

# Integrative landscape of dysregulated signaling pathways of clinically distinct pancreatic cancer subtypes

Musalula Sinkala<sup>1</sup>, Nicola Mulder<sup>1</sup> and Darren Patrick Martin<sup>1</sup>

<sup>1</sup>University of Cape Town, School of Health Sciences, Department of Integrative Biomedical Sciences, Computational Biology Division, Observatory, 7925, South Africa

Correspondence to: Musalula Sinkala, email: smsinks@icloud.com

Keywords: pancreatic cancer; bioinformatics; signal transduction pathways; integrative analysis; network analysis

Received: April 30, 2018

Accepted: June 04, 2018

Published: June 26, 2018

Copyright: Sinkala et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

## ABSTRACT

**Despite modern therapeutic advances, the survival prospects of pancreatic cancer patients have remained poor. Besides being highly metastatic, pancreatic cancer is challenging to treat because it is caused by a heterogeneous array of somatic mutations that impact a variety of signaling pathways and cellular regulatory systems. Here we use publicly available transcriptomic, copy number alteration and mutation profiling datasets from pancreatic cancer patients together with data on disease outcomes to show that the three major pancreatic cancer subtypes each display distinctive aberrations in cell signaling and metabolic pathways. Importantly, patients afflicted with these different pancreatic cancer subtypes also exhibit distinctive survival profiles. Within these patients, we find that dysregulation of the phosphoinositide 3-kinase and mitogen-activated protein kinase pathways, and p53 mediated disruptions of cell cycle processes are apparently drivers of disease. Further, we identify for the first time the molecular perturbations of signalling networks that are likely the primary causes of oncogenesis in each of the three pancreatic cancer subtypes.**

## INTRODUCTION

Pancreatic cancer is the most lethal form of cancer. It has an extremely poor prognosis with less than 20% of patients surviving for more than one year following diagnosis [1, 2]. Factors contributing to reduced survival rates are the difficulty of diagnosing the disease during its early stages, the rapid progression of tumours with few specific associated symptoms, and the diversity of responses that different forms of pancreatic cancer have to anticancer drugs [3, 4]. Despite progress having been made towards understanding the histological phenotypes and molecular mechanisms at play, positive responses to conventional chemotherapy regimens have remained infrequent, and the overall survival rates of patients have not substantially improved [5].

A significant challenge to achieving better treatment outcomes has been the heterogeneity of pancreatic cancers. Underlying this heterogeneity is the vast array of somatic mutations that are acquired during oncogenesis, and the varied effects that these mutations have on cell signalling pathways [6, 7]. Recent analyses of genomic sequence

datasets from patients with advanced disease have identified potential activating mutations, many of which occur in genes encoding proteins that might be suitable drug targets [1, 8]. In this regard the discovery of mutation hot-spots in various signalling kinases has already prompted the development of highly selective kinase inhibitors that are capable of specifically killing pancreatic cancer cells. Although the antitumor activities of some of these kinase inhibitors have been strong, they have rarely been long-lasting, with the targeted cancers frequently developing resistance [7]. There is, therefore, a pressing need to identify additional potential drug targets amongst the dysregulated signalling and metabolic pathway components that differentiate pancreatic cancer subtypes. Used in conjunction with kinase-inhibitors, novel drugs targeting these pathway components could yield pancreatic cancer therapies with longer lasting effectiveness.

Aiding in the discovery of novel drug targets has been the use of next-generation sequencing based analytical methods that simultaneously identify mutations in sequences and quantify the expression of all the cellular genes that might have an impact on cancer progression.

In this regard, the Cancer Genome Atlas (TCGA) project, has performed a systematic genomic, transcriptomic and proteomic characterisation of matched healthy and cancerous tissue samples from thousands of individuals afflicted with a variety of cancers [9]. This data, together with matched clinical information is publicly available and includes data for 185 pancreatic cancer patients. Combining data on transcription levels, gene mutations, copy number alteration, protein expression levels and clinical information, the intention of such datasets is to uncover causal relationships between specific genetic and/or cellular aberrations and the onset of disease [10].

Recent genomic studies using such datasets have both provided insights into the biological heterogeneity of pancreatic cancer and identified genomic aberrations that may be of therapeutic and prognostic value [1, 8, 10]. These studies have identified somatic mutations in the proto-oncogene *KRAS* as a hallmark of pancreatic cancers in that more than 90% of pancreatic cancer cases have mutations in this gene [8, 10, 11]. Several other mutations are also strongly associated with the onset of pancreatic cancers, including homozygous deletions in the *TP53*, *SMAD4*, and *CDKN2A* tumour suppressor genes [10, 11]. Alteration in the *KRAS*, *TP53*, *SMAD4* and *CDKN2A* are considered as the critical drivers of pancreatic tumorigenesis and altered signalling through *KRAS* and p53 is associated with varied treatment response to therapy and disease outcomes [11–13]. Nonetheless, as in other cancers, genetic alterations and variations in gene expression also occur in many other genes. Specifically, genes involved in the RB, beta-catenin, PI3K-Akt, and NOTCH pathways, commonly exhibit alterations that likely contributed to tumour development and progression [2, 8, 11]. Further, while tumours displaying *KRAS*-pathway alterations either alone or in combination with *TP53* pathway alterations have a poor prognosis, it has been shown that tumours with more complex pathway disruptions tend to have even poorer outcomes [12, 14].

Notably, these studies have highlighted the alterations in key genes that function in these various signaling pathways. However, these studies have relied on smaller and/or less detailed datasets than those which are now available and have therefore failed to comprehensively define the specific signaling network perturbations that arise during different forms of pancreatic cancer. Here we explore the molecular characteristics of the three major pancreatic cancer subtypes and define the altered signaling pathways and subcellular process that, besides differentiating these subtypes and potentially being the underlying drivers of oncogenesis, also present a variety of potential prognostic biomarkers and drug targets.

## RESULTS

We assembled a TCGA pancreatic ductal adenocarcinoma (PDAC) dataset comprising clinical information for 185 patients together with their associated

cellular transcription data (based on RNA sequencing (RNA-Seq)), protein expression data (based on reverse phase protein array (RPPA)), and information on genomic mutations and copy number alterations (CNA). We performed, survival, clustering, and integrative pathway and network analyses of these diverse data types, both to classify the pancreatic cancers into different subtypes, and to reveal their clinical characteristics and the potential underlying causes of oncogenesis in of these different subtypes.

### Pancreatic cancer subtypes display distinctive clinical outcomes

Based on the mRNA transcription data we identified three major mRNA expression profiles using unsupervised hierarchical clustering (Figure 1A). Upon returning only exemplars for each profile, three PDAC subtypes were identified as: (1) quasi-mesenchymal PDAC (QM-PDAC; 35 samples), (2) classical PDAC (C-PDAC; 87 samples), and exocrine-like PDAC (EL-PDAC; 48 samples) [15, 17]. These three subtypes were associated with distinct overall survival, and duration of disease-free survival (Figure 1A and 1B). Specifically, both overall survival and the duration of disease-free survival were shorter for patients with the QM-PDAC subtype, intermediate for patients with the C-PDAC subtype and longer for patients with the EL-PDAC subtype. We observed a similar trend for treatment outcomes, with patients having QM-PDAC and EL-PDAC respectively displaying the worst and best outcomes (Figure 1C and 1D). Further, we did not observe significant associations between age, gender, or diabetes with the distribution of the PDAC subtypes (Supplementary Figure 1C).

### Gene expression and pathway characteristics of different PDAC subtypes

We compared gene expression profiles between all pairs of PDAC subtypes and identified genes that were differentially expressed within the tumours of each PDAC subtype (see Supplementary File 1). Using gene set enrichment analysis (GSEA), we established that the genes which were differentially expressed between the subtypes were involved in a variety of different signalling pathways [16]. Compared with tumours of the other subtypes, those of the QM-PDAC subtype displayed elevated transcription levels for genes involved in the epidermal growth factor receptor (EGFR) signalling pathway, the transforming growth factor-beta (TGF- $\beta$ ) signalling pathway, the phosphoinositide 3-kinase-mechanistic target of rapamycin (PI3K-mTOR) oncogenic pathway, the mitogen-activated protein kinase (MAPK) oncogenic pathway and among others (Figure 2A). Dysregulation of the EGFR signalling pathway has previously been linked to tumour aggressiveness and reduced patient survival in various cancers including those of the breast and lung

[17, 18]. The PI3K-mTOR pathway was inactive in EL-PDAC tumours but was activated in C-PDAC and QM-PDAC tumours. In other cancers, including those of the breast, gastrointestinal tract and prostate, activation of this pathway has been previously associated with significantly decreased 5-year survival rates [19].

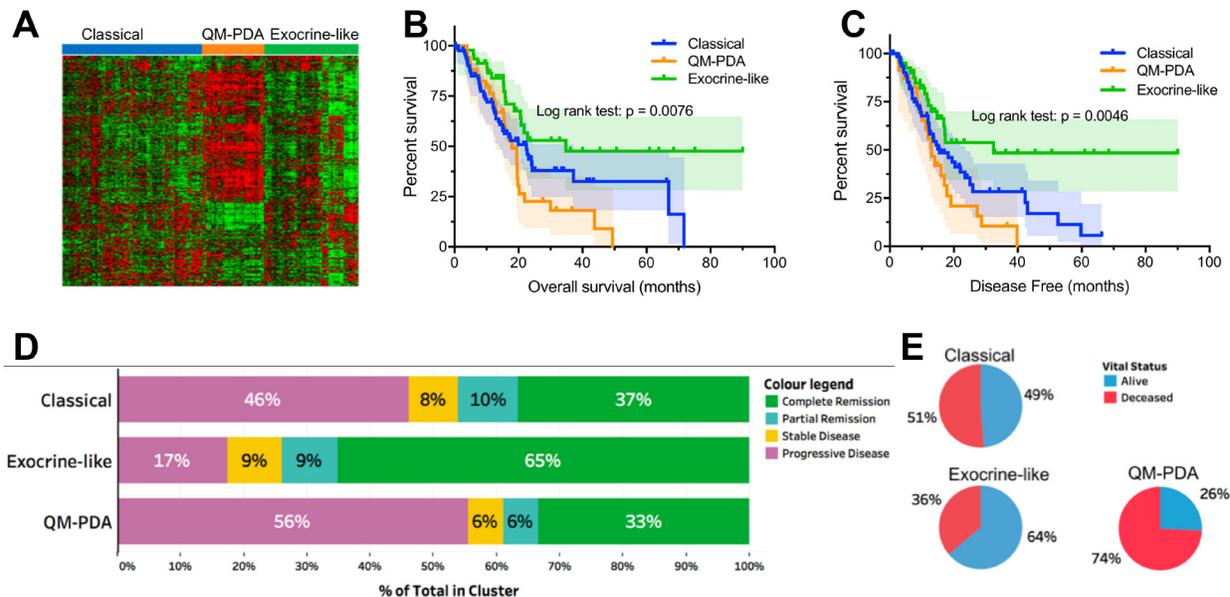
Further, we observed reduced expression of genes involved in electron transport chain and oxidative phosphorylation in the QM-PDAC tumours and, to a lesser degree, in the C-PDAC tumours (Figure 2A and Figure 3A). Since patients with QM-PDAC and C-PDAC tumours exhibited the worst clinical outcomes, these findings are consistent with the hypothesised link between the Warburg effect (characterized by decreased mitochondrial respiration and increased glycolytic activity) and tumour aggressiveness [20, 21]. It is well established that hypoxia-inducible factor 1 (HIF-1), which regulates glucose homeostasis by controlling the expression of multiple glycolytic genes and glucose transporters [22, 23], drives the Warburg phenotypes of various cancers, including those of the lungs and clear renal cells [22, 24].

