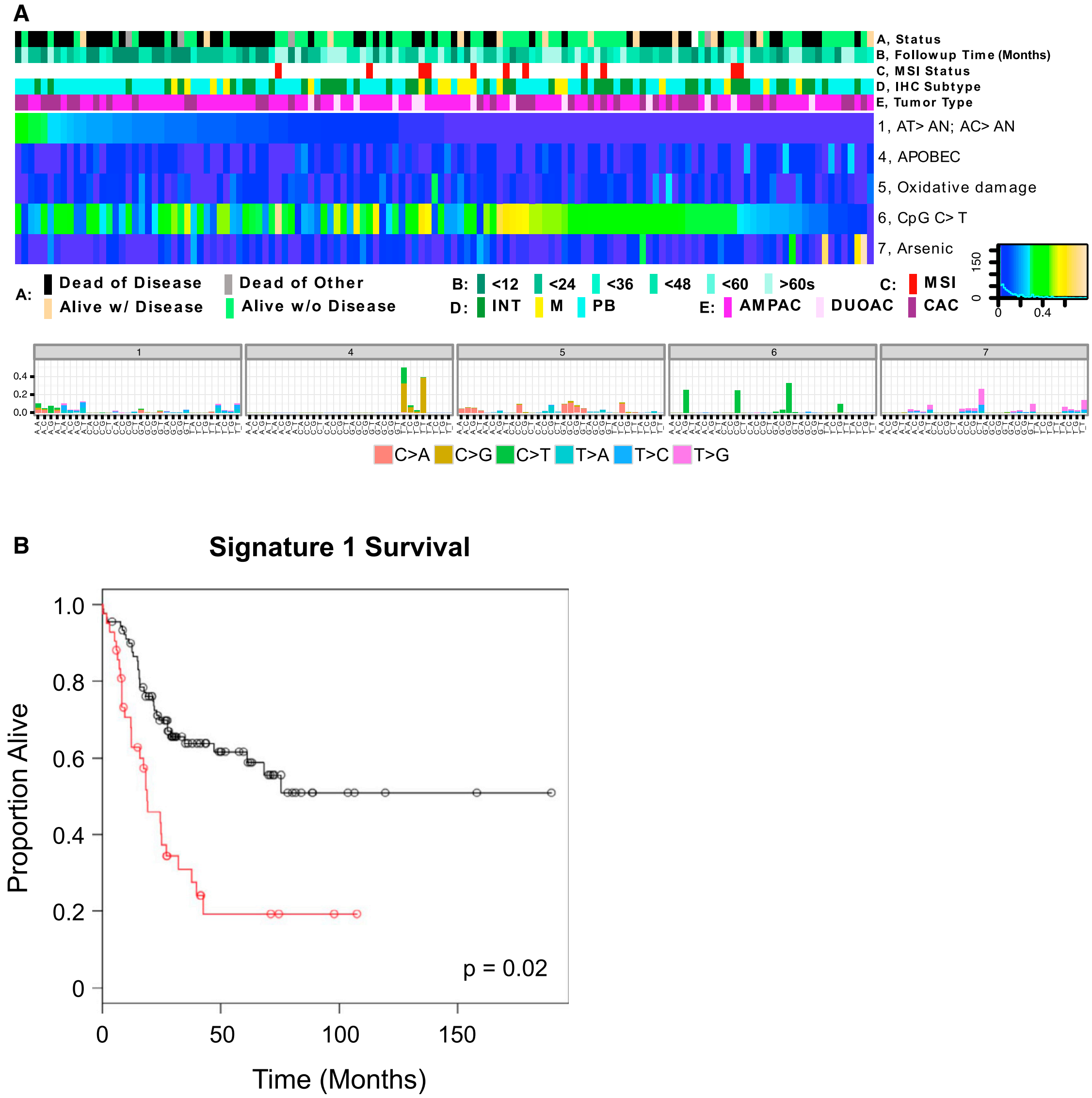


Figure 2. Mutation Signature in Periampullary Tumors

(A) Heatmap of five dominant mutation signatures from NMF analysis of mutation spectrum for each subject. Intensity indicates the proportion of mutations for that subject attributed to the indicated signature. Subjects are sorted first by signature 1, then signature 6 from the highest to the lowest value. Only signatures with high penetrance are shown.

(B) Kaplan-Meier curve of survival in this cohort stratified by signature 1 levels (high, red line: signature 1 component >10% of all mutations; low, black line: otherwise, multivariate Cox proportional hazards p = 0.001).

See also Figures S2A and S2B and Table S2.

Considering the 44 CAC alone, four genes were significantly mutated in this cancer: *TP53*, *KRAS*, *SMAD4*, and *CDKN2A* with the highest mutation incidence in *TP53*. Whereas intrahepatic CAC tumors frequently harbor *BAP1*, *IDH1*, and *IDH2* (Nakamura et al., 2015), these were absent with the exception of a single *IDH1* hotspot mutation in the periampullary CAC. This is in agreement with Nakamura et al. (2015), where no IDH1 mutations could be detected among 74 extrahepatic tumors (compared to a 5% mutation rate in intrahepatic tumor) and a less than 3% BAP1 mutation rate was found in extrahepatic tumors (compared to a 12.4% mutation rate in intrahepatic tumor).

Alteration of Key Signaling Pathways

The significantly mutated genes defined five pathways in periampullary tumors: *TP53*/cell division, *RAS/PI3K*, *WNT*, *TGF-β*, and chromatin remodeling pathways. We combined the point mutations and copy-number alterations (CNA) changes at the gene level within these five pathways to assess the impact of these pathways among the three anatomical sites (Figures

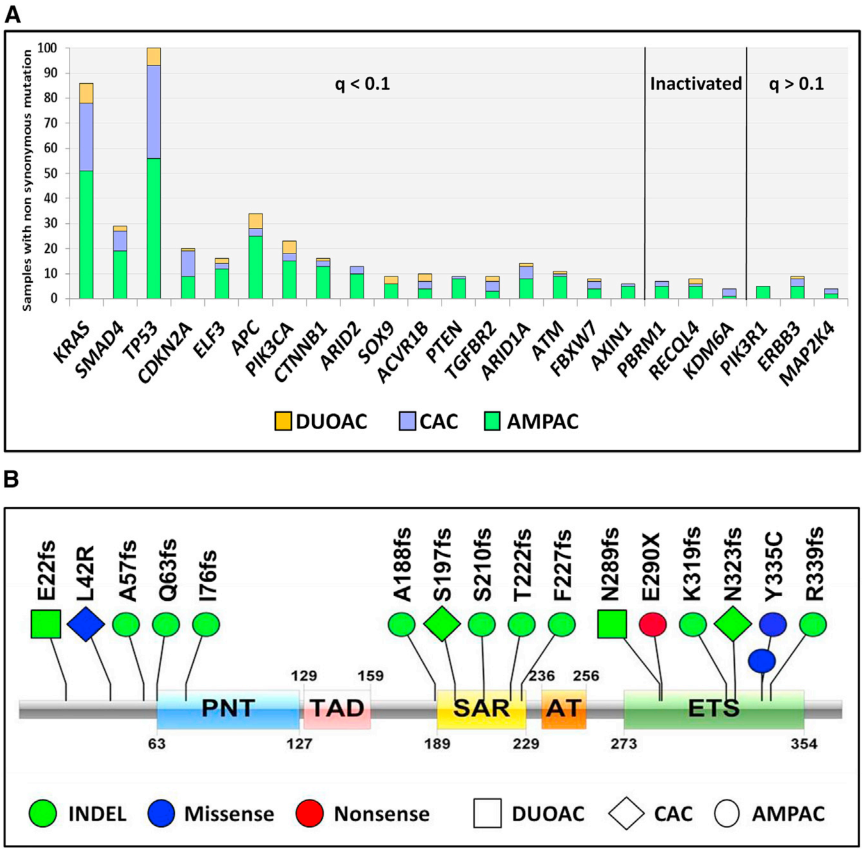


Figure 3. Significantly Mutated Genes in Non-MSI Periampullary Tumors

(A) Significantly mutated genes are displayed by FDR value (MutSigCV). Genes with FDR < 0.1 are located in the left panel, genes with FDR > 0.1 but significantly inactivated are in the middle panel, and genes slightly under the significant threshold of the significantly mutated gene (SMG) list are in the right panel. The amount of samples for each tumor type is stacked.

