Proteotranscriptomic classification and characterization of pancreatic neuroendocrine neoplasms

KC Yang[1,2], S Kalloger[3,4], J Aird[3], M Lee[5], C Rushton[2], SE Spencer Miko[1], KL Mungall[1], AJ Mungall[1], J Xu[1,2], J Karasinska[4], S Colborne[1], RD Morin[1,2], JM Loree[5], MA Marra[1,6], DJ Renouf[4,5], GB Morin[1,6], DF Schaeffer[3,4], SM Gorski[1,2,7]

[1]Canada's Michael Smith Genome Sciences Centre, BC Cancer, Vancouver, BC, Canada. [2]Department of Molecular Biology & Biochemistry, Simon Fraser University, Burnaby, BC, Canada. [3]Division of Anatomical Pathology, Vancouver General Hospital, Vancouver, BC, Canada. [4]Pancreas Centre BC, Vancouver, BC, Canada. [5]Division of Medical Oncology, BC Cancer, Vancouver, Canada. [6]Department of Medical Genetics, University of British Columbia, Vancouver, BC, Canada. [7]Centre for Cell Biology, Development, and Disease, Simon Fraser University, Burnaby, BC, Canada.

Background

Pancreatic neuroendocrine neoplasms (PNENs) are biologically and clinically heterogeneous neoplasms with variable patient outcomes. Little is known about the molecular differences between PNENs and the biological significance of such divergence. Our study aims to uncover the molecular factors that underlie the clinical and biological heterogeneity among PNENs for a better understanding and potential classification of this disease.

Experimental Approach

We used a multi-omics approach to profile and characterize the molecular landscapes of primary PNEN specimens. Formalin-fixed paraffin-embedded primary tumour specimens from 84 patients with PNEN were procured for the study and split into two cohorts for discovery (DISC) and validation (VALI) purposes. Next-generation sequencing profiled the exome and transcriptome, and quantitative mass-spectrometry profiled the global proteome of specimens. Non-negative matrix factorization was used to identify subgroups, followed by differential analysis to identify subgroup-specific features. Potential associations between the identified subgroups and known clinicopathological characteristics were evaluated.

