A cross-species analysis in pancreatic neuroendocrine tumors reveals molecular subtypes with distinctive clinical, metastatic, developmental, and metabolic characteristics.

Anguraj Sadanandam[1,2,3,#], Stephan Wullschleger[2], Costas A. Lyssiotis[4,a], Carsten Grötzinger[5], Stefano Barbi[6], Samantha Bersani[6], Jan Körner[5], Ismael Wafy[2], Andrea Mafficini[6], Rita T. Lawlor[6], Michele Simbolo[6], John M. Asara[7], Hendrik Bläker[8], Lewis C. Cantley[4], Bertram Wiedenmann[5], Aldo Scarpa[6,#], Douglas Hanahan[2,#]

[1]Swiss Institute of Bioinformatics (SIB), Lausanne, Switzerland

[2]Swiss Institute for Experimental Cancer Research (ISREC), Swiss Federal Institute of Lausanne (EPFL), Lausanne, Switzerland

[3]Division of Molecular Pathology, Institute of Cancer Research (ICR), London, UK

[4]Meyer Cancer Center, Weill Cornell Medical College, New York, USA

[5]Department of Hepatology and Gastroenterology, Charite, Campus Virchow-Klinikum, University Medicine, Berlin, Germany

[6]ARC-Net Research Centre and Department of Pathology and Diagnostics, University and Hospital Trust of Verona, Verona, Italy

[7]Department of Medicine, Beth Israel Deaconess Medical Center, Boston, Massachusetts, USA

[8]Institut für Pathologie, Charite, Campus Virchow-Klinikum, University Medicine, Berlin, Germany

[a]Current address: Department of Molecular and Integrative Physiology and Department of Internal Medicine, Division of Gastroenterology, University of Michigan, Ann Arbor, MI, USA

[#] Correspondence to:

Douglas Hanahan, Ph.D.

Swiss Institute for Experimental Cancer Research (ISREC),

Swiss Federal Institute of Lausanne (EPFL), Lausanne

EPFL SV ISREC CMSO

Station 19, Batiment SV

Lausanne, CH-1015

Switzerland

Ph: +41-21-693-0657

douglas.hanahan@epfl.ch

Anguraj Sadanandam, Ph.D.

Division of Molecular Pathology

Institute of Cancer Research (ICR)

15 Cotswold Road

Sutton, Surrey SM2 5NG

United Kingdom

Ph: +44-20-8915-6631

anguraj.sadanandam@icr.ac.uk

Aldo Scarpa, M.D., Ph.D.

ARC-NET Research Center for Applied Research on Cancer

Department of Pathology and Diagnostics

University and Hospital Trust of Verona

Piazzale L.A. Scuro, 10, 37134 Verona

Italy

Ph: +39-045-812-4043

aldo.scarpa@univr.it

Running Title: A cross-species pancreatic neuroendocrine tumor subtypes

Type of manuscript: Research article.

Keywords: Rip1-Tag2, subtypes, Ki67 index, NET grades, biomarkers, transcriptome, integrated omics analysis, proteomics, mutations, metabolomics, histopathology of PanNETs, cell-of-origin, tumor subtypes, EMT, tumor progression pathway, pyruvate metabolism, flux assay, patient samples, angiogenesis, genetically engineered mouse model, personalized medicine.

Conflict of Interest

L.C.C. owns equity in, receives compensation from, and serves on the Board of Directors and Scientific Advisory Board of Agios Pharmaceuticals. Agios Pharmaceuticals is identifying metabolic pathways of cancer cells and developing drugs to inhibit such enzymes to disrupt tumor cell growth and survival.

Abstract.

Seeking to assess the representation and instructive value of an engineered mouse model of pancreatic neuroendocrine tumors (PanNET) for its cognate human cancer, we profiled and compared mRNA and microRNA transcriptomes of tumors from both. Mouse PanNET tumors could be classified into two distinctive subtypes, well-differentiated islet/insulinoma tumors (IT) and poorly differentiated tumors associated with liver metastases, dubbed metastasis-like primary (MLP). Human PanNETs were independently classified into these same two subtypes, along with a third, specific gene mutation-enriched subtype. The MLP subtypes in human and mouse were similar to liver metastases in terms of microRNA and mRNA transcriptome profiles and signature genes. The human/mouse MLP subtypes also similarly expressed genes known to regulate early pancreas development, whereas the IT subtypes expressed genes characteristic of mature islet cells, suggesting different tumorigenesis pathways. Additionally, these subtypes exhibit distinct metabolic profiles marked by differential pyruvate metabolism, substantiating the significance of their separate identities.

Statement of Significance. This study involves a comprehensive cross-species integrated analysis of multi-omics profiles and histology to stratify PanNETs into subtypes with distinctive characteristics. We provide support for RIP1-Tag2 mouse model as representative of its cognate human cancer with prospects to better understand PanNET heterogeneity and consider future applications of personalized cancer therapy.

Introduction.

Genetically engineered mouse (GEM) models of human cancer have fueled progress in understanding mechanisms of tumor development and progression in different organs, induced by various driving oncogenes and/or loss of tumor suppressors (1). Among these, the RIP1-Tag2 (RT2) mouse model, in which pancreatic neuroendocrine tumors (PanNET) are induced by expression of the SV40 T-antigen oncogenes in insulin-producing islet ß cells (2, 3), has proved a valuable prototype for studying the stepwise progression of multistage tumorigenesis. For example, this model has revealed the angiogenic switch (4), the importance of attenuating apoptosis (5), and the determinants of progression to an invasive growth phenotype (6-9). The RT2 model has also proved to be tractable for preclinical trials of targeted therapies. For example, preclinical trials of angiogenesis inhibitors targeting the VEGF signaling pathway (10-15) predicted efficacy and incentivized clinical trials (16) that led to the approval of sunitinib for treating human PanNET. These results notwithstanding, a persistent question concerns the extent to which these tumors, induced by a viral oncogene that abrogates the Tp53 and Rb tumor suppressors, represent human PanNET. This study has sought to address the question via independent profiling followed by comparative analysis (cross-filtering) of the mRNA and microRNA (miR) transcriptomes of tumors from the mouse model and from human patients.

Human PanNET can be categorized based on World Health Organization (WHO)'s classification into comparatively benign, well-differentiated neuroendocrine tumors (WD-NET, also known as NET Grade G1, with 0 to 2% Ki67 cellular proliferation marker-based immunolabeling; or G2, with 3 to 20% Ki67 labeling) that can be functional or non-

functional, secreting in the former case insulin (insulinoma) or other islet cell hormones. Aggressive, poorly differentiated neuroendocrine carcinomas (PD-NEC or NEC G3, with >20% Ki67 labeling), on the other hand, are mostly non-functional, and are defined by loss or marked down-regulation of the islet cell hormone genes that define their origins (17-21). However, it is increasingly being recognized that not all G3 neuroendocrine neoplasms are poorly differentiated. In fact, a proportion of G3 aggressive tumors display well differentiated morphology (22).

While the majority of the PanNETs are sporadic, a certain fraction represents familial hereditary disease that result from inactivating mutations in the multiple endocrine neoplasia type 1 (*MEN1*) gene (23, 24). Notably, exome sequencing of human PanNET has confirmed somatic inactivation of *MEN1* in about 40% of cases and revealed a spectrum of somatic mutations in genes associated with chromatin remodeling (*DAXX*, encoding death-domain-associated protein, and *ATRX*, encoding α-thalassemia/mental retardation syndrome X-linked protein) and in negative regulators of the PI3K/mTOR pathway (*TSC2*, encoding tuberous sclerosis protein, and *PTEN*, encoding a lipid phosphatase that modulates PI3K signaling) (23, 24). However, these tumor genome analyses have not provided clarity into the underlying determinants that correspond to the varying degrees of malignancy observed in human PanNET.

The RT2 model develops PanNETs of varying malignancy, including both adenomatous (encapsulated tumors with well defined margins which are less invasive into the surrounding exocrine pancreas) islet tumor (IT) and invasive carcinomas (IC), the latter being defined in part by loss of E-cadherin expression (6) and desmosomes (25). Expression of the SV40

large T antigen hybrid oncogene in the ~400 pancreatic islets elicits a highly synchronous tumorigenesis pathway: hyperplastic/dysplastic islets begin arising at 3 weeks of age, of which about 25% switch on angiogenesis by 6-9 weeks, followed by the formation of PanNET; at end stage of 14-16 weeks, every mouse has 2 to 10 solid tumors (2, 3). This synchronicity might have predicted homogeneity in histologic and genetic phenotypes, in contrast to the *bona fide* human cancer. Interestingly, however, there is evidence for genetic and phenotypic heterogeneity. First, genome profiling by array CGH has revealed that developing PanNETs in this model acquire a range of chromosomal aberrations (26, 27). Second, profiling of the miR transcriptome during the stereotypical stages of tumorigenesis in this pathway revealed, in addition to stepwise changes in miR expression during progression, the existence of two distinctive tumor subtypes, defined by markedly distinct miR signatures. The majority of tumors profiled fell into a class that was similar to the earlier tumor development stages (hyperplastic/dysplastic islets and angiogenic islets), and were designated as canonical IT. The remaining primary tumors clustered together with a very distinct miR profile, and had clear similarity to a small sample set of liver metastases. This latter set, defined by its miR signature as a distinct subtype, was dubbed 'Metastasis-Like Primary' (or met-like primary, MLP) (28). Herein, we establish the existence of human miR PanNET subtypes that resemble those seen in mice by characterizing the human and mouse PanNET transcriptomes. We further use this information to define associated phenotypic characteristics that are suggestive of differing metabolism and distinct cellular phenotypes or tumorigenesis pathways of PanNETs.

Results.

miR profiling of human PanNET reveals distinctive molecular subtypes similar to mouse PanNET

Given the identification of PanNET subtypes in the mouse model by miR profiling (28), we began by similarly analyzing miR profiles from 40 human PanNETs, including liver or lymph node (LN) metastases, designated as our core clinical miR dataset (from a previous report (29)). We chose 88 variable miRs that have standard deviation (SD) greater than 0.5 across samples and asked whether the tumors could be clustered into distinctive subtypes using the consensus clustering-based non-negative matrix factorization (NMF) method, as we have described previously (30-32). The human PanNETs clustered into three distinct subsets with an NMF-based cophenetic coefficient (cc) – indicative of a robust number of clusters (**Supplementary Figure 1A**) – being highest for $k=3$. These clusters were named miR-cluster-1, miR-cluster-2, and miR-cluster-3. To identify subtype-specific miR signatures, we performed Significance Analysis of Microarrays (SAM) (33), as described in our previous publications (31, 32). We identified a 30 miR signature (PanNETassigner-miR) that was differentially expressed across the three subtypes (**Figure 1A and Supplementary Table 1a**). The three miR PanNET subtypes were subsequently validated with an independent human PanNET miR profiling dataset (n=50; miR validation dataset), applying the PanNETassigner-miR signatures and NMF analysis (**Supplementary Figure 1B and C and Supplementary Table 1b**).

Next, we cross-filtered the mouse and human miR profiles to assess possible similarities between the two miR subtypes seen in the mouse (28) and the three miR clusters found in human PanNET. We performed this analysis by mapping the mouse MLP signature miRs

from Olson *et al.* (28) onto the human PanNET data. Indeed, eight mouse MLP- and metastases-specific miRs (miR-137, -132, -181a2, -181a1, -23b, -27b, -24-1, and -24-2) were highly expressed in the majority of the samples in miR-cluster-1 and -2. Among these eight miRs, miR-137, the most highly enriched miR in the mouse MLP subtype, was highly expressed in the human miR-cluster-2 (**Figure 1B**). Given the implication that the mouse model was in fact recapitulating some of the heterogeneity seen in human disease, according to this metric, we went on to profile the mRNA transcriptomes of both human and mouse PanNETs, as well as the premalignant stages in the mouse model.

mRNA profiling confirms the existence of molecular subtypes in human PanNET

Human mRNA transcriptomes (n=86; a core clinical gene expression dataset that primarily included non-functional PanNETs, insulinomas and normal pancreatic islet samples) obtained from a previous study (34) were analyzed using NMF consensus clustering analysis. This dataset includes 4 normal islets, 7 liver and lymph node metastases, and 75 primary tumors with or without associated metastasis. This dataset was analyzed after selecting variable genes that were defined to have SD>0.8, which revealed five clusters of PanNET (cc highest for *k=5;* **Figure 2A, Supplementary Figure 2A and B and Supplementary Table 1c**). One of these clusters consisted in large part (28%) of normal human islet samples and was therefore termed "normal islet-like". This cluster was not considered further as a tumor subtype in our current study. Among the four other subtypes, the one termed "islet/insulinoma tumors" (IT) consisted primarily of less aggressive insulinomas (n=16) that expressed insulinoma-associated genes, such as *INS*, *IAPP* and *INSM1*. None of these tumors from this subtype were associated with metastases.

Two of the other subtypes consisted primarily of samples (n=31) categorized as non-functional that were associated (in more than 50% of primary tumor samples) with metastasis. As such, we termed these groups MLP-1 and MLP-2 using nomenclature from our work classifying tumors from the RT2 mouse model (28). In general, MLP-1 and MLP-2 share signatures that are enriched for genes associated with fibroblasts/stroma, stem cells and hypoxia (**Figure 2B**; additional information about these genes is available in **Supplementary Table 1d**). While these subtypes have subtle differences in their constitution (e.g. 45% of the MLP-2 samples were from liver/LN metastases, in contrast to only 5% of the MLP-1 subtype samples; **Figure 2A**), on the whole, there were more similarities than differences. As such, MLP-1 and MLP-2 were subsequently considered as a single MLP PanNET subtype, thereby reducing the total number of human PanNET transcriptome subtypes to three. Three primary PanNETs with functional symptoms of insulinoma, gastrinoma, and Cushing's syndrome (secretes ectopic adrenocorticotropic hormone; ACTH) were classified as MLP-1 and MLP-2, respectively. These tumors were metastatic and hence, probably denote that rarely functional PanNETs with malignant phenotype can be classified into MLP subtype. The third distinctive subtype was predominantly composed of non-functional PanNETs (n=25), of which ~16% of primary tumors were associated with distant metastasis (**Figure 2A**). Additionally, this subtype shared many genes with the IT subtype while also being moderately associated with metastasis. As such, it was provisionally named the "intermediate subtype".