(B) *ELF3* inactivating mutations were distributed along the entire gene characteristic of a tumor suppressor ($q < 1.6 \times 10^{-11}$). All the mutations found in the study are represented in the figure, each mutation being found in one patient.

See also Tables S3A and 3B–S3D.

4A and 4B). The similarities and differences in gene mutations per tumor types and subtypes are illustrated in Figures S3A and S3B.

The *WNT* pathway was mutated in 46% of patients overall but was clearly differentially mutated across the three tumor types, being more frequently mutated in DUOAC (72%) than in AMPAC (49%) or CAC (30%) (chi-square $p < 0.05$) (Tables S4A and S4B). This predominance of *WNT* pathway mutation in DUOAC was due mainly to more frequent mutations of *APC* and *SOX9*. Whereas the *TP53*, *RAS*, *TGF-β* signaling and chromatin remodeling pathways are deregulated in many tumor types, the *WNT* pathway deregulation is reported only in gastrointestinal tumors (Biankin et al., 2012; Cancer Genome Atlas, 2012). We reasoned that grouping the patients by our histological classification might enrich *WNT* mutation in the intestinal subtype relative to the pancreatobiliary subtype. As expected, the intestinal subtype had 67% *WNT* pathway alterations compared to pancreatobiliary with 30% *WNT* alterations, very close to the *WNT* frequency based on anatomical site (Figures 4B, S3A, and S3B; Tables S4A and S4B). Although we observe a gradient of *WNT* pathway disruption in tumors as their anatomical site moves away from the GI tract, *WNT* mutation is still frequent in CAC, or "pancreatobiliary" subtype tumors.

TGFBR2 was also more frequently mutated in DUOAC than AMPAC and CAC, but this may have been secondary to MSI, which was in higher proportion in DUOAC. *TGFBR2* harbors an A homopolymer run of eight bases that is a frequent target of mutation in MSI patients, and 5 of the 12 *TGFBR2* mutations were at this site. Interestingly, *SMAD4,* a gene frequently mutated in PDAC, was the most commonly mutated gene of the *TGF-β* pathway in AMPAC and CAC, the tissue sites in closest proximity to the pancreas.

Mutant *KRAS* was the major *RAS* signaling oncogene in all three tumor types. Overall, the RTK/RAS/PI3K pathway was activated in all periampullary patients at a statistically similar rate ranging from 84% to 94% among the three tumor types (Tables S4A and S4B).

Alterations in the SWI/SNF chromatin remodeling pathway were observed most frequently in *ARID1A* and *ARID2.* Overall, mutations in the SWI/SNF complex were equally frequent in the three tumor types.

Pathway Mutation Correlates with Disease Outcome

Multivariate analyses on the periampullary tumors as a group showed mutations in the *TGF-β* pathway are associated with better overall survival (multivariate Cox proportional hazard $p = 0.0059$, HR = 0.42) independent of stage, gender, subtype, and MSI status (multivariate Cox proportional hazard $p = 0.029$). Mutations in the *PI3K* pathway were also associated with better overall survival (multivariate-Cox proportional hazards $p = 0.036$, HR = 0.43) (Figure S3C). Mutations in *TP53*, *KRAS*, *WNT*, and chromatin remodeling pathways showed no significant association with outcomes in multivariate modeling. Interestingly, *TGF-β* pathway mutations were also negatively associated with mutation signature 1 (multivariate ANOVA $p = 0.02$), possibly explaining the association with outcomes. However, the contribution of signature 1 to outcomes was still significant when considering *TGF-β* pathway mutations in the model, indicating that these two effects are not entirely redundant.

RNA Expression

RNA expression was analyzed in 28 AMPAC and 2 DUOAC. Due to the high frequency of mutation in *WNT* and the current development of therapeutic agents targeting β-catenin, we evaluated

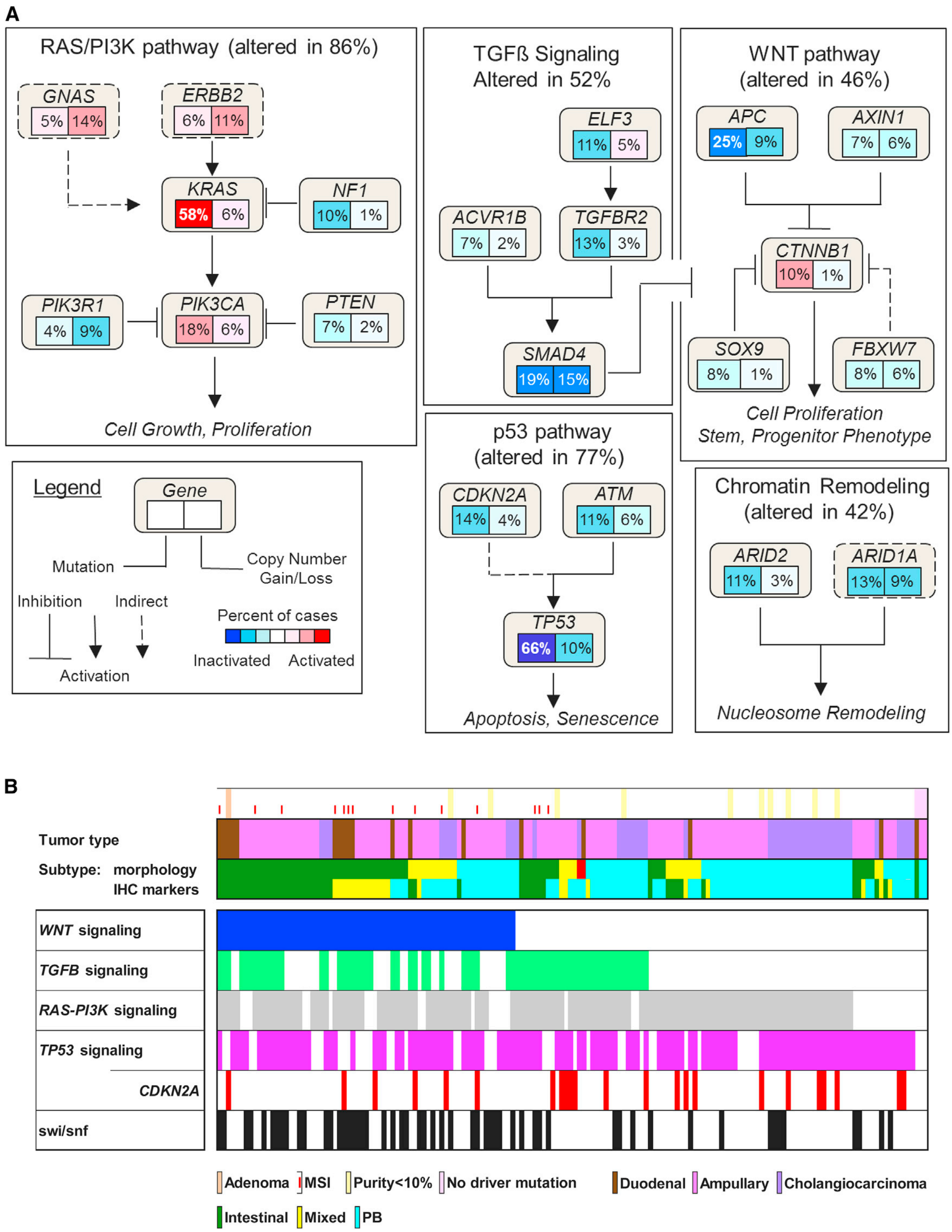


Figure 4. Major Altered Pathways in Periampullary Tumors

(A) Frequency of changes defined by somatic mutations or copy-number loss or gain is expressed as a percentage of cases for each gene. Inactivation (blue) or activation (red) is graded in intensity by percent of patients affected.

(B) Genetic alterations in the significantly mutated genes grouped by pathway are illustrated for each patient. Note *WNT* and *PI3K* signaling pathways could be found in the three tumor types and in each of their subtypes, including the pancreatobiliary subtype.

See also Figures S3A and S3B, wherein mutations in each gene are grouped by tumor type and subtype, and Figure S3C and Tables S4A and S4B.