Four Proteotranscriptomic PNEN Subgroups

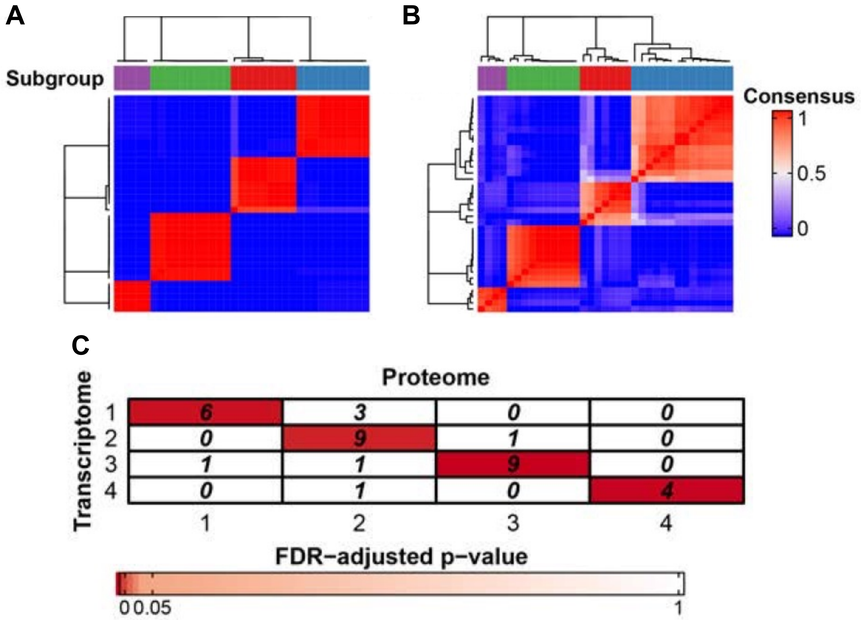

Figure 1. Unsupervised clustering analysis of transcriptome and proteome data identifies four distinct PNEN subgroups. Unsupervised clustering analysis using consensus non-negative matrix factorization was performed using the top 25% variably expressed mRNAs (A) or variably abundant proteins (B) from the DISC specimens and suggested an optimal rank of 4 based on high cophenetic and silhouette coefficients. Comparison between the mRNA- and protein- based clustering results showed significant overlap in the subgroup assignments of the DISC specimens (C). Analysis of the VALI specimens was performed using the same workflow and further supported the existence of the four subgroups (not shown)

Clinicopathological Associations

Characteristic	PDX1-high	Stromal/ Mesenchymal	Alpha cell-like	Proliferative	p-value
Sex					
Female	13	12	17	6	0.69
Male	7	13	11	5	
Functional					
Functional	5	1	3	1	0.47
Nonfunctional	10	11	11	7	
All-cause mortality					
Censored	17	19	22	4	3.3×10^{-2}
Deceased	3	6	6	7	
Histological Differentiation					
Well-differentiated	20	24	28	6	
Poorly-differentiated	0	1	0	5	5.2×10^{-5}
Ki67					
< 3%	14	12	14	1	
3 – 20%	6	12	14	5	4.7×10^{-4}
> 20%	0	1	0	5	
All-time Metastases					
No	10	15	13	4	0.59
Yes	9	10	15	7	

p-values were obtained from Fisher's exact test with Monte Carlo simulation.

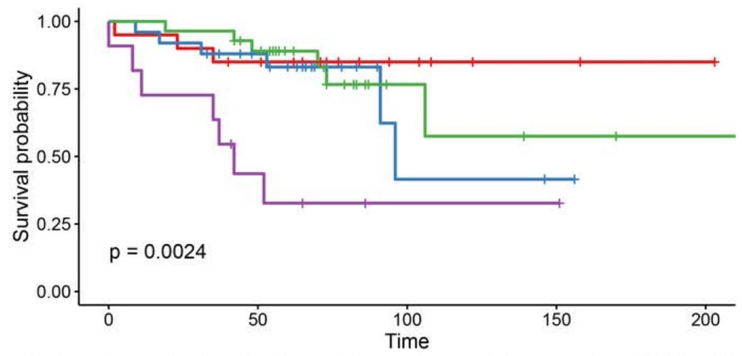

Figure 2. Overall survival probability of the patients of the complete (DISC + VALI) cohort of PNENs stratified by PNEN subgroups. Logrank test was used to evaluate the statistical significance of the differences in survival probability.

Cellular Differences between Subgroups

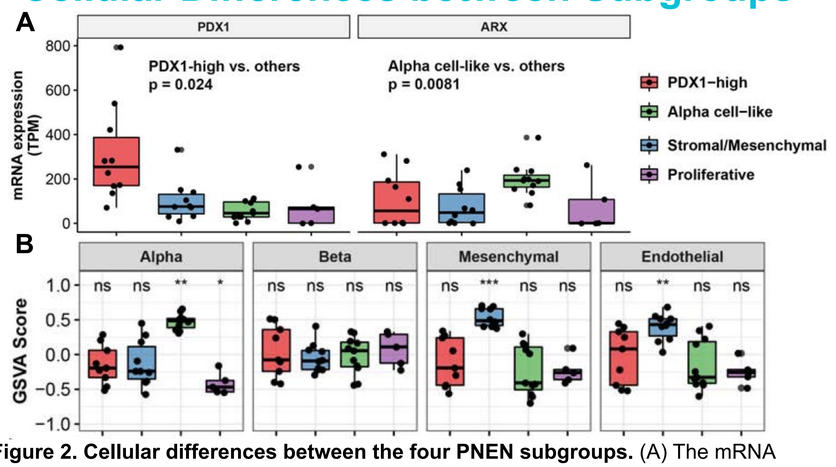

Figure 2. Cellular differences between the four PNEN subgroups. (A) The mRNA expression of PDX1 and ARX between the four subgroups. The FDR-adjusted p-value from differential expression analysis is indicated for each gene. (B) Transcriptomic similarity of each DISC specimens to pancreatic alpha, beta, mesenchymal and endothelial cells based on gene set variation analysis. Statistical significance was computed using Wilcoxon test between each subgroup and the entire cohort.