To validate the robustness of these aforementioned subtypes in other PanNET datasets, we derived gene signatures (PanNETassigner-mRNA; 221 genes) specific to each of the

subtypes using SAM (**Figure 2A and Supplementary Table 1c**). We then used another PanNET dataset (n=29; gene expression validation dataset) and mapped the PanNETassigner-mRNA signatures onto the new dataset and performed NMF analysis (**Supplementary Figure 2B-C and Supplementary Table 1e**). Concordantly, in this validation dataset we observed all the five (including normal-like and two MLP) mRNA expression subtypes.

Comparison of miR and mRNA subtypes in human PanNET

Next, we sought to associate the human PanNET mRNA and miR transcriptome subtypes using the matched samples (n=36) from the core clinical gene (mRNA) expression dataset and the miR dataset. Interestingly, each of the three miR subtypes was significantly enriched in one of the three mRNA transcriptome subtypes (**Figure 2C-D**). Tumors from miR-cluster-1 were significantly enriched in the intermediate subtype; tumors from miR-cluster-2 enriched in the MLP-1 sub-subtype; and finally, miR-cluster-3 enriched in the IT subtype (**Figure 2C-D**). Due to low sample size in MLP-2 subtype, it is not significantly associated with any of the miR subtypes. However, both of the two MLP-2 samples belong to miR-cluster-2 (**Figure 2D**). Given the regulatory nature of miR for mRNA, concordance between miR and mRNA subtypes is not unexpected, but serves to further validate the existence of distinct subtypes of human PanNET.

Transcriptome subtypes of mouse PanNET and its neoplastic stages

Having established the existence of three human PanNET subtypes, we next investigated the mRNA expression profiles from the progressive stages of disease in RT2 mice, including

pools of normal (n=3), hyperplastic/dysplastic (n=3) and angiogenic (n=3) islets, as well as individual tumors (n=10) and liver macrometastases (mets; n=3). As expected, hierarchical clustering (genes selected with SD>0.8 across samples and average linkage method applied) grouped these samples according to the stage of disease progression, with all of the normal, hyperplastic and angiogenic islets grouped into one main cluster, while tumors and mets grouped into a second main cluster (**Figure 2E and Supplementary Table 1f**). The tumor samples were further split into two distinct clusters, one of which shared similarities with the metastasis samples. Based on our current understanding, we expected these tumors to be the MLPs, with the remaining tumors being ITs. In an effort to identify distinguishing characteristics between these mouse tumor subtypes, we performed differential gene expression analysis using the SAM analysis and identified 394 genes (m-PanNET-assigner-mRNA) that were differentially expressed (**Figure 2F and Supplementary Table 1g**). These data indicate that both human PanNET and RT2 tumors can be classified into distinct subtypes based on their miR or mRNA expression profiles.

Cross-species analysis of human and mouse PanNET

Next, we sought to further investigate the similarities between the mouse and human PanNET subtypes, this time at the mRNA level. To this end, we first compared mRNA expression profiles of the distinct subtypes of human PanNET and RT2 tumors. Human PanNET mRNA expression profiles from the MLP, IT and intermediate subtypes and mouse PanNET mRNA expression profiles from the MLP and IT subtypes were combined and analyzed using two different batch correction methods after selecting those genes that had a SD>0.8 from each dataset. First, we applied the ComBat ("combatting batch effects when combining batches of

gene expression microarray data") method (35) directly to the SD selected datasets to remove batch effects, followed by hierarchical clustering. Secondly, we applied distance weighted discrimination (DWD) (36) analysis after normalizing the data (described in the methods section below), followed by hierarchical clustering to validate the results from ComBat. Both of these analyses revealed that the majority of mouse and human MLPs clustered together, and that the mouse and human ITs clustered together. None of the mouse tumors clustered with human tumors of the intermediate subtype (**Figure 2G** and **Supplementary Table 1h** for ComBat analysis; **Supplementary Figure 3** and **Supplementary Table 1i** for DWD analysis). In **Figure 2G**, six human MLP tumors clustered more closely than most MLPs with the human and mouse ITs, and these are suspected to represent a subset of human MLPs with increased IT-specific genes (similar to the subset referred to below as insulin-high MLP mouse and human tumors). Similarly, four human ITs clustered more closely with the majority of the human and mouse MLPs, again suggesting that on rare occasion ITs can be aggressive and metastatic. In sum, this cross-species analysis clearly demonstrates concordance between the two human and mouse PanNET subtypes - IT and MLP.

Mutational associations in PanNET subtypes

The intermediate human subtype is not observed in the RT2 mouse model (**Figure 2G**). One possibility is that this subtype has a molecular pathogenesis different from the concomitant loss of function in the *TP53* and *RB* tumor suppressor pathways (as is suspected for this and many forms of human NET), but rather through the loss of the *MEN1* tumor suppressor. In order to address this hypothesis, we performed targeted next-generation DNA sequencing using 65 human tumor specimens representing all three subtypes. Indeed, we observed that

50% (n=24) of the intermediate tumors were *MEN1* mutant, whereas less than 13% of tumors from each of the other two subtypes had mutations in *MEN1* ($p<0.05$; Fisher's exact test; **Figure 3A**). These data provide a possible explanation for the absence of intermediate subtype tumors amongst those samples from the mouse model, and provide insight into the etiology of intermediate subtype human PanNET tumors. As such, we have renamed this subtype as "Men1-like" to reflect this association. We infer that *Men1* mutations are not occurring in either IT or MLP subtypes in the RT2 mouse model of PanNET, as evidenced by the lack of detectable copy number losses by comparative genomic hybridization (CGH) on mouse chromosome 19 where *Men1* resides (26, 27). Notably, most MEN1-related human PanNETs have reported loss of heterozygosity detectable by CGH (37).

Also of note, 42% (n=24) and 28% (n=25) of the human Men1-like/intermediate and MLP subtype tumors, respectively, had recently identified PanNET-specific mutations in *DAXX/ATRX* (24), whereas none of the IT tumors exhibited this mutation ($p<0.05$; Fisher's exact test; **Figure 3A**), confirming recent reports that *DAXX/ATRX* mutations are enriched in more aggressive non-functional PanNETs (24, 38). We observed that 21% (n=24) of the intermediate/MEN1-like tumors had both *MEN1* and *DAXX/ATRX* mutations (29% (n=24) of the intermediate/MEN1-like tumors did not have either *MEN1* or *DAXX/ATRX* mutations). In RT2 mice, we did not detect *DAXX/ATRX* mutations, as assessed by DNA sequencing of 24 RT2 mouse tumor samples, including 14 MLPs and 10 ITs.

In addition, we evaluated the association of the mTOR pathway genes *TSC2* and *PTEN*, which are mutated in PanNETs, with the human transcriptome subtypes (24). We observed that both MLP (24%; n=25) and IT (19%; n=16) subtypes are associated with *TSC2* or *PTEN*

mutations. In contrast, only 4% (n=24) of the intermediate/MEN1-like tumors have *TSC2* or *PTEN* mutations (**Figure 3A**). Similar to the mTOR pathway genes, the DNA repair pathway gene *ATM* also is also mutated in PanNETs (24). Again, MLP (16%; n=25) and IT (25%; n=16) subtypes are identifiable (not statistically significant) with *ATM* mutations, whereas, only 4% (n=24) of the intermediate/MEN1-like tumors have this mutation (**Figure 3A**). However, these associations of mTOR pathway genes (*TSC2* and *PTEN)* or *ATM* mutations with subtypes are not statistically significant (**Figure 3A**), which may be attributed to the limited sample size of tumors.

Overall, this integrated analysis revealed that *MEN1* mutations are significantly enriched in intermediate subtype (which is renamed accordingly as the MEN1-like subtype); *DAXX/ATRX* mutations significantly associated with the MLP and MEN1-like/intermediate subtype; and *TSC2/PTEN/ATM* mutations associated (not statistically significant) with the MLP and IT subtypes. These findings warrant further validation using increased sample size.

Subtype-specific associations in tumor functionality, biomarker expression, proliferative index and vascularity

Next, we sought to histologically characterize the transcriptome subtypes of PanNET based on tumor differentiation state and insulin protein expression. We identified three histopathological subtypes (histotypes) of PanNET in the mouse model, with the following distinctive characteristics: *a)* moderately- to well-differentiated tumors that stained highly positive for insulin, ostensibly the IT subtype (**Figure 3B, top**), *b)* less differentiated tumors that stained variably positive for insulin expression, ascribed as insulin-high MLP (Ins-hi

MLP; **Figure 3B, middle**) and *c)* a group of poorly-differentiated RT2 tumors that did not express insulin, and were therefore designated as insulin-low MLP (Ins-lo MLP; **Figure 3B, bottom**). Notably, tumors with Ins-lo MLP characteristics have recently been identified in the RT2 model of PanNET, and referred to as poorly-differentiated invasive carcinoma (PDIC) (39). Given the many characteristics this classification shares with the one we define herein as Ins-lo MLP, it is considered and renamed as such in the present study.

These three mouse PanNET histotypes cannot be linked directly *per se* with the transcriptome subtypes. In order to investigate a possible association, we identified a small number of genes that readily distinguish mouse MLP from IT tumors based on our transcriptome profiling analyses (**Figure 2E**). These included *Enpp2* (autotaxin, ATX), *Sema3e* (semaphorin 3e), *Epha3* (ephrin A3)*, Gpm6a* (glycoprotein m6a) and the genes encoding for the metabolic enzymes *Hk1* (hexokinase) and *Mct1* (*Slc16a1;* monocarboxylate transporter; **Figure 3C**); all are expressed at higher levels in MLP than IT. Seeking to associate the PanNET histotypes with the molecular subtypes, we performed laser capture microdissection (LCM) from tissue sections of mouse RT2 tumors (*n=5*) that were either positive or negative for insulin protein expression by immunohistochemistry (IHC). After microdissection, expression of the *Ins1* and *Enpp2* genes in LCM tumors was assessed using quantitative (q)RT-PCR to associate the different histotypes with the delineated transcriptome subtypes. Four of the five tumors expressed *Ins1*, of which two had relatively low *Enpp2* expression, identifying them as IT; the two others had high *Enpp2* expression, and were identified as Ins-hi MLP. In addition, one tumor had low *Ins1* expression and high *Enpp2* expression, consistent with the Ins-lo MLP subtype (**Figure 3D**). Next we performed microarray analysis of these LCM tumor samples followed by hierarchical clustering using a selected set of genes representing IT and

MLP transcriptome subtypes. Concordant with the RT-PCR results, the two Ins-hi MLP samples clustered with Ins-lo MLP whereas the two IT tumors clustered separately. **Figure 3E** shows relatively higher expression of MLP markers in Ins-hi and Ins-lo MLP tumors compared to IT tumors; in contrast, IT markers were more highly expressed in IT compared to Ins-hi and Ins-lo MLP tumors. Furthermore, in order to identify and establish IHC markers for these subtypes, we evaluated *Enpp2* as a marker by IHC for Enpp2 protein expression. Consistent with the gene expression data from **Figure 3D and E**, we observed significantly increased Enpp2 protein expression in Ins-hi and Ins-lo MLP, whereas there was decreased Enpp2 protein expression in the IT subtype (**Figure 3F**).

In addition to insulin, we observed that both the IT and Ins-hi MLP tumors expressed higher levels of chromogranin A (*Chga*) and secretogogin (*Scgn*; **Figure 3E**) compared to the Ins-lo MLP tumors, suggesting that the insulin secretory pathway is still active in the IT and Ins-hi MLP tumors, consistent with their insulin IHC staining in **Figure 3F**. Hence, by gene expression profile and histology, the mouse Ins-lo MLP and IT subtypes are evidently similar to the human MLP (incorporating both human MLP-1 and MLP-2 sub-subtypes) and IT subtypes, respectively. Overall, these results clearly link the transcriptome and histotypes of human and mouse PanNET and demonstrate the existence of two robust subtypes in common: IT and MLP.

We further corroborated ENPP2 as a marker in human PanNETs by analyzing *ENPP2* expression in our core clinical gene expression dataset. We found that tumors of the MLP-2 subtype have relatively high *ENPP2* gene expression compared to MLP-1 and IT. In addition, MLP-2 subtype tumors have relatively low gene expression of insulin (**Supplementary**

Figure 4). Next, we tested ENPP2 protein expression in human PanNETs using a tissue microarray (TMA; *n=20*) and IHC. Interestingly, 11 human PanNETs were negative for insulin, but positive for ENPP2 (**Figure 3G and Supplementary Table 3a**). These tumors are similar to the mouse Ins-lo MLPs, and probably represent human Ins-lo MLPs. However, some of these human *ENPP2* high and insulin negative tumors could represent intermediate/MEN1-like subtype tumors (**Supplementary Figure 4**). In addition, we observed 2 human PanNETs in our TMA were positive for insulin, but negative for ENPP2, probably representing the IT subtype. There were also 3 human PanNETs that were positive for both insulin and ENPP2, which likely represent the ins-hi MLP subtype (**Supplementary Table 3a**). However, an increased sample size is required for definitive evaluation of the human PanNET subtype protein markers by IHC. Collectively, these results suggest that ENPP2 may serve as a biomarker that – along with insulin - can be used to distinguish human and mouse PanNET subtypes (and even sub-subtypes of MLP).

