Introduction

Pancreatic ductal adenocarcinoma (PDAC) is the most common cancer arising from the pancreas. PDAC is believed to develop from morphologically distinct non-invasive precursor lesions; namely, pancreatic intraepithelial neoplasia (PanIN), intraductal papillary mucinous neoplasm (IPMN) and mucinous cystic neoplasm (MCN), in order of decreasing frequency (1). Despite its low incidence, PDAC is currently the fourth leading cause of cancer mortality in the United States and represents a major health burden. The low survival rate and poor prognosis are largely due to its aggressive and metastatic nature, disappointing response to therapy, and late diagnosis as a consequence of lack of effective screening strategies and absence of specific symptoms at early stages (2,3).

Thus, there is an urgent need to understand PDAC at the molecular and genomic level in order to improve overall survival. Although KRAS mutations are almost universal in PDAC, this is an otherwise genetically heterogeneous disease that displays a complex mutational landscape. TP53, CDKN2A and SMAD4 are the most commonly mutated oncogenic drivers after KRAS, but a growing number of other low-prevalence mutated genes have been identified in various genomic studies (4).

Gene expression profiling has been instrumental for subtype classification in cancer. Variance in gene expression patterns among cancer subtypes reflects fundamental differences in tumors at the molecular level. Identification of tumor subtypes by transcriptomic analysis has biological, prognostic and predictive value, can enable patient stratification and reveal potential therapeutic vulnerabilities (1). Several groups have used transcriptomic analysis, alone or in combination with other modalities, to identify biologically and clinically distinct subsets of PDAC. For example, Moffitt et al. identified two major subtypes (classical and basal-like) (5), Collisson et al. identified three (quasimesenchymal, classical, and exocrine like) (6), and Bailey et al. identified four [squamous, pancreatic progenitor, aberrantly differentiated endocrine exocrine
Table 1 Lineage specifier expression in PDAC subtypes, effect in pancreatic cancer progression and prognostic value

<table>
<thead>
<tr>
<th>Transcription factor</th>
<th>Classical/progenitor subtype</th>
<th>Basal-like/squamous subtype</th>
<th>Effect</th>
<th>Prognostic value</th>
</tr>
</thead>
<tbody>
<tr>
<td>FoxA1/FoxA2</td>
<td>+</td>
<td>−</td>
<td>Dichotomous</td>
<td>Undetermined</td>
</tr>
<tr>
<td>HNF4α</td>
<td>+</td>
<td>−</td>
<td>Tumor suppressor</td>
<td>Yes</td>
</tr>
<tr>
<td>HNF1α</td>
<td>+</td>
<td>−</td>
<td>Dichotomous</td>
<td>Yes</td>
</tr>
<tr>
<td>PDX1</td>
<td>+</td>
<td>−</td>
<td>Dichotomous</td>
<td>Yes</td>
</tr>
</tbody>
</table>

PDAC, pancreatic ductal adenocarcinoma.

(ADEX), and immunogenic) (7). Most recently, the Cancer Genome Atlas Research Network reported an integrated genomic, transcriptomic and proteomic analysis of an additional 150 PDAC samples (8). The authors also revisited the three prior genomic analyses and found significant overlap between the squamous (7) and basal-like (5) subtypes and also between the pancreatic progenitor (7) and the classical subtypes described in Collisson et al. (6) and Moffitt et al. (8). The authors argue that PDAC can fundamentally be classified into two major subtypes based on cancer cell autonomous properties: basal-like/squamous and classical/pancreatic progenitor. They further suggest that additional subtypes identified to date (exocrine-like, ADEX and immunogenic) reflect either contributions from non-neoplastic tissue or less common subtypes not included in their cohort.