In support of our findings that QM-PDAC tumours have a Warburg phenotype, we found elevated levels of *HIF1A* and concomitantly lower levels of its corepressor, *SIRT6* (Figure 3B) [25]. Also, we found higher transcript levels of *SLC2A1* and *HK2* in C-PDAC tumours relative to EL-PDAC tumours and the highest *SLC2A3* transcript levels in QM-PDAC tumours. The transcription of *SLC2A1*,

*SLC2A3* and *HK2* is up-regulated by HIF1A [25]. Whereas *SLC2A1* and *SLC2A3* respectively encode the glucose transporters, GLUT1 and GLUT3, *HK2* encodes hexokinase II. Interestingly, among the class 1 GLUT transporters GLUT3 has the highest affinity for glucose and among the hexokinase isoforms, hexokinase II has the highest catalytic efficiency. Further, GLUT 3 and hexokinase II are both reportedly elevated in various cancers [26–28]. This is particularly significant as the combined action of GLUT3 and hexokinase II should afford tumour cells preferential access to available glucose for energy production via glycolysis.

We also found that the transcript levels of certain key glycolytic pathway enzymes varied between PDAC subtypes: these included the transcript levels of pyruvate kinase (*PKM*), lactate dehydrogenase (*LDHA*) and pyruvate dehydrogenase complex kinase-1 (*PDK1*) which were highest in C-PDAC tumours and lowest in EL-PDAC tumours (Figure 3B). Recently, studies have shown that the HIF-1A induced expression of *PDK1* limits the oxidation of pyruvate to acetyl-CoA by inhibiting the pyruvate dehydrogenase complex in cancers of the breast and kidney [29, 30]. Accordingly, we suggest that in QM-PDAC and C-PDAC tumours, the upregulation of *PDK1* and *LDHA* would likely favour the conversion of glucose-derived pyruvate to lactate, thereby promoting the Warburg effect in these PDAC subtypes (Figure 3C).

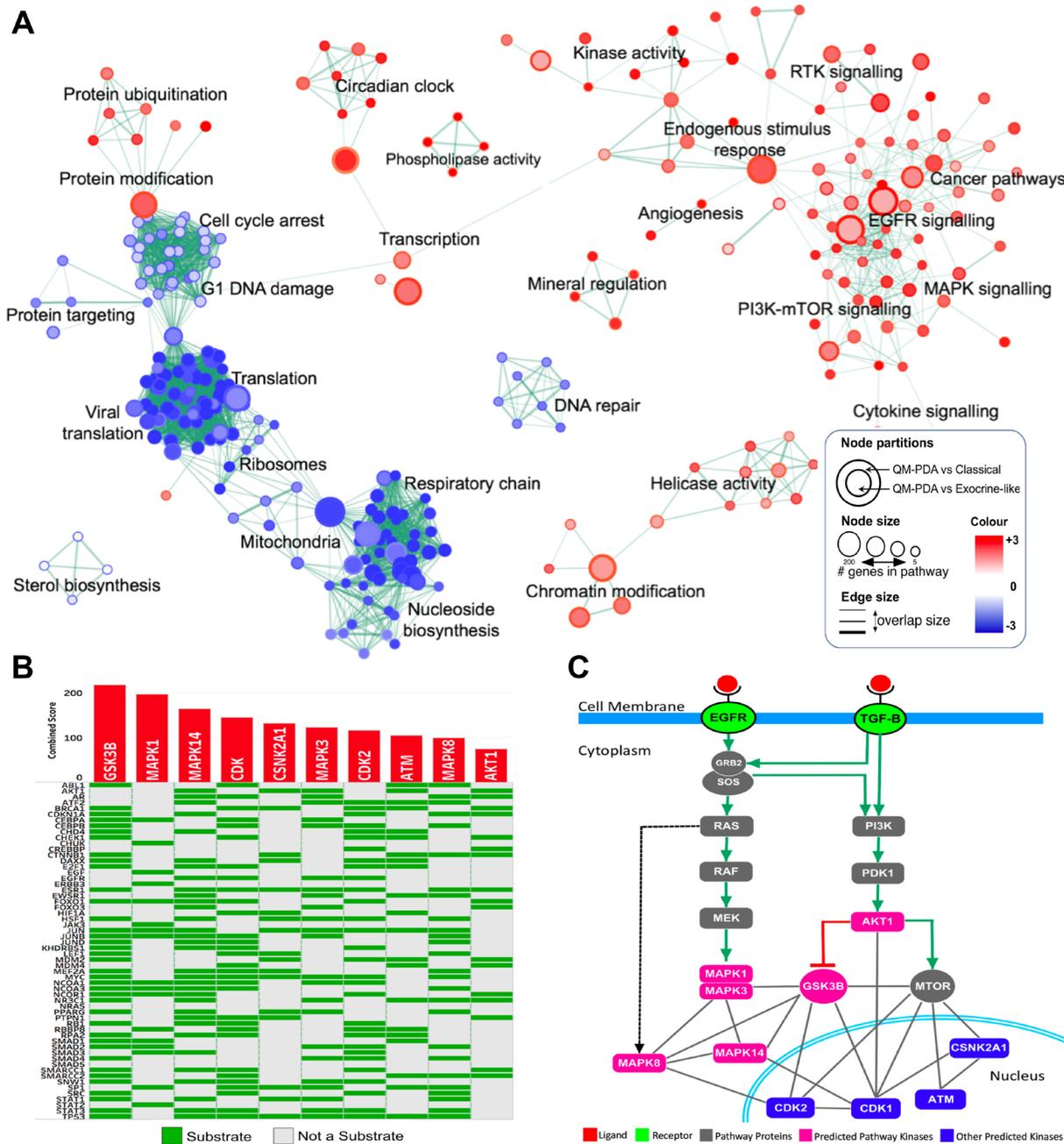
To further investigate the degree of EGFR and TGF- $\beta$  signalling pathway activation in the different



**Figure 1:** (A) Clustering of mRNA expression data identified three major pancreatic cancer subtypes, each with distinct expression patterns. (B) Kaplan–Meier Curves: overall patient survival periods were lower for patients with QM-PDAC and highest for those with EL-PDAC. Pairwise comparisons showed statistically significant differences between: C-PDAC vs. EL-PDAC ( $\chi^2 = 4.4$ ,  $p = 0.036$ ) and QM-PDAC vs. EL-PDAC ( $\chi^2 = 9.7$ ,  $p = 0.002$ ). (C) Kaplan–Meier Curves: disease-free survival months were lower for patients with QM-PDAC and highest for those with EL-PDAC: C-PDAC vs. EL-PDAC ( $\chi^2 = 5.3$ ,  $p = 0.02$ ) and QM-PDAC vs. EL PDAC ( $\chi^2 = 9.2$ ,  $p = 0.002$ ). (D) Vital statistics after the first course of treatment; only a quarter of QM-PDAC patients were alive compared with nearly half of C-PDAC patients and two-thirds of EL-PDAC patients. Odds ratios (95% CI): C-PDAC vs QM-PDAC = 2.7 (1.158–6.568), C-PDAC vs EL-PDAC = 0.541 (0.261–1.122), EL-PDAC vs QM-PDAC = 5.51(1.95–13.33). (E) Treatment outcomes after the first course of therapy were most favourable for EL-PDAC patients and least favourable for QM-PDAC patients.

PDAC subtypes, we mapped mRNA expression levels onto these pathways. We observed that there were higher mRNA levels for genes involved in these pathways in the

QM-PDAC tumours than in the C- and EL-PDAC tumours (Supplementary Figures 3 and 4). Here and subsequently, we opted to compare QM-PDAC to C-PDAC and EL-



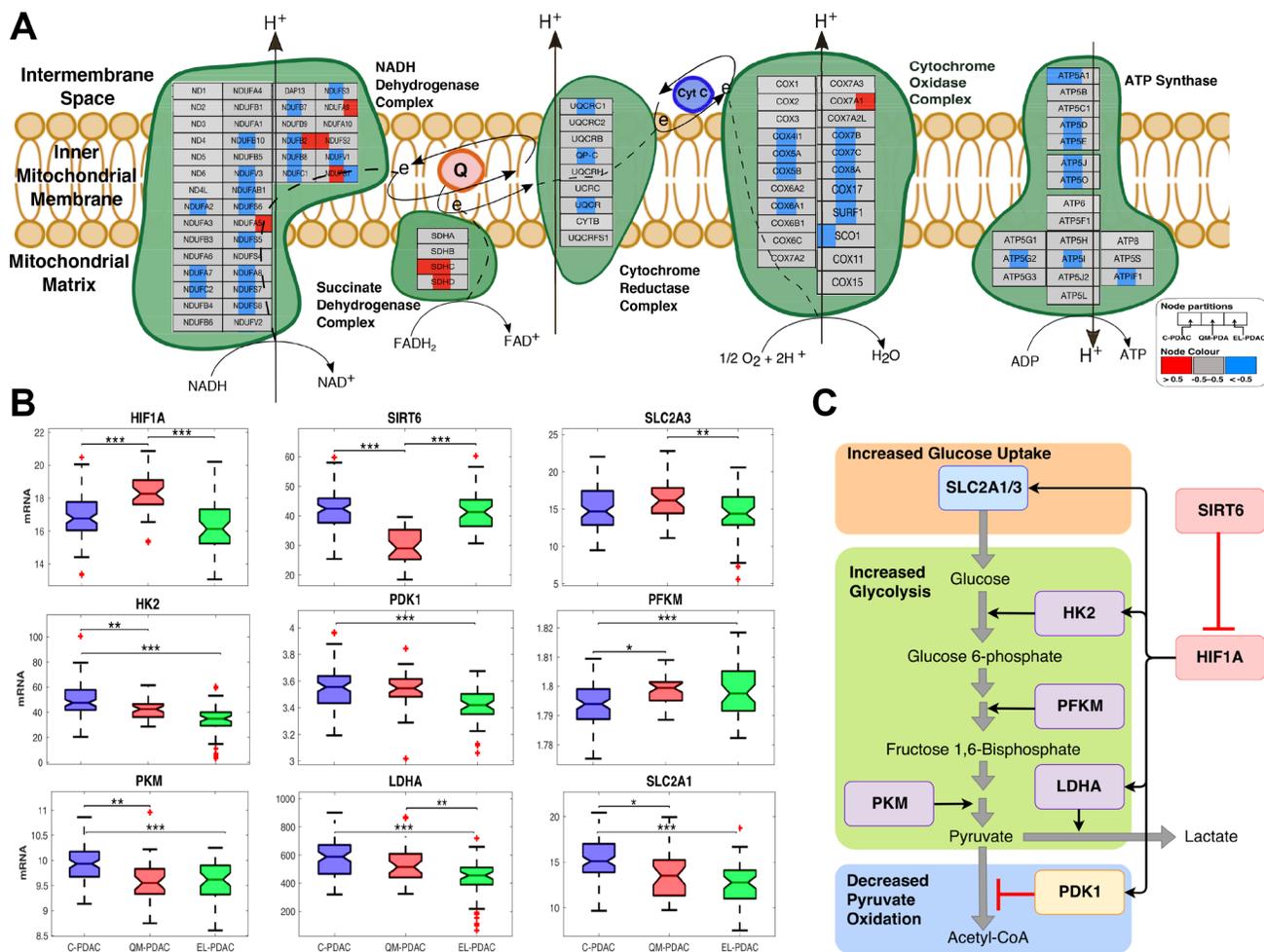
**Figure 2:** (A) Enrichment Map of QM-PDAC vs C-PDAC and EL-PDAC tumours: GSEA was used to obtain enriched gene ontology (GO)-terms that were visualised using the Enrichment Map plug-in for Cytoscape. Each node represents a GO-term with similar nodes clustered together and connected by edges with the number of known interactors between the nodes being represented by the thickness of edges. The size of each node denotes the gene set size for each specific node GO-term. A map comparing C-PDAC and EL-PDAC tumours is shown in Supplementary Figure 2. (B) Expression2Kinases solution: Heat map showing the top ten predicted kinases ranked according to their combined statistical score based on the number of substrates they phosphorylate within a protein-protein interaction subnetwork. Along the rows of the heatmap are proteins which are the substrates for kinases given along the columns of the heatmap. (C) Mapping of the top ten predicted kinases onto simplified models of the EGFR and TGF-β signaling pathways. Six of the top ten ranked kinases (pink nodes) fall within these two pathways whereas the other predicted proteins are involved either directly in the cell cycle, or in the regulation of the cell cycle (blue nodes). We have provided simplified EGFR and TGF-β pathways with mapped mRNA expression levels in Supplementary Figure 5B.