To further characterize the phenotypes of distinct PanNET subtypes in RT2 mice, we evaluated the tumor vasculature by immunostaining for CD31. As presented in **Figure 4A** (top panel), and consistent with the clinical presentation of human IT and our previous studies in the RT2 model, mouse IT samples were highly vascularized. Interestingly, the MLP tumors were comparatively less vascularized. IT samples contained about 40-60% more CD31 positive cells than Ins-hi and Ins-lo MLP tumors (**Figure 4B**).

We additionally assessed cancer cell proliferation in human and mouse tumors by Ki67 staining, which is clinically used for assessing NET grades (17-21). We observed a statistically significant increase in Ki67 staining in the mouse Ins-lo MLP tumors compared

to the IT samples, whereas the differences with Ins-hi MLP tumors were not significant (**Figure 4A, bottom panel and 4C**). This characteristic of increased Ki67 staining in the RT2 Ins-lo MLP sub-subtype is also evident in the human tumors, where MLP tumors (equivalent to mouse Ins-lo tumors) exhibit a substantially higher Ki67 staining compared to Men1-like/intermediate and insulinoma subtype PanNETs (less than 3%), albeit lower overall (**Figure 4D**). The difference in frequencies of Ki67 positive cells between mouse and human PanNETs may be attributable to the effects of the driving oncogene - SV40 large T-antigen – in the mouse PanNET cancer cells.

Association of NET grades with PanNET transcriptome subtypes

In order to ascertain whether the transcriptome subtypes are associated with the WHO classification of NET grades, we sought to associate the core clinical mRNA dataset with NET grade information. A majority of the IT (81%) and intermediate (75%) subtypes are identifiable as NET G1, while the remainder are NET G2. In contrast, 60% of the MLP subtype PanNETs are NET G2, and 16% are NET G3; the remaining 24% are NET G1 (**Figure 4E**). Thus lower grade (including NET G2) human PanNETs are heterogeneous, variably associating with all three transcriptome subtypes, whereas high grade NET G3 tumors are exclusively associated with the MLP subtype.

Genes regulating early pancreas development distinguish MLP from IT

Given the highly distinctive mRNA and miR profiles of the PanNET subtypes, we hypothesized that these subtypes may be the result of tumors arising from different cellular

origins. To identify the putative cell-of-origin for the IT and MLP subtypes, we examined the expression of various pancreatic stem/progenitor- and differentiation-specific genes. In both human and mouse PanNETs, the MLP samples were enriched for expression of pancreatic progenitor-specific genes, such as *Hnf1b* (40) and *Gata6* (41), and the stem cell factor *Lin28b* (42) (*Lin28b* significant only in mouse PanNETs), relative to the IT samples (**Figure 5A-B**). Collectively, this pattern of gene expression suggests that the MLP subtypes share characteristics with β-cell precursors or immature β-cells. Conversely, the IT subtype differentially expresses a number of mature β-cell–specific genes, including *Pdx1*, *Ins1*, *Nkx2-2*, *Pax6*, *Mafa*, *Nkx6-2*, *Gck* (*Gk*) and *Slc2a2* (*Glut2*), and the insulinoma-specific gene *Insm1* (**Figure 5A-B**). These results provide evidence that the IT subtype resembles mature β-cells.

To further investigate a possible stem/progenitor pancreatic cell origin of MLPs, we compared the gene expression profiles of the PanNET subtypes with those of stem/progenitor islet cells from early pancreatic developmental stages. Islet ß-cells originate during two sequential regulatory phases of pancreas development, termed the primary and secondary transitions. The primary transition occurs between e9.5 and e12.5, whereas the secondary transition begins at e13.5 and ends at e15.5. During the primary transition, the developing pancreas primarily contains multi-lineage precursor cells and a small number of immature β cells. These precursor and immature cells undergo a differentiation process during the secondary transition to commit to separate pancreatic cell lineages, including one that produces the islet β-cells (43). We used the ComBat algorithm to merge the gene expression profiles from RT2 tumors (IT and MLP) with a published dataset (GSE8070) consisting of serial samples collected from the dorsal pancreas of the mouse embryo at each day from

e12.5 to e16.5. We observed that the samples from e12.5 to e14.5, which are part of the late primary transition and early secondary transition, clustered tightly with the MLP and its associated liver metastasis samples, whereas the majority of the IT samples clustered with the e15.5 and e16.5 samples that are representative of the late secondary transition period (**Figure 5C**). It is notable that the RIP-Tag2 transgene is activated at e9.5 (44) and is demonstrably expressed in the multi-potential progenitors; its expression in immature ß cells can be surmised by its activation during the primary transition. These results indicate that MLPs share characteristics with immature β-cells of the pancreas and presumably with the recently described rare adult pancreatic stem/progenitor cells (45). In contrast, the IT cancer cells are more closely related to mature islet ß cells of the pancreas than to these progenitors. Moreover, a set of epithelial-mesenchymal transition (EMT)-specific genes – including *Vim* - was highly expressed in MLPs compared to ITs (**Figure 5B**). The expression of EMT genes in MLP in **Figure 5B** along with our previous report of EMT in RT2 mouse cell lines (28) suggest that a dedifferentiation process is probably also taking place in a subset of tumors that progressed from IT into MLP tumors.

To gain further insights into the canonical IT multistage tumorigenesis pathway, we used the ComBat based merging of data to compare the gene expression profiles of samples from the different stages of tumor progression with profiles (from Chun *et al*. (25)) of microdissected non-invasive IT and the broadly invasive IC2 that had been microdissected from tissue sections of RT2 tumors. Interestingly, we observed that MLP tumors have similarities to those of the microdissected invasive regions of IC2 tumors (**Supplementary Figure 5A**). Similar to MLP, the microdissected invasive IC2s showed increased expression of EMT gene *Vim*. On the other hand, the non-invasive IT tumors showed increased expression of mature ß

cell genes such as *Ins1*, *Pdx1* and *Insm1* (**Supplementary Figure 5B**). These data suggest that IC2s arising via progression from ITs switch on an invasive program similar to that of MLP tumors.

In order to clarify whether both Ins-hi and Ins-lo MLP tumors have increased expression of β-cell precursors genes, we further examined the LCM-derived tumor samples described in **Figure 3D-E**. Interestingly, we observed that only the Ins-lo MLP tumors have increased expression of majority of the β-cell precursor genes compared to the Ins-hi MLP and IT tumors (**Figure 5D**). The above results motivated us to further examine the histological differences in hyperplastic/dysplastic islets in the RT2 mouse model, a premalignant stage we had historically considered to be homogeneous. We analyzed 8 to 10 week old mice, and observed heterogeneity in this pre-malignant stage: some hyperplastic/dysplastic islets had characteristics of Ins-lo MLP, whereas the majority had the expected characteristics of Ins-hi MLP or IT (**Figure 5E**). Although dysplastic islets can be detected with poorly differentiated features and reduced insulin secretion, similar to Ins-lo MLP tumors, they are infrequent (2.8% of 145 premalignant hyperplastic islet lesions evaluated from 12 mice of 8 and 10 week old had this phenotype) (**Figure 5E**). Overall, these results, along with the above mentioned analysis comparing PanNET tumors with the stem/progenitor cells of the pancreas, suggest the existence of an early and separate cell- and/or pathway-of-origin for pro-metastatic Ins-lo MLP tumors, whereas Ins-hi MLP arise either directly from mature β cells or via progression of ß cell-derived IT tumors.

Differences in central carbon metabolism illustrate subtype-specific similarities to distinct stages of β-cell development

Among the panel of genes differentially expressed between the MLP and IT subtypes was a set involved in central carbon metabolism. Tumors from the IT subtype demonstrably expressed metabolic enzymes of the β-cell lineage, including the glucose transporter *Glut2* (*Slc2a2*), glucose 6-phosphatase (*G6pc2*) and glucokinase (*Gck)*. In contrast, the lactate transporter *MCT1 (SLC16a1)* and hexokinase (*Hk1)*, which are not expressed in mature β-cells (46), were considerably enriched in both mouse and human samples from the MLP subtype (**Figure 6A**). In addition to differences in expression of enzymes involved in glucose uptake and glycolysis, human and murine IT tumors exhibited increased expression of pyruvate carboxylase (*PC*) and cytoplasmic malic enzyme 1 (*ME1*), relative to MLP subtype tumors (**Figure 6A**). Expression of these enzymes is consistent with pyruvate cycling, a phenomenon observed in mature β-cells and insulinomas that facilitates the generation of energy in the mitochondria to sustain glucose-stimulated insulin secretion (47-49) (**Figure 6B**). The differential expression of metabolic genes in the IT and MLP subtypes was further validated at the protein expression level. Global, mass spectrometry-based proteomic analyses of mouse IT and Ins-lo MLP tumor samples confirmed that Glut2 and Pc were highly expressed in IT but not in Ins-lo MLP, whereas Mct1 and Hk1 protein levels were increased in Ins-lo MLP compared to IT (**Supplementary Table 3b**). In addition, protein expression of Hk1 and Pc protein in the Ins-lo MLP subtype was confirmed by immunoblot and IHC (**Figure 6C and D**).

Given these observations, we predicted that the metabolic phenotypes of these two subtypes would be similarly distinct. To study PanNET metabolism *in vitro*, we sought to identify mouse PanNET cell lines that represent each subtype. To this end, we classified cell lines

derived from RT2 mouse tumors using two genes that are among those most highly expressed in mouse MLP samples relative to IT samples (*Enpp2* and *Hk1)* and three other genes that conversely are highly expressed in IT cell lines compared to MLP (*Gck*, *Pc*, and *Glut2 (Slc2a2))* (**Figure 6E top panel).** qRT-PCR analyses revealed that two RT2 PanNET-derived cell lines, βTC1b and βTC1e (50), have MLP subtype characteristics, displaying high expression of *Enpp2* and *Hk1* in both the cell lines and low expression of *Gck*, *Pc*, and *Glut2* in both the cell lines. In contrast, two other lines, βTC4 and βTC3, had low expression of the MLP genes and high expression of the IT genes (except for *Hk1* that is heterogeneous in these two cell lines), and are thus representative of the IT subtype (**Figure 6E top panel**). The MLP subtype of βTC1b and the IT characteristics of βTC4 were further validated by immunoblot analyses confirming that the MLP cell line exhibited low expression of Pc and high expression of Hk1 compared to the IT cell line (**Figure 6E bottom panel**).

With IT and MLP cell lines in hand, we performed metabolomic profiling using a liquid chromatography tandem mass spectrometry (LC-MS/MS) based metabolomics platform (51). Consistent with our expectation, we found that the IT (βTC4 and βTC3) and MLP (βTC1b and βTC1e) cells had distinctive global metabolic profiles (**Supplementary Figure 6A-B**). Among these differences, we observed an increase in the relative abundance of metabolites involved in pyruvate cycling, including aspartate, fumarate, and malate (**Figure 6F**).

To further explore differences in metabolites as well as features of central carbon metabolism between these subtype-representative cell lines, we traced glucose metabolism using isotopically carbon-13 (^{13}C) labeled glucose in cell lines representing the IT (ßTC4) and MLP (ßTC1b) subtypes. Cells were grown overnight in media containing uniformly ^{13}C-labeled

glucose to achieve steady state saturation of label, and then subjected to targeted LC-MS/MS, as above. Here, incorporation of ^{13}C was monitored at each carbon position for metabolites in pyruvate cycling and the TCA cycle. Indeed, we observed an elevation in abundance of the glucose-derived M+3 isotopomer in several TCA intermediates in the IT cells (**Figure 6G**), indicative of pyruvate entry into the TCA cycle via pyruvate carboxylase (Pc). In addition, we observed that the IT cell line had a five-fold increased dependence on glucose for proliferation (**Figure 6H**). Glucose-derived carbon may be particularly important to these cells for energy generation via pyruvate cycling, consistent with the metabolic characteristics of mature β-cells, which harness such energy to produce and release insulin (52). The increased pyruvate cycling observed in IT cells provides evidence for the retention of mature β-cell metabolism in ITs, in contrast to the distinctive metabolic characteristics of the MLP cell line. Thus, the distinctive metabolic phenotypes of these subtypes of PanNET further delineate them as such.

Discussion

The low incidence of PanNET has made it difficult to correlate apparent heterogeneity in clinical presentation with responses to therapies (23). By using dual miR and mRNA transcriptome profiling analysis, we have revealed the existence of three distinctive molecular subtypes and associated biomarkers in human PanNET. Remarkably, despite its rapid and synchronous ontogeny, a PanNET GEM model (RT2) is demonstrably recapitulating two of these subtypes based on analysis of the transcriptome, the metabolome, selected protein expression, histologic and phenotypic presentation, including distant metastasis. Hence, the distinct pathophysiology and molecular mechanisms behind two of the

three human PanNET subtypes, MLP and IT, can be further studied using the RT2 model as a surrogate for human (both insulinoma and non-functional) tumors. In addition, the RT2 model reveals distinctive angiogenic phenotypes between PanNET subtypes, with decreased blood vessels in MLP compared to IT, predicting analogous differences in human PanNETs that are worthy of future investigation. Although the RT2 model does not represent the MEN1-like/intermediate subtype, due to apparent necessity of MEN1 mutations to elicit this tumor type, its selective presence in a discrete subtype of human PanNET tumors provides clarity in regard to heretofore unrecognized heterogeneity among non-functional PanNETs. The characteristics of the distinct subtypes of human and mouse PanNETs are summarized in **Figure 7A-B**.