One major difference between these two major subtypes is the relative levels of several lineage specifiers, i.e., transcription factors that regulate endodermal development and differentiation. Classical/progenitor subtypes express high levels of a transcriptional network pivotal for pancreatic endodermal determination, including genes such as FOXA2, FOXA3, HNF4A, HNF4G, HNF1A, HNF1B and PDX1 (7). In contrast, the basal-like/squamous subtype exhibits loss of endodermal identity by hypermethylation and downregulation of these lineage specifiers and concomitant upregulation of pathways driven by MYC, ΔNp63 and TGF-β (7).

The basal-like/squamous subtype confers poor prognosis, with a median survival significantly shorter than the classical/progenitor subtype (5,7). This raises the possibility that these lineage specifiers are not merely markers of PDAC subtype, but might regulate the malignant potential of this disease. Here we review the literature on the role of relevant lineage specifiers in normal pancreas and PDAC pathogenesis, with a specific focus on those that are highly expressed in the classical/progenitor subtype relative to the basal/squamous subtype (Table 1).

**FoxA1 and FoxA2**

The winged helix transcription factors FoxA1 and FoxA2 together control the development and morphogenesis of both the exocrine and endocrine pancreas. FoxA1 and FoxA2 are expressed in the foregut endoderm prior to pancreatic development and persist in all islet and acinar cells into adulthood (9). Pancreata of Foxa1 and Foxa2-null mice showed diminished PDX1 expression, intense pancreatic hypoplasia, hyperglycemia, impaired acinar and islet cell content, and subsequent death. Endoderm-specific genetic deletion of Foxa2 in mice provoked severe hypoglycemia, impaired α-cell differentiation and early death (9). Interestingly, FoxA1 and FoxA2 are expressed in the foregut endoderm prior to multiple relevant lineage specifiers, such as PDX1, HNF1β, PTF1α, HNF4α, etc. (10). In fact, it has been shown that FoxA1 and FoxA2 participate in the transcriptional activation of Pdx1 (11), as well as HNF4A, HNF1A, HNF1B and HNF4G in the pancreas and other endodermal organs (12-14), suggesting a role as upstream transactivators in the transcriptional hierarchy.

Although FoxA1 and FoxA2 proteins are pivotal regulators of normal development of different endoderm-derived organs such as pancreas, liver, lungs and prostate (15), their involvement in cancer remains largely unexplored. Nevertheless, some progress has been made elucidating the role of these important pioneer factors in pancreatic carcinogenesis. For example, Mishra and Guda found that FOXA1 (as well as HNF4A and HNF1A) are hypermethylated in primary PDAC tumors versus normal samples in their analysis of previously published TCGA studies (16). An additional report identified FoxA1 as one of the transcription factors that characterize the classical PDAC gene signature described by Collisson et al. (6). The authors showed that FoxA1 levels and enhancer-associated binding is dramatically downregulated in high-grade compared to low-grade human PDAC cell lines (17). Moreover, an immunohistochemical analysis of 25
human PDAC specimens showed that FoxA1 and FoxA2 were expressed in normal pancreatic epithelium, acinar to ductal metaplasia, PanINs and well differentiated invasive PDACs. In contrast, FoxA1 and FoxA2 were downregulated in poor to undifferentiated PDACs (18). In human PDAC cell lines, FoxA1 and FoxA2 impose a transcriptional program that sustains an epithelial phenotype while antagonizing the epithelial to mesenchymal transition (EMT), in part by positively regulating E-cadherin expression (18). However, this study did not directly address the impact of altered FoxA1/2 expression on the overall tumorigenicity of these cell lines.

With regard to mechanisms by which FoxA1 and FoxA2 levels might be regulated in PDAC, two studies have shown that FoxA1 and/or FoxA2 are downregulated during TGF-β-induced EMT in PDAC cell lines (18,19). In the course of elucidating the role of GATA6 in pancreatic oncogenesis, FOXA2 was shown to be a GATA6 target gene and to positively correlate with GATA6 levels in PDAC samples (20). ChIP-Seq data in this study also suggest that GATA6 and FoxA2 may cooperatively regulate a pro-epithelial/anti-mesenchymal gene expression program in PDAC (20). Finally, miR-199a can inhibit FOXA2 expression in human PDAC cells by directly binding its 3' UTR (21). FoxA2 suppression in a panel of human PDAC cell lines (using siRNA, CRISPR, and miR199a expression) increased cellular growth, invasion, migration and colony formation, as well as tumor growth in xenografts (21).