PDAC tumours (referred to collectively as the "other" subtypes) because QM-PDAC has the most distinct molecular signature of the three PDAC subtypes. Many kinases have been previously implicated in carcinogenesis and, accordingly, we identified variations between the PDAC subtypes in the mRNA levels of the various kinases that are involved in the EGFR and TGF- $\beta$  pathways [31]. We used an unbiased computational approach called Expression2Kinases (X2K) to identify the kinases that might be driving the hyperactivation of the EGFR and TGF- $\beta$  pathways in QM-PDAC tumours. Among the top ten kinases identified were six well-documented oncogenes (*AKT1*, *GSK3B*, *MTOR*, *MAPK1*, *MAPK14*, and *MAPK7*), all of which are involved in the EGFR and

TGF- $\beta$  pathways (Figure 2B and 2C) [32]. Also present in the top ten list were *CDK1*, *CDK2* and *ATM*: kinases which are involved in cell cycle control.

### The mutational landscape of PDAC subtypes

We evaluated the scope of genomic alterations in PDAC subtypes by focusing on the types of genetic changes that are known to promote oncogenesis. Specifically, these encompassed gain-of-function mutations in oncogenes (OGs), amplification of OGs, loss-of-function mutations in tumour suppressor genes (TSGs), and deletions in TSGs. Across all the PDAC subtypes we found that, as has been reported elsewhere, *KRAS*, *TP53*, *CDKN2A*, *SMAD4*,



**Figure 3:** (A) PDAC subtype-specific electron transport chain activity: a comparison of electron transport chain activity between pancreatic cancer subtype based on mRNA expression data. Node denote genes— left section = C-PDAC, middle section = QM-PDA and right section = EL-PDAC tumours. Node are coloured based on overall subtype mRNA-expression z-score (blue = low, grey = no change, and red = high). Edges represent various types of protein interaction (refer to legend to full notations for all edges). (B) PDAC subtype-specific transcript levels of Warburg effect associated mRNA: levels were compared between the tumour subtypes using one-way analysis of variance. The data are where transformed using the Box-Cox transformation. \*\*\*, \*\*, and \*, denote pairwise student *t*-test statistical significance for *p* values of 0.001, 0.01 and 0.05, respectively. On each box, the central mark indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers, and the outliers are plotted individually using the '+' symbol. (C) Control of glucose metabolism by HIF1A: SLC2A1, SCL2A3, HK2, LDHA, PDK1 are positively regulated (black arrows) by HIF1A. PDK1 negatively regulates (red blunt arrows) the conversion of pyruvate to acetyl-CoA.

and *CDKN2B* were the most commonly altered genes (Figure 4A) [1, 2, 8].

SMAD4 signals through the canonical TGF- $\beta$  pathway and therefore deletions in *SMAD4* should limit signalling through this pathway [33]: a factor that may seem inconsistent with our earlier finding that genes in this pathway display elevated levels of transcription (Figure 2C and Supplementary Figure 4). However, SMAD4 loss does not initiate tumorigenesis in human pancreatic cancers [33–35]. Further, in pancreatic tumours displaying either *SMAD4* deletions or SMAD4 under-expression, ligand stimulation of the TGF- $\beta$  pathways activates non-canonical pathways including the MAPK, p38/JNK, and PI3K-mTOR pathways which can function independently of SMAD signalling [33, 36, 37]. Moreover, we found that among the PDAC subtypes, QM-PDAC had the lowest transcript levels of genes in the SMAD gene family (*SMAD2/3/4/7*; Supplementary Figure 4). Therefore, consistent with other studies, these results indicated an association between reduced SMAD expression and poor survival [36, 38].

Whereas we observed higher gene deletion frequencies in EL-PDAC tumours than in QM-PDAC and C-PDAC tumours, QM-PDAC and C-PDAC tumours displayed higher gene amplification frequencies (Figure 4C). Across all the PDAC subtypes, 118 genomic alterations were observed, mostly impacting genes involved in diverse cell signalling pathways (Figure 4B and 4D); a finding consistent with the hypothesis that, like most other cancers, PDAC is primarily a consequence of disrupted signal transduction pathways (Figure 4D, 4E) [11, 39].

We uncovered little overlap in the signaling pathways that were impacted by genetic alterations that were observed in only one of the PDAC subtypes (Figure 4E and 4F, also see Supplementary Figure 6). We found that 62% of the pathways affected by mutations were altered in tumours belonging to at least two PDAC subtypes, whereas only 18%, 11%, and 8% of altered pathways were unique to C-PDAC, QM-PDAC, and E-PDAC tumours, respectively. This suggests that only a small proportion of mutations contribute to differences in the signaling pathway perturbations that are seen between the subtypes. Conversely, it therefore also suggests that most of the oncogenesis promoting mutations perturb pathways that are active in all three of the PDAC subtypes.

### Integrative pathway analysis

We used the co-occurring mutated driver pathway (CoMDP) mathematical algorithm to discover *de novo*, two co-occurring pathways that may be driving the progression of PDAC; the first pathway involved the genes *KRAS*, *COL4L4* and *FBWX7*, and the second involved the genes *SMAD4*, *TP53*, *PHF24*, *PRG4*, *PI3KCA*, *RPTOR* and *EP300* (Figure 5A) [40]. We expanded the CoMDP solution pathways using known protein-protein interactions to generate a network enriched with *MAPK*, *PI3K*, and

*TP53* pathway members (Supplementary Figure 7). Two of the genes in the CoMDP solution pathways, *PRG4* and *PHF24*, did not map to any signal transduction pathway, emphasising the fact that the roles of some potentially key proteins in oncogenesis still need to be defined. In particular, virtually nothing is known about *PHF24* [41].

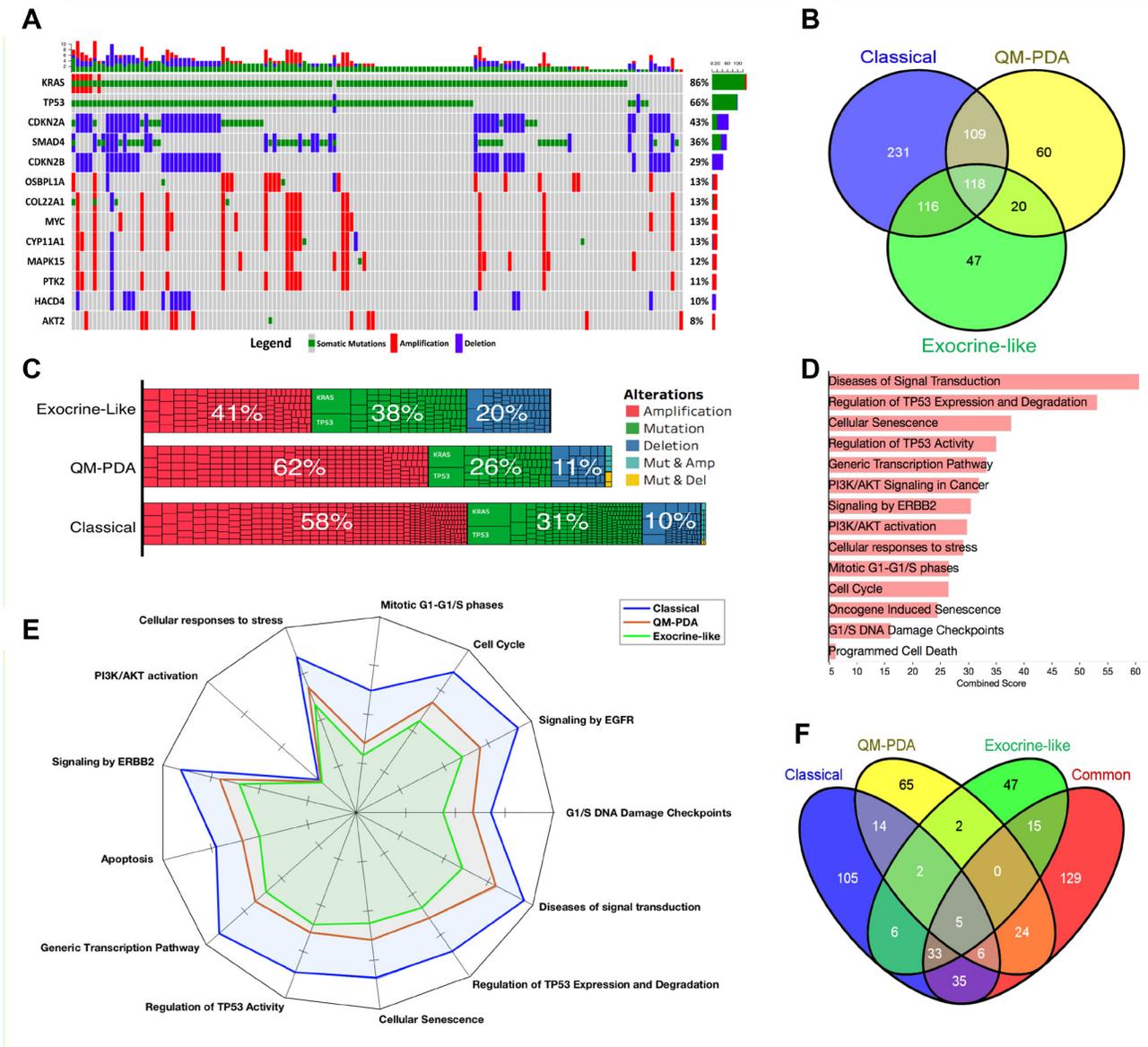
To further investigate the degrees of alteration that are evident in the expanded CoMDP solution networks for the different PDAC subtypes, we mapped a combined dataset of mRNA transcript levels, protein expression levels, mutations and CNA onto these networks. We found that alterations in p53 and cell cycle checkpoint pathway genes were most apparent in C-PDAC tumours (Figure 5B and 5C), whereas alterations in specific MAPK and PI3K-mTOR pathway genes were more apparent in QM-PDAC tumours (Figure 5D and 5E). We found that the PI3K-mTOR and MAPK pathways were altered in 93% of all PDAC tumours. Alterations, included the activation of, among others, the *PI3KCA* (in 13% of tumours), *AKT* (in 27%) and *BRAF* (in 9%) genes, and the inactivation of the *PTEN* (in 4% of tumours), *TSC1/2* (in 7%) and *FOXO3* (in 5%) genes. Alterations in the *PI3KCA* oncogene and its negative regulator, *PTEN*, occur in cancers of the colon, breast, and prostate: cancers where the co-occurrence of *PI3KCA* and *PTEN* mutations appears to both drive oncogenesis, and reduce anticancer drug sensitivity [10, 42]. Furthermore, we observed that *PIK3R1* (the regulatory subunit of PI3K) was inactivated in 6% of PDAC tumours; inactivated *PIK3R1* promotes the phosphorylation of *AKT*, which itself promotes oncogenesis as it activates numerous OGs and inhibits TSGs within the cell [43, 44].