Although treatment decisions for PanNET patients are based on NET grade, which is assessed in part using Ki67 labeling, there is heterogeneity in patient responses to standard-of-care therapies (23, 53). This heterogeneity in NET G2, for example, may be explained by the observation that G2 tumors are distributed across all three transcriptome subtypes, with highest prevalence in the metastasis-associated MLP subtype. In addition, the observation that there is a higher density of CD31 positive microvessels in the less aggressive mouse IT subtype compared to the MLP subtypes corroborates data indicating that increased microvessel density is associated with better histoprognostic factor (and negatively correlated with tumor progression), and *vice versa* in PanNET patients (54-57). This variable association of transcriptome subtypes with histological grades and microvessel density in regard to therapeutic responses should be further assessed in additional patient samples.

Previous efforts to molecularly sub-classify other human tumor types have implicated distinctive cells of origin/phenotypes in several instances (32, 58, 59). In the current study, expression of progenitor-specific genes in the MLP subtype implicates a stem cell/progenitor origin. Additionally, however, the variable expression of certain pancreatic progenitor genes amongst the mouse MLP samples (**Figure 5B**) implies that tumors of this subtype may originate from more than one progenitor cell, a possibility consistent with the observation that the SV40 oncogene is expressed in multiple β-cell precursors of the embryonic mouse pancreas (44). Further support for the proposition that MLP tumors are not derived from differentiated β-cells comes from the observation that the mature β-cell-specific transcription factor *MafA* is poorly expressed in MLP tumors (**Figure 5B**) (41, 43).

The data from this and a previous report (28) implicate two pathways to invasive carcinomas, which either have the complete MLP signature, or a subset of it. Tumors with a partial MLP signature are envisaged to arise out of the canonical islet tumorigenesis pathway that leads to IT, which then progress to IC2/MLP-ins-lo tumors. The second pathway, already invasive in small islet-like lesions (**Figure 5E**), has the full MLP signature, including expression of pancreatic progenitor genes, and we hypothesize that it originates from islet progenitor cells. The alternative pathways to MLP in RT2 mouse are presented in **Figure 7B**.

Beyond molecular profiles of stem cell- and lineage-specific gene expression, we also observed subtype-correlated variation in metabolic enzymes, and in metabolic activities, providing additional evidence in support of different cells-of-origin/phenotype for IT and MLP tumors. During terminal β-cell differentiation, GCK – with its lower affinity for glucose

- replaces HK1; the K_m of GCK is in the physiological range for blood glucose, and thus GCK can act as a sensor, relating glucose concentration in the circulation to insulin production and secretion. Similarly, expression of the MCT family of cation transporters is turned off in terminally differentiated β-cells, since lactate/pyruvate uptake would feed directly into the tricarboxylic acid (TCA) cycle, bypassing the glucose sensing feature of GCK (46). Thus MLP express HK1 and MCT1, much like progenitors, whereas IT express GCK but not HK1 or MCT1, much like normal islet ß cells. Consistent with this line of reasoning, IT tumors express other β-cell restricted genes, whereas MLP tumors do not (**Figures 5A-B, 6A, 6C-D and Supplementary Table 3b**). Moreover, our metabolomic analysis (both metabolite profiles in **Figure 6F** and **Supplementary Figure 6** and ^{13}C-labeling experiments in **Figure 6G**) of PanNET cancer cells revealed distinct metabolic profiles for IT and MLP cell lines, which provide further support for distinctive cells of origin. Namely, we observed enriched activity in pyruvate cycling, a β-cell phenomenon (47-49), in the IT subtype both in terms of metabolic enzyme expression and metabolite profiles. Taken together, our data suggest that islet tumors have mature β-cell characteristics, whereas MLP tumors have islet cell precursor characteristics. A revised summary of tumor progression in the RT2 model (3), reflective of what our results suggest for the human IT and MLP subtypes, and now including information regarding the cell-of-origin/phenotype, is presented in **Figure 7B**.

Human PanNETs show considerable genetic heterogeneity and malignant potential. Notably, Jiao, *et al*. recently sequenced the exomes of 10 non-familial PanNETs, seeking to identify the most frequent mutations in this tumor type (24); their analysis revealed a variety of low frequency mutations with potential functional roles, but no prevalent mutations in commonly

known driving oncogenes or tumor suppressor genes (TSGs). Now, we have stratified human and mouse PanNETs into subtypes, utilizing a wealth of data from miR, mRNA, mutation and metabolome profiling. Furthermore, our current study has revealed that *MEN1* mutant tumors are significantly associated with the intermediate human subtype we defined; additional studies will be required to determine if those MEN-like/intermediate subtype tumors that have not suffered *MEN1* loss-of-function mutations have found alternative means to abrogate the *MEN1* TSG pathway. Notably, Jiao, *et al.* reported new inactivating mutations in *DAXX/ATRX* in non-functional PanNETs (24). Now we demonstrate that tumors carrying these mutations can be alternatively segregated into MEN-like/intermediate and MLP subtypes with distinct transcriptomic profiles. Our finding that 21% of MEN1-like/intermediate subtype tumours also contained *DAXX/ATRX* alterations suggests that these mutations are probably late events in PanNET pathogenesis driven by *MEN1* loss, consistent with the detection of *DAXX/ATRX* mutations in PanNET patients with *MEN1* syndrome by de Wilde *et al.* (60).

In order to apply these transcriptomic subtypes clinically, it will be important to identify robust biomarkers and assays that can discriminate each of these subtypes similar to that in our previous colorectal cancer study (32). Based on the analysis of both human and mouse tissue, we suggest ENPP2 and INS as possible biomarkers that can distinguish IT and MLP subtypes via IHC and RT-PCR assays. These biomarkers warrant prospective and retrospective validation. Additional biomarkers, alongside these two, will be required to unambiguously identify the MEN1-like/intermediate subtype.

In summary, we have used an integrated multi-omics data analysis approach that includes transcriptomics (mRNA and miR), mutatomics (selected mutations) and metabolomics, along with histology and clinical information, to evaluate PanNET from patient samples and a GEM model as well as derivative cancer cell lines. This cross-species analysis has revealed and characterized previously unrecognized subtypes of PanNET in both mice and humans, and assigned to these tumor subtypes different mutations and phenotypic, clinical and pathological properties that underlie the heterogeneous biology of this disease – knowledge that may prove applicable to development of subtype-selective (personalized) therapies.

Materials and Methods.

PanNET patient samples and RNA isolation for validation studies. The miR expression validation (n=50) and matched gene expression validation (n=29; except one unmatched sample) datasets (**Supplementary Figures 1B-C and 2B-C**) involved a total of 51 primary and metastatic PanNET samples. For gene expression analysis, total RNA was isolated from tumor tissues using Ambion RiboPure kit (Life Technologies, Darmstadt, Germany) as per the manufacturer's instructions. For miR analysis, tumor tissues were sectioned using cryotome, every sixth section was evaluated using hematoxylin and eosin (H&E) staining, and total RNA was isolated using miRNeasy kit (Qiagen, Hombrechtikon, Switzerland) as per the manufacturer's instructions. This study was conducted in accordance with the Declaration of Helsinki. It was approved by the ethics committee of Charité - Universitätsmedizin Berlin, and informed consent was obtained from all patients.

RT2 transgenic mouse model of PanNET. The breeding and generation of RT2 mice has been described previously (2). While majority of the mice in this study were on a C57Bl/6 background, which is typically used for this model, we also crossed C57BL/6J (B6) RT2

mice with A/J (B6AF1) or DBA/2J (B6D2F1) mice to produce F1 mice. All animal research was performed under the auspices of animal protocols approved by Experience sur animaux (EXPANIM) – Service de la consommation et des Affaires veterinaries (SCAV) in Switzerland.

Isolation of RT2 tumor and pre-malignant lesions. RT2 mouse PanNET tumors, liver metastases, normal, hyperplastic, and angiogenic islets were dissected out or isolated as described previously (25, 28, 61).

RT2 mouse tumor cell lines: βTC cell lines were previously derived from RT2 mouse model (50). These cell lines (passages between 20 and 30) were authenticated by growth, morphological characteristics and insulin gene and protein expression. These were also tested for mycoplasma.

Processing of microarray data and subtype identification. The core clinical gene and miR expression datasets were from Missiaglia, *et al.* (GEO Omnibus ID - GSE73338) (34) and Roldo, *et al*. (GSE73350) (29). Gene expression was assessed using Affymetrix GeneChip® Human and Mouse Gene 1.0 ST arrays (Affymetrix, Santa Clara, CA, USA; GSE73339 - patient gene expression validation samples and GSE73514 - mouse RT2 samples). Human microarrays were analyzed using the statistical framework R and Bioconductor (62) as described earlier (31, 32) and mouse microarrays using the R package aroma.affymetrix (63). miR expression was assessed using Agilent Human miRNA Microarrays (based on miRBase release 16; Agilent Technologies, Santa Clara, CA, USA; for patient miR validation samples; GSE73367). Normalization of the miR profile data were performed as described (64).

For subtype identification, the variable genes (from the core clinical gene expression dataset) with SD>0.8 or miRs (from the core clinical miR expression dataset) with SD>0.5 from the normalized data were median-centered across patient samples for each gene, and used for subtype identification by applying the NMF algorithm (30) followed by SAM analysis (33) to identify significantly differentially expressed features (mRNA or miRNA), as described previously (31, 32). For (both gene and miR) validation datasets, SAM genes from the core clinical datasets were mapped, followed by NMF analysis for subtype identification.

Cross-species analysis. ComBat ("sva" Biocoductor package) (35) or DWD (36, 65) was used to correct the batch effects and merge human and mouse gene expression data after selecting variable genes with SD>0.8 across samples from each dataset. For merging the human and mouse data, we used official gene names from HUGO Gene Nomenclature Committee (HGNC) as references.

Cross-comparison of gene and miR subtypes. The hypergeometric test was performed to reconcile gene and miR subtypes, as described (66).

Laser capture microdissection. Snap frozen pancreatic sections (10μm) from 15-week old RT2 mice were mounted onto PET-membrane slides (MMI). Sections were incubated with ice-cold 70% ethanol and stained with cresyl violet. Individual tumors were microdissected using a PALM laser dissecting microscope (Zeiss, Feldbach, Switzerland) followed by RNA isolation and hybridization to Affymetrix GeneChip Mouse Gene 1.0 ST arrays.

Immunohistochemistry (IHC). IHC was performed as described in our previous studies (32, 67, 68) using the primary antibodies listed in **Supplementary Information**. Ki67 scoring was performed as previously described (34). For mouse samples, ImageJ software was used to count the total number of Ki67 positive cells normalized to hematoxylin-stained cells. The number of CD31 positive blood vessels in random 20× magnification fields was determined independently by two investigators, and averaged.

Metabolic profiling. To characterize central carbon metabolism, targeted LC-MS/MS was performed (51). Briefly, βTC cells were grown in DMEM (25mM glucose, 2mM glutamine, without pyruvate) to 80% confluency, and metabolites were extracted using 80% methanol at dry ice temperatures. Metabolite fractions were normalized to protein concentration from a biological replicate processed in parallel. For ^{13}C-labeling experiments, cells were grown in DMEM (without pyruvate, glucose and glutamine) in the presence of 10% dialyzed FBS, supplemented with ^{13}C-glucose (25mM; Cambridge Isotope Laboratories, Tewksbury, MA) for 20hrs.

Next-Generation Sequencing. Human DNA samples were analyzed with a panel targeting all coding sequences of *MEN1*, *ATRX*, *DAXX*, *PTEN*, *TSC2* and *ATM*. Next-generation targeted sequencing was performed as previously described (69).

Gene-specific primers used for qRT-PCR are provided in **Supplementary Table 4.**

R scripts associated with data analysis are available from github (https://github.com/syspremed/PanNETassigner).

A detailed description of materials and methods are available in the supplementary information.

Acknowledgements

We thank I. Michael for insightful comments on the manuscript, P. Olson (Pfizer) for preparing RNA samples from the mouse tumors for gene expression profiling while he was working in Hanahan lab. We thank K. Homicsko (EPFL) and O. Michielin (SIB) for helpful scientific discussions, E. Drori, L. Ostos, M. Schnetz, M. Peng, R. Prakash, B. Torchia and other Hanahan lab members (EPFL) for help with mouse colonies and other experiments, and P. Poudel and G. Nyamundanda (ICR) for help with validating the R scripts, data and certain results, and making an R package. We would like to extend our thanks to the following core-facilities: the Histology Core Facility, EPFL, Switzerland; Lausanne Genomic Technologies Facility, Center for Integrative genomics (CIG), University of Lausanne, Lausanne, Switzerland; the Mass Spectrometry Facility, EPFL (including R. Hamelin for processing samples and help with the data analysis); and Genomics Core Laboratory (University of California San Francisco), San Francisco, USA, for their services. A.Sa. was partially supported by a US Department of Defense Postdoctoral Fellowship (BC087768). C.A.L. was partially supported by a PanCAN-AACR Pathway to Leadership award and a Dale F. Frey award for Breakthrough Scientists from the Damon Runyon Cancer Research Foundation (DFS-09-14). The work on mutational analysis of human PanNET was supported by the Italian Cancer Genome Project (FIRB RBAP10AHJB), Associazione Italiana Ricerca Cancro

(grant n. 12182) and Fondazione Italiana Malattie Pancreas – Ministero Salute (CUP_J33G13000210001) to A.Sc. L.C.C. was supported by P01 grant CA117969. While A.Sa. performed the majority of the research at the EPFL, a portion of the work was also performed at the ICR, London. We acknowledge NHS funding to the NIHR Biomedical Research Centre at The Royal Marsden and the ICR to ASa. This study was primarily supported by a Swiss National Science Foundation research grant to D.H.