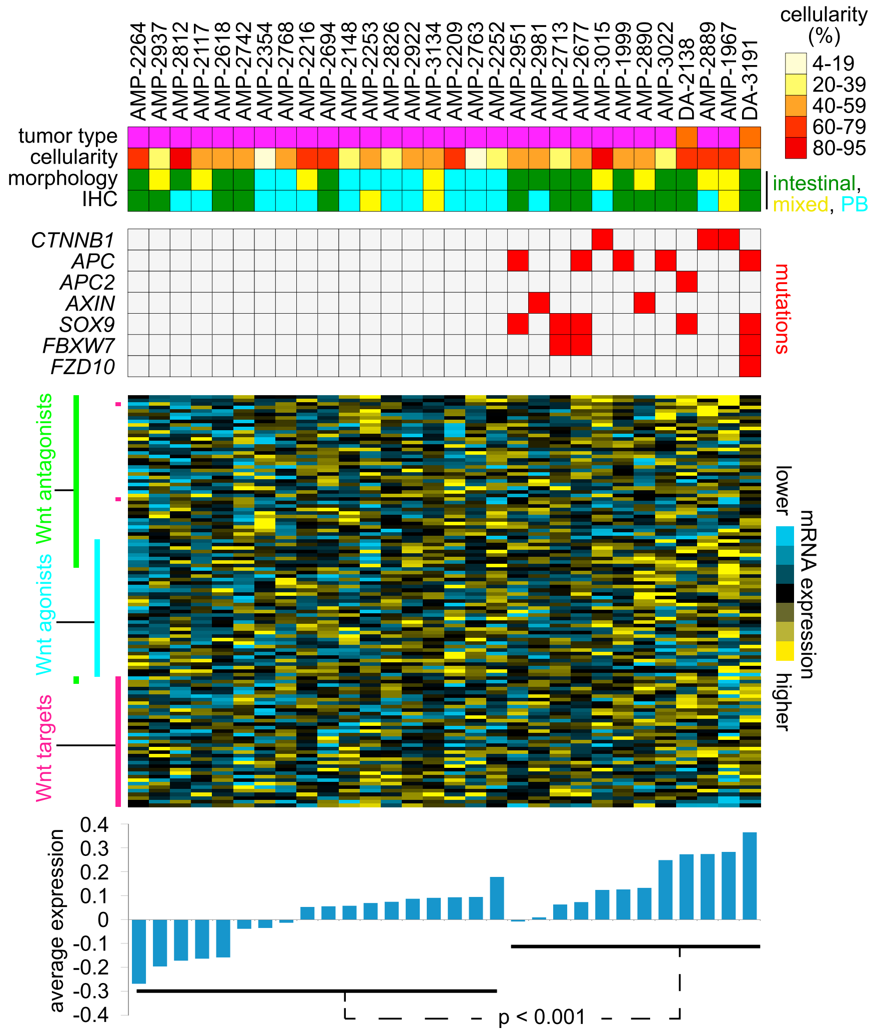


Figure 5. Relative RNA Expression of *WNT* Antagonist, Agonist, and Target Genes

Tumors were split between those with and without *WNT* canonical pathway mutations as shown in the mutation panel. The level of RNA expression for each gene can be visualized in the heat map and the average expression of all the genes is summarized in the lower panel. See also Figure S4 for fusions.

the expression data using a previously developed *WNT* signature that included *WNT* antagonist, *WNT* agonist, and *WNT* target genes (Donehower et al., 2013). An increase in expression in these three gene groups as a result of the *WNT* pathway deregulation was noticed in colorectal cancer (Donehower et al., 2013). This could be explained by the fact that *CTNNB1* activation resulted in an increased expression of targeted genes and the unrestricted *WNT* signaling set up a negative feedback loop of the *WNT* antagonist genes attempting to shut down signaling. In this study, we divided the patients into *WNT* mutated and those without (Figure 5, mutation panel). We then looked at the relative RNA expression in the two tumor groups for *WNT* antagonists, *WNT* agonists, and *WNT* targets (Figure 5, middle panel). The tumors with *WNT* mutations trend significantly toward higher overall *WNT* gene expression ($p < 0.001$) (Figure 5, lower panel). The *WNT* gene expression profile was also increased in some of the *WNT* non-mutated patients, indicating that some other mechanism affecting the *WNT* pathway might be at play.

Fusion analysis identified two noticeable non-recurrent fusions: *SLC45A3-ELK4* used as a prognosis marker in prostate cancer, where its expression is elevated (Kumar-Sinha et al., 2012; Ren et al., 2014), and a LINE-*MET* fusion in a patient without any *KRAS* or *TP53* driving mutations and a high *MET* expression (Figure S4; Table S5). LINE element insertions are found in PDAC, colon, hepatocellular, esophageal, and lung carcinoma (Paterson et al., 2015; Rodić et al., 2015).

Copy-Number Alteration

The majority of CNAs involved entire chromosomes or chromosome arms as opposed to focal events, which are common in gastrointestinal tumors. Arm-level deletions outnumbered amplifications across all tumors (Figure 6A). The three tumor types shared four arm-level amplifications and nine arm-level deletions. AMPAC shared amplification of 1q and deletion of 1p and 8p CAC (Table S6A). AMPAC shared no events specifically with DUOAC, making AMPAC marginally more similar to CAC in its CNA pattern. AMPAC also had two unique amplications on 5p and 6p, whereas 3q amplification was unique to CAC, and 6p was unique to DUOAC.

A combined GISTIC analysis revealed as expected a focal deletion of 9p23.1, involving *CDKN2A* (Table S6B). A focal deletion in chromosome 9 removed the promotor and 5′ end of *KDM4C* (Figure 6B). Although present in every tumor type, it was only statistically significant in AMPAC (Table S6B). This deletion resulted in a significant decrease in expression of *KDM4C* as well as the upstream *UHRF2* (Figure 6B, inserts). Interestingly, overexpression of both genes has been associated with a pro-growth effect on colon cancer cells (*KDM4C*) and a much lower disease-free survival and overall survival in patients with colon cancer (*UHRF2*) (Lu et al., 2014; Kim et al., 2014). *KDM4C* forms also complexes with β-catenin (Kim et al., 2014; Yamamoto et al., 2013).

DISCUSSION

This study compares the genetic constitution of ampullary cancer with two nearby tumor types with pathologic classification. Historically, ampullary cancers have been classified as intestinal