Alterations in the p53 and cell cycle pathways are frequent in cancer, and here we found that such changes were apparent in 85% of the tumours examined [45]. In addition to activated *MDM2* and *MDM4* (which both inhibit p53 activity), *MYC*, and *CDK2/4/6*, we found that p53 was inactivated in 58% of all tumours. Similarly, *CDKN2A*, *CDKN2B* and *ATM* (a kinase that activates p53) were inactivated in 35%, 42% and 3% of all tumours, respectively [45, 46]. This suggests that, within the p53 and cell cycle checkpoint pathways, hyperactivated OGs such as *MYC*, *MDM2*, *CDKs* and inactivated TSGs such as *TP53*, *CDKN2A* and *CDKN2B*, may act together to promote oncogenesis both by limiting the repair of damaged DNA, and by permitting affected cells to proliferate uncontrollably through the inhibition of apoptosis [45, 46].

Consistent with previous observations, we found mutations, CNAs and changes in mRNA transcription and protein expression levels for proteins that participate in various signalling pathways that have previously been associated with pancreatic cancer [2, 8, 10]. Specifically, we observed alterations in the Notch (61%), apoptosis (32%) and NF- $\kappa$ B (19%) pathways (Supplementary Figure

8). Also, we found a variety of alterations in 41 genes that encode the ATP-binding cassette (ABC) transporter proteins in 72%, 82% and 79% of all tumours that belong to the C-PDAC, QM-PDAC and EL-PDAC subtypes, respectively (Supplementary Figure 8D). The most altered ABC transporter gene in any PDAC subtype was *ABCC9*; found altered in 21% of all in QM-PDAC tumours. Overall, we observed that 78% of all PDAC tumours harboured a

genetic alteration in at least one ABC transporter gene. ABC transporter-mediated energy-dependent efflux of a multitude of unrelated classes of anticancer drugs across membranes is a major cause of multidrug resistance and chemotherapeutic failures during cancer therapy [47, 48]. Therefore, future efforts to determine tumour cell ABC transporter gene mutations that accentuate the activities of their encoded transporters are expected to guide precision medicine [48].

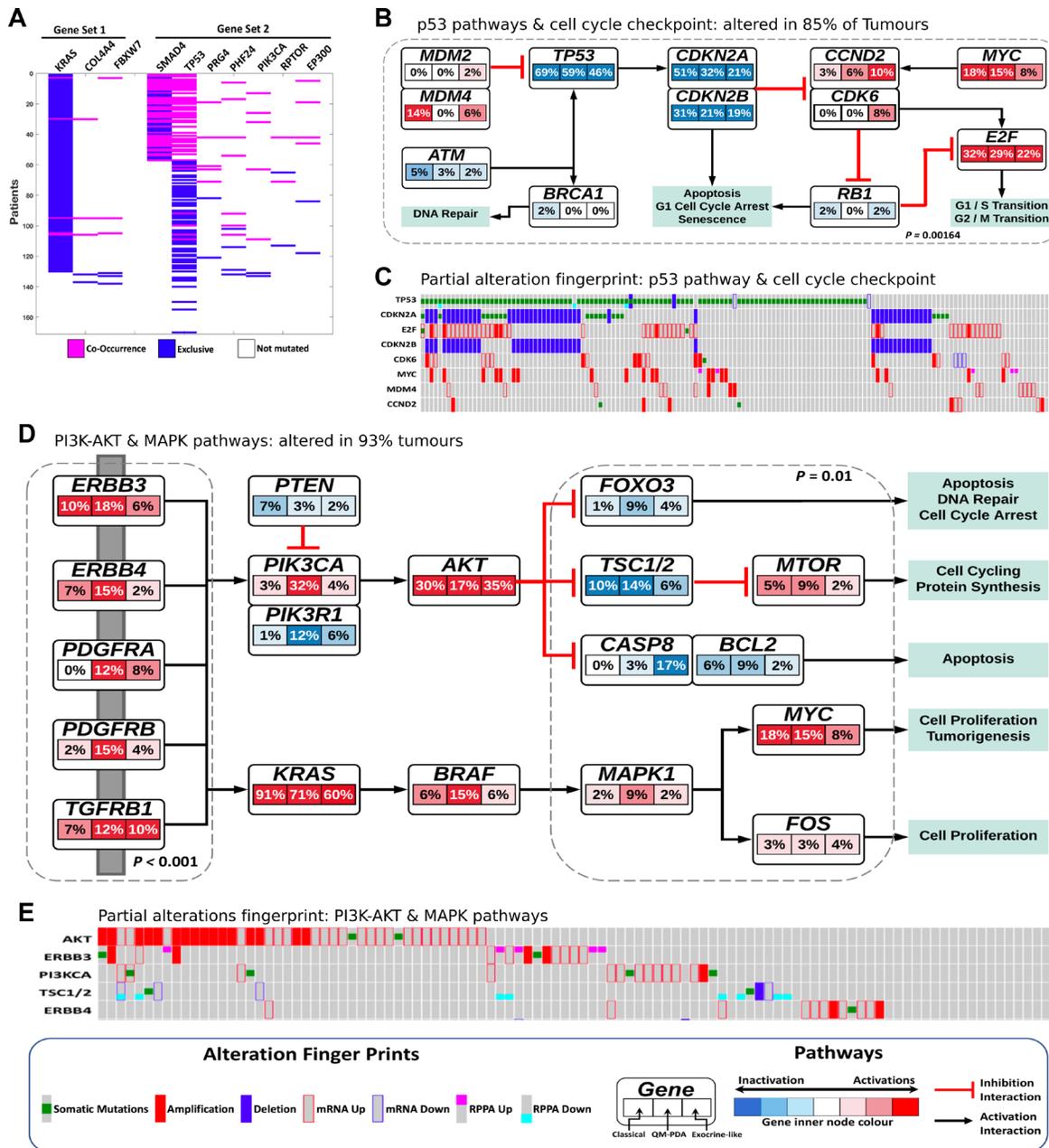


**Figure 4: Mutation and gene copy number analyses.** (A) Genes with the most alterations in PDAC tumours. The only genetic alterations considered are mutations in, and amplifications of known oncogenes, and mutations in, and deletions of, known tumour suppressor genes. (B) The distribution of alterations among the three PDAC subtypes. Refer to Supplementary File 5 for details concerning alterations in each set. (C) The extent of genomic alterations expressed as a percentage of total numbers of alterations found within the tumours of each PDAC subtype. Each cell in the bar-grid represents a mutant gene. (D) Reactome pathway enrichment results of the 118 genes that are commonly altered in tumour cells of all three PDAC subtypes. Refer to Supplementary File 6 for the complete list of Reactome pathways that represent significantly enriched genetic alterations in the different PDAC subtypes. (E) The predicted extent of mutation-induced pathway dysregulation for the different PDAC subtypes. (F) The distribution of mutation-induced pathway dysregulations for mutations specifically associated with particular PDAC subtypes and common pathway enrichment. The non-overlapping pathways are those disrupted only in single PDAC subtypes (see Supplementary File 6).

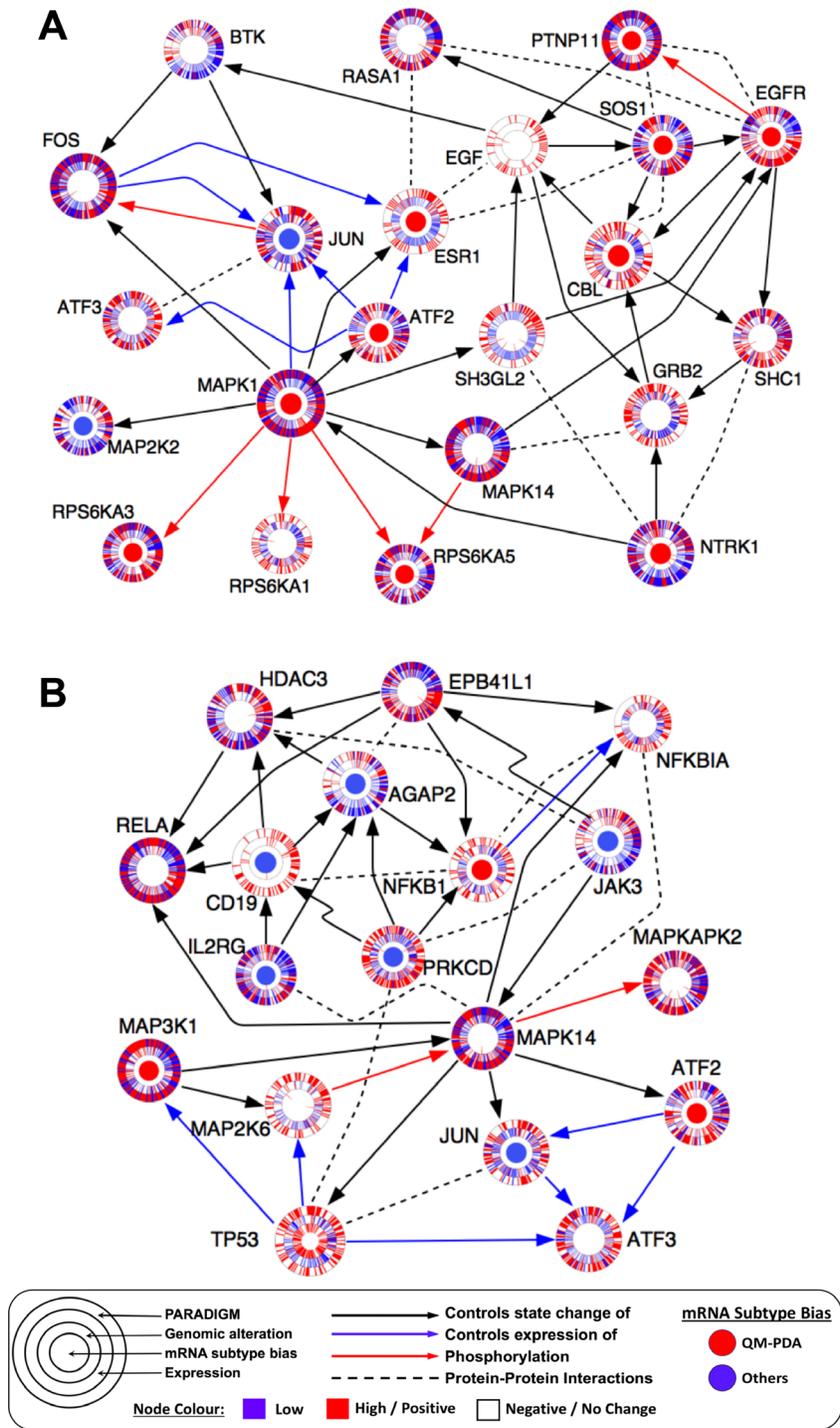
## Connectivity of genomic alterations to transcription factors and their pathway activities in pancreatic cancer

To link genomic changes to transcriptional events, we applied the Tied Diffusion Through Interacting

Event (TieDIE) approach to reveal a protein interaction subnetwork that connects altered genes to transcription factors and their putative targets [49]; a network referred to below as the TieDIE subnetwork. Additionally, we used the PARADIGM-shift algorithm to infer pathway activity levels of all proteins that are known to participate



**Figure 5:** (A) The two co-occurring pathways in PDAC that are predicted to drive oncogenesis based on the mutational landscape representing two common driver pathways. (B) Alterations in p53 and cell cycle checkpoint pathways.  $p$ -value = C-PDAC vs Other subtypes, calculated using Fisher's exact test. Red indicates activating genetic alterations whereas blue indicates inactivating alterations. Darker shades correspond to higher alteration frequencies. Each node within the pathway represent a gene and the highlighted segments within each node and the percentage representing the alteration in the three PDAC subtype: C-PDAC, QM-PDAC and EL-PDAC from left, centre, and right, respectively. (C) The pattern of genetic alteration in selected genes that encode proteins involved in the p53 and cell cycle checkpoint pathways. (D) Alterations in the MAPK, RTKs and PI3K signaling pathways.  $p$ -values = QM-PDAC vs Other subtypes, calculated using Fisher's exact test. The connectivity of network components was extracted from Reactome Pathways, BioGrid and the literature. (E) The pattern of genetic alterations in selected genes that encode components of the MAPK and PI3K-mTOR pathways.