Contributions

A.Sa. and D.H. conceived the hypothesis, designed the experiments, interpreted results and wrote the manuscript. A.Sa. performed most of the bioinformatics and wet-lab experiments and prepared the figures. S.W. performed metabolite isolation, DNA isolation from mouse tumors and normal pancreas for targeted mutation profiling, proteomic profiling and analysis, microdissection and biomarker validation studies, participated in significant discussions, critically reviewed the manuscript and helped with figure preparation. C.A.L. performed the metabolic studies, prepared its associated figures and critically reviewed the manuscript. S.Ba. helped with curating the human core clinical gene and miR expression data sets and Ki67 IHC staining data; C.G. and J.K. prepared human PanNET samples and RNA for validation of microarray experiments; I.W. performed the microdissection of mouse tumor samples followed by RNA isolation; S.Be. and M.S. designed and performed next-generation sequencing and analysis on human and mouse samples; A.M. performed bioinformatic analysis of next-generation sequencing; R.T.L. coordinated patients and sample data management and supervised ethical protocols; J.M.A. supervised the LC-MS/MS experiments; H.B. performed and evaluated Ki67 IHC staining in human PanNET samples

from the validation gene expression dataset; L.C.C. participated in evaluative discussions; and B.W. supervised the RNA preparation from patient samples and information for the validation datasets. A.Sc. supervised the curation of the core clinical data sets, mutation data and patient information. The project was supervised by D.H.

References

1. Hanahan D, Wagner EF, Palmiter RD. The origins of oncomice: a history of the first transgenic mice genetically engineered to develop cancer. Genes & development. 2007;21:2258-70.
2. Hanahan D. Heritable formation of pancreatic beta-cell tumours in transgenic mice expressing recombinant insulin/simian virus 40 oncogenes. Nature. 1985;315:115-22.
3. Hanahan D. Dissecting multistep tumorigenesis in transgenic mice. Annual review of genetics. 1988;22:479-519.
4. Folkman J, Watson K, Ingber D, Hanahan D. Induction of angiogenesis during the transition from hyperplasia to neoplasia. Nature. 1989;339:58-61.
5. Christofori G, Naik P, Hanahan D. A second signal supplied by insulin-like growth factor II in oncogene-induced tumorigenesis. Nature. 1994;369:414-8.
6. Perl AK, Wilgenbus P, Dahl U, Semb H, Christofori G. A causal role for E-cadherin in the transition from adenoma to carcinoma. Nature. 1998;392:190-3.
7. Joyce JA, Baruch A, Chehade K, Meyer-Morse N, Giraudo E, Tsai FY, et al. Cathepsin cysteine proteases are effectors of invasive growth and angiogenesis during multistage tumorigenesis. Cancer cell. 2004;5:443-53.
8. Lopez T, Hanahan D. Elevated levels of IGF-1 receptor convey invasive and metastatic capability in a mouse model of pancreatic islet tumorigenesis. Cancer cell. 2002;1:339-53.
9. Li L, Hanahan D. Hijacking the neuronal NMDAR signaling circuit to promote tumor growth and invasion. Cell. 2013;153:86-100.
10. Paez-Ribes M, Allen E, Hudock J, Takeda T, Okuyama H, Vinals F, et al. Antiangiogenic therapy elicits malignant progression of tumors to increased local invasion and distant metastasis. Cancer cell. 2009;15:220-31.
11. Pietras K, Pahler J, Bergers G, Hanahan D. Functions of paracrine PDGF signaling in the proangiogenic tumor stroma revealed by pharmacological targeting. PLoS medicine. 2008;5:e19.
12. Casanovas O, Hicklin DJ, Bergers G, Hanahan D. Drug resistance by evasion of antiangiogenic targeting of VEGF signaling in late-stage pancreatic islet tumors. Cancer cell. 2005;8:299-309.
13. Bergers G, Hanahan D. Combining antiangiogenic agents with metronomic chemotherapy enhances efficacy against late-stage pancreatic islet carcinomas in mice. Cold Spring Harbor symposia on quantitative biology. 2002;67:293-300.

14. Bergers G, Song S, Meyer-Morse N, Bergsland E, Hanahan D. Benefits of targeting both pericytes and endothelial cells in the tumor vasculature with kinase inhibitors. The Journal of clinical investigation. 2003;111:1287-95.
15. Chiu CW, Nozawa H, Hanahan D. Survival Benefit With Proapoptotic Molecular and Pathologic Responses From Dual Targeting of Mammalian Target of Rapamycin and Epidermal Growth Factor Receptor in a Preclinical Model of Pancreatic Neuroendocrine Carcinogenesis. J Clin Oncol. 2010.
16. Raymond E, Dahan L, Raoul JL, Bang YJ, Borbath I, Lombard-Bohas C, et al. Sunitinib malate for the treatment of pancreatic neuroendocrine tumors. The New England journal of medicine. 2011;364:501-13.
17. Rindi G, Wiedenmann B. Neuroendocrine neoplasms of the gut and pancreas: new insights. Nat Rev Endocrinol. 2011;8:54-64.
18. Bosman FT. World Health Organization., International Agency for Research on Cancer. WHO classification of tumours of the digestive system. 4th ed Lyon: International Agency for Research on Cancer. 2010.
19. Sobin LH, Gospodarowicz MKMK, Wittekind CC. TNM classification of malignant tumours. 2009.
20. Modlin IM, Kidd M, Latich I, Zikusoka MN, Shapiro MD. Current status of gastrointestinal carcinoids. Gastroenterology. 2005;128:1717-51.
21. Scarpa A, Mantovani W, Capelli P, Beghelli S, Boninsegna L, Bettini R, et al. Pancreatic endocrine tumors: improved TNM staging and histopathological grading permit a clinically efficient prognostic stratification of patients. Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc. 2010;23:824-33.
22. Velayoudom-Cephise FL, Duvillard P, Foucan L, Hadoux J, Chougnet CN, Leboulleux S, et al. Are G3 ENETS neuroendocrine neoplasms heterogeneous? Endocrine-related cancer. 2013;20:649-57.
23. de Wilde RF, Edil BH, Hruban RH, Maitra A. Well-differentiated pancreatic neuroendocrine tumors: from genetics to therapy. Nat Rev Gastroenterol Hepatol. 2012;9:199-208.
24. Jiao Y, Shi C, Edil BH, de Wilde RF, Klimstra DS, Maitra A, et al. DAXX/ATRX, MEN1, and mTOR pathway genes are frequently altered in pancreatic neuroendocrine tumors. Science (New York, NY. 2011;331:1199-203.
25. Chun MG, Hanahan D. Genetic deletion of the desmosomal component desmoplakin promotes tumor microinvasion in a mouse model of pancreatic neuroendocrine carcinogenesis. PLoS genetics. 2010;6.
26. Hodgson G, Hager JH, Volik S, Hariono S, Wernick M, Moore D, et al. Genome scanning with array CGH delineates regional alterations in mouse islet carcinomas. Nature genetics. 2001;29:459-64.
27. Hager JH, Hodgson JG, Fridlyand J, Hariono S, Gray JW, Hanahan D. Oncogene expression and genetic background influence the frequency of DNA copy number abnormalities in mouse pancreatic islet cell carcinomas. Cancer research. 2004;64:2406-10.
28. Olson P, Lu J, Zhang H, Shai A, Chun MG, Wang Y, et al. MicroRNA dynamics in the stages of tumorigenesis correlate with hallmark capabilities of cancer. Genes & development. 2009;23:2152-65.
29. Roldo C, Missiaglia E, Hagan JP, Falconi M, Capelli P, Bersani S, et al. MicroRNA expression abnormalities in pancreatic endocrine and acinar tumors are associated with distinctive pathologic features and clinical behavior. J Clin Oncol. 2006;24:4677-84.

30. Brunet JP, Tamayo P, Golub TR, Mesirov JP. Metagenes and molecular pattern discovery using matrix factorization. Proceedings of the National Academy of Sciences of the United States of America. 2004;101:4164-9.
31. Collisson EA, Sadanandam A, Olson P, Gibb WJ, Truitt M, Gu S, et al. Subtypes of pancreatic ductal adenocarcinoma and their differing responses to therapy. Nature medicine. 2011;17:500-3.
32. Sadanandam A, Lyssiotis CA, Homicsko K, Collisson EA, Gibb WJ, Wullschleger S, et al. A colorectal cancer classification system that associates cellular phenotype and responses to therapy. Nature medicine. 2013:doi:10.1038/nm.3175.
33. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. Proceedings of the National Academy of Sciences of the United States of America. 2001;98:5116-21.
34. Missiaglia E, Dalai I, Barbi S, Beghelli S, Falconi M, della Peruta M, et al. Pancreatic endocrine tumors: expression profiling evidences a role for AKT-mTOR pathway. J Clin Oncol. 2010;28:245-55.
35. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics (Oxford, England). 2007;8:118-27.
36. Benito M, Parker J, Du Q, Wu J, Xiang D, Perou CM, et al. Adjustment of systematic microarray data biases. Bioinformatics (Oxford, England). 2004;20:105-14.
37. Pieterman CR, Conemans EB, Dreijerink KM, de Laat JM, Timmers HT, Vriens MR, et al. Thoracic and duodenopancreatic neuroendocrine tumors in multiple endocrine neoplasia type 1: natural history and function of menin in tumorigenesis. Endocrine-related cancer. 2014;21:R121-42.
38. Cao Y, Gao Z, Li L, Jiang X, Shan A, Cai J, et al. Whole exome sequencing of insulinoma reveals recurrent T372R mutations in YY1. Nature communications. 2013;4:2810.
39. Hunter KE, Quick ML, Sadanandam A, Hanahan D, Joyce JA. Identification and characterization of poorly differentiated invasive carcinomas in a mouse model of pancreatic neuroendocrine tumorigenesis. PloS one. 2013;8:e64472.
40. Jorgensen MC, Ahnfelt-Ronne J, Hald J, Madsen OD, Serup P, Hecksher-Sorensen J. An illustrated review of early pancreas development in the mouse. Endocrine reviews. 2007;28:685-705.
41. Benitez CM, Goodyer WR, Kim SK. Deconstructing pancreas developmental biology. Cold Spring Harbor perspectives in biology. 2012;4.
42. Shyh-Chang N, Daley GQ. Lin28: primal regulator of growth and metabolism in stem cells. Cell stem cell. 2013;12:395-406.
43. Hang Y, Stein R. MafA and MafB activity in pancreatic beta cells. Trends in endocrinology and metabolism: TEM. 2011;22:364-73.
44. Alpert S, Hanahan D, Teitelman G. Hybrid insulin genes reveal a developmental lineage for pancreatic endocrine cells and imply a relationship with neurons. Cell. 1988;53:295-308.
45. Smukler SR, Arntfield ME, Razavi R, Bikopoulos G, Karpowicz P, Seaberg R, et al. The adult mouse and human pancreas contain rare multipotent stem cells that express insulin. Cell stem cell. 2011;8:281-93.
46. Quintens R, Hendrickx N, Lemaire K, Schuit F. Why expression of some genes is disallowed in beta-cells. Biochemical Society transactions. 2008;36:300-5.

47. Jensen MV, Joseph JW, Ronnebaum SM, Burgess SC, Sherry AD, Newgard CB. Metabolic cycling in control of glucose-stimulated insulin secretion. American journal of physiology. 2008;295:E1287-97.
48. Lu D, Mulder H, Zhao P, Burgess SC, Jensen MV, Kamzolova S, et al. 13C NMR isotopomer analysis reveals a connection between pyruvate cycling and glucose-stimulated insulin secretion (GSIS). Proceedings of the National Academy of Sciences of the United States of America. 2002;99:2708-13.
49. Pongratz RL, Kibbey RG, Shulman GI, Cline GW. Cytosolic and mitochondrial malic enzyme isoforms differentially control insulin secretion. The Journal of biological chemistry. 2007;282:200-7.
50. Efrat S, Linde S, Kofod H, Spector D, Delannoy M, Grant S, et al. Beta-cell lines derived from transgenic mice expressing a hybrid insulin gene-oncogene. Proceedings of the National Academy of Sciences of the United States of America. 1988;85:9037-41.
51. Yuan M, Breitkopf SB, Yang X, Asara JM. A positive/negative ion-switching, targeted mass spectrometry-based metabolomics platform for bodily fluids, cells, and fresh and fixed tissue. Nature protocols. 2012;7:872-81.
52. Sugden MC, Holness MJ. The pyruvate carboxylase-pyruvate dehydrogenase axis in islet pyruvate metabolism: Going round in circles? Islets. 2011;3:302-19.
53. Oberg K, Akerstrom G, Rindi G, Jelic S, Group EGW. Neuroendocrine gastroenteropancreatic tumours: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2010;21 Suppl 5:v223-7.
54. Marion-Audibert AM, Barel C, Gouysse G, Dumortier J, Pilleul F, Pourreyron C, et al. Low microvessel density is an unfavorable histoprognostic factor in pancreatic endocrine tumors. Gastroenterology. 2003;125:1094-104.
55. Couvelard A, O'Toole D, Turley H, Leek R, Sauvanet A, Degott C, et al. Microvascular density and hypoxia-inducible factor pathway in pancreatic endocrine tumours: negative correlation of microvascular density and VEGF expression with tumour progression. British journal of cancer. 2005;92:94-101.
56. Takahashi Y, Akishima-Fukasawa Y, Kobayashi N, Sano T, Kosuge T, Nimura Y, et al. Prognostic value of tumor architecture, tumor-associated vascular characteristics, and expression of angiogenic molecules in pancreatic endocrine tumors. Clin Cancer Res. 2007;13:187-96.
57. Zhang J, Jia Z, Li Q, Wang L, Rashid A, Zhu Z, et al. Elevated expression of vascular endothelial growth factor correlates with increased angiogenesis and decreased progression-free survival among patients with low-grade neuroendocrine tumors. Cancer. 2007;109:1478-86.
58. Visvader JE. Cells of origin in cancer. Nature. 2011;469:314-22.
59. Verhaak RG, Hoadley KA, Purdom E, Wang V, Qi Y, Wilkerson MD, et al. Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. Cancer cell. 2009;17:98-110.
60. de Wilde RF, Heaphy CM, Maitra A, Meeker AK, Edil BH, Wolfgang CL, et al. Loss of ATRX or DAXX expression and concomitant acquisition of the alternative lengthening of telomeres phenotype are late events in a small subset of MEN-1 syndrome pancreatic neuroendocrine tumors. Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc. 2012;25:1033-9.