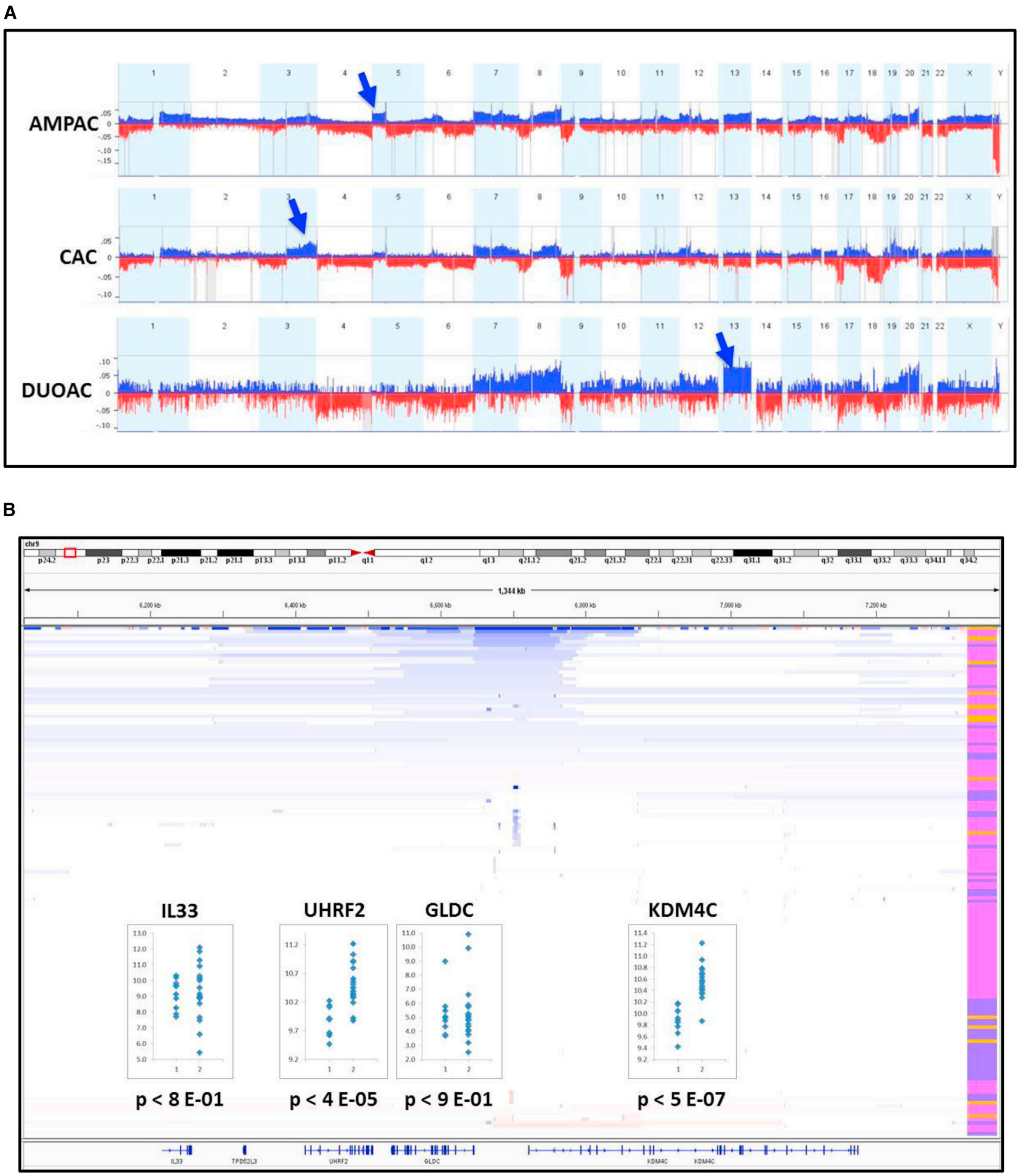


Figure 6. Copy-Number Alteration

(A) Nexus GISTIC analysis of copy-number alteration by anatomical site. Upper blue panel shows copy-number gains and the lower red panel shows copy-number losses for each tumor type. Blue arrows demark changes characteristic of a given anatomical site.

(legend continued on next page)

or pancreatobiliary subtypes based on immunohistochemistry and/or cellular morphology. The genomic analysis mirrored these results by the finding that some ampullary tumors exhibit properties of intestinal tumors such as microsatellite instability, *ELF3* mutation, and disruption of *WNT* signaling.

We found that the classification approaches of the three periampullary tumors are often discordant with one another. No unique molecular characteristics were specifically associated with one tumor subtype or one tumor type. Interestingly patients from each tumor type and subtype exhibited alterations in *WNT* pathway genes, including nearly one-fourth of the CAC tumors and one-fourth of the pancreatobiliary tumors. Other studies using subtype classification different from ours report *WNT* pathway mutations in AMPAC pancreatobiliary (PB) subtype (Achille et al., 1996; Hechtman et al., 2015). Transcriptional changes in AMPAC tumors in *WNT* signaling genes was increased, as expected, in tumors with *WNT* mutation, reinforcing a molecular dichotomy. With half of the patients across the three tumor types harboring *WNT* mutation, this could impact greatly the choice of treatment since several *WNT* pathway targeted therapies are in development. Ampullary, duodenal and distal bile duct adenocarcinoma could be regarded as a *WNT* ± entity from the perspective of treatment. Thus, the molecular data suggest that clinical testing for *WNT* signaling status might be beneficial to patients in the near future, making this a stepping stone to personalized medicine.

The identification of *ELF3* as a significantly mutated gene with an inactivating mutation pattern is also of interest. It was reported at lower frequency in bladder and biliary tract cancers, but not in any other cancer so far. *ELF3* encodes an ETS-domain transcription factor. By interacting with promoter regions, *ELF3* is implicated in the regulation of several genes during epithelial cell differentiation (Oliver et al., 2012). One of the genes transactivated by *ELF3* is *TGFBR2*, a prime initiator of *TGF-β* signaling, a pathway with a dual role in tumorigenesis, suppressing tumor progression at early stages but enhancing invasion and metastasis at later stages (Roberts and Wakefield, 2003). The tumor-suppressor antiproliferative function of *ELF3* was previously noted in studies on colorectal, prostate, and oral squamous cancer cells (Iwai et al., 2008; Lee et al., 2003, 2008; Shatnawi et al., 2014) and more recently in biliary tract cancer cell line (Nakamura et al., 2015). Such studies showed that *ELF3* directly binds to the promoter region of EGR1 (Lee et al., 2008) and TGFBR2 (Lee et al., 2003), increasing the transcription of these two tumor suppressor genes in CRC, whereas *ELF3* binding to androgen receptor (*AR*) (Shatnawi et al., 2014) and matrix metalloproteinase-9 (*MMP9*) (Iwai et al., 2008) promoters suppressed the transcriptional activity of these tumor growth- and invasiveness-promoting genes in prostate and squamous cancers, respectively. However, recent observations also suggested an oncogenic functional role in CRC development when *ELF3* is amplified, and its upregulated expression correlated with cancer progression and decreased patient survival (Wang et al., 2014). A *WNT*-independent *CTNNB1* transactivation facilitating tumor development was also reported (Wang et al., 2014). Such dual function has also been observed in breast and prostate cancer (Longoni et al., 2013; Oliver et al., 2012; Shatnawi et al., 2014). It could be argued that when *ELF3* inactivation occurs early in tumor development, it provides a moderate growth advantage by suppressing *TGF-β* signaling. The fact that we found *ELF3* mutation in a duodenal adenoma with intraepithelial neoplasia and dysplasia components (DUOAC 707) and that 75% of the tumors with *ELF3* mutation were lower-grade tumors (stage I or II) could support this hypothesis. The *ELF3* functional switch might depend on tumor stage and expression of other factors and/or be associated with its expression level, some genes being transactivated only when *ELF3* is overexpressed. In any case, *ELF3* is implicated in the development of periampullary tumors, and its exact functional role during periampullary tumor development will need to be investigated further.

EXPERIMENTAL PROCEDURES

Clinical Data

A total of 160 tumors (98 AMPAC, 44 CAC, and 18 DUOAC) were collected by the different groups participating in this study: Australian Pancreatic Cancer Genome Initiative (APGI), Baylor College of Medicine Elkins Pancreatic Center (BCM) as a member of The Cancer Research Banking, MD Anderson Cancer Center (MDA), and Technical University of Dresden (TUD). Ethical approval was obtained from each of these institution's research ethic boards. All patients underwent surgical pancreatoduodenectomy with curative intent without known residual disease. Clinical data variables including race, sex, age, familial history, operative procedure, pathological findings, and survival from the date of initial surgery to the date of death or last follow up are presented in Table S1A.

Tumor Classification

A section of the tumor was fixed in formaldehyde and embedded in paraffin (FFPE). H&E-stained and immunohistochemistry slides from the FFPE tissue were examined by the pathologists from the original site of collection to confirm diagnosis of the specific tumor section and to grade expression of subtype markers. All slides were then centrally reviewed by a single pathologist (A.G.) who was blinded to all clinical, molecular, and pathological data at the time of review and scoring. The distinction between pancreatic, biliary, ampullary or intestinal carcinoma was based on the anatomical site from which the carcinoma was thought to arise using the guidelines recommended in the AJCC seventh edition 2009 staging manual (Edge et al., 2009).

Histology and Morphology

Tumors were classified as pancreatobiliary, intestinal, or mixed morphological subtype based on the cellular morphology. A morphology similar to colorectal adenocarcinoma (tall often pseudostratified columnar epithelium with oval nuclei forming elongated glands) was defined as intestinal type. Morphology similar to pancreatobiliary carcinoma (small solid nest of cells with rounded nuclei surrounded by desmoplastic stroma and forming simple or branching rounded glands) was defined as pancreatobiliary type. Mixed histological types contained a mixture of both intestinal and pancreatobiliary types with 80% or less of the cells with either morphology. Grade of differentiation was also noted as well as presence of adenoma, signet-ring cells, or mucinous cells (Table S1B).