**Figure 6:** (A) MAPK heat diffusion sub-network: pathway extracted from the TieDIE subnetwork using heat diffusion analysis from the MAPK1 network node. (B) p53 heat diffusion sub-network: pathway extracted from the TieDIE solution network using heat diffusion analysis from the TP53 network node. Each node indicates a pathway protein shown as concentric rings. The inner node denotes differential mRNA expression (Benjamini-Hochberg adjusted  $p < 0.05$ ) bias for genes when comparing the QM-PDAC subtype tumours to those of the other PDAC subtypes (red = QM-PDAC bias, blue = bias towards other subtypes). The second ring indicates the presence of genomic alterations for that gene in each patient's tumour, with each patient's tumour being denoted by a spoke within the ring. The third ring shows mRNA expression levels for each tumour sample (red = high, blue = low). The outer ring indicates the PARADIGM inferred pathway activity for that protein in each tumour sample (red = high, blue = low). Arrows indicate known protein-protein interactions extracted from UCSC Super pathway, KEA, ChEA or inferred from the literature. We have attempted to make the visualisation clearer by omitting some interactions between some network nodes.

in various signaling pathways in each of the three PDAC subtypes [50]. Furthermore, using heat diffusion analysis from the MAPK1 and TP53 nodes of the TieDIE subnetwork, we extracted two pathways that recapitulated signaling via the MAPK and p53 pathways.

The MAPK1 network was enriched with proteins whose associated mRNA transcription levels were significantly higher in QM-PDAC tumours compared to other PDAC subtypes. These proteins included EGFR (a receptor of the EGFR pathway that we found activated in QM-PDAC tumours), SOS1 and GRB2 (both of which are upstream signalling proteins in the canonical MAPK signalling pathway; Figure 6A) [51]. Also, the MAPK network connected TFs that are induced upon activation of the MAPK pathway, to proteins which are known to promote oncogenesis (such as FOS, JUN, ATF2 and ESR1). This inferred connectivity was further supported by the PARADIGM analysis which predicted that ESR1, JUN, GRB2 and CBL would have high degrees of activity [51, 52].

The p53 network, on the other hand, was more prominent in C-PDAC and EL-PDAC tumours, and it connected signalling proteins to various TFs that are known to promote carcinogenesis, including ATF3, PRKCD, NFKB1 and NFKBIA [53–55]. These TFs were also predicted by PARADIGM to have a high degree of activity (Figure 6B).

Collectively these analyses emphasise that certain pathways may be more prominent than others in the different PDAC subtypes. Differences between the PDAC subtypes in the activity of specific signal transduction proteins suggests that some of these proteins could be targets of PDAC subtype-specific anti-cancer drugs (see Supplementary Figure 10 and Supplementary Table 2).

## DISCUSSION

Through comprehensive transcriptomic and integrative profiling of pancreatic cancer, we have uncovered various functional alterations and signaling pathway perturbations and revealed how these alterations and perturbations might be associated with clinically relevant differences between patients with different PDAC subtypes. In particular, the discovery that QM-PDAC tumours are characterised by what is likely to be ESR1 and NTRK1 transcription factor-mediated over-activation of genes associated with the EGFR and TGF- $\beta$  pathways (Figure 6A), provides a rationale to target these tumours with drugs that either downregulate ESR1 and NTRK1, or inhibit EGFR and TGFBR2 (Supplementary Figure 10) [56, 57].

Furthermore, we find that, in general, PDAC is characterised by pervasive RTK, MAPK and PI3K-mTOR alterations [1, 8, 58, 59] and that over 90% of the potentially oncogenic alterations occur in genes that are directly involved in the RTK, MAPK and PI3K-mTOR

signalling pathways. It is well established that PDAC tumours frequently respond to MAPK and/or PI3K-mTOR pathway inhibitors [58, 60]. While cases where these inhibitors have failed to provide therapeutic benefits have highlighted the heterogeneity of PDAC, they have also emphasised the importance of finding additional drugs that are either more generally applicable to PDAC treatment, or which can be used to target the signalling pathways that are most relevant for specific PDAC subtypes [60–62]. We identified that the most prominent cell signalling changes in EL-PDAC and C-PDAC tumours were within the p53 and cell cycle checkpoint pathways; hinting that these tumours might respond to cell cycle inhibitors. Consistent with our findings, other PDAC studies have also reported co-occurring mutations in genes involved in the p53 and cell cycle pathways. Collectively these studies provide a rationale for potentially treating PDAC using an approach that synchronously targets all of these pathways [63–65]. Furthermore, we have uncovered other receptors, intermediary signal transduction proteins and TF targets that may drive oncogenesis: all of which could be targeted by drugs designed to specifically treat QM- C- or EL-PDAC tumours.

Alterations in metabolism and cellular bioenergetics are hallmarks of cancer cells and represent an active area of research that is anticipated to yield novel anti-cancer drugs that could be used in combination with targeted-therapies or chemotherapy [66–68]. Here, we found that QM-PDAC and, albeit to a lesser extent, C-PDAC tumours exhibit a Warburg metabolic phenotype (Figure 3) [64]. Associations between the Warburg phenotype and both increased disease aggressiveness and poorer clinical outcomes have been previously reported [64, 67]. As expected, we observed decreased overall survival and a shorter duration of disease-free survival in patients with QM- and C-PDAC tumours (i.e. tumours with the Warburg phenotype) relative to patients with EL-PDAC tumours (i.e. those without the Warburg phenotype). In this regard, scrutinising the metabolic differences between PDAC tumour subtypes is likely to yield further leads for the development of novel therapeutic approaches.

We observed that most of the genomic alterations which are found within PDAC tumour cells are found in tumours belonging to all three of the defined PDAC subtypes. This finding suggests that improved responses to targeted-therapies may be achievable by systematic targeting of hub kinases within the multiple alternative signalling pathways that enable cancer cells to frequently acquire resistance [69, 70].

By integrative analyses of genomic, transcriptomic, and proteomic data, we have uncovered novel signalling pathway aberrations that exist in PDAC tumour cells at the DNA, mRNA and protein levels. Although the mutational landscape of a particular tumour could influence drug efficacy, recent studies of a range of other cancers employing approaches similar to those used here have

identified an array of altered pathways, each of which could present either new cancer drug targets or clinically relevant biomarkers of disease [9, 71]. Altogether, our analyses have revealed widespread signaling network perturbations in PDAC subtypes, many of which could likely impact treatment outcomes and which are therefore also potential targets for novel anticancer drugs.

## MATERIALS AND METHODS

We obtained data for 185 PDAC patients involved in the TCGA project. Besides treatment outcomes these data include: whole exome sequence (WES;  $n = 185$ ), transcriptome data (determined using RNAseq;  $n = 179$ ); DNA copy number and mutation data ( $n = 179$ ), and targeted proteome data (determined using RPPA;  $n = 123$ ). Not all types of data were available for all patients because of assay failures, incomplete specimen availability and quality of issues with certain samples. All data used in our analyses are available from the TCGA website; <https://portal.gdc.cancer.gov/repository>.

### Transcriptome-based classification

We performed unsupervised hierarchical clustering on RNA-seq data to identify three distinct PDAC subtypes. Before clustering, we removed data for unexpressed genes and genes that exhibited little variation between patients. To return only exemplars for each cluster, we applied an anomaly detection algorithm based on an approximate Gaussian distribution [72]. Finally, we further validated the consistency of tumours within each cluster using a support vector machine classifier which yielded an average 10-fold cross validation classification accuracy of 95.5% over ten models (Supplementary Figure 1A). Using the transcriptomic classification framework established by Collision *et al.* [15], we classified the pancreatic cancer clusters as C-PDAC, QM-PDAC and EL-PDAC; respectively corresponding to clusters 1, 2 and 3 [15]. We have summarised the distribution of tumour grades across these PDAC subtypes in Supplementary Figure 1B.

### Treatment outcomes

We integrated mRNA expression-based classification of PDAC subtypes with clinical information to review tumour characteristics specific to each of the PDAC subtypes. The Kaplan–Meier method was used to estimate overall survival and the duration of disease-free survival in a pairwise manner between subtypes [73]. Furthermore, the Fisher exact test was used evaluate associations between tumour subtypes and various clinical variables including treatment outcomes at the first, and later courses of treatment.

## Differential gene expression, functional and pathways analyses

The identification of differentially expressed genes was performed in MATLAB using an implementation based on the negative binomial model (see Supplementary File 1) [74, 75]. Gene set enrichment analysis (GSEA) was employed to extract knowledge of overrepresented Gene Ontology (GO) terms for various functional processes and signalling pathways between molecular subtypes [16]. Complete GSEA results are provided in Supplementary Files 2A–2C, for C-PDAC vs QM-PDAC, C-PDAC vs EL-PDAC and QM-PDAC vs EL-PDAC, respectively. Visualisation of significantly enriched GO terms of functional process and signalling pathways between subtypes was done in the Cytoscape plugin, Enrichment Map [76]. Furthermore, the mapping of gene expression levels onto template WikiPathways of the EGFR and TGF- $\beta$  signalling pathways and the electron transport chain was done using the software PathVisio 3 (See Figure 3A, Supplementary Figure 3 and 4) [77]. For this, we used z-score normalised expression data categorised into three levels: 1) Low for z-scores below  $-0.5$ ; 2) no change for z-scores between  $-0.5$  to  $0.5$ ; and 3) high for z-score above  $0.5$ . The highlighted scale was chosen to consistently capture variations in gene expression across entire pathways.

### Prediction of regulator kinases

We computationally predicted upstream regulatory kinases that likely effect the observed differences in the gene expression signatures between QM-PDAC and the other PDAC subtypes using Expression2Kinases (X2K) [78]. X2K employs a reverse engineering network-based approach to predict upstream regulatory kinases based on prior knowledge. We obtained a list of differentially expressed genes between QM-PDAC and the others PDAC subtypes: 242 up-regulated genes and 1011 down-regulated genes based empirical Bayes statistics. Using this gene list, we predicted upstream regulatory TFs that are likely to be responsible for the observed changes in gene expression using the Chromatin Immunoprecipitation (ChIP) Enrichment Analysis (ChEA; 2016) [79]. In the next analysis, we linked the top 10 predicted TFs to upstream regulatory mechanisms by generating a TF-intermediate protein-protein interaction sub-network based on prior knowledge (Supplementary Figure 5A). Finally, we analysed the sub-network for enriched targets of known protein kinases that are likely to phosphorylate proteins within the sub-network using the Kinase Enrichment Analysis (KEA; 2015) [80]. See supplementary file 3 for a full list of computationally predicted kinases and their rankings.