61. Parangi S, Dietrich W, Christofori G, Lander ES, Hanahan D. Tumor suppressor loci on mouse chromosomes 9 and 16 are lost at distinct stages of tumorigenesis in a transgenic model of islet cell carcinoma. Cancer Res. 1995;55:6071-6.
62. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, et al. Bioconductor: open software development for computational biology and bioinformatics. Genome biology. 2004;5:R80.
63. Bengtsson H, Simpson K, Bullard J, Hansen K. aroma.affymetrix: A generic framework in R for analyzing small to very large Affymetrix data sets in bounded memory. Tech Report #745 Department of Statistics, University of California, Berkeley, February 2008 2008.
64. Pradervand S, Weber J, Thomas J, Bueno M, Wirapati P, Lefort K, et al. Impact of normalization on miRNA microarray expression profiling. RNA (New York, NY. 2009;15:493-501.
65. Herschkowitz JI, Simin K, Weigman VJ, Mikaelian I, Usary J, Hu Z, et al. Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. Genome biology. 2007;8:R76.
66. Sadanandam A, Wang X, de Sousa EMF, Gray JW, Vermeulen L, Hanahan D, et al. Reconciliation of classification systems defining molecular subtypes of colorectal cancer: Interrelationships and clinical implications. Cell cycle. 2014;13.
67. Sadanandam A, Varney ML, Singh S, Ashour AE, Moniaux N, Deb S, et al. High gene expression of semaphorin 5A in pancreatic cancer is associated with tumor growth, invasion and metastasis. International journal of cancer. 2010;127:1373-83.
68. Sadanandam A, Sidhu SS, Wullschleger S, Singh S, Varney ML, Yang CS, et al. Secreted semaphorin 5A suppressed pancreatic tumour burden but increased metastasis and endothelial cell proliferation. British journal of cancer. 2012.
69. Amato E, Molin MD, Mafficini A, Yu J, Malleo G, Rusev B, et al. Targeted next-generation sequencing of cancer genes dissects the molecular profiles of intraductal papillary neoplasms of the pancreas. The Journal of pathology. 2014.

Figure Legends.

Figure 1. microRNA (miR) expression subtypes in human PanNET and their similarity to mouse MLP PanNET. Heatmaps of human PanNET samples showing **(A.)** three miR expression subtypes as defined using NMF and **(B.)** the expression of previously identified RT2 mouse MLP-specific miRs (from Olson, *et al*. (28)) in the three human PanNET

subtypes defined in **A**. In both cases the columns represent individual human tumors. In **A.** the rows indicate 30 differentially expressed miRs from 40 samples. In the rainbow bar below the heatmap, red indicates elevated expression, blue decreased and white no change. (**Supplemental Table 1a** also lists the miRs in the same order as the columns in part (**A.**)). The miRs shown in (**A.**) and **Supplementary Table 1a** constitute a signature, designated the 'PanNETassigner-miR signature' (genes were selected using SAM analysis with a delta value Δ=1.3, and a false discovery rate FDR<0.05).

Figure 2. Gene expression subtypes and cross-species analysis of human and mouse PanNETs. (A.) Heatmap of human PanNET samples showing subtypes defined by gene expression differences, and the association of these subtypes with distant metastasis and clinical syndrome. The subtypes are defined by sets of genes – signatures – that are listed in the same order in **Supplementary Table 1c,** and labeled as **'**PanNETassigner-mRNA signatures' (genes were selected using SAM analysis with Δ=11.1 and FDR<0.05). ACTH represents ectopic adrenocorticotropic hormone-secreting tumor. **(B.)** The heatmap shows genes (listed in **Supplementary Table 1d**) that were commonly expressed/repressed in human MLP-1 and MLP-2 subtypes, and inversely expressed in tumors of the other two subtypes (genes were selected using SAM analysis with Δ=1.7 and FDR<0.05). **(C.)** Enrichment of miR subtype samples in gene expression subtype samples was assessed using the hypergeometric test (red represents a significant enrichment of samples between miR and gene expression subtypes with FDR<0.05). **(D.)** A bar-plot displays the total number of miR subtype samples in each gene expression subtype. **(E-F.)** Heatmaps of RT2 mouse PanNET samples showing gene expression profile differences. (**E.**) The distinct stages of tumorigenesis in the RT2 mouse model are shown, including IT, MLP and liver metastasis

(mets) samples. Genes (908 genes selected with SD>0.8, and highest variable probe across samples in the cases where multiple probes per gene were present) and samples are listed in the same order in **Supplementary Table 1f.** (**F.**) mRNAs that clearly delineate the two subtypes - IT and MLP - are shown. The rows indicate 394 differentially expressed genes (genes were selected using SAM analysis with Δ=0.6 and FDR<0.05) amongst the profiled genes of 10 samples. These genes define the "m-PanNETassigner-mRNA" signature, with the individual genes listed in the same order in **Supplementary Table 1g**. **(G.)** Heatmap showing cross-species analysis and association of human and mouse PanNET subtypes using ComBat-merged gene expression profiles of distinct stages of RT2 mouse (n=10) and patient (n=72) PanNET samples. In all the heatmaps, the rows indicate differentially expressed genes (**Supplementary Table 1h**) amongst the profiled gene expression data. In the rainbow bar beneath the heatmaps, red indicates elevated expression, blue decreased expression, and white no change.

Figure 3. Mutation changes and biomarker expression in human and mouse PanNET subtypes. (A.) A summary of *MEN1*, *DAXX/ATRX*, mTOR pathway (*PTEN/TSC2*) and *ATM* gene mutations in human PanNET subtypes. *P* values correspond to Fisher's exact test. 'ns' represents not significant. **(B.)** Visualization of different histological subtypes of RT2 tumors with H&E staining, and immunostaining for the driving oncoprotein (SV40 T-antigen) and insulin. Images are representative of the analysis from 60 tumor-bearing RT2 B6, B6AF1 and B6D2F1 mice (see Methods section for different strains of the RT2 mouse model). Scale bar represents 200 μm. **(C.)** Heatmap showing median expression of MLP-specific genes (biomarkers) in distinct stages of the RT2 mouse model of PanNET. N – normal islet, H – hyperplastic/dysplastic islet and A – angiogenic islet. **(D-E.)** Expression of PanNET subtype

genes in 5 different LCM-derived samples from tissue sections of tumors representing IT, Ins-hi MLP and Ins-lo MLP subtypes using (**D.**) RT-PCR (*Ins1* and *Enpp2*) and (**E.**) microarray (multiple PanNET subtype genes) methodologies. In (**D.**) * represents statistical p value < 0.05 as measured using Student's t-test comparing relative *Ins1* or *Enpp2* expression in IT vs. Ins-hi and Ins-lo MLP tumors. **(F.)** Visualization of different histological subtypes of RT2 tumors within a single pancreas by immunostaining for Enpp2 and insulin protein expression. Images are representative of the analysis from 8 tumor-bearing RT2 mice. Scale bar represents 50 μm. **(G.)** Enpp2 and insulin protein expression in human PanNETs as assessed using IHC. DAPI (4',6-diamidino-2-phenylindole) staining represents cellular nucleus. Scale bar represents 50 μm.

Figure 4. Vascular and proliferative characteristics of human and mouse PanNETs. (A.) CD31 (top panel) and Ki67 (bottom panel) staining in RT2 tumor subtypes. Images are representative of the analysis of 60 tumor-bearing RT2 B6, B6AF1 and B6D2F1 mice (see Methods section for different strains of the RT2 mouse model). In top panel, blue represents CD31, pink represents nucleus and scale bar represents 200 μm and in bottom panel, red represents Ki67, blue represents nucleus and scale bar represents 100 μm. **(B-C.)** Quantification of CD31 positive blood vessels (**B.**) and Ki67 positive tumor cells (**C.**) in different subtypes of RT2 tumor subtypes. The quantifications are representative of the analysis of 10 random fields from 10 different tumor-bearing RT2 B6, B6AF1 and B6D2F1 mice from the analysis illustrated in (A.) **(D.)** Percentage of Ki67 positive cells in human PanNET subtypes (from the samples of the core clinical gene expression dataset; see Methods section for detailed description about quantitation). * represents statistical p value < 0.05 as measured using Student's t-test comparing IT vs. Ins-hi MLP and IT vs. Ins-lo MLP.

Error bars in the graphs represent SD. **(E.)** Identifiable percentage and summary of NET grades in human PanNET subtypes. NET grades are significantly ($p<0.05$; Fisher's Exact test) associated with PanNET subtypes.

Figure 5. Stem cell and differentiated phenotypes of PanNET subtypes. (A-B.) Heatmaps showing differential expression of genes that distinguish islet progenitor cells from differentiated ß cells, as well as metabolic pathway genes (including metabolic genes that are disallowed in mature ß cells (46)) showing differential expression, in PanNET samples from **(A.)** human and **(B.)** mouse. Human PanNET data is from gene expression microarrays whereas mouse PanNET data is from RT-PCR analysis. The relative (to house-keeping gene PPIA) gene expression from mouse PanNET RT-PCR data were multiplied by 10,000 units before performing the heatmap analysis. Gray color in the heatmap of (B.) represents not available values. **(C.)** Heatmap comparing the ComBat-merged gene expression profile datasets of IT and MLP subtypes from the RT2 mouse model vs. embryonic pancreas (from stages e12.5 to e16.5; data from GSE8070). **(D.)** Heatmap showing differential expression of genes (measured using microarrays) distinguishing islet progenitor cells from differentiated ß cells, as well as variably expressed metabolic pathway genes, in mouse LCM-derived tissue samples from IT, Ins-hi MLP, and Ins-lo MLP PanNET tumor sections, from **Figure**s **3D-E**. **(E.)** Hyperplastic/dysplastic islets in pancreas tissue sections from RT2 mice showing regions of Ins-lo (top) and Ins-hi (bottom) islets, as visualized with H&E and insulin staining. Images are representative of the analysis from 12 tumor-bearing RT2 mice (8-10 weeks of age). Scale bar represents 50 μm.

Figure 6. Metabolic characteristics of PanNET subtypes. (A.) Heatmap showing median expression of metabolic genes that are differentially expressed in human and mouse IT vs. MLP subtypes. The comparisons involved 5 IT and 5 MLP tumors from the mouse model, and 16 IT and 31 MLP from humans. **(B.)** A simplified model of glycolysis and the tricarboxylic acid cycle, overlaid with the pyruvate cycle (arrows in green), highlighting genes upregulated (in red) and downregulated (in blue) in the IT subtype compared to the MLP subtype. **(C-D.)** Protein expression in IT and Ins-lo MLP determined by immunoblot for Hk1, Pc and insulin (**C**.) and IHC for Hk1 and insulin (**D**.). DAPI staining represents cellular nucleus. Scale bar represents 50 μm in (D.). **(E.)** Expression of MLP-specific genes in mouse PanNET cancer cell lines (ßTC) representing different subtypes of PanNET. The gene expression is relative to the expression of a housekeeping gene, *Rpl13a* (top panel). Immunoblot analysis of ßTC cells using antibodies against Pc, Hk1 and ß-actin (bottom panel). **(F-G.)** Relative abundance of pyruvate cycle metabolites in mouse PanNET subtype-specific cell lines cultured under normal conditions (**F.**) or in the presence of (**G.**) ^{13}C-labelled glucose. Metabolites in (**F.**) are median-centered across samples for each metabolite and this was done separately for two different (batches of) metabolic profiling experiments either using ßTC1b (MLP)/ßTC4 (IT) or ßTC1e (MLP)/ßTC3 (IT); labeled metabolites in (**G.**) are normalized to their characteristically lower level in the MLP subtype. The M+3 isotopomer of Asp, Mal, Fum, Cit and Iso are biosynthesized from glucose by way of Pc. The M+5 species of Cit and Iso are M+3 labeled by Pc and further enriched by glucose-derived Ac-CoA (M+2) through pyruvate dehydrogenase. **(H.)** Proliferation of IT vs. MLP cancer cell lines in the absence of glucose in the media. * (comparison between ßTC1e and ßTC3) or ** (comparison between ßTC1b and ßTC4) represents statistical p value < 0.05 as measured using Student's t-test. Error bars in the graphs represent SD or standard error of the mean.

Figure 7. A new view of PanNET tumor heterogeneity. **(A.)** Summary of human PanNET subtypes and their characteristics. **(B.)** A revised schematic of the parallel pathways of PanNET tumorigenesis, evidently with distinctive cells-of-origin, leading to the IT and MLP subtypes as revealed in the RT2 mouse model, and inferred in the human from the concordance of transcriptomic and phenotypic data presented for mouse and human PanNET.