Although more work needs to be done, the above studies are generally consistent with a model in which loss of FoxA1/2 activity might impair an endodermal differentiation program and enhance tumorigenesis. However, a recent report shows that the situation is likely to be more complex than this simple model would suggest. Specifically, Roe et al. used an innovative organoid culture system to identify genes implicated in metastasis during PDAC progression (22). Using ChIP-Seq for histone H3 lysine 27 acetylation (H3K27ac, a marker of active enhancers and promoters) on a panel of organoids derived from primary and metastatic tumors, they found that the metastatic process is promoted by a dramatic reprogramming of enhancer activity (22). FoxA1, in cooperation with GATA5, was identified as a central driver of enhancer reprogramming, activating a foregut differentiation program and promoting a more invasive and anchorage-independent behavior in vitro as well as an increased metastatic potential in vivo without inducing EMT (22). Thus, it appears that FoxA activity can either promote or restrain tumorigenesis, depending on the specific context. Additional studies are clearly needed to obtain a better understanding of this dichotomy.

**HNF4α**

The nuclear receptor superfamily member HNF4α is a master regulator of epithelial differentiation in multiple tissues, including the gastrointestinal tract and liver (23-26). In the pancreas, HNF4α is vital for β cell function through the regulation of several genes involved in metabolism-secretion coupling (27) and glucose stimulated insulin secretion (28). Mutations in this gene have been associated with Mature-Onset Diabetes of the Young 1 (MODY1), a monogenic autosomal dominant non-insulin-dependent diabetes mellitus type I (29). More recently, it was suggested that loss of HNF4α function alters endoplasmic reticulum Ca2+ homeostasis and insulin release in β-cells of patients with MODY1 (30).

Although HNF4α has been implicated in the pathogenesis of different malignancies (31) and the regulation of multiple biological processes that influence tumor progression, such as proliferation (32-34), apoptosis (35), metabolism (36,37) and inflammation (38,39), little is known about its role in PDAC. Kim et al. successfully developed a model to study early pancreatic cancer progression by generating induced pluripotent stem cells (iPSC) from primary human PDAC (40). This PDAC-iPSC line generates endodermal ductal structures reminiscent of PanIN and further progresses to invasive pancreatic cancer in vivo when implanted in mice (40). Protein profiling from PanIN organoid cultures revealed a network centered in HNF4α. Using IHC, the authors showed that HNF4α expression is absent or barely detectable in normal human or mouse pancreatic ducts, weak in PanIN1, moderate in PanIN2, and stronger and uniform in PanIN3 and well differentiated PDAC (40). Consistent with published genomics studies, HNF4α is essentially undetectable in undifferentiated or poorly differentiated PDAC (40). This study suggested that the dynamic expression of HNF4α and its target genes at different stages of PDAC progression might become a useful diagnostic and prognostic biomarker of pancreatic malignancy (40).

Additional studies have shown that HNF4α levels may have an impact on the malignant potential and therapeutic response of PDAC. EMT in pancreatic cancer has been linked to a greater proclivity for migration and invasion, gain of stemness and increased anticancer treatment resistance (41). Viotti et al. carried out an siRNA screen to identify epigenetic factors responsible for EMT.
in pancreatic cancer, demonstrating that the histone methyltransferase SUV420H2 acts as an orchestrator of epithelial/mesenchymal states (42). SUV420H2 induces silencing of key drivers of epithelial cell identity, such as HNF4α and FoxA1, and its attenuation diminishes invasion, migration, stem cell characteristics, and chemoresistance in human pancreatic cancer cells (42). Additionally, the authors showed correlation between EMT and increased SUV420H2 expression during the pathological progression of PDAC (42).