## Mutation and copy number alteration analyses

We evaluated the scope of genomic alterations in PDAC subtypes using significantly mutated genes and copy number alteration identifications obtained from MutSigCV and GISTIC2.0 outputs, respectively [81, 82]. Data for the genomic alteration analysis was processed as follows: oncogenes (OGs) and tumour suppressor genes (TSGs) in the samples were annotated using information from multiple sources. These include the Sanger Consensus Cancer Gene Database (699 OGs and TSGs), the UniProt Knowledgebase (304 OGs and 741 TSGs), the TSGene database (1,219 TSGs) and the ONGene database (725 OGs) [34, 86–88] [32, 83–85]. Collation of data from these sources yielded a list of 3,688 OGs and TSGs, representing 2,773 unique genes (969 OGs and 1,804 TSGs). We utilised this list of OGs and TSGs to extract genetic changes anticipated to have a potential impact on the oncogenesis of pancreatic cancer. Explicitly, we returned only gain-of-function mutations and gene amplifications for known OGs. Also, for known TSGs, we returned loss-of-function genetic changes that involve mutations and deletions. Using this processed data, we identified frequently altered genes that likely have detrimental impacts concerning pancreatic carcinogenesis (Figure 4). We compared gene mutations between the PDAC subtypes to generate lists of mutations that are common among subtypes or unique to particular subtypes (see Supplementary File 4). Using these lists, we performed a Reactome pathway enrichment analysis by querying Enrichr either with genes that were consistently altered in tumours of all three PDAC subtypes, or with genes that were altered in only one of the PDAC subtypes (see Supplementary File 5 and Supplementary Figure 6A–6C) [86].

## Integrative analysis of expression and genomic alterations

### Identification of Co-occurring driver pathways

To discover driver pathways based on the patterns of mutations associated with PDAC, we applied the CoMDP algorithm which employs a mathematical programming method to identify *de novo* driver pathways in cancer from mutation profiles [40]. Briefly, this method identifies pathways that have a set of mutated genes with both high coverage (i.e. present in the tumours of multiple individuals) and high exclusivity, and the pathways exhibit a statistically significant co-occurrence pattern. Using mutation data, we ran the CoMDP test with  $K = 5$  to 11 ( $K$  equals the total gene set size) to return mutated driver pathways for all  $K$  values (Supplementary Table 1). Genes in the CoMDP solution for  $K = 10$  were connected using experimentally verified protein-protein interactions to generate an intermediary network (Supplementary Figure 7). The solution network was enriched with members of the PI3K, MAPK, p53 and cell cycle regulation pathways. To visualise the extent of pathway aberration at DNA, mRNA

and protein levels, we mapped genomic alteration data, mRNA transcript abundance data and protein expression data onto the network. For the genomic alteration data, we only considered gain-of-function mutations and gene amplifications for the OGs, and loss-of-function mutations and deletions for the TSGs. For the transcription and protein expression data we only considered OGs that had a degree of upregulation indicated by a  $>2$  Z-score and for the TSGs a degree of downregulation indicated by a  $<-2$  X-score. The generated combined dataset was mapped on signaling pathways over-represented in the CoMDP solution network expanded using a prior-knowledge network. Additionally, plots of alteration patterns in genes among tumours were generated using the R package complex heatmaps [40, 87]. Mapping of alterations onto genes in pathways shown in Supplementary Figure 8A–8C were done using the software PathwayMapper [88].

## Inferring gene activity from pathway analysis of copy number and expression data

We used PARADIGM-shift, a probabilistic graphical model approach that infers the activity of signalling pathway proteins by detecting differences in the expected activity of a protein on its downstream target relative to what is expected given its upstream modulator [50]. We ran PARADIGM with default settings using three datasets as inputs: (i) a dataset including only statistically significant CNA as determined by GISTIC2, (ii) a normalised gene expression dataset matching the CNA input file, and (iii) a custom UCSC Pathway formatted file. Pathway information of known gene interactions was created from various sources including Reactome pathways, KEGG Pathways, the KEA database, the ChEA database and the UCSC Super pathway [79, 80, 89, 90]. PARADIGM predicted integrated pathway levels results are provided in Supplementary File 6.

## Identification of genomic perturbation associated with transcriptional changes

Genomic perturbations in PDAC subtypes were connected to associated transcriptional changes using TieDIE [49]. This method uses a heat diffusion process to identify relevant pathways that might be altered in tumours. To reveal sub-networks that distinguish QM-PDAC from the other PDAC subtypes, using genes that we found altered in at least 5% of all tumours, we generated a ranked list of genes that were differentially mutated between QM-PDAC tumours and those of the other PDAC subtypes using the Fisher's exact test. The resulting genes are assumed to be responsible for the distinctive molecular signatures between subtypes—these were used as upstream inputs in TieDIE. A downstream input file was generated by computationally identifying the TFs that are most likely to be responsible for the difference in the

transcriptome signatures between the QM-PDAC tumours and those of the other PDAC subtypes. The upstream and downstream input files, together with a custom super pathway, were used in TieDIE to compute a subnetwork connecting genomic alterations to transcriptional events (Supplementary Figure 9). We performed a secondary heat diffusion query on the TieDIE solution sub-network from the MAPK1 and TP53 nodes. Both MAPK1 And TP53 were flagged as being of likely importance based on both the numerous alterations of these genes within PDAC tumours and the pathway analysis that we had previously performed. These analyses produced two subnetworks that recapitulate signaling through the MAPK1 and TP53 pathways to their downstream TFs.

### Statistical analyses

Except were stated otherwise all statistical analyses were performed in MATLAB 2017b. The Fisher's exact test was used assess associations between categorical variables. The Wilcoxon rank sum test and Kruskal–Wallis test or independent sample Student *t*-test and One-way ANOVA were used for continuous variables were appropriate. Statistical tests were considered significant at  $p < 0.05$  for single comparisons, and Benjamini-Hochberg adjusted *p*-values for multiple comparisons.

### Ethics approval

The University of Cape Town; Health Sciences Research Ethics Committee (HREC) IRB00001938 approved the protocol of this study.

### Abbreviations

TCGA: The Cancer Genome Atlas; PDAC: Pancreatic Ductal Adenocarcinoma; QM-PDAC: quasi-mesenchymal PDAC; EL-PDAC: exocrine-like PDAC; C-PDAC: classical PDAC; mRNA: Messenger Ribose Nucleic Acid; RPPA: Reverse Phase Protein Array; RNA-Seq: RNA-sequencing; CNA: Copy Number Alteration; EGFR: Epidermal Growth Factor Receptor; GSEA: Gene Set Enrichment Analysis; TGF- $\beta$ : Transforming Growth Factor–Beta PI3K: Phosphoinositide 3-Kinase; mTOR: Mammalian Target of Rapamycin; MAPK: Mitogen Activated Protein Kinase; TFs: Transcription Factors; OGs: Oncogenes; TSGs: Tumour Suppressor Genes.

### Author contributions

Conceptualization, M.S., N.M., and D.P.M.; Methodology, M.S., N.M., and D.P.M.; Formal Analysis, M.S., and D.P.M.; Writing – Original Draft, M.S., and

D.P.M.; Writing – Review & Editing; M.S., N.M., and D.P.M.; Visualisation, M.S.; Supervision, N.M., and D.P.M.

### CONFLICTS OF INTEREST

The authors declare that they have no competing interests.

### FUNDING

The H3ABioNet NIH Common Fund funded this study, grant number: U41HG006941.

### REFERENCES

1. Bailey P, Chang DK, Nones K, Johns AL, Patch AM, Gingras MC, Miller DK, Christ AN, Bruxner TJ, Quinn MC, Nourse C, Murtaugh LC, Harliwong I, et al. Genomic analyses identify molecular subtypes of pancreatic cancer. *Nature*. 2016; 531:47–52. <https://doi.org/10.1038/nature16965>.
2. Waddell N, Pajic M, Patch AM, Chang DK, Kassahn KS, Bailey P, Johns AL, Miller D, Nones K, Quek K, Quinn MC, Robertson AJ, Fadlullah MZ, et al. Whole genomes redefine the mutational landscape of pancreatic cancer. *Nature*. 2015; 518:495–501. <https://doi.org/10.1038/nature14169>.
3. Fisher R, Pusztai L, Swanton C. Cancer heterogeneity: implications for targeted therapeutics. *Br J Cancer*. 2013; 108:479–85. <https://doi.org/10.1038/bjc.2012.581>.
4. Martin JD, Fukumura D, Duda DG, Boucher Y, Jain RK. Reengineering the Tumor Microenvironment to Alleviate Hypoxia and Overcome Cancer Heterogeneity. *Cold Spring Harb Perspect Med*. 2016; 6. <https://doi.org/10.1101/cshperspect.a027094>.
5. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2013. *CA Cancer J Clin*. 2013; 63:11–30. <https://doi.org/10.3322/caac.21166>.
6. De Sousa E Melo F, Vermeulen L, Fessler E, Medema JP. Cancer heterogeneity—a multifaceted view. *EMBO Rep*. 2013; 14:686–95. <https://doi.org/10.1038/embor.2013.92>.
7. Hidalgo M, Cascinu S, Kleeff J, Labianca R, Lohr JM, Neoptolemos J, Real FX, Van Laethem JL, Heinemann V. Addressing the challenges of pancreatic cancer: future directions for improving outcomes. *Pancreatology*. 2015; 15:8–18. <https://doi.org/10.1016/j.pan.2014.10.001>.
8. Witkiewicz AK, McMillan EA, Balaji U, Baek G, Lin WC, Mansour J, Mollaei M, Wagner KU, Koduru P, Yopp A, Choti MA, Yeo CJ, McCue P, et al. Whole-exome sequencing of pancreatic cancer defines genetic diversity and therapeutic targets. *Nat Commun*. 2015; 6:6744. <https://doi.org/10.1038/ncomms7744>.