(B) Focal deletion in the promotor region and at the 5′ end of *KMD4C* impacts its and *UHRF2* expression. Human Omni 2.5 SNP array results were analyzed in IGV (Integrative Genomics Viewer, Broad Institute). Deletions are in blue, and amplifications are in red. Gene expression was analyzed by dividing the samples into two groups: samples with (1) or without (2) focal deletion. The color ladder on the right indicates the tumor type (pink, AMPAC; purple, CAC; and orange, DUOAC). See also Tables S6A and S6B.

Immunohistochemistry Staining

FFPE sections were stained with antibodies against MUC1 and CDX2 (see Supplemental Experimental Procedures). This two-antibody panel has previously been validated by our group to predict prognosis in ampullary carcinoma (Chang et al., 2013), and the methods we used were the same as employed in that study. Briefly, expression was evaluated by estimating the percentage of positively stained carcinoma cells and the intensity of the staining (0 absent, 1+ weak, 2+, 3+ strong). H scores were calculated for both markers by multiplying the percentage of stained cells by the intensity of the staining. The ratio of the CDX2/MUC1 H score defined the subtypes: a ratio of 2 and above and smaller than 0.5 were considered intestinal and pancreatobiliary, respectively. Intermediate values were associated to a mixed subtype (Table S1B).

Nucleic Acid Isolation

Samples were retrieved and had full face sectioning performed in OCT embedding media to verify the presence of carcinoma in the sample to be sequenced and to estimate the percentage of malignant epithelial nuclei in the sample relative to stromal nuclei. Macrodissection was performed if possible to excise areas of non-malignant tissue. DNA and RNA extraction was performed at the center of collection following their own protocol with all samples being tracked using unique identifiers though out the process (see Supplemental Methods). DNA was shipped and quantified at BCM-Human Genome Sequencing Center (HGSC) using the PicoGreen DNA Assay.

SNP Array Assays

SNP arrays were processed at the HGSC for each sample using the Illumina Infinium LCG Assay according to the manufacturer's guidelines. Specifically, assays were performed with Human Omni2.5-8 BeadChips (Illumina, catalog no. WG-311-2513), interrogating 2.5 million SNP loci with a MAF detection limit of 1% (see Supplemental Experimental Procedures). SNP calls were collected using Illumina's GenomeStudio software (version 2011.1) in which standard SNP clustering and genotyping were performed with the default settings recommended by the manufacturer. Data from samples that met a minimum SNP call rate of 0.9 were considered passing and were included in subsequent analyses. Results were analyzed on Nexus (BioDiscovery).

Sequencing

Library preparation, whole (Bainbridge et al., 2011) and targeted exome capture, and regular and ultra-deep sequencing on HiSeq 2000 platform are detailed in Supplemental Experimental Procedures. In brief, 152 samples were whole-exome sequenced and their mutations validated with a custom design targeted exome capture. The targeted capture consisted of a panel of 71 genes covering 0.25 Mb, and the probes were designed by Nimblegen (genes are listed in Supplemental Experimental Procedures). These genes were selected on the basis they were significantly mutated and/or had high impact in the development of AMPAC, PDAC, DUOAC, CAC, and other pancreatic tumor types. The selective targeted capture was also used in discovery on eight samples received at a later date (six samples) or of low purity (two samples with <10% tumor). The mutations identified with the targeted capture were validated with ultra-deep (single-molecule reconstruction) sequencing.

Data Analysis

Primary Data Analysis

Initial sequence analysis was performed using the HGSC Mercury analysis pipeline (https://www.hgsc.bcm.edu/software/mercury). First, the primary analysis software on the instrument produces .bcl files that are transferred off-instrument into the HGSC analysis infrastructure by the HiSeq Real-time Analysis module. Once the run is complete and all .bcl files are transferred, Mercury runs the vendor's primary analysis software (CASAVA), which demultiplexes pooled samples and generates sequence reads and base-call confidence values (qualities). The next step is the mapping of reads to the GRCh37 Human reference genome (http://www.ncbi.nlm.nih.gov/projects/genome/assembly/grc/human/) using the Burrows-Wheeler aligner (Li and Durbin, 2009) (BWA; http://bio-bwa.sourceforge.net/) and producing a BAM (Li et al., 2009) (binary alignment/map) file. The third step involves quality recalibration (using GATK; DePristo et al., 2011; https://www.broadinstitute.org/gatk/) and, where necessary, the merging of separate sequence-event BAMs into a single sample-level BAM. BAM sorting, duplicate-read marking, and realignment to improve in/del discovery all occur at this step.

Cancer Data Analysis

Primary BAM files were separately run through Atlas-SNP (Shen et al., 2010), Atlas-Indel, and PInDel (Ye et al., 2009). Data were aggregated for each tumor/normal pair, and variants were cross-checked for each tissue pair. Variant annotation was performed using Annovar (Wang et al., 2010), COSMIC (Forbes et al., 2011), and dbSNP (Sherry et al., 2001). Variant filtering was performed to remove low-quality variants. Cohort-level data processing was performed to remove additional false somatic calls by filtering against a cohort of normal tissues.

Ultra-deep Sequencing Analysis

Duplicate reads were aggregated and consensus variants were defined as the variant being present in 90% of the reads contributing to both halves of the duplex molecule. Subsequent filtering was employed to remove variants in which there was either mapping error (tested using BLAST) or sequence error non-consensus block rate of 50% (Altschul et al., 1990). Variants detected in this way were annotated using ANNOVAR, COSMIC, and dbSNP annotations (Forbes et al., 2011; Sherry et al., 2001; Wang et al., 2010).

Mutational Signature

Mutation signatures were generated from a set of over 6,000 somatic mutations across a range of cancer types using non-smooth non-negative matrix factorization (nsNMF) (K.R.C., unpublished data; Pascual-Montano et al., 2006). The solution resulted in 21 distinct mutational signatures, similar to those previously reported (Alexandrov et al., 2013; Gaujoux and Seoighe, 2010), many of which could be correlated with previously published mutational modes, including APOBEC, UV radiation exposure, *POLE* hypermutation (Lawrence et al., 2013), and CpG island mutation. Mutations for this cohort where compared against the solved NMF to generate a mutational decomposition for each of the tumor samples. Samples were aggregated and compared using hierarchical clustering and other correlative statistics to clinical covariates.

Tumor Purity and Normalization Mutation Rate

Tumor purity was estimated using ASCAT and the tumor variant allelic fraction of driver genes. The average of both analyses was plotted against the number of mutations in each tumor, and the slope value was used to approximate the number of mutations that would have been identified in 100% tumor cellularity (Figure S1A).

Significantly Altered Genes

Several approaches were taken to dissect genes and pathways which were mutated more often than by chance in this dataset. We used the final MAF file (Tables S6A and 6B) to calculate significantly mutated genes using MutSig-CV and an inactivation bias test (Lawrence et al., 2013; Totoki et al., 2014).

Microsatellite Instability Coefficient

For details, see Supplemental Experimental Procedures.

Multivariable Cox Analysis

Cox proportional hazards analysis was performed using the survival (Therneau, 2000) package in R (R Development Core Team, 2008). We included age at diagnosis, gender, stage, grade, tumor type, histologic subtype (IHC), and mutation status (*WNT*, *TGFB*, *TP53*, *KRAS*, *PI3K*, and chromatin remodeling) in the multivariate Cox analysis. Country of origin and ethnicity were not included as covariates since they had no associated effect with survival.

RNA-Seq

Total RNA was prepared using the AllPrep RNA/DNA isolation kit (QIAGEN). RNA integrity was confirmed (RIN > 7.0) on a Bioanalyzer (Agilent). RNA-seq libraries were prepared using the TruSeq Stranded total RNA LT library kit (Illumina) following the manufacturer's instructions. 100-base-pair-end sequencing was then performed to a minimum depth of 50 million reads of each sample on an Illumina HiSeq2000 sequencer.

Transcript Expression Analysis

Gene Expression of the WNT Pathway

The profiles in the RNA-seq dataset were quantile normalized, and log-transformed expression values were then centered to SDs from the median across sample profiles. Tumors were split into two groups (those with *WNT* canonical pathway mutations and those without such mutations) and were scored for relative activity in the *WNT* pathway. The gene signature score within each tumor profile was defined as the average of the centered values for the *WNT* signature genes.