Based on the Collisson et al. PDAC subclassification and the suggested differential chemotherapy response between the classical and quasimesenchymal subtypes (6), Noll et al. interrogated the contribution of different cytochrome P450 (43) enzymes, which play a major role in xenobiotic metabolism, in drug resistance in PDAC (44). In this study, HNF4α and NR1I2 were responsible for mediating basal and drug-induced expression of CYP3A5, respectively, and the conferred drug resistance (44). Thus, interfering with the regulation of CYP3A5 mechanism by these transcription factors offers a potential mechanism to overcome basal and acquired drug resistance in PDAC. Interestingly, HNF4α can be pharmacologically targeted by small molecule antagonists recently described in the preclinical setting (45,46), offering the possibility of suppressing HNF4α transcriptional activity.

Finally, HNF4α levels can be modulated by specific drugs, which may enable therapeutic strategies that alter the differentiation state of PDAC cells. In a mouse model of IPMN/PDAC driven by oncogenic KRASG12D and deletion of Brg1 (47), the bromodomain inhibitor JQ1 concomitantly reduces tumor burden and attenuates expression of HNF4α and PDX1 while increasing the expression of mature murine duct markers, possibly through reactivation of a cellular differentiation program (47). Treatment of the human PDAC cell lines PANC-1 and CFPAC-1 with cantharidin, a vesicant produced by beetles, as well as its derivative norcantharidin, impaired proliferation and downregulated putative stem cell markers (including HNF4α, MYC, CD44, SOX9, HMGA2, EPCAM, EPCAM and SOX2) while strengthening the cytotoxicity of gemcitabine and erlotinib in co-treatment experiments (48). These results show potential therapeutic promise for the eradication of pancreatic CSC, which are a small subset of malignant cells capable of self-renewal and are thought to play a critical role in regulating pancreatic cancer progression, metastasis, and drug resistance (48). Additionally, a recent study identified a PDAC side population in 52 human PDAC resected tumors (49). Gene expression analysis by microarray revealed not only an enrichment in pancreatic CSC markers and a chemoresistance-associated phenotype but also upregulation of molecular networks associated with cellular growth and proliferation, cell-cell signaling and amino acid metabolism, which includes HNF4α (49).

**HNF1α**

HNF1α is a homeodomain transcription factor specifically involved in later islet cell formation and function. HNF1α mutation is one of the most prevalent causes of transcription factor autosomal dominant diabetes. It regulates the expression of genes involved in glucose, amino acid, and tricarboxylic acid metabolism and is necessary for mitochondrial function in pancreatic islet cells (43,50,51).

Given the important role of HNF1α in pancreas development, several reports have explored its function in the pathogenesis of pancreatic malignancy. Evaluation of a large cohort of pancreatic cancer patients identified HNF1α/β and their predicted targets to be globally downregulated in PDAC samples compared to normal pancreas tissues (52). Using transcriptome analysis on a series of human cell lines, PDAC and normal pancreas samples, Hoskins et al. found gene expression changes associated with upregulation of genes involved in EMT as well as repression of transcription factors that participate in pancreatic development and homeostasis, where HNF1α and HNF1α-regulated networks were the most significantly affected (53). Furthermore, HNF1α expression is lost in a subset of human PDAC samples and tissue microarrays as detected by immunohistochemistry (53).

HNF1α may serve as a prognostic biomarker and predictor of therapy response in PDAC (54). Biopsies of two independent cohorts of resected patients were stratified based on HNF1α and KRT81 immunostaining. The HNF1α-positive group showed the best, and KRT81-positive the worst, patient survival (54). In a third unresectable cohort of patients undergoing FOLFIRINOX or gemcitabine therapy, KRT81-positive patients had a negative prognosis compared to HNF1α-positive patients, while within the latter group FOLFIRINOX-treated patients displayed a better response than patients under gemcitabine treatment (54).