9. Weinstein JN, Collisson EA, Mills GB, Shaw KR, Ozenberger BA, Ellrott K, Shmulevich I, Sander C, Stuart JM, Cancer Genome Atlas Research Network. The Cancer Genome Atlas Pan-Cancer analysis project. *Nat Genet.* 2013; 45:1113–20. <https://doi.org/10.1038/ng.2764>.
10. Biankin AV, Waddell N, Kassahn KS, Gingras MC, Muthuswamy LB, Johns AL, Miller DK, Wilson PJ, Patch AM, Wu J, Chang DK, Cowley MJ, Gardiner BB, et al. Pancreatic cancer genomes reveal aberrations in axon guidance pathway genes. *Nature.* 2012; 491:399–405. <https://doi.org/10.1038/nature11547>.
11. Jones S, Zhang X, Parsons DW, Lin JC, Leary RJ, Angenendt P, Mankoo P, Carter H, Kamiyama H, Jimeno A, Hong SM, Fu B, Lin MT, et al. Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. *Science.* 2008; 321:1801–6. <https://doi.org/10.1126/science.1164368>.
12. Muzumdar MD, Chen PY, Dorans KJ, Chung KM, Bhutkar A, Hong E, Noll EM, Sprick MR, Trumpp A, Jacks T. Survival of pancreatic cancer cells lacking KRAS function. *Nat Commun.* 2017; 8:1090. <https://doi.org/10.1038/s41467-017-00942-5>.
13. Morton JP, Timpson P, Karim SA, Ridgway RA, Athineos D, Doyle B, Jamieson NB, Oien KA, Lowy AM, Brunton VG, Frame MC, Evans TR, Sansom OJ. Mutant p53 drives metastasis and overcomes growth arrest/senescence in pancreatic cancer. *Proc Natl Acad Sci U S A.* 2010; 107:246–51. <https://doi.org/10.1073/pnas.0908428107>.
14. Collins MA, Bednar F, Zhang Y, Brisset JC, Galban S, Galban CJ, Rakshit S, Flannagan KS, Adsay NV, Pasca di Magliano M. Oncogenic Kras is required for both the initiation and maintenance of pancreatic cancer in mice. *J Clin Invest.* 2012; 122:639–53. <https://doi.org/10.1172/JCI59227>.
15. Collisson EA, Sadanandam A, Olson P, Gibb WJ, Truitt M, Gu S, Cooc J, Weinkle J, Kim GE, Jakkula L, Feiler HS, Ko AH, Olshen AB, et al. Subtypes of pancreatic ductal adenocarcinoma and their differing responses to therapy. *Nat Med.* 2011; 17:500–3. <https://doi.org/10.1038/nm.2344>.
16. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A.* 2005; 102:15545–50. <https://doi.org/10.1073/pnas.0506580102>.
17. Okita R, Maeda A, Shimizu K, Nojima Y, Saisho S, Nakata M. PD-L1 overexpression is partially regulated by EGFR/HER2 signaling and associated with poor prognosis in patients with non-small-cell lung cancer. *Cancer Immunol Immunother.* 2017; 66:865–76. <https://doi.org/10.1007/s00262-017-1986-y>.
18. De Robertis M, Loiacono L, Fusilli C, Poeta ML, Mazza T, Sanchez M, Marchionni L, Signori E, Lamorte G, Vescovi AL, Garcia-Foncillas J, Fazio VM. Dysregulation of EGFR Pathway in EphA2 Cell Subpopulation Significantly Associates with Poor Prognosis in Colorectal Cancer. *Clin Cancer Res.* 2017; 23:159–70. <https://doi.org/10.1158/1078-0432.CCR-16-0709>.
19. Ocana A, Vera-Badillo F, Al-Mubarak M, Templeton AJ, Corrales-Sanchez V, Diez-Gonzalez L, Cuenca-Lopez MD, Seruga B, Pandiella A, Amir E. Activation of the PI3K/mTOR/AKT pathway and survival in solid tumors: systematic review and meta-analysis. *PLoS One.* 2014; 9:e95219. <https://doi.org/10.1371/journal.pone.0095219>.
20. Gatenby RA, Gillies RJ. Why do cancers have high aerobic glycolysis? *Nat Rev Cancer.* 2004; 4:891–9. <https://doi.org/10.1038/nrc1478>.
21. Gogvadze V, Zhivotovsky B, Orrenius S. The Warburg effect and mitochondrial stability in cancer cells. *Mol Aspects Med.* 2010; 31:60–74. <https://doi.org/10.1016/j.mam.2009.12.004>.
22. Zhao T, Zhu Y, Morinibu A, Kobayashi M, Shinomiya K, Itasaka S, Yoshimura M, Guo G, Hiraoka M, Harada H. HIF-1-mediated metabolic reprogramming reduces ROS levels and facilitates the metastatic colonization of cancers in lungs. *Sci Rep.* 2014; 4:3793. <https://doi.org/10.1038/srep03793>.
23. Zhong L, D'Urso A, Toiber D, Sebastian C, Henry RE, Vadysirirack DD, Guimaraes A, Marinelli B, Wikstrom JD, Nir T, Clish CB, Vaitheesvaran B, Iliopoulos O, et al. The histone deacetylase Sirt6 regulates glucose homeostasis via Hif1alpha. *Cell.* 2010; 140:280–93. <https://doi.org/10.1016/j.cell.2009.12.041>.
24. Semenza GL. HIF-1 mediates the Warburg effect in clear cell renal carcinoma. *J Bioenerg Biomembr.* 2007; 39:231–4. <https://doi.org/10.1007/s10863-007-9081-2>.
25. Zwaans BM, Lombard DB. Interplay between sirtuins, MYC and hypoxia-inducible factor in cancer-associated metabolic reprogramming. *Dis Model Mech.* 2014; 7:1023–32. <https://doi.org/10.1242/dmm.016287>.
26. Muzi M, Freeman SD, Burrows RC, Wiseman RW, Link JM, Krohn KA, Graham MM, Spence AM. Kinetic characterization of hexokinase isoenzymes from glioma cells: implications for FDG imaging of human brain tumors. *Nucl Med Biol.* 2001; 28:107–16.
27. Mathupala SP, Ko YH, Pedersen PL. Hexokinase II: cancer's double-edged sword acting as both facilitator and gatekeeper of malignancy when bound to mitochondria. *Oncogene.* 2006; 25:4777–86. <https://doi.org/10.1038/sj.onc.1209603>.
28. Simpson IA, Dwyer D, Malide D, Moley KH, Travis A, Vannucci SJ. The facilitative glucose transporter GLUT3: 20 years of distinction. *Am J Physiol Endocrinol Metab.* 2008; 295:E242–53. <https://doi.org/10.1152/ajpendo.90388.2008>.
29. Dupuy F, Tabaries S, Andrzejewski S, Dong Z, Blagih J, Annis MG, Omeroglu A, Gao D, Leung S, Amir E, Clemons M, Aguilar-Mahecha A, Basik M, et al. PDK1-Dependent

- Metabolic Reprogramming Dictates Metastatic Potential in Breast Cancer. *Cell Metab.* 2015; 22:577–89. <https://doi.org/10.1016/j.cmet.2015.08.007>.
30. Ma X, Li C, Sun L, Huang D, Li T, He X, Wu G, Yang Z, Zhong X, Song L, Gao P, Zhang H. Lin28/let-7 axis regulates aerobic glycolysis and cancer progression via PDK1. *Nat Commun.* 2014; 5:5212. <https://doi.org/10.1038/ncomms6212>.
  31. Gross S, Rahal R, Stransky N, Lengauer C, Hoefflich KP. Targeting cancer with kinase inhibitors. *J Clin Invest.* 2015; 125:1780–9. <https://doi.org/10.1172/JCI76094>.
  32. Forbes SA, Beare D, Gunasekaran P, Leung K, Bindal N, Boutselakis H, Ding M, Bamford S, Cole C, Ward S, Kok CY, Jia M, De T, et al. COSMIC: exploring the world's knowledge of somatic mutations in human cancer. *Nucleic Acids Res.* 2015; 43:D805–11. <https://doi.org/10.1093/nar/gku1075>.
  33. Yang G, Yang X. Smad4-mediated TGF-beta signaling in tumorigenesis. *Int J Biol Sci.* 2010; 6:1–8.
  34. Blobel GC, Schiemann WP, Lodish HF. Role of transforming growth factor beta in human disease. *N Engl J Med.* 2000; 342:1350–8. <https://doi.org/10.1056/NEJM200005043421807>.
  35. Malkoski SP, Wang XJ. Two sides of the story? Smad4 loss in pancreatic cancer versus head-and-neck cancer. *FEBS Lett.* 2012; 586:1984–92. <https://doi.org/10.1016/j.febslet.2012.01.054>.
  36. Zhang YE. Non-Smad pathways in TGF-beta signaling. *Cell Res.* 2009; 19:128–39. <https://doi.org/10.1038/cr.2008.328>.
  37. Lee MK, Pardoux C, Hall MC, Lee PS, Warburton D, Qing J, Smith SM, Derynck R. TGF-beta activates Erk MAP kinase signalling through direct phosphorylation of ShcA. *EMBO J.* 2007; 26:3957–67. <https://doi.org/10.1038/sj.emboj.7601818>.
  38. Shugang X, Hongfa Y, Jianpeng L, Xu Z, Jingqi F, Xiangxiang L, Wei L. Prognostic Value of SMAD4 in Pancreatic Cancer: A Meta-Analysis. *Transl Oncol.* 2016; 9:1–7. <https://doi.org/10.1016/j.tranon.2015.11.007>.
  39. Dreesen O, Brivanlou AH. Signaling pathways in cancer and embryonic stem cells. *Stem Cell Rev.* 2007; 3:7–17.
  40. Zhang J, Wu LY, Zhang XS, Zhang S. Discovery of co-occurring driver pathways in cancer. *BMC Bioinformatics.* 2014; 15:271. <https://doi.org/10.1186/1471-2105-15-271>.
  41. Nguyen DT, Mathias S, Bologna C, Brunak S, Fernandez N, Gaulton A, Hersey A, Holmes J, Jensen LJ, Karlsson A, Liu G, Ma'ayan A, Mandava G, et al. Pharos: Collating protein information to shed light on the druggable genome. *Nucleic Acids Res.* 2017; 45:D995–D1002. <https://doi.org/10.1093/nar/gkw1072>.
  42. Thorpe LM, Yuzugullu H, Zhao JJ. PI3K in cancer: divergent roles of isoforms, modes of activation and therapeutic targeting. *Nat Rev Cancer.* 2015; 15:7–24. <https://doi.org/10.1038/nrc3860>.
  43. Lin Y, Yang Z, Xu A, Dong P, Huang Y, Liu H, Li F, Wang H, Xu Q, Wang Y, Sun D, Zou Y, Zou X, et al. PIK3R1 negatively regulates the epithelial-mesenchymal transition and stem-like phenotype of renal cancer cells through the AKT/GSK3beta/CTNNB1 signaling pathway. *Sci Rep.* 2015; 5:8997. <https://doi.org/10.1038/srep08997>.
  44. Cheung LW, Mills GB. Targeting therapeutic liabilities engendered by PIK3R1 mutations for cancer treatment. *Pharmacogenomics.* 2016; 17:297–307. <https://doi.org/10.2217/pgs.15.174>.
  45. Joerger AC, Fersht AR. The p53 Pathway: Origins, Inactivation in Cancer, and Emerging Therapeutic Approaches. *Annu Rev Biochem.* 2016; 85:375–404. <https://doi.org/10.1146/annurev-biochem-060815-014710>.
  46. Visconti R, Della Monica R, Grieco D. Cell cycle checkpoint in cancer: a therapeutically targetable double-edged sword. *J Exp Clin Cancer Res.* 2016; 35:153. <https://doi.org/10.1186/s13046-016-0433-9>.
  47. Li W, Zhang H, Assaraf YG, Zhao K, Xu X, Xie J, Yang DH, Chen ZS. Overcoming ABC transporter-mediated multidrug resistance: Molecular mechanisms and novel therapeutic drug strategies. *Drug Resist Updat.* 2016; 27:14–29. <https://doi.org/10.1016/j.drup.2016.05.001>.
  48. Kathawala RJ, Gupta P, Ashby CR Jr, Chen ZS. The modulation of ABC transporter-mediated multidrug resistance in cancer: a review of the past decade. *Drug Resist Updat.* 2015; 18:1–17. <https://doi.org/10.1016/j.drup.2014.11.002>.
  49. Paull EO, Carlin DE, Niepel M, Sorger PK, Haussler D, Stuart JM. Discovering causal pathways linking genomic events to transcriptional states using Tied Diffusion Through Interacting Events (TieDIE). *Bioinformatics.* 2013; 29:2757–64. <https://doi.org/10.1093/bioinformatics/btt471>.
  50. Ng S, Collisson EA, Sokolov A, Goldstein T, Gonzalez-Perez A, Lopez-Bigas N, Benz C, Haussler D, Stuart JM. PARADIGM-SHIFT predicts the function of mutations in multiple cancers using pathway impact analysis. *Bioinformatics.* 2012; 28:i640–i6. <https://doi.org/10.1093/bioinformatics/bts402>.
  51. Zhang W, Liu HT. MAPK signal pathways in the regulation of cell proliferation in mammalian cells. *Cell Res.* 2002; 12:9–18. <https://doi.org/10.1038/sj.cr.7290105>.
  52. Fang JY, Richardson BC. The MAPK signalling pathways and colorectal cancer. *Lancet Oncol.* 2005; 6:322–7. [https://doi.org/10.1016/S1470-2045\(05\)70168-6](https://doi.org/10.1016/S1470-2045(05)70168-6).
  53. Cheng CW, Su JL, Lin CW, Su CW, Shih CH, Yang SF, Chien MH. Effects of NFKB1 and NFKBIA gene polymorphisms on hepatocellular carcinoma susceptibility and clinicopathological features. *PLoS One.* 2013; 8:e56130. <https://doi.org/10.1371/journal.pone.0056130>.
  54. Royds JA, Dower SK, Qwarnstrom EE, Lewis CE. Response of tumour cells to hypoxia: role of p53 and NFkB. *Mol Pathol.* 1998; 51:55–61.