Figure 1

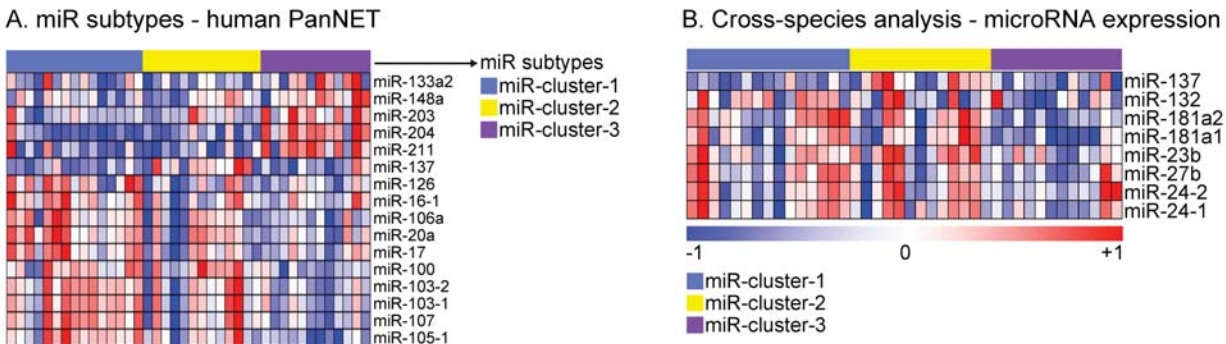


Figure 2

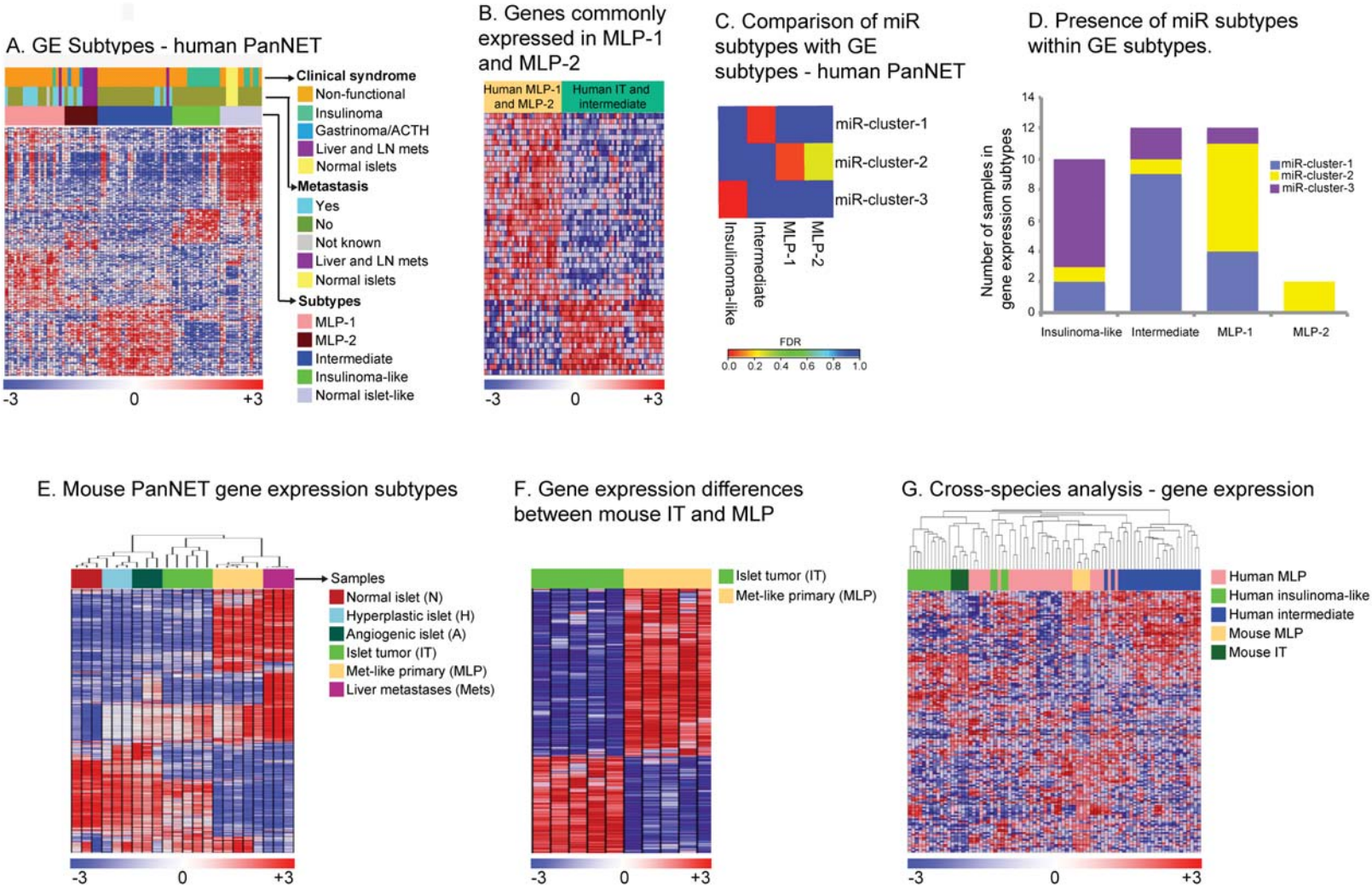
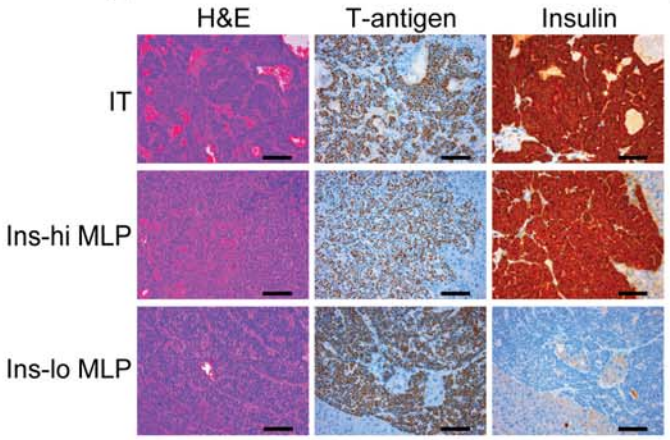


Figure 3

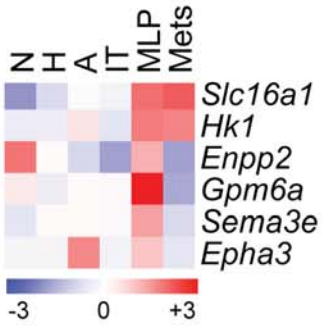
A. Association of TSG mutations with molecular subtype in human PanNET tumors

	Intermediate (%)	MLP (%)	Insulinoma-like (%)	p-value
MEN1	50	12	12.5	<0.05
DAXX/ATRX	42	28	0	<0.05
PTEN/TSC2	4	24	19	ns
ATM	4	16	25	ns

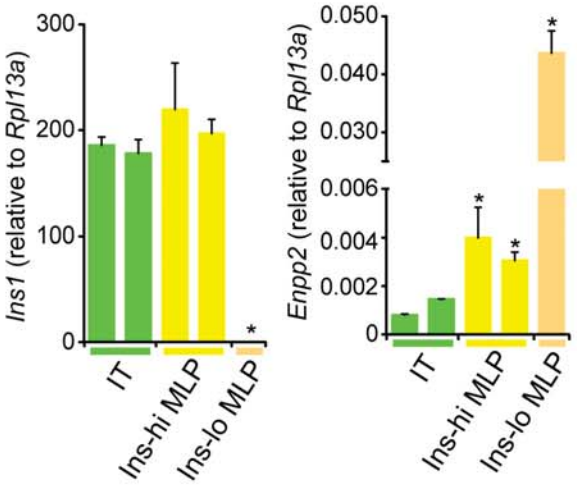
B. Subtypes of mouse PanNET and their histology



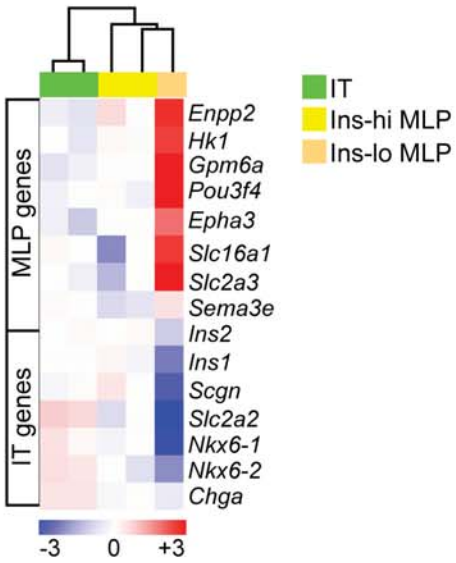
C. MLP-specific genes in mouse tumors



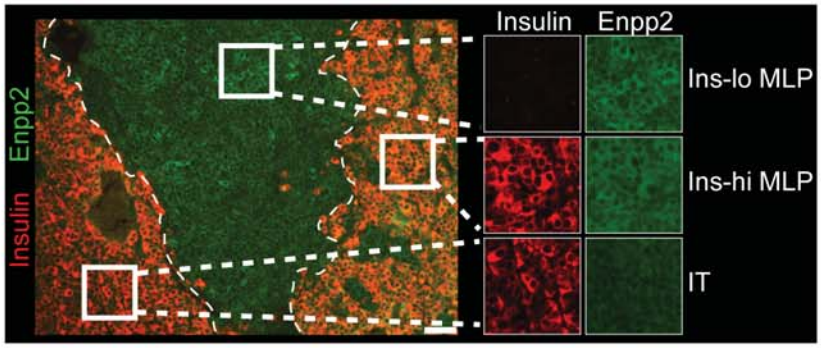
D. Gene expression analysis of LCMD mouse tumors



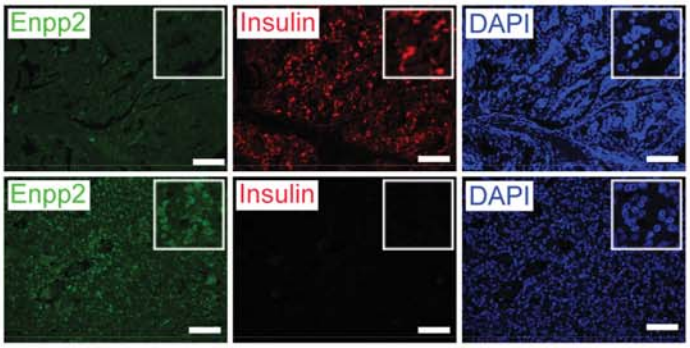
E. Microarray analysis of LCMD mouse tumors



F. Enpp2 (Autotaxin) as a maker for MLP



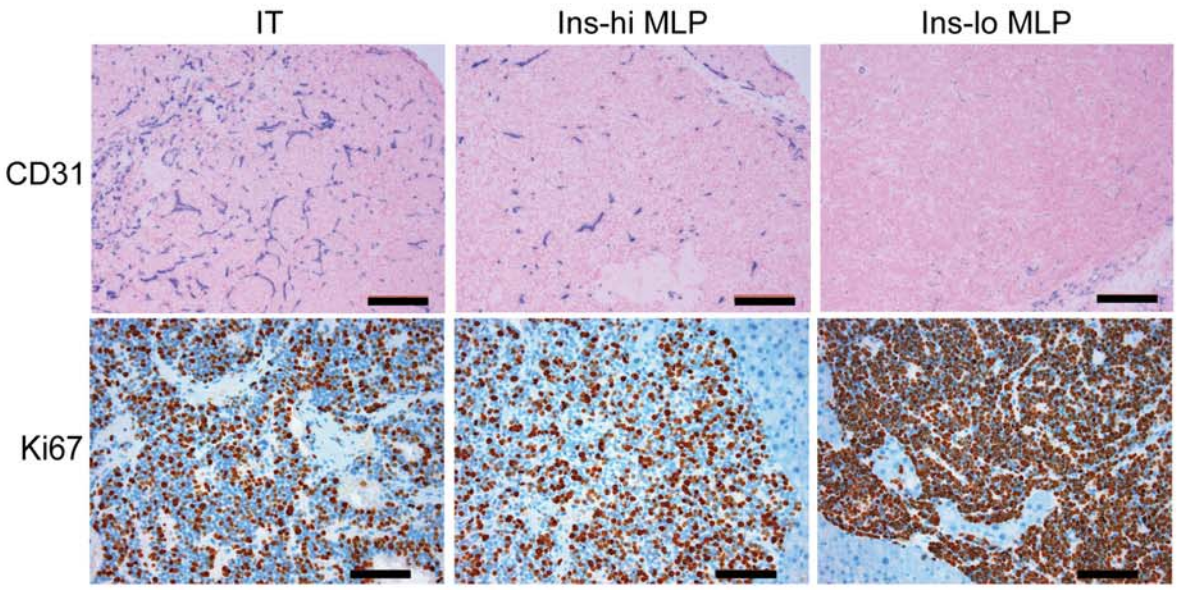
G. Enpp2 expression in human PanNET



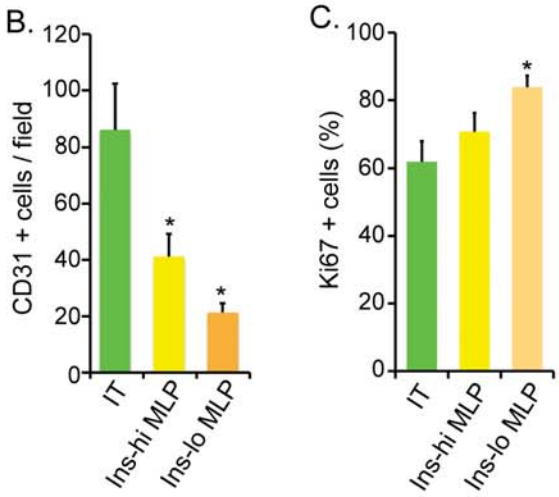
Author Manuscript Published OnlineFirst on October 7, 2015; DOI: 10.1158/2159-8290.CD-15-0068
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