Fusion Analysis

The deFuse software version 0.6.1 (McPherson et al., 2011) with default settings was used to detect fusion genes. The deFuse results were further filtered by removing identified read through fusions, selecting coding regions, selecting in-frame (open reading frame) genes and selecting samples with a deFuse confidence score of >80%. This filtering resulted in a list of candidate fusion genes. To characterize these candidate fusion genes, we took each spanning junction read and using the BLAT tool in UCSC genome browser examined where the reads mapped. The fusions that mapped with 100% identity to each part of the identified fusion (gene1 or gene2) were selected for further analysis. This filter removed genes that mapped to multiple locations. Next, each RNA BAM from candidate fusion genes was examined in IGV, looking for stacked soft clipped reads, changes in coverage, at the identified fusion breakpoints. The sequence of each soft clipped read was brought into the UCSC genome browser and mapped using BLAT. Only fusions that had reads that matched (100%) the identified fusion genes were considered further.

ACCESSION NUMBERS

The accession number for the sequence data reported in this paper is dbGAP: PRJNA280134.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, four figures, and six tables and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2015.12.005.

CONSORTIA

The members of Australian Pancreatic Cancer Genome Initiative are Andrew V. Biankin, Amber L. Johns, Amanda Mawson, David K. Chang, Christopher J. Scarlett, Mary-Anne L. Brancato, Sarah J. Rowe, Skye H. Simpson, Mona Martyn-Smith, Michelle T. Thomas, Lorraine A. Chantrill, Venessa T. Chin, Angela Chou, Mark J. Cowley, Jeremy L. Humphris, Marc D. Jones, R. Scott Mead, Adnan M. Nagrial, Marina Pajic, Jessica Pettit, Mark Pinese, Ilse Rooman, Jianmin Wu, Jiang Tao, Renee DiPietro, Clare Watson, Angela Steinmann, Hong Ching Lee, Rachel Wong, Andreia V. Pinho, Marc Giry-Laterriere, Roger J. Daly, Elizabeth A. Musgrove, and Robert L. Sutherland (Garvan Institute of Medical Research, Australia); Sean M. Grimmond, Nicola Waddell, Karin S. Kassahn, David K. Miller, Peter J. Wilson, Ann-Marie Patch, Sarah Song, Ivon Harliwong, Senel Idrisoglu, Craig Nourse, Ehsan Nourbakhsh, Suzanne Manning, Shivangi Wani, Milena Gongora, Matthew Anderson, Oliver Holmes, Conrad Leonard, Darrin Taylor, Scott Wood, Christina Xu, Katia Nones, J. Lynn Fink, Angelika Christ, Tim Bruxner, Nicole Cloonan, Felicity Newell, John V. Pearson, Peter Bailey, Michael Quinn, Shivashankar Nagaraj, Stephen Kazakoff, Nick Waddell, Keerthana Krisnan, Kelly Quek, and David Wood (Queensland Centre for Medical Genomics and Institute for Molecular Biosciences, Australia); Jaswinder S. Samra, Anthony J. Gill, Nick Pavlakis, Alex Guminski, and Christopher Toon (Royal North Shore Hospital, Australia); Ray Asghari, Neil D. Merrett, Darren Pavey, and Amitabha Das (Bankstown Hospital, Australia); Peter H. Cosman, Kasim Ismail, and Chelsie O'Connnor (Liverpool Hospital, Australia); Vincent W. Lam, Duncan McLeod, Henry C. Pleass, Arthur Richardson, and Virginia James (Westmead Hospital, Australia); James G. Kench, Caroline L. Cooper, David Joseph, Charbel Sandroussi, Michael Crawford, and James Gallagher (Royal Prince Alfred Hospital, Australia); Michael Texler, Cindy Forest, Andrew Laycock, Krishna P. Epari, Mo Ballal, David R. Fletcher, and Sanjay Mukhedkar (Fremantle Hospital, Australia); Nigel A. Spry, Bastiaan DeBoer, and Ming Chai (Sir Charles Gairdner Hospital, Australia); Nikolajs Zeps, Maria Beilin, and Kynan Feeney (St John of God Healthcare, Australia); Nan Q Nguyen, Andrew R. Ruszkiewicz, Chris Worthley, Chuan P. Tan, and Tamara Debrencini (Royal Adelaide Hospital, Australia); John Chen, Mark E. Brooke-Smith, and Virginia Papangelis (Flinders Medical Centre, Australia); Henry Tang and Andrew P. Barbour (Greenslopes Private Hospital, Australia); Andrew D. Clouston and Patrick Martin (Envoi Pathology, Australia); Thomas J. O'Rourke, Amy Chiang, Jonathan W. Fawcett, Kellee Slater, Shinn Yeung, Michael Hatzifotis, and Peter Hodgkinson (Princess Alexandria Hospital, Australia); Christopher Christophi, Mehrdad Nikfarjam, and Angela Mountain; Victorian Cancer Biobank (Australia); James R. Eshleman, Ralph H. Hruban, Anirban Maitra, Christine A. Iacobuzio-Donahue, Richard D. Schulick, Christopher L. Wolfgang, Richard A. Morgan, and Mary Hodgin (Johns Hopkins Medical Institutes, USA); Aldo Scarpa, Rita T. Lawlor, Stefania Beghelli, Vincenzo Corbo, Maria Scardoni, and Claudio Bassi (ARC-Net Centre for Applied Research on Cancer, Italy); Margaret A. Tempero (University of California, San Francisco, USA); David K. Chang (Garvan Institute of Medical Research, Australia; University of Glasgow, UK; and Greater Glasgow and Clyde National Health Service, UK); Sean M. Grimmond and Craig Nourse (University of Glasgow, UK; Queensland Centre for Medical Genomics and Institute for Molecular Biosciences, Australia); Elizabeth A. Musgrove (University of Glasgow, UK); Marc D. Jones (Garvan Institute of Medical Research, Australia; and, University of Glasgow, UK); and Nigel B. Jamieson, Fraser R. Duthie, and Janet S. Graham (University of Glasgow, UK; and Greater Glasgow and Clyde National Health Service, UK).

AUTHOR CONTRIBUTIONS

The research network comprises Baylor College of Medicine Human Genome Sequencing Center, Australian Pancreatic Cancer Genome Initiative, MD Anderson Cancer Center, TU Dresden, and Glasgow as part of the International Cancer Genome Consortium study. Each center contributed biospecimens collected at affiliated hospitals and processed at each biospecimen core resource centre. Data generation and analyses were performed by the Human Genome Sequencing Center. Investigator contributions are as follows: project leader: M.C.G., C.P., R.G., M.J.O., S.M.G., A.V.B., D.A.W., and R.A.G.; writing team: M.C.G. and D.A.W.; bioinformatics/databases: J.D., K.W., O.A.H., L.X., M.D., and P.B.; sequencing: D.M.M, H.D., S.L.L., M.B., J.H., Y.H., and H.H.D.; data analysis: M.C.G., K.R.C., D.K.C., L.A.D., C.J.C, E.S., N.D., J.S.S., J.V.P., and C.S.; surgery: D.K.C, W.E.F, R.G., N.B.J., and G.V.B.; sample collection and clinical annotation: A.L.J., C.P., S.E.H., and A.M.; sample processing and quality control: M.C.G., A.L.J., C.P., R.G., S.L.L., and H.W.; and pathology assessment: A.J.G., M.M.I., H.W., D.A., K.O., R.H.H., and F.R.D.