Lee et al. carried out a genome-wide association study
to identify candidate single nucleotide polymorphisms (SNP) conferring susceptibility to pancreatic cancer by analyzing almost half a million SNPs from 4,000 pancreatic cancer patients and control subjects (55). This *in silico* analysis identified 18 candidate SNPs, 11 genes (among them, HNF1α and HNF4γ), and 30 pathways that could contribute to pancreatic cancer risk (55).

Some functional studies of HNF1α have indicated that it can act as a tumor suppressor in PDAC. One recent report showed that siRNA-mediated HNF1α loss of function caused augmented proliferative capacity and diminished apoptosis (56). In a complementary study, Yu and collaborators explored the role of HNF1α in the regulation of the long non-coding RNA CASC2 (57). HNF1α activates CASC2 transcription, which in turn can positively regulate PTEN. HNF1α overexpression in PDAC cell lines induces PTEN expression, decreases p-Akt and p-mTOR and reduces proliferation, while CASC2 knockdown caused the opposite effect (57). In PDAC HNF1α levels positively correlated with CASC2 expression, and low CASC2 as well as HNF1α levels associated with a poorer patient outcome and shorter overall survival (57).

In contrast, one report has indicated that HNF1α is required for pancreatic CSC function. Specifically, Abel and coworkers showed that HNF1α and its target gene levels are increased in pancreatic CSC and 3D tumorspheres, while HNF1α knockdown reduced PDAC growth, apoptosis, tumorsphere formation and pancreatic CSC (58). Notably, HNF1α silencing in patient derived xenografts hindered tumor progression *in vivo* (58). Lastly, the authors established an HNF1α-dependent gene signature that correlates with poor patient prognosis (58). The HNF1α oncogenic role identified in this study is in clear contrast with its tumor suppressive function shown in previous reports (53,56). Different scenarios were discussed as the source of these seemingly antagonistic functions by the authors (58), such as technical differences in virtue of an unusually elevated level of baseline apoptosis found in control cells in Luo et al. (56), differential role of HNF1α in different PDAC molecular subtypes or the dynamics of its expression as previously identified for PDX1 (59). As discussed below, PDX1 levels and stage specific expression as well as the transcriptional network governed by PDX1 may explain its dichotomous role in pancreatic cancer. Further investigation needs to be done to address whether these discrepancies are reflecting real differential biological properties of HNF1α in the progression of pancreatic disease.

**PDX1**

The homeodomain transcription factor PDX1 is a critical transcriptional regulator involved in development of pancreatic buds and expressed in pancreatic progenitor populations (9). It has fundamental roles in the formation of all pancreatic cell types, as it is expressed throughout the earliest multipotent pancreatic progenitor cell pool, and then in a more restricted manner within all developing and adult islet insulin+ β cells, where it regulates islet β cell function, as well as in a small proportion of islet somatostatin+ δ cells (43). In the adult pancreas, PDX1 levels are elevated in pancreatic regeneration and injury, such as pancreatitis, pancreatic resection and hypoglycemic challenge, where PDX1 expression positively correlates with proliferation of immature progenitor cells (60,61).

IHC analysis of primary human samples has shown that PDX1 is variably expressed in precursor lesions and PDAC. In Park et al.’s report (62), approximately 40% of PanINs and 35% of IPMNs exhibited PDX1 expression in at least half of their nuclei. In contrast, only 13% of PDACs examined exhibited PDX1 expression in >50% of their nuclei, and PDX1 was completely absent in 60% of PDAC cases (62). This study did not correlate PDX1 levels with grade, and the authors did not observe a significant difference in survival based on PDX1 expression (62).