Results and Significance

We identified four PNEN subgroups with transcriptomic and proteomic differences:

- **A Proliferative subgroup associated with reduced OS and specimens of high Ki67 index or poorly-differentiated histology.**
- **A PDX1-high subgroup with high PDX1 expression.**
- **An Alpha cell-like subgroup with high ARX expression and transcriptomic resemblance to pancreatic alpha cells.**
- **A Stromal/Mesenchymal with transcriptomic resemblance to mesenchymal and endothelial cells.**

Biological Differences between Subgroups

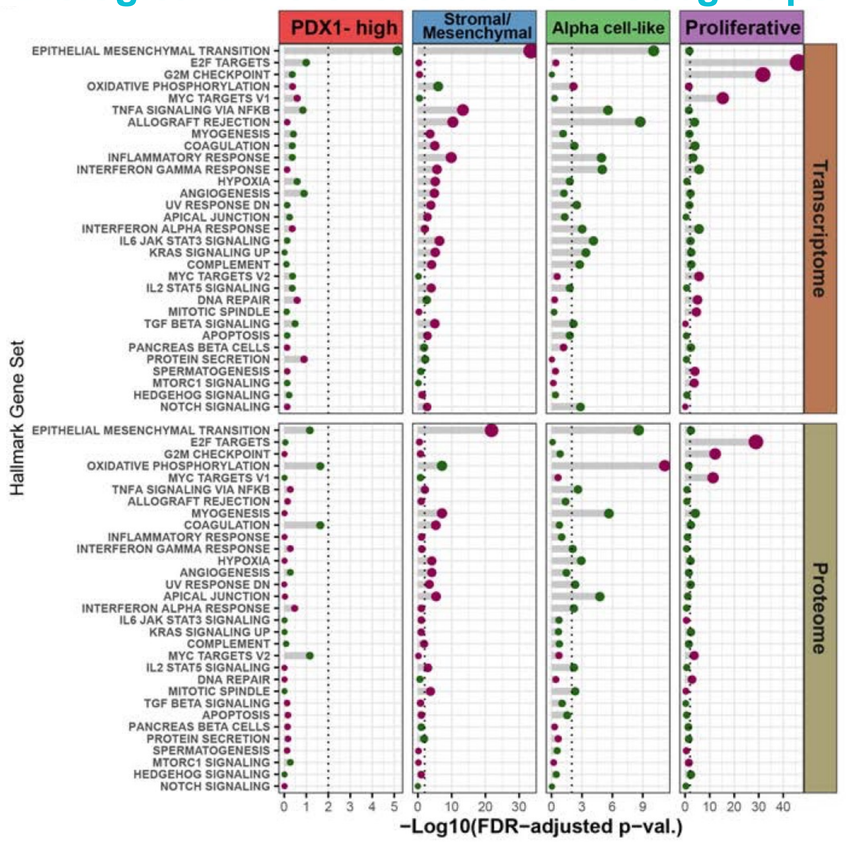

Figure 3. Biological differences between the four PNEN subgroups. Gene set enrichment analysis against MSigDB Hallmark gene sets was performed independently using mRNAs or proteins. Each dot describes the enrichment (magenta) or de-enrichment (green) of each Hallmark gene set (rows) in each of the four PNEN subgroups (columns). Dotted line indicates significance threshold of 0.01. The size of each dot and the length of its trailing grey bar is proportional to the significance of the (de-)enrichment.

Acknowledgements

Research supported by the 2017 Neuroendocrine Tumor Research Foundation – AACR Grant, Grant Number 17-60-33-GORS.