55. Yao L, Wang L, Li F, Gao X, Wei X, Liu Z. MiR181c inhibits ovarian cancer metastasis and progression by targeting PRKCD expression. *Int J Clin Exp Med*. 2015; 8:15198–205.
56. Sun M, Estrov Z, Ji Y, Coombes KR, Harris DH, Kurzrock R. Curcumin (diferuloylmethane) alters the expression profiles of microRNAs in human pancreatic cancer cells. *Mol Cancer Ther*. 2008; 7:464–73. <https://doi.org/10.1158/1535-7163.MCT-07-2272>.
57. Stan SD, Singh SV, Brand RE. Chemoprevention strategies for pancreatic cancer. *Nat Rev Gastroenterol Hepatol*. 2010; 7:347–56. <https://doi.org/10.1038/nrgastro.2010.61>.
58. Soares HP, Ming M, Mellon M, Young SH, Han L, Sinnett-Smith J, Rozengurt E. Dual PI3K/mTOR Inhibitors Induce Rapid Overactivation of the MEK/ERK Pathway in Human Pancreatic Cancer Cells through Suppression of mTORC2. *Mol Cancer Ther*. 2015; 14:1014–23. <https://doi.org/10.1158/1535-7163.MCT-14-0669>.
59. Eser S, Schnieke A, Schneider G, Saur D. Oncogenic KRAS signalling in pancreatic cancer. *Br J Cancer*. 2014; 111:817–22. <https://doi.org/10.1038/bjc.2014.215>.
60. Ning C, Liang M, Liu S, Wang G, Edwards H, Xia Y, Polin L, Dyson G, Taub JW, Mohammad RM, Azmi AS, Zhao L, Ge Y. Targeting ERK enhances the cytotoxic effect of the novel PI3K and mTOR dual inhibitor VS-5584 in preclinical models of pancreatic cancer. *Oncotarget*. 2017; 8:44295–311. <https://doi.org/10.18632/oncotarget.17869>.
61. Ciuffreda L, Del Curatolo A, Falcone I, Conciatori F, Bazzichetto C, Cognetti F, Corbo V, Scarpa A, Milella M. Lack of growth inhibitory synergism with combined MAPK/PI3K inhibition in preclinical models of pancreatic cancer. *Ann Oncol*. 2017; 28:2896–8. <https://doi.org/10.1093/annonc/mdx335>.
62. Pettazzoni P, Viale A, Shah P, Carugo A, Ying H, Wang H, Genovese G, Seth S, Minelli R, Green T, Huang-Hobbs E, Corti D, Sanchez N, et al. Genetic events that limit the efficacy of MEK and RTK inhibitor therapies in a mouse model of KRAS-driven pancreatic cancer. *Cancer Res*. 2015; 75:1091–101. <https://doi.org/10.1158/0008-5472.CAN-14-1854>.
63. Stojanovic N, Hassan Z, Wirth M, Wenzel P, Beyer M, Schafer C, Brand P, Kroemer A, Stauber RH, Schmid RM, Arlt A, Sellmer A, Mahboobi S, et al. HDAC1 and HDAC2 integrate the expression of p53 mutants in pancreatic cancer. *Oncogene*. 2017; 36:1804–15. <https://doi.org/10.1038/onc.2016.344>.
64. Gurpinar E, Vousden KH. Hitting cancers' weak spots: vulnerabilities imposed by p53 mutation. *Trends Cell Biol*. 2015; 25:486–95. <https://doi.org/10.1016/j.tcb.2015.04.001>.
65. Weissmueller S, Machado E, Saborowski M, Morris JP 4th, Wagenblast E, Davis CA, Moon SH, Pfister NT, Tschaharganeh DF, Kitzing T, Aust D, Markert EK, Wu J, et al. Mutant p53 drives pancreatic cancer metastasis through cell-autonomous PDGF receptor beta signaling. *Cell*. 2014; 157:382–94. <https://doi.org/10.1016/j.cell.2014.01.066>.
66. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011; 144:646–74. <https://doi.org/10.1016/j.cell.2011.02.013>.
67. Pavlova NN, Thompson CB. The Emerging Hallmarks of Cancer Metabolism. *Cell Metab*. 2016; 23:27–47. <https://doi.org/10.1016/j.cmet.2015.12.006>.
68. Wilde L, Roche M, Domingo-Vidal M, Tanson K, Philp N, Curry J, Martinez-Outschoorn U. Metabolic coupling and the Reverse Warburg Effect in cancer: Implications for novel biomarker and anticancer agent development. *Semin Oncol*. 2017; 44:198–203. <https://doi.org/10.1053/j.seminoncol.2017.10.004>.
69. Ramos P, Bentires-Alj M. Mechanism-based cancer therapy: resistance to therapy, therapy for resistance. *Oncogene*. 2015; 34:3617–26. <https://doi.org/10.1038/onc.2014.314>.
70. Obenauf AC, Zou Y, Ji AL, Vanharanta S, Shu W, Shi H, Kong X, Bosenberg MC, Wiesner T, Rosen N, Lo RS, Massague J. Therapy-induced tumour secretomes promote resistance and tumour progression. *Nature*. 2015; 520:368–72. <https://doi.org/10.1038/nature14336>.
71. Brat DJ, Verhaak RG, Aldape KD, Yung WK, Salama SR, Cooper LA, Rheinbay E, Miller CR, Vitucci M, Morozova O, Robertson AG, Noushmehr H, Laird FP, et al. Cancer Genome Atlas Research Network. Comprehensive, Integrative Genomic Analysis of Diffuse Lower-Grade Gliomas. *N Engl J Med*. 2015; 372:2481–98. <https://doi.org/10.1056/NEJMoa1402121>.
72. Noto K, Brodley C, Slonim D. FRaC: a feature-modeling approach for semi-supervised and unsupervised anomaly detection. *Data Min Knowl Discov*. 2012; 25:109–33. <https://doi.org/10.1007/s10618-011-0234-x>.
73. Goel MK, Khanna P, Kishore J. Understanding survival analysis: Kaplan-Meier estimate. *Int J Ayurveda Res*. 2010; 1:274–8. <https://doi.org/10.4103/0974-7788.76794>.
74. Robinson MD, Smyth GK. Moderated statistical tests for assessing differences in tag abundance. *Bioinformatics*. 2007; 23:2881–7. <https://doi.org/10.1093/bioinformatics/btm453>.
75. Anders S, Huber W. Differential expression analysis for sequence count data. *Genome Biol*. 2010; 11:R106. <https://doi.org/10.1186/gb-2010-11-10-r106>.
76. Merico D, Isserlin R, Stueker O, Emili A, Bader GD. Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. *PLoS One*. 2010; 5:e13984. <https://doi.org/10.1371/journal.pone.0013984>.
77. Kutmon M, van Iersel MP, Bohler A, Kelder T, Nunes N, Pico AR, Evelo CT. PathVisio 3: an extendable pathway analysis toolbox. *PLoS Comput Biol*. 2015; 11:e1004085. <https://doi.org/10.1371/journal.pcbi.1004085>.
78. Chen EY, Xu H, Gordonov S, Lim MP, Perkins MH, Ma'ayan A. Expression2Kinases: mRNA profiling linked to

- multiple upstream regulatory layers. *Bioinformatics*. 2012; 28:105–11. <https://doi.org/10.1093/bioinformatics/btr625>.
79. Lachmann A, Xu H, Krishnan J, Berger SI, Mazloom AR, Ma'ayan A. ChEA: transcription factor regulation inferred from integrating genome-wide ChIP-X experiments. *Bioinformatics*. 2010; 26:2438–44. <https://doi.org/10.1093/bioinformatics/btq466>.
80. Lachmann A, Ma'ayan A. KEA: kinase enrichment analysis. *Bioinformatics*. 2009; 25:684–6. <https://doi.org/10.1093/bioinformatics/btp026>.
81. Lawrence MS, Stojanov P, Polak P, Kryukov GV, Cibulskis K, Sivachenko A, Carter SL, Stewart C, Mermel CH, Roberts SA, Kiezun A, Hammerman PS, McKenna A, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature*. 2013; 499:214–8. <https://doi.org/10.1038/nature12213>.
82. Mermel CH, Schumacher SE, Hill B, Meyerson ML, Beroukhi R, Getz G. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol*. 2011; 12:R41. <https://doi.org/10.1186/gb-2011-12-4-r41>.
83. UniProt Consortium T. UniProt: the universal protein knowledgebase. *Nucleic Acids Res*. 2017; 45:D158–D69. <https://doi.org/10.1093/nar/gkw1099>.
84. Zhao M, Kim P, Mitra R, Zhao J, Zhao Z. TSGene 2.0: an updated literature-based knowledgebase for tumor suppressor genes. *Nucleic Acids Res*. 2016; 44:D1023–31. <https://doi.org/10.1093/nar/gkv1268>.
85. Liu Y, Sun J, Zhao M. ONGene: A literature-based database for human oncogenes. *J Genet Genomics*. 2017; 44:119–21. <https://doi.org/10.1016/j.jgg.2016.12.004>.
86. Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, Clark NR, Ma'ayan A. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics*. 2013; 14:128. <https://doi.org/10.1186/1471-2105-14-128>.
87. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*. 2016; 32:2847–9. <https://doi.org/10.1093/bioinformatics/btw313>.
88. Bahceci I, Dogrusoz U, La KC, Babur O, Gao J, Schultz N. PathwayMapper: a collaborative visual web editor for cancer pathways and genomic data. *Bioinformatics*. 2017; 33:2238–40. <https://doi.org/10.1093/bioinformatics/btx149>.
89. Fabregat A, Sidiropoulos K, Garapati P, Gillespie M, Hausmann K, Haw R, Jassal B, Jupe S, Korninger F, McKay S, Matthews L, May B, Milacic M, et al. The Reactome pathway Knowledgebase. *Nucleic Acids Res*. 2016; 44:D481–7. <https://doi.org/10.1093/nar/gkv1351>.
90. Wixon J, Kell D. The Kyoto encyclopedia of genes and genomes—KEGG. *Yeast*. 2000; 17:48–55.