Transcriptomic and Spliceosomic Landscapes of Pancreatic Neuroendocrine Tumors Generated through Oxford Nanopore Technology Sequencing

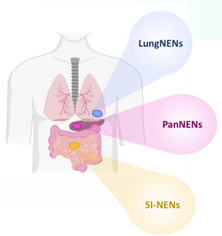
ENETS EUROPEAN NEUROENDOCRINE TUMOR SOCIETY
21ST ANNUAL ENETS CONFERENCE
13 - 15 MARCH 2024, VIENNA, AUSTRIA

b52momom@uco.es @HYCLab

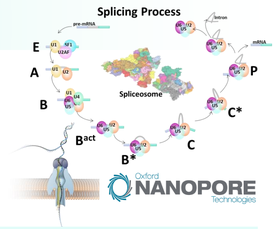
María Trinidad Moreno-Montilla[1,2,3]; Garan Jones[4,6]; Aaron Jeffries[4,6]; Ricardo Blázquez-Encinas[1,2,3], Rosie Bamford[4,6]; Víctor García-Vioque[1,2,3], Sergio Pedraza-Arévalo[1,2,3], Alejandro Ibáñez-Costa[1,2,3]; Chrissie Thirlwell[4,5,6]; Justo P. Castaño[1,2,3,7]

[1]Maimonides Biomedical Research Institute of Cordoba (IMIBIC), Cordoba, Spain. [2]Department of Cell Biology, Physiology and Immunology, University of Cordoba, Cordoba, Spain. [3]Reina Sofia University Hospital, Cordoba, Spain. [4]University of Exeter Medical School, St Luke's Campus, University of Exeter, Exeter, UK. [5]UCL Cancer Institute, London, UK. [6]University of Exeter College of Medicine and Health, University of Exeter, Exeter, UK. [7]CIBER Physiopathology of Obesity and Nutrition (CIBERobn).

BACKGROUND



Neuroendocrine tumors (NETs) are a rare and heterogeneous group of malignancies, arising from cells of the diffuse endocrine system, mainly from gastrointestinal tract, pulmonary and pancreatic tissues, whose incidence is strongly increased during the last four decades. Nevertheless, its molecular pathogenesis and development persist incompletely understood. While the genetic landscape of some NETs types has been reported, their transcriptomic knowledge remains clearly insufficient. Increasing evidence indicates that, in cancer, pathogenic transcriptomic changes often derive from alterations in the splicing process, which can generate abnormal isoforms with oncogenic potential. This information in NETs is scarce, particularly in Pancreatic NETs (PanNETs). Long-read Oxford Nanopore Technology (ONT) has emerged as the greatest sequencing tool to deeply study transcriptomics due to the length of the sequences obtained.



HYPOTHESIS

We propose that employing ONT for transcriptomic analysis of NETs will uncover novel splicing alterations and abnormal isoforms pivotal to the development and pathogenesis of these tumors

OBJECTIVES

- To sequence for the first time PanNETs using ONT sequencing
- To explore the transcriptomic and spliceosomic landscape of these tumors
- To identify novel markers and actionable targets for NETs

METHODS

Human Samples

- 3 high-quality PAN-NET samples
- 5 medium quality PAN-NET samples
- 3 pancreatic controls

RNA extraction and quality control

AllPrep® DNA/RNA/Protein Mini Kit (QIAGEN)

Qubit Fluorometer

Tape Station

Bioanalyzer (Agilent Technologies)

Library-Prep

cDNA-PCR Sequencing Kits

Sequencing

Bioanalyzer

PromethION Flow Cell

Data Analysis

Quality Analysis: EPI2ME nextflow

Transcript isoform report: pycoQC

- *minimap2*
- *pychopper*
- *Stringtie*
- *gffcompare*
- *JAFFAL*

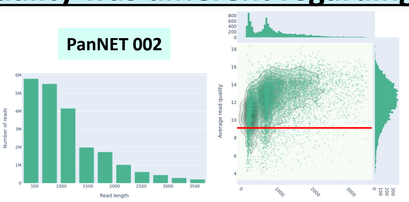
RESULTS

ONT sequencing quality was different regarding the RIN and DIN numbers

Sample ID	RIN	DIN
PanNET 002	8.4	4.4
PanNET 003	4.2	3.0
PanNET 004	4.2	4.0
PanNET 005	5.6	4.2
PanNET 006	4.7	4.6
PanNET 007	5.1	4.5
PanNET 008	4.8	3.6
PanNET 011	8.3	
Control 1	6.8	
Control 2	7.5	
Control 3	7.8	