The functional role of PDX1 in pancreatic malignancy has been extensively reviewed recently (60,61), where the main conclusion was that PDX1 seems to exert primarily oncogenic functions. However, a recent comprehensive analysis of PDX1 in a genetically engineered mouse model demonstrated that this transcription factor has distinct, stage-specific roles in this disease (59). First, the authors ablated a conditional allele of *Pdx1* in adult murine acinar cells using a tamoxifen-inducible Cre*ER* system. *Pdx1* deletion enhanced acinar to ductal metaplasia *in vivo* during acute pancreatitis and *in vitro*, suggesting that PDX1 plays a role in maintaining acinar cell identity. This was further demonstrated by transcriptomic analysis, which showed that *Pdx1* deletion impaired the expression of several acinar specific genes in these cells *in vitro*. Moreover, *Pdx1* deletion accelerated *Kras*G12D-induced acinar to ductal metaplasia and PanIN formation in this mouse model (59). However, inhibition of PDX1 expression in established mouse and human PDAC cell lines generally impaired proliferation and cell survival *in vitro* and xenograft growth *in vivo* (59). These results are in consonance with other studies that have suggested that PDX1 has oncogenic properties in

© Journal of Gastrointestinal Oncology. All rights reserved.
PDAC (63,64). Using both RNA-Seq and ChIP-Seq, it was shown that PDX1 regulates a distinct set of genes in normal acinar cells compared with established PDAC, which likely accounts for its differential roles in tumor initiation and maintenance (59).

Adding further complexity to the picture, the authors found that low PDX1 expression correlates with reduced survival in two independent patient cohorts comprising all molecular PDAC subtypes identified in Bailey et al. (7) and Collisson et al. (6). These data are in agreement with genomics studies showing that the basal-like/squamous subtype, which has low levels of PDX1, confers a worse prognosis (5,7). To reconcile these correlations with their experimental observations, the authors asked whether a subset of PDAC cells could survive PDX1 inhibition and adapt to a PDX1-low state. Interestingly, they found that approximately 3 weeks after induction of hairpin targeting PDX1, a population of PDX1 low cells emerged with proliferative potential similar to controls (59). These data suggest that the dependence of PDAC cells on PDX1, and likely other lineage specifiers, is not absolute, and that particular selective pressures may lead to the emergence of subpopulations in which downregulation of these lineage specifiers leads to a more poorly differentiated, malignant phenotype.

Conclusions
Recent high throughput genomic approaches have identified distinct molecular subtypes of pancreatic cancer and relevant molecular players associated with different biological aspects of pancreatic cancer cells. These subtypes are characterized by differential expression of a number of lineage specifiers that play a role in pancreatic development and differentiation (Figure 1). This review presents the work carried out by different groups that collectively deepens our knowledge of the pivotal role that these pancreatic endodermal determinants play in PDAC progression. Several of these lineage specifiers appear to play context-specific, dichotomous roles in the progression of disease, and additional studies are clearly needed to elucidate the molecular basis of this phenomenon. Moreover, we anticipate that a detailed understanding of these lineage specifiers and their target genes in PDAC will facilitate the identification of new biomarker and therapeutic strategies for this deadly disease.

Acknowledgements
Funding: EL Snyder was supported in part by a Career Award for Medical Scientists from the Burroughs Wellcome Fund, a V Scholar Award, the Huntsman Cancer Foundation and the NIH (R01CA212415 and funds associated with P30CA042014 awarded to the Nuclear Control Program at the Huntsman Cancer Institute).

Footnote
Conflicts of Interest: The authors have no conflicts of interest to declare.

References
5. Moffitt RA, Marayati R, Flate EL, et al. Virtual microdissection identifies distinct tumor- and stroma-


61. Fendrich V, Lauth M. The role of pancreatic and duodenal homeobox 1 as a therapeutic target in pancreatic cancer.

Cite this article as: Camolotto SA, Belova VK, Snyder EL. The role of lineage specifiers in pancreatic ductal adenocarcinoma. J Gastrointest Oncol 2018;9(6):1005-1013. doi: 10.21037/jgo.2018.05.04