Figure 4

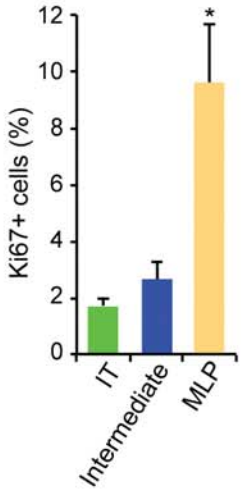
A. CD31 staining in subtypes of mouse tumors



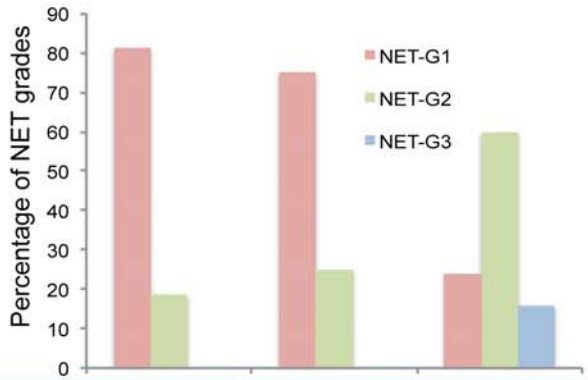
B. - C. Quantification of CD31 and Ki67 staining in subtypes of mouse tumors



D. Ki67 staining in human PanNET subtype tumors

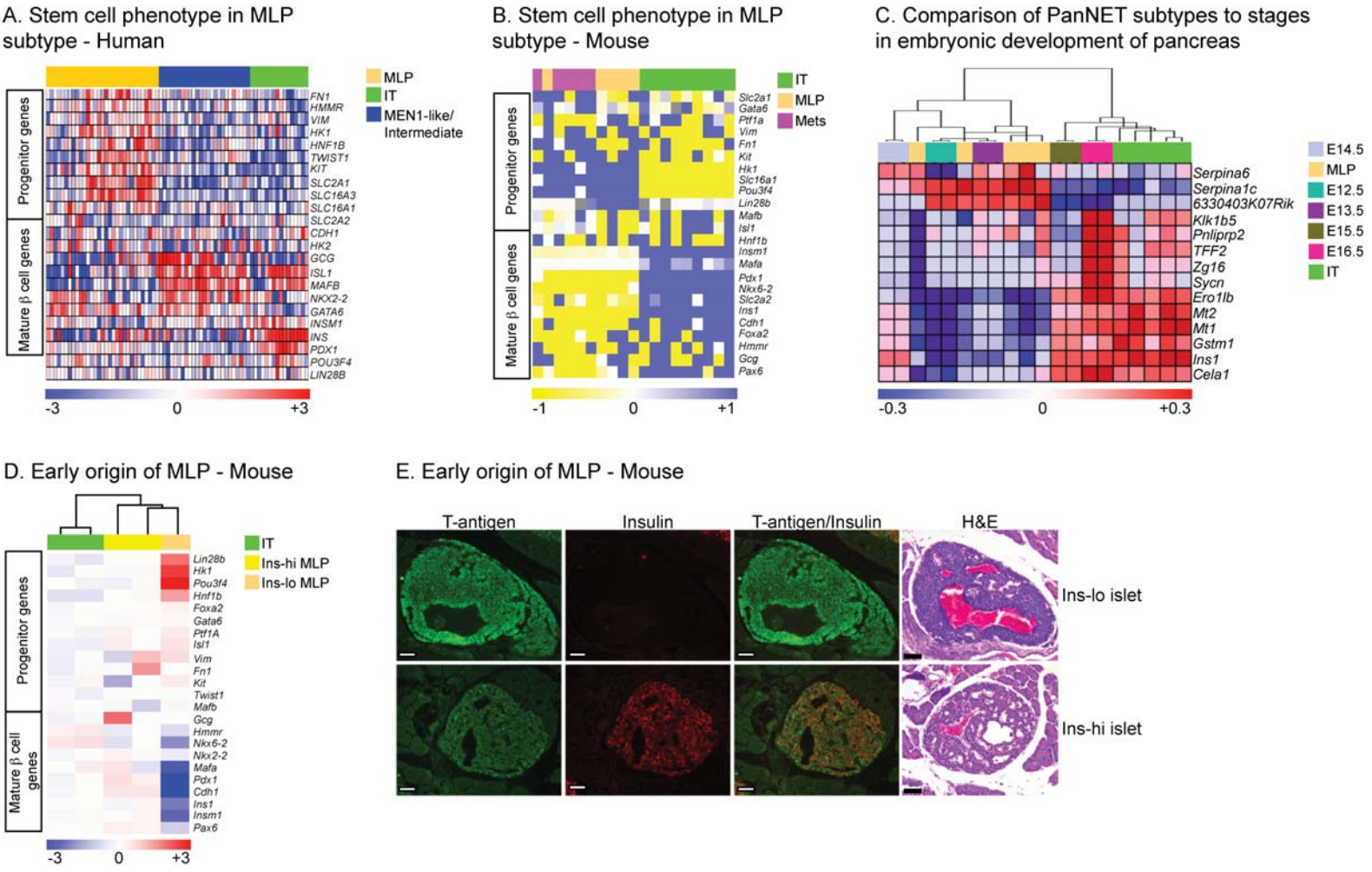


E. Association of grades with subtypes



Subtype/Grade	IT	Intermediate	MLP
NET-G1	13	18	6
NET-G2	3	6	15
NET-G3	0	0	4

Figure 5



Author Manuscript Published OnlineFirst on October 7, 2015; DOI: 10.1158/2159-8290.CD-15-0068
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

Figure 6

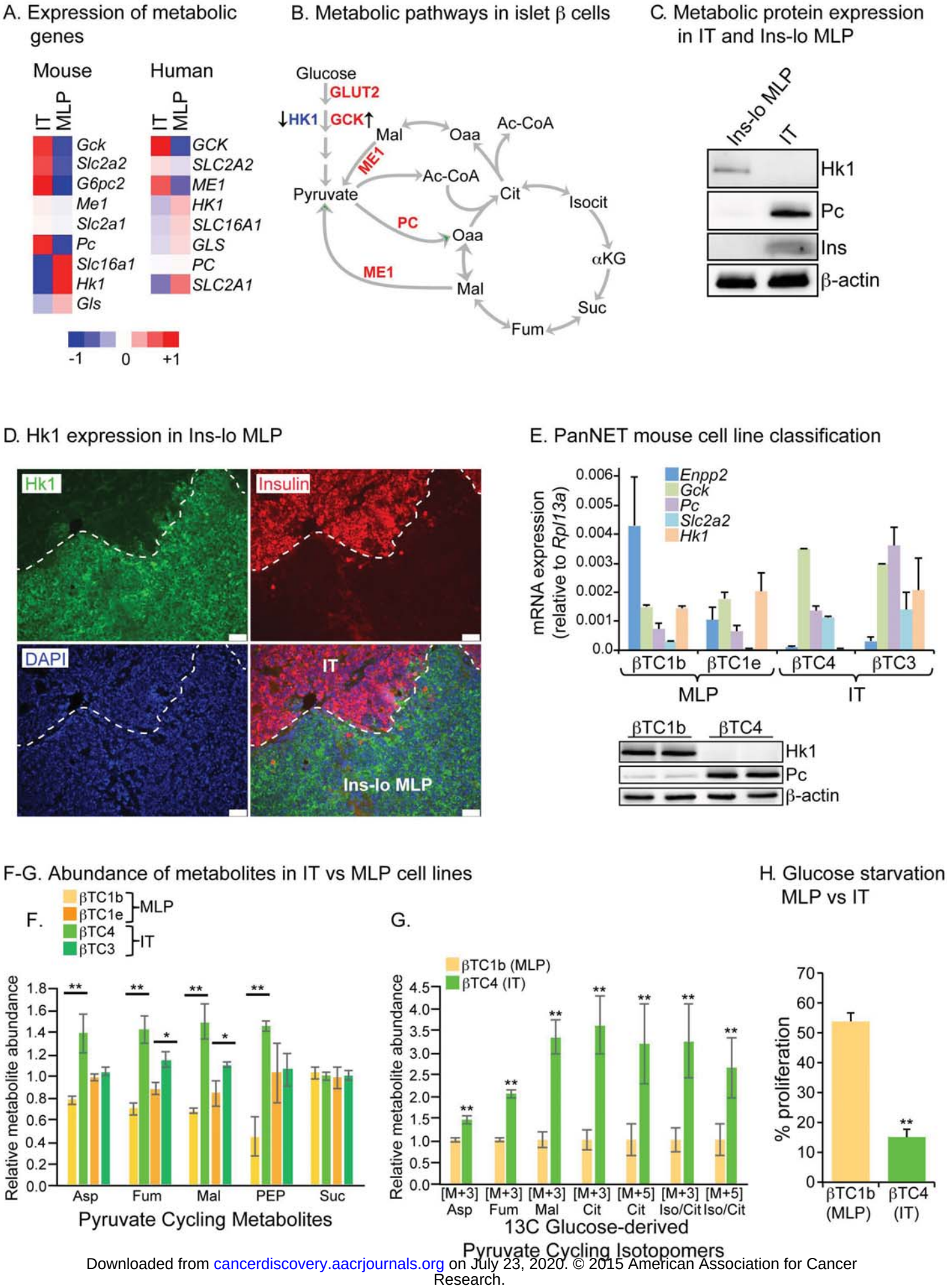
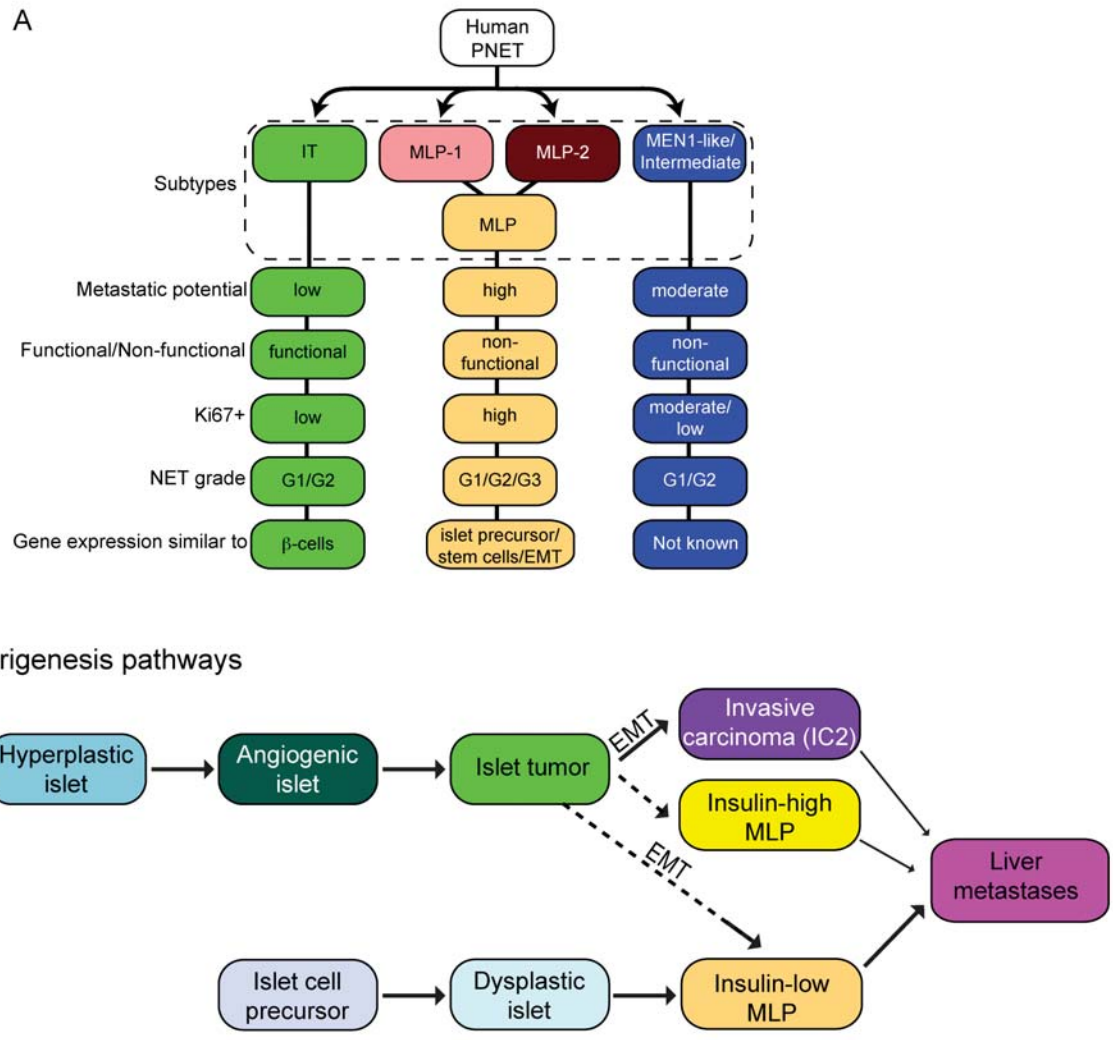


Figure 7



AACR American Association for Cancer Research

CANCER DISCOVERY

A cross-species analysis in pancreatic neuroendocrine tumors reveals molecular subtypes with distinctive clinical, metastatic, developmental, and metabolic characteristics

Anguraj Sadanandam, Stephan Wullschleger, Costas A. Lyssiotis, et al.

Cancer Discov Published OnlineFirst October 7, 2015.