ACKNOWLEDGMENTS

We acknowledge the following funding support: HGSC-BCM: NHGRI U54 HG003273 and CPRIT grant RP101353-P7 (Tumor Banking for Genomic Research and Clinical Translation Site 1); D.W.A. CPRIT grant RPI21018; M.M.I. and C.J.C.: Dan L. Duncan Cancer Center NIH P30 Cancer Center support grant (P30 CA125123) supporting the BCM Human Tissue Acquisition and Pathology Core and the Biostatistics and Bioinformatics Core; M.J.O.: Kavanagh Family Foundation. Australian Pancreatic Cancer Genome Initiative: National Health and Medical Research Council of Australia (NHMRC; 631701, 535903, 427601); Queensland Government (NIRAP); University of Queensland; Institute for Molecular Bioscience; Cancer Research UK (C596/A18076, C29717/A17263); University of Glasgow; Australian Government: Department of Innovation, Industry, Science and Research (DIISR); Australian Cancer Research Foundation (ACRF); Cancer Council NSW: (SRP06-01, SRP11-01. ICGC); Cancer Institute NSW: (10/ECF/2-26; 06/ECF/1-24; 09/CDF/2-40; 07/CDF/1-03; 10/CRF/1-01, 08/RSA/1-15, 07/CDF/1-28, 10/CDF/2-26,10/FRL/2-03, 06/RSA/1-05, 09/RIG/1-02, 10/TPG/1-04, 11/REG/1-10, 11/CDF/3-26); Garvan Institute of Medical Research; Avner Nahmani Pancreatic Cancer Research Foundation; Howat Foundation; R.T. Hall Trust; Petre Foundation; Philip Hemstritch Foundation; Gastroenterological Society of Australia (GESA); American Association for Cancer Research (AACR) Landon Foundation – INNOVATOR Award; Royal Australasian College of Surgeons (RACS); Royal Australasian College of Physicians (RACP); Royal College of Pathologists of Australasia (RCPA); Italian Ministry of Research (Cancer Genome Project FIRB RBAP10AHJB);

Associazione Italiana Ricerca Cancro (12182); Fondazione Italiana Malattie Pancreas – Ministero Salute (CUP_J33G13000210001); Wilhelm Sander Stiftung 2009.039.2; NIH grant P50 CA62924. See the Supplemental Information for further acknowledgments.

Received: April 14, 2015
Revised: October 30, 2015
Accepted: November 19, 2015
Published: January 21, 2016

REFERENCES

Achille, A., Scupoli, M.T., Magalini, A.R., Zamboni, G., Romanelli, M.G., Orlandini, S., Biasi, M.O., Lemoine, N.R., Accolla, R.S., and Scarpa, A. (1996). APC gene mutations and allelic losses in sporadic ampullary tumours: evidence of genetic difference from tumours associated with familial adenomatous polyposis. Int. J. Cancer *68*, 305–312.

Adsay, V., Ohike, N., Tajiri, T., Kim, G.E., Krasinskas, A., Balci, S., Bagci, P., Basturk, O., Bandyopadhyay, S., Jang, K.T., et al. (2012). Ampullary region carcinomas: definition and site specific classification with delineation of four clinicopathologically and prognostically distinct subsets in an analysis of 249 cases. Am. J. Surg. Pathol. *36*, 1592–1608.

Alexandrov, L.B., Nik-Zainal, S., Wedge, D.C., Aparicio, S.A., Behjati, S., Biankin, A.V., Bignell, G.R., Bolli, N., Borg, A., Børresen-Dale, A.L., et al.; Australian Pancreatic Cancer Genome Initiative; ICGC Breast Cancer Consortium; ICGC MMML-Seq Consortium; ICGC PedBrain (2013). Signatures of mutational processes in human cancer. Nature *500*, 415–421.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. *215*, 403–410.

Amptoulach, S., Josefsson, A., Kavantzas, N., and Kalaitzakis, E. (2011). Adenocarcinoma of the ampulla of Vater: does the histopathologic type matter? Scand. J. Gastroenterol. *46*, 1401–1403.

Bainbridge, M.N., Wang, M., Wu, Y., Newsham, I., Muzny, D.M., Jefferies, J.L., Albert, T.J., Burgess, D.L., and Gibbs, R.A. (2011). Targeted enrichment beyond the consensus coding DNA sequence exome reveals exons with higher variant densities. Genome Biol. *12*, R68.

Biankin, A.V., Waddell, N., Kassahn, K.S., Gingras, M.C., Muthuswamy, L.B., Johns, A.L., Miller, D.K., Wilson, P.J., Patch, A.M., Wu, J., et al.; Australian Pancreatic Cancer Genome Initiative (2012). Pancreatic cancer genomes reveal aberrations in axon guidance pathway genes. Nature *491*, 399–405.

Cancer Genome Atlas, N.; Cancer Genome Atlas Network (2012). Comprehensive molecular characterization of human colon and rectal cancer. Nature *487*, 330–337.

Cerami, E., Gao, J., Dogrusoz, U., Gross, B.E., Sumer, S.O., Aksoy, B.A., Jacobsen, A., Byrne, C.J., Heuer, M.L., Larsson, E., et al. (2012). The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. *2*, 401–404.

Chang, D.K., Jamieson, N.B., Johns, A.L., Scarlett, C.J., Pajic, M., Chou, A., Pinese, M., Humphris, J.L., Jones, M.D., Toon, C., et al. (2013). Histomolecular phenotypes and outcome in adenocarcinoma of the ampulla of vater. J. Clin. Oncol. *31*, 1348–1356.

DePristo, M.A., Banks, E., Poplin, R., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis, A.A., del Angel, G., Rivas, M.A., Hanna, M., et al. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat. Genet. *43*, 491–498.

Donehower, L.A., Creighton, C.J., Schultz, N., Shinbrot, E., Chang, K., Gunaratne, P.H., Muzny, D., Sander, C., Hamilton, S.R., Gibbs, R.A., and Wheeler, D. (2013). MLH1-silenced and non-silenced subgroups of hypermutated colorectal carcinomas have distinct mutational landscapes. J. Pathol. *229*, 99–110.

Edge, S.E., Byrd, D.R., and Compton, C.C. (2009). AJCC Cancer Staging Manual, Seventh Edition (New York: Springer).

Ehehalt, F., Rümmele, P., Kersting, S., Lang-Schwarz, C., Rückert, F., Hartmann, A., Dietmaier, W., Terracciano, L., Aust, D.E., Jahnke, B., et al. (2011). Hepatocyte nuclear factor (HNF) 4α expression distinguishes ampullary cancer subtypes and prognosis after resection. Ann. Surg. *254*, 302–310.

Forbes, S.A., Bindal, N., Bamford, S., Cole, C., Kok, C.Y., Beare, D., Jia, M., Shepherd, R., Leung, K., Menzies, A., et al. (2011). COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. Nucleic Acids Res. *39*, D945–D950.

Gao, J., Aksoy, B.A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S.O., Sun, Y., Jacobsen, A., Sinha, R., Larsson, E., et al. (2013). Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci. Signal. *6*, pl1.

Gaujoux, R., and Seoighe, C. (2010). A flexible R package for nonnegative matrix factorization. BMC Bioinformatics *11*, 367.

Hechtman, J.F., Liu, W., Sadowska, J., Zhen, L., Borsu, L., Arcila, M.E., Won, H.H., Shah, R.H., Berger, M.F., Vakiani, E., et al. (2015). Sequencing of 279 cancer genes in ampullary carcinoma reveals trends relating to histologic subtypes and frequent amplification and overexpression of ERBB2 (HER2). Mod. Pathol. *28*, 1123–1129.

Heinrich, S., and Clavien, P.A. (2010). Ampullary cancer. Curr. Opin. Gastroenterol. *26*, 280–285.

Iwai, S., Amekawa, S., Yomogida, K., Sumi, T., Nakazawa, M., Yura, Y., Nishimune, Y., and Nozaki, M. (2008). ESE-1 inhibits the invasion of oral squamous cell carcinoma in conjunction with MMP-9 suppression. Oral Dis. *14*, 144–149.

Kim, T.D., Fuchs, J.R., Schwartz, E., Abdelhamid, D., Etter, J., Berry, W.L., Li, C., Ihnat, M.A., Li, P.K., and Janknecht, R. (2014). Pro-growth role of the JMJD2C histone demethylase in HCT-116 colon cancer cells and identification of curcuminoids as JMJD2 inhibitors. Am. J. Transl. Res. *6*, 236–247.

Kumar-Sinha, C., Kalyana-Sundaram, S., and Chinnaiyan, A.M. (2012). SLC45A3-ELK4 chimera in prostate cancer: spotlight on cis-splicing. Cancer Discov. *2*, 582–585.

Lawrence, M.S., Stojanov, P., Polak, P., Kryukov, G.V., Cibulskis, K., Sivachenko, A., Carter, S.L., Stewart, C., Mermel, C.H., Roberts, S.A., et al. (2013). Mutational heterogeneity in cancer and the search for new cancer-associated genes. Nature *499*, 214–218.