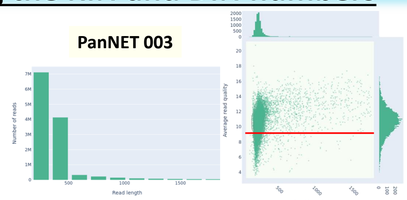
☐ High-Quality
☐ Medium-Quality

PanNET 002



Sample	Number of reads (K)	Total base (M)	Mean read length	Mean read quality
PanNET 002	22401	23384	777	12.3

PanNET 003



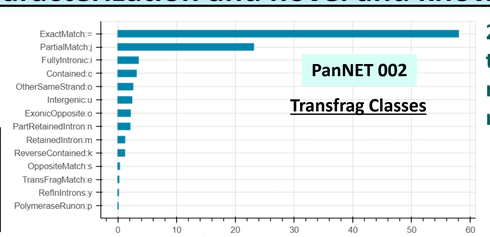
Sample	Number of reads (K)	Total base (M)	Mean read length	Mean read quality
PanNET 003	12727	4817	251	10.7

Sequencing quality was different between the samples depending on the RNA and DNA integrity. It is shown in these two examples: number, length and basecalling of reads were much higher in PanNET 002 than in the 003. The red line shows sequencing quality 9, all reads below this line must be discarded. This approach was done with all the sequenced samples.

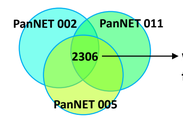
Transcriptome characterization and novel and known transcriptomic variants

PanNETs	002	005	011
Total genes	23976	27434	22523
Total transcripts	67569	61995	65599
Max trans. len	18084	15026	22742
Min trans. len	47	12	40

Controls	1	2	3
Total genes	21350	20745	24237
Total transcripts	55217	56255	68438
Max trans. len	19415	14011	19048
Min trans. len	45	53	45

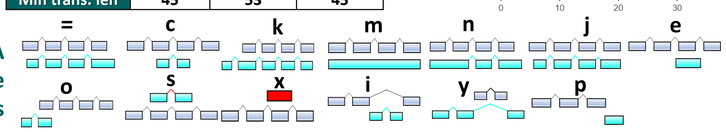


2306 transcripts were found in common between the three high-quality PanNETs, 10 of them have never been described before so these could be novel PanNETs specific transcripts.



10 genes with novel transcripts

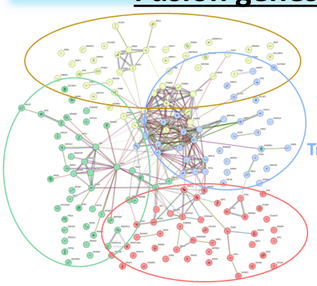
SLC23A2 (j)	*CERS5* (j)
ZNF260 (j)	*SMN1* (j)
SLC37A4 (c)	*SPART* (j)
COMMD6 (c)	*MCPH1-DT* (j)
ITFG2 (j)	*MT-TP* (x)



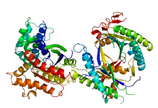
GffCompare tool permits the identification of multiple transcripts fragments (transfrag) comparing to a reference annotation. '=' exact match; 'c' at least one intron is contained;

'm' all introns retained; 'n' at least one intron retained; 'j' at least one junction included; 'e' single exon partially covering an intron; 'o' other strand overlap with an exon; 's' intron match in the opposite strand; 'x' exon overlap in the opposite strand; 'I' contained in a reference intron; 'y' contains the reference in an intron; 'p' polymerase run-on, no overlap.

Fusion genes identification through ONT



ONT allows the identification of fusion transcripts. We found 166 fusion transcripts with the gene *GNAS* in the three tumor samples that were absent in the controls.



GNAS gene: Guanine Nucleotide Binding Protein (G Protein), Alpha Stimulating Activity Polypeptide 1

CONCLUSIONS

The quality of the sequencing was tightly dependent on quality of the RNA. It is necessary to explore whether these analysis could be performed in the rest of the samples and compare the results obtained.

ONT allowed the discovery of known and novel transcripts, splicing isoforms and fusion genes that may be relevant in NETs.

ONT leverages the precision and depth of transcriptomic knowledge extracted from NETs, which can help to better understand their molecular architecture and, ultimately, explain their clinical behavior.