Lawrence, M.S., Stojanov, P., Mermel, C.H., Robinson, J.T., Garraway, L.A., Golub, T.R., Meyerson, M., Gabriel, S.B., Lander, E.S., and Getz, G. (2014). Discovery and saturation analysis of cancer genes across 21 tumour types. Nature *505*, 495–501.

Lee, H.J., Chang, J.H., Kim, Y.S., Kim, S.J., and Yang, H.K. (2003). Effect of ets-related transcription factor (ERT) on transforming growth factor (TGF)-beta type II receptor gene expression in human cancer cell lines. J. Exp. Clin. Cancer Res. *22*, 477–480.

Lee, S.H., Bahn, J.H., Choi, C.K., Whitlock, N.C., English, A.E., Safe, S., and Baek, S.J. (2008). ESE-1/EGR-1 pathway plays a role in tolfenamic acid-induced apoptosis in colorectal cancer cells. Mol. Cancer Ther. *7*, 3739–3750.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754–1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.

Longoni, N., Sarti, M., Albino, D., Civenni, G., Malek, A., Ortelli, E., Pinton, S., Mello-Grand, M., Ostano, P., D'Ambrosio, G., et al. (2013). ETS transcription factor ESE1/ELF3 orchestrates a positive feedback loop that constitutively activates NF-κB and drives prostate cancer progression. Cancer Res. *73*, 4533–4547.

Lu, S., Yan, D., Wu, Z., Jiang, T., Chen, J., Yuan, L., Lin, J., Peng, Z., and Tang, H. (2014). Ubiquitin-like with PHD and ring finger domains 2 is a predictor of survival and a potential therapeutic target in colon cancer. Oncol. Rep. *31*, 1802–1810.

Martinez, V.D., Thu, K.L., Vucic, E.A., Hubaux, R., Adonis, M., Gil, L., MacAulay, C., Lam, S., and Lam, W.L. (2013). Whole-genome sequencing analysis identifies a distinctive mutational spectrum in an arsenic-related lung tumor. J. Thorac. Oncol. *8*, 1451–1455.

McPherson, A., Hormozdiari, F., Zayed, A., Giuliany, R., Ha, G., Sun, M.G., Griffith, M., Heravi Moussavi, A., Senz, J., Melnyk, N., et al. (2011). deFuse: an algorithm for gene fusion discovery in tumor RNA-Seq data. PLoS Comput. Biol. *7*, e1001138.

Morini, S., Perrone, G., Borzomati, D., Vincenzi, B., Rabitti, C., Righi, D., Castri, F., Manazza, A.D., Santini, D., Tonini, G., et al. (2013). Carcinoma of the ampulla of Vater: morphological and immunophenotypical classification predicts overall survival. Pancreas *42*, 60–66.

Nakamura, H., Arai, Y., Totoki, Y., Shirota, T., Elzawahry, A., Kato, M., Hama, N., Hosoda, F., Urushidate, T., Ohashi, S., et al. (2015). Genomic spectra of biliary tract cancer. Nat. Genet. *47*, 1003–1010.

Oliver, J.R., Kushwah, R., and Hu, J. (2012). Multiple roles of the epithelium-specific ETS transcription factor, ESE-1, in development and disease. Lab. Invest. *92*, 320–330.

Pascual-Montano, A., Carazo, J.M., Kochi, K., Lehmann, D., and Pascual-Marqui, R.D. (2006). Nonsmooth nonnegative matrix factorization (nsNMF). IEEE Trans. Pattern Anal. Mach. Intell. *28*, 403–415.

Paterson, A.L., Weaver, J.M., Eldridge, M.D., Tavaré, S., Fitzgerald, R.C., and Edwards, P.A.; OCCAMs Consortium (2015). Mobile element insertions are frequent in oesophageal adenocarcinomas and can mislead paired-end sequencing analysis. BMC Genomics *16*, 473.

R Development Core Team (2008). R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing).

Ren, G., Zhang, Y., Mao, X., Liu, X., Mercer, E., Marzec, J., Ding, D., Jiao, Y., Qiu, Q., Sun, Y., et al. (2014). Transcription-mediated chimeric RNAs in prostate cancer: time to revisit old hypothesis? Omics *18*, 615–624.

Roberts, A.B., and Wakefield, L.M. (2003). The two faces of transforming growth factor beta in carcinogenesis. Proc. Natl. Acad. Sci. USA *100*, 8621–8623.

Roberts, S.A., Lawrence, M.S., Klimczak, L.J., Grimm, S.A., Fargo, D., Stojanov, P., Kiezun, A., Kryukov, G.V., Carter, S.L., Saksena, G., et al. (2013). An APOBEC cytidine deaminase mutagenesis pattern is widespread in human cancers. Nat. Genet. *45*, 970–976.

Rodić, N., Steranka, J.P., Makohon-Moore, A., Moyer, A., Shen, P., Sharma, R., Kohutek, Z.A., Huang, C.R., Ahn, D., Mita, P., et al. (2015). Retrotransposon insertions in the clonal evolution of pancreatic ductal adenocarcinoma. Nat. Med. *21*, 1060–1064.

Romiti, A., Barucca, V., Zullo, A., Sarcina, I., Di Rocco, R., D'Antonio, C., Latorre, M., and Marchetti, P. (2012). Tumors of ampulla of Vater: A case series and review of chemotherapy options. World J. Gastrointest. Oncol. *4*, 60–67.

Shatnawi, A., Norris, J.D., Chaveroux, C., Jasper, J.S., Sherk, A.B., McDonnell, D.P., and Giguère, V. (2014). ELF3 is a repressor of androgen receptor action in prostate cancer cells. Oncogene *33*, 862–871.

Shen, Y., Wan, Z., Coarfa, C., Drabek, R., Chen, L., Ostrowski, E.A., Liu, Y., Weinstock, G.M., Wheeler, D.A., Gibbs, R.A., and Yu, F. (2010). A SNP discovery method to assess variant allele probability from next-generation resequencing data. Genome Res. *20*, 273–280.

Sherry, S.T., Ward, M.H., Kholodov, M., Baker, J., Phan, L., Smigielski, E.M., and Sirotkin, K. (2001). dbSNP: the NCBI database of genetic variation. Nucleic Acids Res. *29*, 308–311.

Therneau, T.M. (2000). Modeling Survival Data: Extending the Cox Model (Springer-Verlag).

Thompson, E., Meldrum, C.J., Crooks, R., McPhillips, M., Thomas, L., Spigelman, A.D., and Scott, R.J. (2004). Hereditary non-polyposis colorectal cancer and the role of hPMS2 and hEXO1 mutations. Clin. Genet. *65*, 215–225.

Totoki, Y., Tatsuno, K., Covington, K.R., Ueda, H., Creighton, C.J., Kato, M., Tsuji, S., Donehower, L.A., Slagle, B.L., Nakamura, H., et al. (2014). Trans-ancestry mutational landscape of hepatocellular carcinoma genomes. Nat. Genet. *46*, 1267–1273.

Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. *38*, e164.

Wang, J.L., Chen, Z.F., Chen, H.M., Wang, M.Y., Kong, X., Wang, Y.C., Sun, T.T., Hong, J., Zou, W., Xu, J., and Fang, J.Y. (2014). Elf3 drives β-catenin transactivation and associates with poor prognosis in colorectal cancer. Cell Death Dis. *5*, e1263.

Westgaard, A., Pomianowska, E., Clausen, O.P., and Gladhaug, I.P. (2013). Intestinal-type and pancreatobiliary-type adenocarcinomas: how does ampullary carcinoma differ from other periampullary malignancies? Ann. Surg. Oncol. *20*, 430–439.

Yamamoto, S., Tateishi, K., Kudo, Y., Yamamoto, K., Isagawa, T., Nagae, G., Nakatsuka, T., Asaoka, Y., Ijichi, H., Hirata, Y., et al. (2013). Histone demethylase KDM4C regulates sphere formation by mediating the cross talk between Wnt and Notch pathways in colonic cancer cells. Carcinogenesis *34*, 2380–2388.

Ye, K., Schulz, M.H., Long, Q., Apweiler, R., and Ning, Z. (2009). Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. Bioinformatics *25*, 2865–2871.

Zaret, K.S., and Grompe, M. (2008). Generation and regeneration of cells of the liver and pancreas. Science *322*, 1490–1494.